

Machine learning for data-driven primary prevention at population scale

Dissertation

zur Erlangung des Grades eines
Doktors der Naturwissenschaften

Dr. rer. nat.

am Fachbereich Mathematik und Informatik
der Freien Universität Berlin

vorgelegt von

Thore Manuel Buergel

Berlin 2023

Thore Manuel Buergel,

Machine learning for data-driven primary prevention at population scale

Dissertation

© Thore Manuel Buergel 2023

Erstgutachter: Prof. Dr. Roland Eils

Zweitgutachter: Prof. Dr. Tim Landgraf

Tag der Disputation: 31.03.2023

Je größer die Insel des Wissens, desto länger die Küste der Verzweiflung.

*I dedicate this work to my parents,
Siamaris Stella Salcedo de Buergel
and
Erich Kurt Buergel.*

Contents

Abstract	ix
1 Introduction	1
1.1 Motivation	2
1.2 Background	4
1.2.1 Machine learning and neural networks	4
1.2.2 Primary prevention	6
1.2.3 Benefits, harms and guidance of interventions	8
1.2.4 Fundamentals of risk modeling	11
1.2.5 Requirements for clinical risk models	14
1.2.6 Clinical predictors in primary care	16
1.3 State of the art	18
1.3.1 Risk models in primary prevention	18
1.3.2 Polygenic risk scores and integration in primary prevention	22
1.3.3 Metabolomics and disease risk	24
1.3.4 Electronic health records and disease risk	27
1.3.5 Retinal fundus photographs and disease risk	28
1.3.6 Neural networks for image analysis	30
1.3.7 Neural networks for risk modeling	32
1.4 Aim of study	34
1.5 Outline	37
2 Methodology	39
2.1 Neural networks	40
2.1.1 Multi-layer perceptron	40
2.1.2 Convolutional neural networks	42
2.1.3 Residual neural networks and ConvNeXt	44
2.1.4 Activation functions	45
2.1.5 Regularization	47
2.1.6 Normalization	49

Contents

2.2	Survival models	49
2.2.1	Kaplan-Meier Estimator	49
2.2.2	Cox Proportional Hazards model	50
2.2.3	Deep Surv	52
2.2.4	Deep Survival Machines	53
2.2.5	Model calibration	55
2.3	Image analysis	55
2.3.1	Image augmentation methods	55
2.3.2	Canny edge detection operator	57
2.4	Evaluation metrics	58
2.4.1	Concordance index	58
2.4.2	Net reclassification improvement	59
2.4.3	Net benefit	60
2.4.4	Decision curve analysis	62
2.4.5	Calibration curve analysis	63
2.4.6	Relative risk differences	63
2.5	Attribution methods	64
2.5.1	Shapley additive explanations	64
2.5.2	Information bottleneck attribution	65
2.6	Model implementation	67
2.6.1	NeuralCVD	67
2.6.2	Metabolomic state model	68
2.6.3	Medical history model	69
2.6.4	Retinal risk model	70
3	Data sources and preprocessing	73
3.1	Data sources	74
3.1.1	UK Biobank	74
3.1.2	Whitehall II	74
3.1.3	PROspective Study of Pravastatin in the Elderly at Risk	75
3.1.4	Leiden Longevity Study	75
3.1.5	Rotterdam Study	76
3.2	Data preprocessing	76
3.2.1	Extraction of covariates and clinical predictors	76
3.2.2	Calculation of PGS and the PGSMETA Score	78
3.2.3	Processing of NMR Metabolomics data	79

3.2.4	Processing of EHR data	80
3.2.5	Retinal Fundus Photographs	80
3.2.6	Dataset partitions and imputation	81
3.3	Endpoint definitions	83
3.3.1	MACE endpoint	83
3.3.2	Common disease endpoints	83
3.3.3	Phecode endpoints	85
4	A Neural Survival model integrates polygenic information for cardiovascular risk prediction	87
4.1	Experimental Setup	88
4.2	Characteristics of the study population	88
4.3	NeuralCVD outperforms established CVD-Risk models	91
4.4	Assessing the ability of NeuralCVD to integrate PGS information	93
4.5	Individual differences in predicted risk after PGS addition	94
4.6	Discussion	98
5	Metabolomic profiles predict individual multi-disease outcomes	101
5.1	Experimental Setup	102
5.2	Characteristics of the study population	104
5.3	The metabolomic state is associated with event rates and stratifies the population according to disease risk	106
5.4	The metabolomic state shares information with common clinical predictors	109
5.5	The metabolomic state adds predictive value over clinical variables	112
5.6	Discriminative performance translates into potential clinical utility	115
5.7	Identification of disease-specific metabolite profiles	118
5.8	High-risk individuals for type-2 diabetes differ in their personal metabolomic profiles	120
5.9	Discussion	123
6	Medical history predicts future health trajectories over the human phenome	127
6.1	Experimental setup	128
6.2	Characteristics of the study population	129
6.3	The medical history is associated with future disease onset	130
6.4	Medical history yields discriminative improvements over basic predictors	133
6.5	Learned medical risk states are applicable for cardiovascular prevention and public health reactions	134

Contents

6.6	Discussion	138
7	Phenome-wide prediction of disease onset from retinal fundus photographs	141
7.1	Experimental setup	142
7.2	Characteristics of the study population	143
7.3	Phenome-wide association of retinal states with future disease onset	144
7.4	Retinal information yields discriminative improvements over basic predictors	147
7.5	Extracted retinal information is similar to cardiovascular predictors	149
7.6	Image attributions reveal risk-related regions in retinal fundus photographs	151
7.7	Discussion	153
8	Conclusion	157
	Acronyms	165
	Bibliography	169
	Declaration of authorship	195
	Zusammenfassung	197
	Acknowledgements	199
	Supplementary Figures	201
1	A Neural Survival model integrates polygenic information for cardiovascular risk prediction	201
2	Metabolomic profiles predict individual multi-disease outcomes	207
	Supplementary Tables	235
1	Data and Implementation	235
2	A Neural Survival model integrates polygenic information for cardiovascular risk prediction	283
3	Metabolomic profiles predict individual multi-disease outcomes	289
4	Medical history predicts future health trajectories over the human phenome	323
5	Phenome-wide prediction of disease onset from retinal fundus photographs	415

Abstract

Healthcare costs are systematically rising, and current therapy-focused healthcare systems are not sustainable in the long run. While disease prevention is a viable instrument for reducing costs and suffering, it requires risk modeling to stratify populations, identify high-risk individuals and enable personalized interventions. In current clinical practice, however, systematic risk stratification is limited: on the one hand, for the vast majority of endpoints, no risk models exist. On the other hand, available models focus on predicting a single disease at a time, rendering predictor collection burdensome. At the same time, the density of individual patient data is constantly increasing. Especially complex data modalities, such as -omics measurements or images, may contain systemic information on future health trajectories relevant for multiple endpoints simultaneously. However, to date, this data is inaccessible for risk modeling as no dedicated methods exist to extract clinically relevant information. This study built on recent advances in machine learning to investigate the applicability of four distinct data modalities not yet leveraged for risk modeling in primary prevention. For each data modality, a neural network-based survival model was developed to extract predictive information, scrutinize performance gains over commonly collected covariates, and pinpoint potential clinical utility. Notably, the developed methodology was able to integrate polygenic risk scores for cardiovascular prevention, outperforming existing approaches and identifying benefiting subpopulations. Investigating NMR metabolomics, the developed methodology allowed the prediction of future disease onset for many common diseases at once, indicating potential applicability as a drop-in replacement for commonly collected covariates. Extending the methodology to phenome-wide risk modeling, electronic health records were found to be a general source of predictive information with high systemic relevance for thousands of endpoints. Assessing retinal fundus photographs, the developed methodology identified diseases where retinal information most impacted health trajectories. In summary, the results demonstrate the capability of neural survival models to integrate complex data modalities for multi-disease risk modeling in primary prevention and illustrate the tremendous potential of machine learning models to disrupt medical practice toward data-driven prevention at population scale.

1 Introduction

1.1 Motivation

Since antiquity, the times of Hippocrates, Galen and Bian Que, diagnosis and therapy have been the central elements of medicine. Medicine is commonly considered successful once health, the symptomless status has been restored through a structured diagnostic process followed by disease resolving therapy. Thanks to constant progress over the millenia manifested in the development of ever more precise diagnoses, and ever better treatments, many diseases that once were almost certainly fatal have become curable. However, as positive as this development is, there are strong warning signals indicating the lack of sustainability in a disease and treatment focussed system. Over the past four decades, healthcare costs have been consistently increasing among OECD countries (Joint OECD, EUROSTAT and WHO Health Accounts SHA Questionnaires (JHAQ) 2022), reaching up to 16.9 % of the GDP of the United States (11.2 % and 9.8 % of the GDPs of Germany and UK respectively) in 2018 (Tikkanen and Abrams 2020). While the rise in healthcare costs has partly been linked to population increase and the development and utilization of modern, more expensive techniques, common and age-related diseases have been identified as a major driver of healthcare expenditure (Dieleman et al. 2017). Further, despite projections expecting the increase in health expenditure to stall, its anticipated growth rate still exceeds the growth rate in the general economy within the OECD until 2030 (Lorenzoni et al. 2019). At the same time, many diseases - in the USA nearly 50 % - are considered preventable as they may be attributed to modifiable risk factors (US Burden of Disease Collaborators et al. 2018). Translating this to potential savings, in 2016 alone, the USA spent a staggering amount of \$ 730.4 billion for the treatment of diseases that would have been preventable in the first place (Bolnick et al. 2020). Further, as drastic these financial losses may seem, the economic side alone fails to portray the full picture, as the manifestation of every preventable disease inherently represents unnecessary suffering and lost years of life.

However, despite all these shortcomings, prevention is far from being the paradigm of today's healthcare system. In fact, as a consequence of the centuries-old prioritization of diagnostics and therapeutics, today's healthcare, despite all its progress is rather reactionary in the original sense of the word: The notion of healthy individuals not requiring medical attention is commonly accepted and most individuals would seek medical attention only in response to an illness (see figure 1.1 on the facing page). While preventive programs and health checks are starting to be recommended (Bundesgesundheitsministerium 2022; NHS 2022), in the majority of cases, it is only after disease onset, that the full capabilities of modern healthcare systems are leveraged. In consequence, almost the entire healthcare industry

existing today is oriented towards the end of the causal disease chain: the focus is on the disease as an abnormality to be corrected, not on health as the normality to be maintained and current healthcare delivery systems act in fee-for-service or capitated approaches reimbursing, (re-)actions and treatments, often independent of the individual health outcome (Porter and R.S. Kaplan 2016). In order for prevention to become the paradigm of health-care, significant challenges in reimbursement and care-delivery have to be overcome.

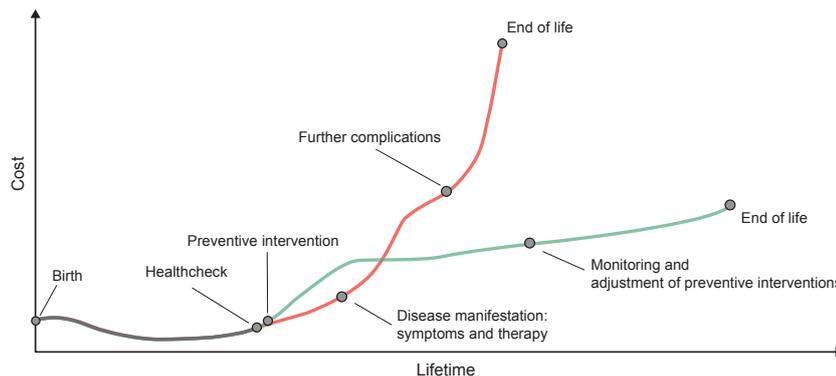


Figure 1.1: The effect of prevention. Schematic depiction of life trajectories for two individuals with preventive intervention (green) and without (red). Without intervention, individuals at risk, likely develop diseases, shortening the lifespan and causing high treatment costs. While preventive intervention causes short-term healthcare costs, it results in long-term benefits such as a longer, disease-free lifespan and lower overall healthcare costs.

An alternative approach to health care organization is described in the concept of value-based health care (Porter 2010). Value-based healthcare aims to establish the maximization of patient value as a shared objective among stakeholders, where patient value is defined as the achieved health benefit over the costs required. Thus, by monitoring individual health outcomes and carefully selecting the optimal treatment, value-based care decreases costs while promoting the overall health of populations. As disease manifestations drive costs, disease mitigation is implicitly valued higher than disease treatment. However, it is hard to reliably estimate and quantify the value of preventive interventions as they often represent long term investments into future health with negative short term financial rewards. With reliable financial value being centered around the end of the causal disease chain, vendors often lack the incentive for preventive care, even in a value-based care model. In comparison, the returns on preventive measures heavily depend on the selection of individuals being recommended for intervention. Intervening in a population of high-risk individuals with a high probability of experiencing disease onset will result in a larger portion of saved costs, compared to promoting measures in a population of low-risk individuals that likely would not have experienced a disease event regardless of the measures taken. Thus, an ac-

1 Introduction

curate selection of individuals for interventions is a decisive factor for economic returns on preventive measures. Disease risk models can inform about an individual's future health trajectory by providing probabilistic estimates for future disease onset from variables at hand. Based on these projections, the preventive potential becomes quantifiable as a function of the estimated risk, available interventions and the anticipated costs resulting from a potential disease onset.

More fundamentally, however, the very medical act of initiating preventive interventions already depends on the anticipation of an individual's future health trajectory. While harmless interventions such as life-style advice may be promoted for wider parts of the general population, targeted prevention measures involving medication or treatment, require precise knowledge on the potential costs, benefits and harms in the selected individuals. This form of holistic risk mitigation has been exemplified in cardiovascular prevention, where the use of disease risk models is recommended by medical guidelines to identify high-risk individuals and guide interventions (National Institute for health and Care Excellence (NICE) 2014). For the majority of diseases, however, the application of risk models is not part of the common clinical practice, as no means of projecting on future disease onset exist. Even for diseases where risk models have been proposed, the adoption into clinical practice is often hampered by the effort and cost to collect the required clinical predictors (Steyerberg, Moons, et al. 2013). Consequently, there is huge potential for novel risk modeling approaches to improve on the existing standard of care and disrupt clinical practice, both by allowing the predictions for so far inaccessible endpoints, and by reducing costs and efforts in predictor collection. By allowing for quantified and targeted risk mitigation, these novel risk models may form the foundation of a systematic roll-out of preventive measures and help to pave the way for a new prevention centered age of value-based care.

1.2 Background

1.2.1 Machine learning and neural networks

Most generally, machine learning describes a branch of statistics concerned with the development or application of computational models able to complete unseen tasks. Machine learning models learn and adapt based on inferred patterns in given data without following explicit instructions. Given a pre-collected dataset, machine learning algorithms may be trained by supervised or unsupervised learning. In unsupervised learning, data is unlabelled, and the algorithm is required to infer patterns relying on similarities and differences between

the individual samples. For example, given a dataset of images of mammals, without further annotation, an unsupervised learning algorithm might be applied to identify distinguishing features between the animals and regroup the photographs accordingly. In medicine, unsupervised learning has been applied to identify new subgroups of coronary artery disease (Flores et al. 2021) or new biomarkers separating healthy and disease populations (Shomorony et al. 2020). While unsupervised learning does not require labels, it is generally considered a more complicated learning problem. Therefore, most machine learning applications that require a high degree of confidence and accuracy are developed by supervised learning. Supervised learning builds on labeled data to optimize a model's parameters to minimize an objective function, commonly measuring the prediction error or goodness of fit between the label and the model's prediction. For example, supervised learning could be applied to a dataset of labeled images of cats and dogs to distinguish between the two animals accurately. In medicine, supervised learning has been utilized for many years and has led to applications including clinical risk prediction models (Goff et al. 2014; SCORE2 working group and ESC Cardiovascular risk collaboration 2021), systems for automated diagnosis and detection (Morey et al. 2020) of diseases and systems for optimization of hospital workflows (Tomašev et al. 2019).

To date, the most commonly applied machine learning method to leverage large and complex datasets are neural networks. While the fundamental concept of neural networks dates back to the 1950s (Rosenblatt 1958) and backpropagation and convolutional neural networks were already established methods in the 1990s (Lecun et al. 1998), it was the increase in available computing power that fostered the development of AlexNet, the epochal breakthrough moment, heralding the deep learning boom of the 2010s (Krizhevsky, Sutskever, and Hinton 2012). Since then, deep learning has become the method of choice for feature extraction, representation learning, and cognition across many scientific domains, from Earth science (Reichstein et al. 2019), to physics (Erdmann et al. 2021) and biology (Webb 2018). Deep learning relies on large neural networks with many layers (hence the term *deep*) to approximate a given function or minimize a specified objective. Importantly, neural networks do not require a manual specification of the features or feature combinations that the model should consider. Instead, a neural network operates by learning weights to process, combine and reweigh information in the input space to best minimize the given objective. Thus, where other machine learning algorithms require analysts to engineer features carefully, neural networks are entirely data-driven. Large deep-learning models may have millions of parameters, allowing them to approximate complex functions and extract specific features from the given inputs. Due to this large number of parameters, deep learning

1 Introduction

models require vast amounts of labeled data, and datasets might well extend to millions of samples for image recognition tasks (Deng et al. 2009; T.-Y. Lin et al. 2014).

In consequence, with the availability of increasingly large labeled datasets, deep learning became applicable for many real-world applications. Notable examples in the scientific medical domain include the detection of skin cancer in photographs (Esteva et al. 2017), the prediction of in-hospital mortality, readmission, and length of stay from large-scale electronic health records data (Rajkomar et al. 2018), and the detection of atrial fibrillation in standard 12-lead electrocardiograms (Attia et al. 2019). In summary, neural networks represent a powerful machine-learning method with a wide range of applications. Especially in prevention, the application of neural networks on complex medical data promises to disrupt the current medical practice by allowing for capable risk models to assess individual risk landscapes over many diseases systematically and simultaneously.

1.2.2 Primary prevention

Prevention, or preventive healthcare, seeks to aid individuals in maintaining their health, promote well-being and thereby avoid disease onset. Preventive actions may include measures such as immunization, screening for the identification of asymptomatic diseases or risk factors, behavioral counseling to motivate lifestyle changes as well as the use of drugs to prevent disease (Fletcher 2022). Leavell and Clark categorized preventive measures into three tiers (Leavell et al. 1953). Primary prevention aims to avoid disease onset by targeting its causal factors. According to this definition, primary prevention measures are undertaken *per se* without indication or subclinical manifestations of disease, such as the administration of vaccines for immunization, but also lifestyle interventions (such as smoking cessation), prophylactic surgery or intervention by medicines (such as antihypertensives and statins). Secondary prevention, on the other hand, is interested in acting upon subclinical and asymptomatic disease manifestations to stop disease progression. Secondary prevention first requires a screening process to detect the subclinical manifestation followed by a treatment intervening on the identified issue. Secondary prevention may, for instance, act on specific risk factors such as high blood pressure, obesity, or smoking. Ultimately, tertiary prevention measures seek to prevent the onset of further complications once an individual has experienced disease onset, for instance, the prevention of additional cardiovascular disease events or the onset of atherosclerosis in an individual with type-2 diabetes. Tertiary prevention commonly includes medical or surgical intervention and thus involves treatment.

Although the categorization into the three levels defined by Leavell et al. is common, the terminology is not definitive. In fact, the lines between the categories have become blurred, with more preventive measures entering into clinical practice (Fletcher 2022). Thus, while primary prevention, in its original definition, referred to unspecific measures such as vaccinations, it may now also include the administration of medicines. Further, in cardiovascular prevention, for instance, definitions are more straightforward: primary prevention refers to the measures administered to prevent the onset of the first cardiovascular event (the primary disease event), including those usually categorized as secondary prevention. Primary prevention in cardiology would include behavioral advice encouraging, for instance, more physical activity and the administration of statins to act on a diagnosed hypertension. Similarly, secondary prevention in cardiovascular disease refers to what Leavell and Clark categorized as tertiary prevention, namely the prevention of disease recurrence or consequential complications, for instance, a second, repeated myocardial infarction (Roe and Ohman 2012). Due to its simplicity, this work will resort to the definitions of primary and secondary prevention as utilized in cardiovascular prevention and the British National Health Service (NHS): “primary prevention means working with partners [...] to prevent disease or injury before it ever occurs.” (Alderwick and Dixon 2019).

Primary prevention is closely linked to primary care as it is commonly the main point of interaction between individuals and their healthcare providers. Primary care refers to family medicine or general practices offering non-specialist healthcare services from professionals, called General Practitioner (GP), Family Physician (FP), or Primary Care Provider (PCP). In the primary care setting, day-to-day healthcare, such as treatment of acute diseases as well as public health measures like vaccinations, are provided, and specialist care often requires a referral and previous examination from PCP. Due to this tight contact between individuals and their PCP, many primary prevention programs, even those involving specialists such as intervention on blood lipids in cardiovascular prevention, are initialized at primary care sites (National Institute for health and Care Excellence (NICE) 2014). Depending on the prevention strategy, prevention at the primary care site may include an examination by the PCP and a prescription of laboratory tests to assess disease risk factors. In many countries, national health strategies endorse risk factor assessment through the promotion of regular health check-ups (Bundesgesundheitsministerium 2022; NHS 2022).

It is important to note, however, that guideline recommendations and the practice of clinically adopted prevention strategies are very heterogeneous between healthcare systems and health outcomes. For instance, while prevention has long been a central element of cardio-

1 Introduction

vascular medicine, and risk factor examination, as well as risk modeling, are integral parts of the cardiovascular preventive strategy, this is far from being representative of the entire disease spectrum. In fact, for the majority of diseases, there exists no preventive strategy nor means of assessing the probability of future disease onset. Thus, even after systematic risk factor collection, the spectrum of diseases for which a targeted intervention is feasible is currently still limited. At the same time, the potential of established interventions is not fully exhausted, as preventable diseases are still highly prevalent (Galea and Maani 2020) a systematic roll-out of available primary prevention strategies is needed to reduce the burden of preventable diseases.

1.2.3 Benefits, harms and guidance of interventions

In medicine, each action, such as a surgical procedure or pharmacological treatment, has benefits and harms. The benefits describe the positive effects of an intervention, thus, for instance, the reduction of risk, decrease in disease burden, or increase in well-being. Harms, on the other hand, describe the detrimental effects of the intervention itself: side effects of medications, pain experienced at the surgery, or psychological burden experienced in the course or consequence of a medical measure. Notably, the interplay between the benefits and harms of an intervention is not static but depends on the individual and their situation. For instance, in secondary prevention and regular treatment, where an individual is subject to a medical condition and has already embarked on a particular course of disease, the harms of the treatment are justified by the experienced suffering and projected detrimental course of the disease, which is to be cured. An example is prophylactic cardioverter-defibrillators, which significantly reduce the risk for mortality and improve the quality of life in patients with coronary artery disease or dilated cardiomyopathy (Theuns et al. 2010) but come with significant harms, as the implantation of the device requires surgery.

In contrast, in primary prevention, individuals lack symptoms and do not suffer. Further, disease trajectories are more uncertain, as there are, at most subclinical manifestations of disease with an often non-deterministic future. Therefore, measures in primary prevention require an understanding of the harms of an intervention, which need to be carefully weighed against its projected benefits. While interventions with less harm can be rolled out more easily, interventions with potentially more considerable harm require a more careful selection of qualifying individuals. For instance, lifestyle advice may yield risk reduction at comparatively low harm. At the same time, more targeted approaches, such as CD3 monoclonal antibodies in type-1 diabetes, may reduce risks more significantly, at potentially more

significant harm (Herold et al. 2019).

Ultimately, the impact of an intervention on a population is determined by its overall benefits and harms. Thus, to maximize the benefits of interventions and balance potential harms and benefits, preventive measures must be recommended to those who would benefit most while sparing others from the negative consequences. For instance, today's medicine offers precise interventions able to significantly reduce the future disease burden when recommended to the correct individuals but of none to negative benefit when applied to an entire population. Therefore, to guide preventive interventions and navigate the interplay between benefits and harms, prevention and screening strategies commonly rely on risk assessment as a form of selection process (see figure 1.2).

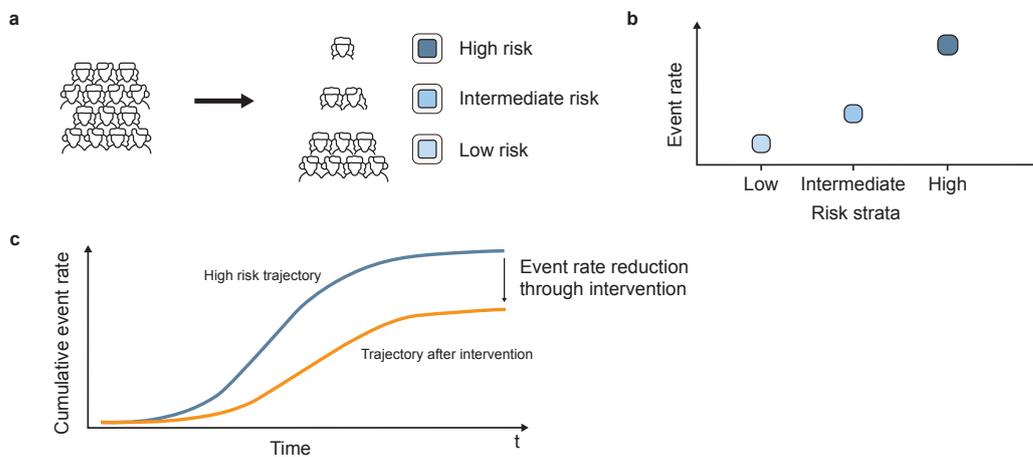


Figure 1.2: Overview on risk stratification. **a**, Risk stratification allows the categorization of individuals into high-, medium-, and low-risk strata. **b**, Risk strata show different event rates, with individuals in the high-risk group experiencing more events. **c**, Schematic depiction of cumulative event trajectories. Intervention on individuals in the high-risk group results in reduced cumulative event trajectories, thus reduced risk. Risk reduction depends on the form of intervention, the prevalence of the health outcome and the quality of the stratification.

Disease risk describes the probability of an individual experiencing disease onset in the future. In the simplest form, risk may be estimated implicitly by examining known factors associated with disease onset, so-called risk factors. For instance, a high Body Mass Index (BMI) is associated with metabolic diseases, cardiovascular complications, and cancers. While the measurement of an elevated BMI is not an exact probability, it already allows a form of risk assessment and may be used for risk stratification. In risk stratification, individuals in a population are categorized by some proxy for disease risk, i.e., a risk estimate or a risk factor value higher than a threshold, with the underlying goal of identifying a sub-

1 Introduction

population with a consistent risk profile (see figure 1.2 on the preceding page). Then, once a high-risk group has been identified, an intervention with substantial harm would be recommended to individuals in a very narrow high-risk subpopulation, minimizing harm over the entire population. Similarly, a less harmful intervention would be recommended for a broader part of the population. Further, relative risks can be computed between strata once a population has been stratified by risk. Relative risks are not probabilities on the individual level but rather provide an estimate of how likely an event is within certain groups of individuals. Thereby, relative risks are utilized to assess differences between groups, scrutinize the stratification, and even estimate risk reductions if one stratum has been treated and another has not. However, the more specific preventive interventions become, the more fine-grained risk stratification is required to guide measures accurately. Personalized prevention takes the balancing efforts a step further by aiming to direct preventive measures on an individual basis to provide the right measure - considering benefits and harms - to the right individual at the right time. However, making decisions on the individual level requires quantifying the absolute individual disease risk. Only estimates of the probability of disease onset on the individual level, i.e., absolute risk estimates, allow a balancing of benefits and harms on the individual level. Personalized prevention is common practice, for instance, in cardiovascular prevention, where risk models are applied to estimate the individual risk for disease onset and to guide interventions (National Institute for health and Care Excellence (NICE) 2014), but is less common for the prevention of other diseases. An overview on the state-of-the-art risk models in cardiovascular prevention and the utilization of risk models in primary prevention more generally is provided in section 1.3.1 on page 18.

Before a prevention strategy can enter clinical practice, reliable estimates of the overall benefits and harms are required. Generally, these estimates are collected in a prospective Randomized Controlled Trial (RCT), where the proposed intervention, along with its guiding criteria (the selection of recommended individuals), is benchmarked against the standard of care. After trial completion, a comparison of relative risks observed in individuals who received the intervention and those treated regularly can then inform on the overall impact of an intervention strategy. Commonly, intervention strategies only enter clinical practice after multiple trials have proven systematic benefits. An example is the management of Low-Density Lipoprotein Cholesterol (LDL-C) with statins, where randomized trials have proven statins to lower cardiovascular disease burden and mortality while not significantly increasing other causes of mortality (Armitage et al. 2019; Collins et al. 2016; Tonelli et al. 2011). In this case, risk models quantifying the absolute risk on an individual level have been utilized as a form of inclusion criteria: individuals with a 10-year risk below 20 % for

developing cardiovascular outcomes such as cardiovascular death or non-fatal myocardial infarction (Tonelli et al. 2011) were selected to investigate the impact of statins in a low-risk population.

In summary, benefits and harms need to be considered to maximize the positive impact of preventive interventions. By applying risk stratification through risk factor assessment or estimating disease risk via dedicated models, appropriate candidate populations are selected to guide interventions and optimally balance benefits and harms. Risk assessment, therefore, is a fundamental requirement for preventive interventions, and the more accurately an individual's future health trajectory can be anticipated, the more fine-grained interventions become applicable.

1.2.4 Fundamentals of risk modeling

Disease risk models provide means to assess individuals' future health trajectories. Generally, survival analysis or risk modeling is a branch of machine learning that seeks to investigate durations or times until the occurrence of an event of interest. Traditionally, risk modeling has been applied to investigate the lifespan of populations from birth to death (Cox 1972), hence the name "survival analysis". Applied to today's medical context and primary prevention, risk modeling intends to answer the question of estimating the duration until the time of occurrence of a specific disease or the time of need for certain procedures based on a given set of information on the individual. Risk models may estimate absolute risk, thus an exact probability for the event of interest at a particular time point, or relative risk, an estimate of how likely an individual is to experience a specific event compared to others in the population. While the latter is sufficient for investigating associations of clinical predictors with disease onset and may even be exploited to identify high-risk individuals within a population, clinical guidelines involving risk models rely on absolute risk estimates (National Institute for health and Care Excellence (NICE) 2014).

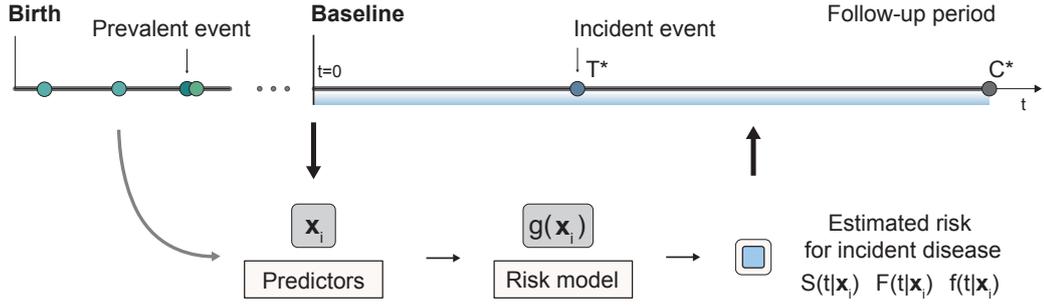


Figure 1.3: Risk models predict incident disease onset. Displayed is a schematic depiction of an individual’s health trajectory in observational data from birth until the end of the follow-up period T . Prevalent events are indicated by green dots, while an incident event at time T^* is indicated by a blue dot. The measurement of clinical predictors \mathbf{x} at baseline t_0 informs on the individual’s health status (gray arrow). Applying a risk model $g(\mathbf{x}_i)$ on the clinical predictors gives an estimate of the risk of incident events. Commonly this estimate is given by one of the relevant survival functions, i.e. $S(t|\mathbf{x})$, $F(t|\mathbf{x})$ or $f(t|\mathbf{x})$.

More generally, risk modeling revolves around time-to-event data, comprising information on whether an event occurred in the time an individual was observed $\Delta \in \{0, 1\}$, information on the duration of observation T , as well as information on the individual, some feature vector \mathbf{x} . Typically, the true event time T^* (i.e., the exact time at which the event occurred) is not observed for all individuals in a population, either because they left the study prior to the occurrence of an event, or because the observation window is not long enough for an event to occur. This phenomenon, noted as $\delta = 0$, is denoted as censoring, and instead of the true event time, a right-censored event time C^* , the time until the censoring occurs, is analyzed: $T = \min C^*, T^*$. While right-censoring is the most relevant, censoring may occur in other forms; for instance, left-censoring would describe individuals joining the study at a later point in time, and interval censoring would describe unobserved time-frames within the observation period. When analyzing time-to-event data, the primary object of interest of risk modeling is to estimate the survival function, typically denoted as $S(t)$, informing on the probability of event-free survival up to a time point t (see figure 1.3). The survival function is complemented by the cumulative risk function $F(t)$, informing on the total accumulated risk up to a time point t .

$$S(t) = P(T > t) = 1 - F(t) \tag{1.1}$$

with $0 \leq S(t) \leq 1$, where $F(t) = \int_0^t f(z)dz$. Thus, $f(t)$ is the incidence function, informing on the probability of occurrence of an event at a specific point in time, t .

Another important concept in risk modeling is the hazard function or hazard rate $h(t)$, defined as the incidence function at a specific time point t , conditioned on survival up to t or later.

$$\lim_{dt \rightarrow 0} \frac{P(t \leq T \leq t+dt)}{dt \times S(t)} = \frac{f(t)}{S(t)} \quad (1.2)$$

The goal of risk modeling is to accurately estimate the survival, cumulative risk and hazard functions for a given individual based on a given set of covariates \mathbf{x} at a desired time point of interest t (see figure 1.3 on the preceding page). Thus, a risk model regresses covariates against events and durations, estimating the survival not only as a function of time, but rather as a function of time conditioned on the covariates, i.e. $S(t|\mathbf{x})$, $F(t|\mathbf{x})$ and $h(t|\mathbf{x})$. Clinically adopted risk models such as the cardiovascular risk scores AHA-ASCVD and ESC-SCORE2 or the CAIDE dementia score rely on linear Cox Proportional Hazards (CPH) models (see section 2.2.2 on page 50, section 1.3.1 on page 18) (Goff et al. 2014; Kivipelto et al. 2006; SCORE2 working group and ESC Cardiovascular risk collaboration 2021).

Information considered in the covariates of a clinical risk model may range from information on the individual's phenotype, including physiological measures, laboratory values, or other risk factors, over sociodemographic values, such as income and diet, to information on the family history and even genetics. However, there is no limitation to what can be considered a covariate, and novel data modalities are regularly exploited for risk modeling. Typically, clinical risk models are developed on data from observational cohorts, comprising necessary sample sizes, baseline measurements (i.e., the covariates \mathbf{x}), and a follow-up period after baseline throughout which observations (T , Δ) were recorded. Commonly, observation windows last 5-10 years or longer, depending on the disease of interest. For infectious diseases, for instance, shorter observation windows are applied than for slow-building diseases such as cardiovascular diseases (Goff et al. 2014; Hippisley-Cox, C.A. Coupland, et al. 2021; SCORE2 working group and ESC Cardiovascular risk collaboration 2021). The baseline defines the time point $t = 0$, thus the time at which the model prediction is made from the collected covariates, \mathbf{x} . With a defined baseline, disease events may be categorized into prevalent events, those observed prior to the baseline measurements and collection of covariates, and incident events, those observed after baseline within the observation period. Outcomes to be predicted by a risk model are incident events, and individuals with prevalent events are commonly excluded from the analysis as the incidence function for secondary events (those events observed for individuals with prevalent events) is much different than for primary events. While models may as well be trained on populations with

prevalent outcomes, the application of these models is not in prediction, but in *prognosis*, the question of how a disease develops over time. Generally, the performance of risk models is dependent on the outcome that is modeled, the population the model is applied to, and the amount of information the predictors contain on the outcome. As the calculation of benefits and harms of interventions may depend on a risk model's output, it is crucial to thoroughly scrutinize clinical risk models under relevant metrics prior to application in clinical practice.

1.2.5 Requirements for clinical risk models

In the primary prevention setting, risk modeling is applied with two motivations. On the one hand, risk modeling is used to stratify a population by disease risk. Here, relative risks are sufficient as they allow medical professionals to define population subgroups with high-risk individuals pinpointing opportunities for preventive intervention. On the other hand, risk modeling may be used for personalized prevention, guiding interventions based on risk profiles of individuals. For this application, reliable absolute risks are required. Considering these two scenarios, three requirements must be met for risk models to be applicable in clinical practice: sufficient discrimination, adequate calibration, and ample clinical utility. *Discrimination* describes a model's ability to stratify a given population, or in other words, rank individuals correctly by their estimated disease risk. The ability to accurately distinguish high- from low-risk individuals is the most fundamental requirement for clinical risk models (Alba et al. 2017). Metrics, such as Harrel's C-index or Concordance Index, allow the assessment of the discriminatory performance in a given population (D'Agostino and Nam 2003). Thereby, the C-index measures the ratio of correctly ranked pairs over the number of all comparable pairs in a population, with the intuition that individuals with earlier events are expected to be assigned higher estimated risks compared to individuals with later events or censored times. A detailed explanation of the C-index and its computation is presented in the methods (see section 2.4.1 on page 58). It is important to be aware that the discriminatory performance of a risk model is population dependent (Longato, Vettoretti, and Di Camillo 2020). Thus, populations may be easier or more difficult to discriminate depending on event and censoring rates and their homogeneity, i.e., the similarity of the individuals in a population. Further, discrimination is upper bound by the predictors collected, the information they contain on the event of interest, and the randomness of events. A population where individuals differ very little from their measured predictors will be harder to discriminate compared to a population where some individuals pose a much higher risk, for instance, due to a prevalent event. Therefore, a model's discriminatory performance requires reassessment prior to applications in unseen populations.

In addition, as ranking is relative within a population, a model's discriminative performance does not inform on the quality of model predictions in terms of actual scale, thus on the accuracy of the predicted absolute risks (Cook and Ridker 2009). For instance, a model could stratify a population perfectly while only predicting cumulative risks close to 0. Therefore, especially concerning personalized prevention, *calibration* is another vital requirement for clinical applicability (Alba et al. 2017). Calibration describes the agreement between a model's estimated risk and the empirical risk observed in the population. The more accurately a model's prediction follows the empirical risk distribution, the more useful the predictions become. A perfectly calibrated model would estimate risks in perfect accordance with the observed risk, i.e., of the individuals with a predicted ten-year risk for ischaemic heart disease of 20 %, about 20 % would experience an event within 10 years. Similarly to discrimination, calibration is population-dependent, too. For instance, when transferring a model from a low-risk to a high-risk population, the model might underestimate disease risk, as it has been developed on less high-risk individuals (Alba et al. 2017). In order to overcome this problem, models can be recalibrated to unseen populations by rescaling their outputs. Methods for the evaluation of model calibration and recalibration of model predictions are presented in section 2.4.5 on page 63 and section 2.2.5 on page 55.

With discrimination and calibration as prerequisites, the decisive requirement for a risk model's clinical application is its potential to impact health trajectories, the *clinical utility*. Besides the model's performance, which is reflected in discrimination and calibration, clinical utility considers the effects of a prediction by factoring in the harms and benefits of an intervention done in consequence of the prediction (also see section 1.2.3 on page 8). In other words, if the model recommends an intervention, are the effects of this intervention overall beneficial or detrimental? The clinical utility of a risk model is dependent on the incidence of the disease of interest in the population and on the benefits and harms of an intervention, represented by the choice of decision threshold at which the intervention is recommended. The disease incidence upper bounds the number of individuals that benefit from an intervention (Vickers and Elkin 2006). In a treat-all scenario, an intervention would be provided for the entire population, with potentially only a few benefiting (maximally the percentage of incident disease), but every treated individual experiencing the harm. The choice of a decision threshold subsets the population into treated and untreated. While the treated definitely experience the harm of the intervention and hopefully its benefits, the untreated definitely do not experience the harm of the intervention but may still be subject to disease onset. At the decision threshold, the benefits and harms of treatment are equal to the benefits and harms of not treating. Therefore, at levels of risk higher than the decision threshold, the benefits outweigh the harms. The overall benefit thus not only depends on the

1 Introduction

chosen decision threshold but also on the ability of the model to stratify the population into high- and low-risk individuals. The measure of net benefit allows for a holistic comparison between models by quantifying their potential population-level impact (Marsh, Janes, and Pepe 2020; Pauker and Kassirer 1980; Vickers, Van Calster, and Steyerberg 2016; Weinstein, Fineberg, Elstein, et al. 1980), and a model is said to be superior to another if its net benefit surpasses the other's net benefit at the chosen decision threshold. Decision curve analysis allows assessing the net benefit of a model continuously over decision thresholds (Vickers and Elkin 2006; Vickers, Van Calster, and Steyerberg 2016). Both methods are introduced in detail in the methods (see section 2.4.2 on page 59 and section 2.4.4 on page 62).

Ultimately, in addition to fulfilling the necessary criteria of discrimination, calibration and clinical utility, a clinical risk model must be adopted into clinical practice to impact care delivery. While the corpus of literature on clinical prediction models is vast, and over the years, many clinical prediction models have been proposed (Damen et al. 2016; Lambert, Abraham, and Inouye 2019), few are ever adopted into clinical practice (Ahmed et al. 2014; Wyatt and Altman 1995). Concerning individual prediction models, adoption may be hampered by the lack of available predictors in the actual care setting (Steyerberg, Moons, et al. 2013) and missing integration with clinical systems and workflows (Sharma et al. 2021; Steyerberg, Moons, et al. 2013). Thus, in order for a model to be utilized in clinical practice, the necessary effort to gather the required information and compute the model output is critical. While some predictors, such as age or biological sex, are readily and ubiquitously available, others, such as blood tests or functional assays, require dedicated measurements. Further, while more clinical predictors usually represent a more comprehensive picture of an individual and consequently allow for more accurate predictions, the increased effort required to facilitate computation may aggravate adoption. In other words, there is a trade-off between the information provided by clinical predictors and the effort required for their collection. Balancing information against effort is crucial in order to render models easily adaptable, and the ideal clinical risk model relies on informative data sources requiring minimal effort in collection.

1.2.6 Clinical predictors in primary care

In primary prevention, the set of commonly assessed predictors is limited due to constraints on the available resources in time and cost in the primary care setting. Thus, risk assessment in primary prevention commonly considers risk factors that are either ubiquitously available or collectible at low cost (Goff et al. 2014; Kivipelto et al. 2006; Lindström and Tuomilehto 2003; SCORE2 working group and ESC Cardiovascular risk collaboration 2021). Ubiquitous

uitously available information refers to predictors such as age or biological sex, which do not require a dedicated measurement. Also, information on an individual's medical history may be considered if the healthcare system allows accessing this data at low thresholds. Factors collectible at low cost may include information on an individual acquirable through questionnaires or examination, such as their socioeconomic status, lifestyle, and medication or family history. Further, physiological measures such as the blood pressure or the Body Mass Index (BMI) may be collected. Depending on national guidelines and their preventive strategies, prevention programs may recommend calculating disease risk scores, for which additional measures, such as laboratory tests assessing blood lipids or urine metabolites, need to be acquired (Hippisley-Cox and C. Coupland 2010; Hippisley-Cox, C. Coupland, and Brindle 2017). For instance, the German national prevention strategy promotes recurrent health checkups in three-year intervals for adults aged 35 or older involving the collection of blood measures for the prevention of heart disease, kidney disease, and type-2 diabetes. The checkup involves the collection of blood measures like lipids and glucose as well as urine measures like glucose, nitrate, and leukocyte counts (Bundesgesundheitsministerium 2022). Similarly, the NHS health check for adults between 40 to 74 aims to identify individuals with high-risk or subclinical manifestations of stroke, kidney disease, heart disease, type-2 diabetes, or dementia (NHS 2022). Especially in cardiovascular prevention, the utilization of risk models forms an integral part of the preventive strategy (National Institute for health and Care Excellence (NICE) 2014). Therefore specific cardiovascular predictors such as total cholesterol or HDL-cholesterol are collected frequently in primary care. Clinically adopted risk models for primary prevention are also available for other diseases such as type-2 diabetes (Lindström and Tuomilehto 2003), common cancers, and dementia (Kivipelto et al. 2006) (see section 1.3.1 on the following page for a detailed overview).

Further, the deep integration of prediction models in prevention strategies is only representative of some of the disease spectrum, and most diseases lack available means of risk assessment. While additional risk models predicting other diseases or leveraging more comprehensive parameter sets are frequently proposed, clinical adoption requires the optimization of the tradeoff between required resources and predictive information (Steyerberg, Moons, et al. 2013). Clinically adopted scores aim to deliver the most accurate predictions from data collectible in the day-to-day primary care operations and are, therefore, often limited in the choice of predictors. At the same time, today's medicine generates increasing amounts of data per individual: Information on an individual's medical history involving diagnoses, procedures, and medications is collected and aggregated in electronic health records, and specialist examinations often produce complex data. One example is reti-

1 Introduction

nal fundus photographs taken at standard ophthalmology visits (MacGillivray et al. 2014). These images are currently only assessed in the context of the specialist’s domain but could carry valuable information on future disease trajectories. Further, in clinical research complexes -omics assays, such as genotyping or 1H-NMR metabolomics, are already applied to measure unseen dimensions of human physiology. These assays promise to hold vital information that cannot be assessed by examining standard clinical risk factors. At the same time, examining complex data modalities could help reduce the required resources in primary care operations by replacing multiple measurements for independent risk factors with a single assay. A better understanding of the information contained in these complex data sources is needed to help public health decision-makers evaluate the potential for inclusion in primary prevention strategies.

1.3 State of the art

The following section introduces the reader to state-of-the-art topics relevant to this work. First, an overview of statistical models applied in primary care for the identification of individuals at highest risk for early intervention is presented. Next, to prepare the reader for the following experimental work, the data modalities examined in this study are introduced. Thus, prior work linking polygenic scores, metabolomics, electronic health records, and retinal fundus photographs to disease onset is described in detail to portray the relevance and potential of each data modalities for risk stratification. Lastly, to inform about the methodological foundation for this work, the state-of-the-art in the two deep learning domains of image analysis and time-to-event modeling is summarized.

1.3.1 Risk models in primary prevention

In cardiovascular disease prevention, research on cardiovascular risk factors and the use of predictive models has a long-standing history and risk assessment through risk models is a central element of preventive guidelines (Goff et al. 2014; Mach et al. 2020; National Institute for health and Care Excellence (NICE) 2014). Beginning with the Framingham Heart Study in the 1980s, followed by many similar investigations, a set of covariates heavily associated with cardiovascular disease onset has been identified. Besides age and biological sex, smoking, overweight and obesity, an unhealthy diet, lack of physical inactivity, dyslipidemia (total cholesterol, HDL cholesterol), and conditions such as hypertension (assessed through systolic blood pressure) and type-2 diabetes have been identified as cardiovascular risk factors (Caldwell et al. 2019; Goff et al. 2014; SCORE2 working group and ESC Cardiovascular risk collaboration 2021; P.W. Wilson et al. 1998; P.W.F. Wilson, Castelli,

and Kannel 1987). Additionally, information on blood pressure treatment and family history of CVD has been included (Hippisley-Cox, C. Coupland, and Brindle 2017). This set of risk factors is commonly exploited in prevention and treatment and used by the vast majority of today's clinically relevant cardiovascular risk models, such as the SCORE-2, ACC/AHA-ASCVD, and QRISK-3. Available and established risk models differ in the exact endpoint definition, characteristics of the study population, predictor sets, and observation windows. The Framingham Risk Score calculates cardiovascular disease risk defined as a 10-year risk for non-fatal myocardial infarction, stable and unstable angina pectoris, or fatal coronary heart disease from nine covariates using a CPH model (P.W. Wilson et al. 1998). Currently recommended in the United States cardiovascular risk prevention guidelines, the ACC/AHA-ASCVD Score utilizes pooled cohort equations (sex-specific CPH models) relying on nine clinical predictors to calculate the 10-year risk for a composite endpoint of fatal coronary heart disease, non-fatal myocardial infarction and fatal or non-fatal stroke (Goff et al. 2014). The European guidelines rely on the ESC-SCORE2 composed of sex-specific and competing risk-adjusted CPH models stratified by regions. The ESC-SCORE2 relies on seven covariates to calculate the 10-year risk for a composite of cardiovascular mortality, non-fatal myocardial infarction, and non-fatal stroke (SCORE2 working group and ESC Cardiovascular risk collaboration 2021). Its predecessor, the ESC-SCORE, relied on the same covariate set and utilized a CPH model but considered only fatal events in the endpoint definition (Conroy 2003). Further, more complex models exist: for instance the QRISK3 score, relying on 22 clinical predictors in a CPH model to calculate the 10-year risk for a composite outcome of coronary heart disease, ischaemic stroke, or transient ischaemic attack (Hippisley-Cox, C. Coupland, and Brindle 2017).

In cancer prevention, risk modeling is much more diverse, as different cancer types have organ- or tissue-specific risk factors. Cancers are characterized by malignant cell growth initiated sporadically but affected by both genetic and lifestyle factors with a tissue and organ-specific context. Thus, proposed clinical risk scores rely on diverse predictor sets involving socio-demographic factors, laboratory assays, and genetic information (Cintolo-Gonzalez et al. 2017; Frampton and Houlston 2017). In general, lifestyle factors, such as smoking, obesity, lack of healthy diet, and physical inactivity, have been linked to multiple cancer types, including the most common cancers (Siegel et al. 2022) lung cancer, breast cancer, prostate cancer and colorectal cancer (Bray et al. 2018; Hippisley-Cox and C. Coupland 2015). In breast cancer prevention, risk models are frequently consulted to both assess a patient's risk of being diagnosed with breast cancer and to guide further measures, such as genetic counseling. Breast cancer-specific risk factors include age at menarche,

1 Introduction

first-term birth, biopsy history, and atypical hyperplasia. These factors are exploited in recommended scores, such as the Gail Score (Gail et al. 1989) and the Breast Cancer Risk Assessment Tool (BCRAT) (Costantino et al. 1999), both relying on CPH models. Further, breast cancer has a strong genetic component, with *BRCA1/2* being known risk alleles. This is considered in recently postulated scores such as the BCRAPRO that, in addition to the clinical predictors in the BCRAT, considers ethnicity, ancestry and extended family history of breast, ovarian, and other cancers (Berry et al. 2002) using a CPH model. A recently proposed score, BOADICEA, additionally considers a Polygenic Score (PGS) based on 313 single-nucleotide polymorphisms along with information on known truncating variants in risk genes and conventional risk factors (A. Lee et al. 2019). Risk factors for lung cancer include the individual's age, biological sex, the intensity or duration of cigarette smoking, medical history and exposures from occupation or area of living (Schwartz and Cote 2016). While lung cancer risk models have been proposed, to date, there is none recommended for clinical practice (Schwartz and Cote 2016). Prostate cancer has a very strong relationship with age, additional risk factors, including ethnicity, genetics, and nutrition (Bratt et al. 2016; J.M. Chan, Gann, and Giovannucci 2005). Models predicting risk for prostate cancer relying on the CPH approach have been proposed (Hippisley-Cox and C. Coupland 2021; Parekh et al. 2006; Steyerberg, Roobol, et al. 2007) and also been explored in clinical practice (Vugt et al. 2012). Risk factors specific for colorectal cancer include red meat consumption, medical history and family history of polyps, colonoscopies, inflammatory bowel disease, medication for post-menopausal hormone, calcium supplementation as well as specific genetic variants (A.T. Chan and Giovannucci 2010; Pearlman et al. 2017). Despite the strong genetic component, the occurrence of colorectal cancer is thought to be mainly sporadic (Cavestro et al. 2018). Several risk models have been proposed (Freedman et al. 2009; Johnson et al. 2013; Zheng et al. 2020); however, screening is currently recommended based on age as sole risk factor, without consideration of risk models (US Preventive Services Task Force et al. 2021).

Cognitive decline or dementia is characterized by substantial impairment in human cerebral capabilities, such as information processing, memory, audiovisual cognition and language (Larson 2019). Dementia is a common disease associated with aging, affecting 47 million individuals in 2015, with projections of strongly increasing prevalence (Livingston et al. 2017). Strong links to vascular conditions have been established: risk factors for dementia include cardiovascular risk factors such as age, hypertension, hypercholesterolemia, physical inactivity, body mass index, and educational level, type-2 diabetes, smoking as well as the prevalence of vascular diseases (Gottesman et al. 2017; Kivipelto et al. 2006;

Luchsinger et al. 2005; Sindi et al. 2015). About 35 % of all prevalent dementias are attributable to modifiable risk factors, such as sociodemographics, physical activity, mental exercise, and smoking (Livingston et al. 2017). Although many risk prediction models for dementia have been proposed, currently, none is recommended for clinical practice (E.Y.H. Tang et al. 2015). Further, as no interventions with proven benefits exist, prevention programs center around modifiable lifestyle risk factors (Larson 2019). Notable models include the CAIDE score estimating the long-term (20-year) risk for dementia incidence (Kivipelto et al. 2006; Sindi et al. 2015) and the THIN score estimating the 5-year risk of a diagnosis of dementia for those aged 60–79 (Walters et al. 2016). The CAIDE score relies on a CPH model using cardiovascular risk factors and information on the education level along with genetic information on the *APOE4* variant carrier status. The THIN score also utilizes a CPH model, however, it solely relies on predictors readily available in primary care (Walters et al. 2016).

Type-2 diabetes mellitus involves hyperglycemia and insulin resistance, driven both by genetic factors involved in glucose metabolism and lifestyle factors such as nutrition and physical activity affecting the metabolism (Paul Robertson 2022). As type-2 diabetes has a high prevalence, affecting around 10 % of adults globally (H. Sun et al. 2022), there has been extensive research on clinical risk factors and prevention measures. Clinical risk factors for type-2 diabetes include family history, ethnicity, obesity, and lifestyle factors, such as physical activity, smoking, nutrition, and sleep duration (G.A. Bello et al. 2019; Chatterjee, Khunti, and Davies 2017). Many prediction models have been developed (Abasi et al. 2012; Schwarz et al. 2009), relying on clinical predictors including age, family history of diabetes, Body Mass Index (BMI), physical activity, blood pressure, High-Density Lipoprotein (HDL), triglycerides, and Impaired Fasting Glucose (IFG). Notable and clinically adopted models include the FINDRISC score (Lindström and Tuomilehto 2003; Saaristo et al. 2005) and a simple model presented in the Framingham Offspring Study (P.W.F. Wilson, Meigs, et al. 2007). The FINDRISC score estimates the 10-year risk of being diagnosed with type-2 diabetes from a set of seven risk factors, including age, antihypertensive medication, history of high blood glucose, body mass index, waist circumference, information on diet and physical activity (Lindström and Tuomilehto 2003; Saaristo et al. 2005). The simple clinical model developed in the Framingham Offspring Study estimates the 7-year risk of incident type-2 diabetes in middle-aged individuals from age, sex, parental history of diabetes, body mass index, and metabolic syndrome traits risk factors (P.W.F. Wilson, Meigs, et al. 2007). Further models, including genetic information, have been proposed (Cornelis et al. 2009; Lyssenko et al. 2008; Meigs et al. 2008), however not

1 Introduction

been adopted in clinical practice as the information gained by genetics over conventional risk factors did not justify population-wide testing (Paul Robertson 2022).

In summary, while risk modeling is a central element of the prevention of cardiovascular diseases, preventive efforts for other diseases are less developed. Although risk models have been proposed for some diseases like dementia and lung cancer, they are not recommended by guidelines and thus are not applied in clinical practice (Schwartz and Cote 2016; E.Y.H. Tang et al. 2015). Further, most models currently applied in clinical practice rely on linear CPH models, imposing strong limitations on the tasks and setups that these models may be used for. For instance, the kind of data modalities that may be considered for risk modeling is limited and specifically, complex measurements of high dimensionality consisting of multiple interrelated values, such as metabolomics measurements, health records or images, are not accessible. Additionally, the linear methodology applied cannot inherently model interactions between covariates, leading to inaccurate risk estimates if non-linearities are present in the data. Further, risk models applied in today's clinical practice commonly assess a single endpoint at a time, rendering systematic risk analysis very tedious. In light of these limitations, there is a great need for a new generation of clinically applicable risk models able to assess complex data sources and multiple endpoints at a given time to disrupt clinical practice and improve on the current standard of care.

1.3.2 Polygenic risk scores and integration in primary prevention

The relationship between human genetics and disease onset has long been of interest to medical professionals, and over time, many disease-related genetic variants have been identified (Sherry et al. 2001). While genetic variation is manifested along multiple dimensions, including epigenetics, histone modifications, and DNA scaffolding, the genotype in the form of Single Nucleotide Polymorphism (SNP) is most commonly assessed. A SNP describes genetic variation at a single position in the DNA sequence (Ma and Zhou 2021). Individual SNPs may be associated with specific phenotypic traits or diseases, and while high-risk variants such as *APOE4* for Alzheimer's or *BRCA2* for breast cancer exist, complex traits are commonly distributed over the entire genome (Boyle, Y.I. Li, and Pritchard 2017). In the past 15 years, with the advent of Next-Generation Sequencing (NGS) and the creation of large-scale genotyping studies, the relationship between genetic variation and complex phenotypes was investigated on a much broader scale through Genome Wide Association Study (GWAS) (Uffelmann et al. 2021). In GWAS, genetic features of samples with- and without a phenotypic trait are compared in a case-control setting to derive SNP-wise effect size estimates. Often, the interpretation of the estimated effects in a biomedical context,

however, is not straightforward, as the variance of a trait is often explained by thousands of (causal) variants. Further, causal variants may be correlated with other causal and non-causal variants in spatial proximity in an effect called linkage disequilibrium. The application of GWAS on populations with increasing sample size has led to the comprehensive catalogization of human genetic variation (Sherry et al. 2001), identifying many SNPs associated with diseases or complex traits (Ganna et al. 2019; International Consortium for Blood Pressure Genome-Wide Association Studies et al. 2011; S.J. v.d. Lee et al. 2018; Yengo et al. 2022).

Polygenic Score (PGS) or Polygenic Risk Score (PRS) build on the knowledge of associated variants to derive a single numerical indicator summarizing the genetic risk of developing a certain disease or trait. In their simplest form, PGS are linear combinations of multiple disease- or trait-associated variants, where the weights are the variants' estimated genetic effect sizes. Commonly, effect sizes of SNPs associated with the trait or disease of interest are estimated in a given population sample and then transferred to another, unseen population to develop the polygenic predictor. In recent years, PGS have been developed for a wide array of disease endpoints, such as Coronary Artery Disease (CAD) (Inouye et al. 2018), stroke (Abraham et al. 2019), chronic kidney disease (A. Khan, Turchin, et al. 2022), type-2 diabetes (Khera, Chaffin, Aragam, et al. 2018; Khera, Chaffin, Wade, et al. 2019), common cancers (Hung et al. 2021; Mavaddat et al. 2019), and phenotypic traits such as drug-susceptibility (Ruderfer et al. 2016; Ward et al. 2018) and obesity (Khera, Chaffin, Wade, et al. 2019). Albeit low overall predictive performance, PGS have been shown to be able to stratify a population by risk for disease onset (Abraham et al. 2019; Inouye et al. 2018; A. Khan, Turchin, et al. 2022; Khera, Chaffin, Aragam, et al. 2018). Despite these advances, the applicability of PGS for clinical practice is disputed (Torkamani, Wineinger, and Topol 2018), and clinical utility has not been demonstrated thus far (Kumuthini et al. 2022). Especially in cardiovascular prevention, where many genetic variants have been known to be associated with cardiovascular diseases (Nikpay et al. 2015) for a long time, advocates of PGS argue that genetic testing may help to identify high-risk individuals at earlier stages (Khera, Chaffin, Aragam, et al. 2018). Thus, the promise of PGS to leverage genetic information in primary prevention for early disease detection has sparked the interest of regulatory authorities (Khouri MJ 2019; UK Department of Health and Social Care 2019). However, the general applicability and benefit of pgs for prevention is hampered by the low information content for the majority of individuals in the population. In the long-tailed pgs distribution, only the individuals in the top percentiles show substantial changes in the associated disease frequencies. Moreover, it is unclear how much information pgs add when combined with commonly assessed cardiovascular risk factors for risk modeling

1 Introduction

and which and how many individuals would benefit from a systematic roll-out of pgs in cardiovascular prevention. Several studies have investigated the possibility of adding pgs to cardiovascular risk scores for primary prevention: Mosley et al. found no additional benefit in adding a pgs against CHD to the features of the AHA/ASCVD pooled cohort equation in two cohorts of US adults (Mosley et al. 2020). Elliot et al. found “significant, yet modest” improvements in discrimination and reclassification after adding their score against CAP to the features of the AHA ASCVD pooled cohort equation and the QRISK3 score in the UK Biobank cohort (Elliott et al. 2020). Sun et al. found only incremental improvements in discrimination over the population but notable reclassification after adding two polygenic scores against coronary heart disease and stroke, respectively, to conventional predictors (L. Sun et al. 2021). With the additional information on genetic predisposition, the authors estimate prevention of additional 7 % CVD events compared to conventional scores based on the altered treatment recommendations (L. Sun et al. 2021). Most recently, Riveros McKay et al. developed a novel integrative risk tool (IRT) combining a novel CAP PGS with the Pooled Cohort Equations (PCE) and QRISK3 score, respectively (Riveros-Mckay Fernando et al. 2021). The authors report a benefit in CAP prediction and estimate a net reclassification improvement of 13.1 % at the 7.5 % 10-year risk threshold, proposing an effect of age and sex on reclassification. Further, a recent meta-analysis identified only “negligible to modest” improvement in discriminative performance upon the addition of PGS to established cardiovascular risk scores (Groenendyk, Greenland, and S.S. Khan 2022). Importantly, all studies investigating the addition of pgs along conventional cardiovascular risk factors relied on linear CPH models, posing strong assumptions and potential limitations on effects between traditional cardiovascular risk factors and the importance of PGS information.

Consequently, although PGS bear an enormous potential for preventive medicine and risk modeling alike, their relationship to clinical phenotypes and known predictors remains elusive (Janssens 2019). Specifically, the incorporation of SNPs acting on known risk factors in PGS has raised concerns about potential biases emerging from joint analysis with those same risk factors (Janssens 2019). This concern calls for tools capable of modeling these complex interactions and a dedicated investigation into the integration of polygenic scores and clinical predictors.

1.3.3 Metabolomics and disease risk

Metabolomics is concerned with analyzing small molecules, metabolites, and the downstream effectors of physiological pathways. The metabolome represents a snapshot of the current state of physiology, combining information on the environment, the proteome, and

the genome (Aderemi et al. 2021). As such, metabolites have long been of interest for biomedical research, and in clinical practice individual metabolites such as glucose are utilized in diagnostics, while others such as cholesterol are established clinical predictors for cardiovascular disease risk (Goff et al. 2014; SCORE2 working group and ESC Cardiovascular risk collaboration 2021). Many other blood metabolites are routinely measured in clinical practice for diagnosis and monitoring. Beyond these established markers, the past decade has witnessed extensive research in the association of individual blood metabolites and their alterations with common disease phenotypes, such as cardiovascular diseases, type-2 diabetes, Alzheimer's, or breast cancer (Holmes et al. 2018; A. Khan, Y. Choi, et al. 2020; Lécuyer et al. 2018; Mahendran, Cederberg, et al. 2013; Mahendran, Vangipurapu, et al. 2013; Stancáková et al. 2012; Tynkkynen et al. 2018; Würtz, Tiainen, et al. 2012). Especially blood lipids have been associated with risk for disease onset: For instance, Holmes et al. found very low-, intermediate-, and low-density lipoprotein particles to be associated with increased risk for myocardial infarction and glycoprotein acetyls, ketone bodies, glucose, and docosahexaenoic acid with increased risk for myocardial infarction, ischemic stroke, and intracerebral hemorrhage (Holmes et al. 2018). Other studies identified an important role of apolipoprotein B and LDL cholesterol in hyperglycemia (Stancáková et al. 2012). Further, lipoproteins have been linked to increased dementia and Alzheimer's risk, while branched-chain amino acids had a protective effect (Tynkkynen et al. 2018).

Recently, however, studies have moved from associating individual markers to examining entire metabolomic profiles. The metabolome has been linked to aging (Ahadi et al. 2020; Alpert et al. 2019), disease onset (Alpert et al. 2019; Schüssler-Fiorenza Rose et al. 2019), and mortality (Deelen et al. 2019), fully appreciating snapshots of the human blood metabolite composition direct reflections of the physiological states (Buergel, Steinfeldt, Ruyoga, et al. 2022). It was only through high-throughput methodologies such as Proton Nuclear Magnetic Resonance ($^1\text{H-NMR}$) spectroscopy, that enabled the assessment of metabolomic profiles at scale. Proton Nuclear Magnetic Resonance ($^1\text{H-NMR}$) spectroscopy, allows a standardized quantification of a multitude of small circulating molecules in the blood at once, while differing from other methods, such as mass spectrometry, by its virtual absence of batch effects, minimal requirements of expensive reagents, and high throughput at comparatively low costs (Markley et al. 2017). In a current assay, 168 original markers are quantified, including amino and fatty acids and metabolites related to carbohydrate metabolism and fluid balance, partly overlapping with conventional clinical predictors, including glucose, albumin, and creatinine (Ala-Korpela et al. 2021; Soininen et al. 2015; Würtz, A.J. Kangas, et al. 2017). Further, the assay has a high resolution of lipoprotein par-

1 Introduction

ticles, resolving their components, and measuring their sizes, and concentrations (Soininen et al. 2015; Würtz, A.J. Kangas, et al. 2017). Over the past years, data generated through this 1H-NMR metabolomics platform has been exploited in multiple studies, which investigated all-cause mortality (Deelen et al. 2019; Fischer et al. 2014), cardiovascular disease (Soininen et al. 2015; Würtz, A.S. Havulinna, et al. 2015), type-2 diabetes (Ahola-Olli et al. 2019; Fizek et al. 2015; Stancáková et al. 2012), Alzheimer's disease (Tynkkynen et al. 2018), and COVID-19 (Julkunen et al. 2021). Based on data collected with the described assay, Deelen et al. developed a risk model for 10-year all-cause mortality relying on a CPH model, identifying 14 circulating metabolites independently associated with all-cause mortality (Deelen et al. 2019). The generalization and robustness of the developed model was demonstrated in an external population. Julkunen et al. relied on a similar setup for the development of their COVID-19 prediction model (Julkunen et al. 2021). The developed model, comprising 25 metabolomic markers, was significantly associated with enhanced susceptibility to severe COVID-19 and pneumonia up to 11 years after the metabolomic measure was taken. While the success of these approaches demonstrates the extent of the information contained in 1H-NMR metabolomics data, approaches not relying on feature selection might increase productivity even further (Buergel, Steinfeldt, Ruyoga, et al. 2022).

A recent metabolomics study, by Pietzner et al., examined the multimorbidity and disease onset from baseline untargeted plasma metabolomics profiling covering >1,000 markers simultaneously (Pietzner et al. 2021). The work demonstrates the existence of shared disease etiologies and the systematic importance of many circulating blood metabolites. If this systematically relevant information was contained in 1H-NMR metabolomic profiles as well, it could be leveraged to assess future health trajectories, providing a holistic view of the combined effects of environment and genetics. However, 1H-NMR metabolomics has not been utilized for the risk prediction of common diseases, and the contained information is yet to be assessed. Especially, an application of multi-disease prediction models could help to identify multi-disease susceptibilities by leveraging the systemic information in metabolomic profiles. Therefore, there is a great need for quantifying the predictive information in 1H-NMR metabolomic data and assessing the added predictive information over established clinical covariates. If resulting risk models had demonstrable clinical utility, the application of 1H-NMR metabolomics could help to streamline clinical workflows by reducing the need for dedicated risk factor collection, facilitating the prevention of many diseases at once (Buergel, Steinfeldt, Ruyoga, et al. 2022).

1.3.4 Electronic health records and disease risk

In modern western medicine, an individual's medical history is a central element for medical decision-making. Thus, most medical decisions on diagnosis, treatment, and prevention of diseases are fundamentally based on an individual's medical history (Hampton et al. 1975). With the widespread digitalization, this information is routinely collected by healthcare providers, insurances, and governmental organizations at a population scale in the form of Electronic Health Record (EHR) data (Clalit Institute 2022; Dk 2016; Health 2021; My Health Record 2022; Wood et al. 2021). These readily accessible records, including diseases, medications, and procedures, are potentially informative about future risk trajectories, but their potential to improve medical decision-making is limited by the human ability to process and understand vast amounts of data (Rush 2019). Thus, routine health records in clinical practice are often perceived as a burden rather than a valuable asset (Furlow 2020).

Organized longitudinally with timestamped information, EHRs typically contain concept codes informing on diagnoses and treatments, such as administered medications or conducted procedures. In addition, EHRs may contain demographic and laboratory measurements as well as hospital or ward-specific information recorded in the context of secondary care. A wide variety of medical vocabulary and coding schemes exist, with common ones being the international classification of diseases (Who 1992) and SNOMED-CT (Donnelly 2006) for disease concepts, LOINC (Huff et al. 1998) for laboratory measurements, RxNorm (Bennett 2012) for medications and OPCS codes (NHS Connecting for Health, Great Britain: Department of Health, and Health and Social Care Information Centre 2006) for procedures. The choice of these coding schemes depends on the care provider and thus requires standardization and mapping efforts between healthcare systems and vendors.

While population scale EHR databases exist, examples of applied methodology are scarce. In the research context, information collected through and stored in EHRs has been leveraged for disease phenotyping (Shang et al. 2021; A.S. Tang et al. 2022), diagnostic screening (Sekelj et al. 2021) and prediction of in-hospital and disease outcomes (Rajkomar et al. 2018). Existing approaches differ greatly in the breadth and depth to which EHR data is analyzed. In the most basic form, methods examining EHR-data extract established risk factors from the records to allow the automatic computation of established risk models (Z. Xu et al. 2022) or leveraged established risk factors with modern, complex methodology (Hill et al. 2019). Due to the complex, sparse and longitudinal nature of EHR data, a great corpus of work is dedicated to developing advanced methods for clustering (Beaulieu-Jones, Greene, and Pooled Resource Open-Access ALS Clinical Trials Consortium 2016) and representation learning for outcome prediction (Y. Li et al. 2020), diagnosis (E. Choi et al. 2018; Y. Li

1 Introduction

et al. 2020; Rajkomar et al. 2018; Solares et al. 2021; J. Zhang et al. 2018) or risk prediction (Bagheri et al. 2020; Langham et al. 2021).

To date, information in EHRs has been used to predict the onset of cardiovascular diseases such as atrial fibrillation (Hill et al. 2019; Tiwari et al. 2020), myocardial infarction (Zhao et al. 2019) and Major Adverse Cardiac Events (MACE) (Bagheri et al. 2020), dementia (Ben Miled et al. 2020; Langham et al. 2021), as well as other common diseases including type-2 diabetes, multiple sclerosis (C.A. Nelson et al. 2022) and cancers (Appelbaum et al. 2021). Further, the information in EHRs has been demonstrated to be relevant for the prediction of in-hospital outcomes, such as in-house mortality, length of stay, and readmission (Rajkomar et al. 2018). However, it is important to note that the utilized information varies greatly between studies due to strong differences in data quality and heterogeneity of data types included in the analyzed EHR datasets (C. Xiao, E. Choi, and J. Sun 2018). While autoregressive (Ayala Solares et al. 2020; C. Xiao, E. Choi, and J. Sun 2018; J. Zhang et al. 2018) and transformer-based architectures (Y. Li et al. 2020; Rasmy et al. 2021) have been applied to entire health records leveraging their longitudinal nature, and other approaches leverage multi-modal data types (Bagheri et al. 2020), the majority of approaches rely on feature selection to reduce the sparsity in the EHR data (Rajkomar et al. 2018; Zhao et al. 2019), even to the point of extracting a static, fixed length vector of information (Hill et al. 2019; Sekelj et al. 2021).

In summary, EHRs have been utilized for a wide range of prediction and diagnosis tasks as well as for the definition of patient’s phenotypes, appreciating EHRs as a proxy for an individual’s medical history. Despite this evidence emphasizing the relevance of the medical information on an individual’s future, the application of EHRs for risk stratification in primary prevention has hardly been investigated. While studies for individual endpoints exist (Goldstein et al. 2017; Hippisley-Cox and C. Coupland 2013), there is a lack of comprehensive assessment of the applicability and utility of the information contained in EHRs. Especially as the information in EHRs is medically relevant, collected routinely at no additional cost, and readily available at primary care sites, this data modality seems predestined to improve primary care by allowing a systematic risk assessment.

1.3.5 Retinal fundus photographs and disease risk

The examination of the human retina is a crucial standard procedure in ophthalmology to inform on diseases of the eye (Kolb, Fernandez, and R. Nelson 1995; Wong, Hubbard, Cruickshanks, et al. 2004). The retina is the only tissue in the human body that allows a noninvasive

examination of the nervous and vascular systems simultaneously (Kolb, Fernandez, and R. Nelson 1995). Therefore, the retina has been analyzed beyond ophthalmology, and associations with the human vascular and neurological conditions (Patton et al. 2005; Rim, G. Lee, et al. 2020; Rim, Teo, et al. 2020; T.Y. Wong, Islam, et al. 2006; T.Y. Wong, Klein, et al. 2003) have been established. Retinal features may inform on subclinical disease years before manifestation (Baker et al. 2008; N. Cheung, Mitchell, and T.Y. Wong 2010; Patton et al. 2005; T.Y. Wong and Mitchell 2007). Of special interest are patterns in the microvascular network such as the geometry of vascular bifurcations, the vascular tortuosity and diameters, nicks and dents of visible blood vessels (such as arteriolar narrowing or arteriovenous nicking), which could inform on the microvascular health (C.Y.-L. Cheung et al. 2011; M.K. Ikram et al. 2006; Liew et al. 2011; Stanton et al. 1995). Retinal microvascular features have been associated with stroke (Baker et al. 2008), coronary heart disease (Liew et al. 2011; T.Y. Wong, Klein, et al. 2003), type-2 diabetes (Wong, Hubbard, Cruickshanks, et al. 2004), and hypertension (M.K. Ikram et al. 2006). In addition, the retinal tissue is a direct extension of the central nervous system (Kolb, Fernandez, and R. Nelson 1995). As the nerves are located in a transparent layer, however, explicit features have been more difficult to define (MacGillivray et al. 2014).

Retinal funduscopy is a noninvasive, simple, and fast assay of the human posterior eye. Photographs are recorded using a low-potency microscope and a camera, flashing through the iris and mapping the 3D inner posterior structure on a 2D image. Photographs are commonly recorded in full color and display the retina in high-resolution (MacGillivray et al. 2014). The onset of deep learning facilitated a great array of studies building large neural networks to analyze and associate retinal fundus photographs with clinical risk factors and disease onset. Recent studies have been exploiting retinal fundus photographs for the prediction of basic demographic markers such as age and biological sex (Nusinovici et al. 2022; Poplin et al. 2018), but also a broad array of vascular and systemic risk factors (Gerrits et al. 2021; Y.D. Kim et al. 2020; Poplin et al. 2018; Rim, G. Lee, et al. 2020), among them traditionally invasive measurements such as the glomerular filtration rate (K. Zhang et al. 2021), a coronary artery calcium score (Rim, C.J. Lee, et al. 2021) or lipid profiles (Rim, G. Lee, et al. 2020). Other work has demonstrated approximations of biological age modeled through retinal fundus photographs to predict mortality (Nusinovici et al. 2022; Zhu et al. 2020). The quality of the estimated risk factors has been shown to be affected by demographics (Gerrits et al. 2021) and prevalent diseases (Y.D. Kim et al. 2020), and the transferability of models to distant populations was found to be limited (Rim, G. Lee, et al. 2020). Further risk factors related to thyroid function, biochemical measures, hematological

1 Introduction

parameters other than hematocrit and hemoglobin, as well as C-reactive protein, could not be predicted from retinal fundus photographs (Rim, G. Lee, et al. 2020), indicating the limits of the information contained in retinal fundus photographs.

In terms of direct relationships to disease, studies have been analyzing retinal fundus photographs with deep learning methods for disease detection (Babenko et al. 2022; B. Li et al. 2021; F. Li et al. 2021; Mitani, Hammel, and Y. Liu 2021; Mitani, A. Huang, et al. 2020; Sabanayagam et al. 2020; K. Zhang et al. 2021), the prediction of disease progression (Arcadu et al. 2019; Peng, Dharssi, et al. 2019; Peng, Keenan, et al. 2020; Yan et al. 2020) and the prediction of disease onset (Bora et al. 2021; Poplin et al. 2018). To date, retinal fundus photographs have been associated with multiple diseases, such as Age-related Macula Degeneration (AMD) (Peng, Dharssi, et al. 2019; Peng, Keenan, et al. 2020; Yan et al. 2020), diabetic retinopathy (Arcadu et al. 2019; Bora et al. 2021), Alzheimer’s and all-cause dementia (C.Y. Cheung, Ran, et al. 2022), anemia (Mitani, A. Huang, et al. 2020), type-2 diabetes (K. Zhang et al. 2021), Major Adverse Cardiac Events (MACE) (Poplin et al. 2018) and Chronic Kidney Disease (CKD) (Sabanayagam et al. 2020; K. Zhang et al. 2021). Poplin et al., for instance, explored the prediction of incident MACE within 2-years after screening. However, their retinal score did not add predictive performance over established cardiovascular risk factors (Poplin et al. 2018). Bora et al. developed a model to predict the development of mild or worse diabetic retinopathy within 2-years after screening in patients with diabetes (Bora et al. 2021). In a two-stage approach, Diaz-Pinto et al. trained a variational autoencoder to reconstruct cardiovascular magnetic resonance images from retinal fundus photographs. Subsequently, they leveraged the reconstructions with a large convolutional neural network to predict cardiovascular risk factors and myocardial infarction (Diaz-Pinto et al. 2022). While these studies have been successfully predicting disease onset from retinal fundus photographs, actual time-to-event modeling from retinal fundus photographs is yet to be conducted. With the retina having been postulated as a broad source of potentially untapped clinical information (MacGillivray et al. 2014), there is a great need for a thorough assessment of the utility of retinal fundus photographs in primary prevention over an overall spectrum of diseases. Specifically, there is great need for quantification of the added predictive information over established clinical covariates to pinpoint opportunities for applications with potential for clinical utility.

1.3.6 Neural networks for image analysis

Even though the methodology of convolutional neural networks was established already in the 1990s (Lecun et al. 1998), it was the creation of AlexNet (Krizhevsky, Sutskever, and

Hinton 2012), that for the first time proved the capabilities of deep neural networks for image analysis and sparked the deep learning boom of the 2010s. Since then, convolutional neural networks have grown increasingly more capable and increasingly large, encompassing hundreds of layers and millions of weights in a single model (He et al. 2016; Z. Liu, Mao, et al. 2022; Simonyan and Zisserman 2014; Xie et al. 2016). While the core methodology of LeCun et al., the convolutional layer, remains effectively unchanged, modern deep neural networks rely on architectural extensions and modifications to improve learning stability, generalization, and parameter efficiency. Notable architectures include the ResNet(X)t architecture (He et al. 2016; Xie et al. 2016) (see section 2.1.3 on page 44), built to circumvent the vanishing gradient problem and allow for even larger networks, the MobileNet (Howard et al. 2017) and EfficientNet (Tan and Le 2019) architectures developed to further optimize parameter requirements and ultimately vision transformers, bridging the generalizability of the transformer architecture with traceable computation of sliding window convolutions (Z. Liu, Y. Lin, et al. 2021). ResNets are powerful image feature extractors that have been applied to many different domains, from general image classification (He et al. 2016) to medical tasks (Guo and Z. Yang 2018). Over the years, many ResNet-based architectures have been proposed (Xie et al. 2016), and both the principle as well as the original architecture continue to be relevant (I. Bello et al. 2021). Most recently, Liu et al. introduced the ConvNeXt architecture (Z. Liu, Mao, et al. 2022), relying on the ResNet principle. The ConvNeXt architecture introduces specific changes to the ResNet architecture, upgrading its performance to outperform much more parameter-heavy neural network architecture, such as vision transformers (Z. Liu, Mao, et al. 2022).

The medical domain has embraced the capabilities of deep learning, and the use of deep neural networks has become the gold standard for medical image analysis. As the image datasets in the medical domain are not as extensive as the large image datasets used to develop deep neural networks in machine learning research, transfer learning is a common practice. In 2017 Esteva et al. proposed a classifier based on Google’s Inception v3 architecture, developed by Szegedy et al. (Szegedy et al. 2015), to distinguish among 757 classes of malignant and benign skin cancer (Esteva et al. 2017). Notably, they utilized a model pre-trained on ImageNet data, only fine-tuning the weights on their specific task (Esteva et al. 2017). In 2018, Poplin et al. a deep neural network of the Inception v3 architecture, pre-trained on ImageNet to estimate cardiovascular risk factors from retinal fundus photographs (Poplin et al. 2018). Another area predestined for the application of deep neural networks is radiology, where models have been developed for mammography scans as early as 2016 (D. Wang et al. 2016), estimating cardiovascular features from magnetic resonance images

1 Introduction

(Ferdian et al. 2020), and the prediction of COVID-19 from chest X-rays (Minaee et al. 2020). Further, deep learning has been applied for cancer detection (Tolkach et al. 2020; Zhang et al. 2019) in pathology slides and even to estimate genetic mutations (Fu et al. 2020).

To date, the ConvNeXt architecture has yet to be applied for the analysis of retinal fundus photographs, leaving an interesting gap for investigation. However, as deep learning is a general tool applicable to any image dataset, the latest methods developed in computer vision will continue to be applied to real-world applications in medicine and beyond. Ultimately, the latest methodology promises analytic advantages, such as fewer data requirements, more stable training, or better feature detection.

1.3.7 Neural networks for risk modeling

Following the great success of deep learning methods in image and natural language processing, neural networks became the method of choice in many scientific fields. Consequently, in recent years, an increasing number of methods for neural network-based survival analysis has been proposed. Among these approaches are semi-parametric and fully parametric models, allowing for discrete and continuous time intervals. DeepSurv, presented by Katzman et al. (Katzman et al. 2017) builds upon existing methodology by modeling the Cox’s log partial likelihood by a neural network as initially proposed by Faraggi and Simon (Faraggi and Simon 1995). While this allows for non-linear interactions between covariates, the model outputs a log partial hazard, which is subject to the proportional hazards assumption. Kvamme et al. alleviated this assumption with the Cox-time model (Kvamme, Borgan, and Scheel 2019), where the log Cox’s partial hazard function is approximated via a neural network accessing both the covariates and the event time. In contrast, Luck et al. modeled the hazard function directly for discrete time intervals and optimized it for tied events using Efron’s approximation (Luck et al. 2017). In contrast to non-linear extensions and modifications of the Cox model, other neural network-based approaches seek to approximate the survival function, incidence function, or even hazard functions directly. For instance, Gensheimer et al. directly estimated the hazard function via maximum likelihood estimation in a discrete-time model (Gensheimer and Narasimhan 2018). These ideas were taken further in the DeepHit model, estimating the survival and incidence functions for discrete time points by incorporating a ranking term into the loss function. DeepHit further offers the flexibility of accounting for competing risks (C. Lee et al. 2018; Rietschel, Yoon, and Van Der Schaar 2018). Flexible parametric assumptions are made in the Deep Survival Machine model proposed by Nagpal et al. where survival and incidence are fitted by a weighted mixture of basis functions parameterized by neural networks (Nagpal, X.R. Li, and Dubrawski 2021).

While this approach theoretically offers possibilities for generative modeling and variational inference by sampling event times from the learned mixture, this has only been realized by Xiu et al. in the Variational Survival Inference model (Xiu, Tao, and Henao 2020). Generative adversarial learning has also been attempted for event time prediction in the work of Chapfuwa et al. by building on the classic accelerated failure time models (Chapfuwa et al. 2018). Leaving the observational study setting of associating measurements at a single baseline with an observed outcome and reverting to the idea of a continuous analysis enabled by longitudinal records, an entirely different category of survival models can be imagined. First attempts to leverage longitudinal data in survival modeling have been made by Ren et al., applying a recurrent neural network to continuously model the hazard rate at each point in time (Ren et al. 2019).

Based on the growing corpus of methods for neural survival models, multiple applications for clinical risk modeling have been developed thus far. DeepSurv was exploited for the prediction of survival of oral cancer patients (D.W. Kim et al. 2019), where the neural survival model outperformed a CPH model on a relatively small dataset of 255 individuals. In another study, DeepSurv outperformed a CPH model for risk modeling of critically-ill Covid-19 patients when applied to a tabular dataset of 1,590 patients and 74 features (Liang et al. 2020). Others used the DeepSurv objective for survival outcome prediction in cervical cancer from tabular features (Matsuo et al. 2019) and for clinical outcome prediction from cancer genomic profiles (Yousefi et al. 2017). In addition, the DeepSurv model was applied in combination with convolutional neural networks to predict cancer outcomes from histology images (Mobadersany et al. 2018) and patient survival from cardiac motion videos (G.A. Bello et al. 2019).

The discrete-time survival loss function specified by Gensheimer et al. was applied for long-term cancer survival prediction from multimodal data, including information from histology images, cancer genomics and individual medical history (Vale-Silva and Rohr 2021). Others applied the method in combination with convolutional neural networks to predict the onset of distant metastasis in a time-to-event analysis of head and neck cancer (Lombardo et al. 2021). DeepSurvivalMachines were applied for long-term post-discharge mortality prediction among ICU patients from demographic data, diagnostic data (ICD-9 codes), and unstructured clinical notes in combination with complex transformer models (Dencker et al. 2022).

1 Introduction

These studies indicate the general applicability of neural network-based survival models on real-world clinical data. Importantly, these models are applicable to any data for which deep learning methods have been developed. This opens up exciting opportunities to explore the utility of complex data sources such as images or -omics data for risk modeling in primary prevention.

1.4 Aim of study

Healthcare costs are systematically on the rise, and current therapy-focused healthcare systems are not sustainable in the long run. At the same time, many diseases are linked to modifiable risk factors, rendering disease prevention a viable instrument for reducing costs and suffering (Galea and Maani 2020). Towards personalized prevention, risk modeling is fundamentally important to stratify populations, identify high-risk individuals and deliver the right interventions at the right time.

In current clinical practice, however, risk modeling is subject to severe limitations, and the guidelines for the utilization of risk models in primary prevention strategies are very heterogeneous. While, for instance, assessing future disease risk through predictive models is an integral part of preventive guidelines for cardiovascular diseases, and established risk models for diabetes, dementia, and certain types of cancers exist, this is not representative of the entire disease spectrum. In fact, for most diseases, there exists no means of assessing the probability of future disease onset. In the primary care setting, the substantial constraints on available resources impose severe limitations on the number and complexity of clinical predictors that may be considered in a given risk model (see section 1.2.5 on page 14). Historically, risk modeling in clinical practice has been limited to predicting the risk for a single endpoint at a time, which increases the burden for a systematic risk assessment considering the limited resources in primary prevention and the necessity to collect additional predictors to predict endpoints. In consequence, only a few of the frequently proposed novel disease risk models are ever adopted into clinical practice (Steyerberg, Moons, et al. 2013), and those that are adopted optimize the trade-off between the resources required for the collection and the information provided by the predictors.

At the same time, the density of individual patient data is constantly increasing, and complex data from -omics or imaging assays are routinely collected in clinical practice or are becoming available in the clinical (research) context. This data may hold valuable information on future health trajectories that could disrupt medical practice and decision-making.

For instance, if complex data modalities contained similar information to clinical predictors used today, the ability to utilize complex and rich data modalities for risk modeling could reduce the number of required predictors and decrease the cost and effort of risk prediction. This would not only foster the adoption of new risk models into clinical practice but also generally promote risk modeling for preventive purposes. Additionally, if the information in these complex data modalities was systemically relevant, and respective methods were available for its extraction, the utilization of complex data could allow risk modeling for a wide array of diseases at once from a single measurement. However, these possibilities and the entire potential of leveraging complex data modalities for risk modeling in primary prevention currently remain untapped due to a lack of dedicated methodology.

Neural survival models represent the state-of-the-art in survival analysis. Compared to approaches used in today's clinical practice, these models provide greater flexibility in both the input and output space. As neural survival models rely on neural networks for feature extraction, they are theoretically capable of integrating any data modality for which deep learning methods exist. Thus, while current approaches are limited to scalar, independent features, neural networks are able to extract meaningful information from complex signals. Similarly, while current approaches are limited to predicting a single endpoint at a time, neural survival models are well capable of predicting multi-disease onset. Thus, with complex data modalities promising to contain systemic information, recent advances in machine learning may help to unlock this potential and create a new generation of risk models for primary prevention to perform a systematic risk assessment for a wide array of diseases simultaneously from a single measurement.

Thus, the aim of this study is to investigate the potential of complex data modalities for risk modeling in primary prevention. Importantly, this study presents the development, training, and validation of novel, neural network-based risk models to extract and integrate information from four complex data modalities so far inaccessible for primary prevention: Polygenic risk scores (section 1.3.2 on page 22), NMR metabolomics (section 1.3.3 on page 24), the previous medical history in the form of electronic health records (section 1.3.4 on page 27) and retinal fundus photographs (section 1.3.5 on page 28). Each of these data modalities is highly relevant and bears the potential to improve risk modeling in primary prevention. For example, genotyping assays allowing the calculation of polygenic risk scores are discussed as a potential addition to primary prevention programs. Here, methods are needed to efficiently integrate genetic information with conventional risk factors to understand the added predictive value in primary care and pinpoint clinical utility (section 1.3.2 on page 22). 1H-

1 Introduction

NMR metabolomics, informing on a multitude of small circulating molecules in the blood, have recently been shown to contain systemic information relevant to multiple diseases simultaneously. Available at scale and low costs, with appropriate methods, these assays could be leveraged as a drop-in replacement for commonly collected covariates to inform on risk trajectories for a whole spectrum of common diseases (section 1.3.3 on page 24). In many modern healthcare systems, electronic health records are routinely collected in the background, summarizing a patient's medical history in the form of codes for observations, diseases, procedures, and medications (section 1.3.4 on page 27). This data is readily available at no additional cost and may contain important information on an individual's future health trajectory. Like health records, retinal fundus photographs are routinely collected in clinical practice and contain biological features associated with cardiovascular, neurological, and metabolic conditions (section 1.3.5 on page 28). With dedicated tools, the relevance of this information and its potential clinical utility could be scrutinized over a broad set of clinical endpoints. For all these data modalities, an investigation of the shared and added information over commonly collected clinical predictors and laboratory values is imperatively required to allow for an informed and structured decision on public health guidelines.

In exploring these four examples, this study has three concrete objectives: First, this study aims to challenge existing strategies for clinical disease risk estimation by investigating applications of neural network-based approaches for time-to-event regression on real-world medical data. Second, building on these methods, this study aims to develop dedicated risk models to extract relevant information from "modern" complex data modalities (i.e., genetics, metabolomics, health records, and retinal fundus photographs). Third, examining the developed models' predictions, this study aims to contribute to the understanding of each data modality's information content to allow an assessment of clinical utility. Specifically, this involves scrutinizing the learned information against established clinical predictors and quantifying additive information. Given the four independent data modalities, this study is divided into four experiments, each presenting an independent model's development and evaluation. With the entirety of its four parts, this study provides a blueprint for the structured process of investigating the potential of complex data sources and a reference point for the future development of dedicated tools leveraging complex data modalities for clinical risk modeling.

1.5 Outline

The previous sections aimed at providing the reader with an understanding of this work's greater context and background as well as the scientific state-of-the-art; the following sections will present details of the applied methodology and a description as well as a discussion of the experiments conducted. The methods chapter (section 2 on page 39) aims to equip the reader with an understanding of the applied methodology and its mathematical foundations. Specific emphasis is placed on methods related to deep learning (section 2.1 on page 40), neural survival models (section 2.2 on page 49), utilized evaluation metrics (section 2.4 on page 58) and attribution methods (section 2.5 on page 64). Further, details on the implementation of the developed risk models are provided (section 2.6 on page 67). Subsequently, the data chapter of this study (section 3 on page 73) provides the details of all cohorts utilized in this study (section 3.1 on page 74). Further, the chapter contains details on the data preprocessing applied for each data modality (section 3.2 on page 76), the definition and extraction of predictor sets, as well as the specifications of the endpoints for each experiment (section 3.3 on page 83). Next, the results chapter (section 4 on page 87) is divided into four experiments, each of which comprises an independent piece of original research. In the first experiment, presented in section 4 on page 87, the applicability of neural networks to the problem of cardiovascular risk modeling was investigated. A novel cardiovascular risk score, NeuralCVD, is presented and examined for its ability to integrate genetic information in the form of Polygenic Score (PGS) with conventional cardiovascular predictors. It is shown that NeuralCVD was able to inherently model an attenuating effect between the observed cardiovascular risk phenotype and polygenic information, confirming a previously suggested hypothesis of information transitioning from genotype to phenotype. Based on the results of this part of the study an article has been published in *The Lancet Digital Health* (Buerger, Steinfeldt, Looock, et al. 2022). The second experiment (section 5 on page 101) investigated the applicability of Proton Nuclear Magnetic Resonance (¹H-NMR) metabolomics as a single-domain multi-disease assay for risk modeling in primary prevention. Relying on a deep learning approach ¹H-NMR metabolomic profiles were associated with the risk of onset for 24 common disease endpoints. The added information over established clinical predictors was quantified, and model examination revealed both global as well as individual-level metabolite-disease associations. The results demonstrated the systemic information contained in ¹H-NMR metabolomic profiles and showcased the applicability of this data for risk modeling in primary prevention. Based on the results of this part of the study an article has been published in *Nature Medicine* (Buerger, Steinfeldt, Ruyoga, et al. 2022). Subsequently, in the third experiment, the predictive information con-

1 Introduction

tained in electronically versioned medical history was investigated over the entire human phenome (section 6 on page 127). A neural survival model was trained to predict risk for future disease onset for ~2,000 endpoints simultaneously. For a wide spectrum of diseases, the learned risk states were additive to basic demographic predictors (age and biological sex), indicating potential clinical utility. The applicability of the learned risk states for risk stratification in primary care was showcased at two examples: primary prevention of cardiovascular diseases and repurposing of the risk states to react to emerging health threats, represented by COVID-19. The experiment demonstrated the vast amount of predictive information that is contained in individuals' electronic medical history and its potential for risk stratification in primary prevention. In the fourth experiment (section 7 on page 141), the phenome-wide investigation of predictive information for incident disease onset is continued with the data modality of retinal fundus photographs. Relying on a large convolutional architecture, retinal fundus photographs were associated with the risk of onset for 1,171 endpoints simultaneously. Similar to the experiment involving the medical history, the added information over baseline demographic predictors, age, and biological sex was quantified to pinpoint diseases where retinal fundus photographs could provide relevant information. A dedicated investigation into the application of retinal information to simplify risk stratification in the primary prevention of cardiovascular diseases did not reveal significant benefits. Ultimately, attribution maps were computed, highlighting parts of the image that most positively and negatively influenced predictions. This experiment identified endpoints for which retinal fundus photographs contain predictive information to guide further investigations and the development of dedicated risk models. Ultimately, the conclusion chapter (section 8 on page 157) summarizes and contextualizes the results of the four experiments and their implications for the future of risk modeling in primary care, inviting the reader to think along the lines of prevention into a future of digital and efficient medicine.

2 Methodology

2.1 Neural networks

2.1.1 Multi-layer perceptron

Neural networks or artificial neural networks represent the core method of deep learning. A Neural Network (NN) or Artificial Neural Network (ANN) consists of layers of nodes, which individually process information with a set of trainable weights to produce an individual output. The fundamental idea behind the structure of a singular node has been proposed already in 1958 by Frank Rosenblatt (Rosenblatt 1958) as the perceptron. Following this definition, in today's practice, a node is defined by a linear operation between a set of learnable parameters θ , where $\mathbf{w}, b \in \theta$ and an input vector \mathbf{x} of length D which produces a scalar output z that is transformed by a non-linearity σ to form the output y of length C :

$$n_{\theta}(\mathbf{x}) = \sigma(\mathbf{w} \cdot \mathbf{x} + b) = \sigma(z) = \hat{y} \quad (2.1)$$

A layer in a multi-layer perceptron may thereby consist of multiple nodes with individual sets of weights, which may be defined by extending equation (2.1) to a weight matrix \mathbf{W} , where $\mathbf{W} \in \mathbb{R}^{n_{nodes} \times n_{inputs}}$ resulting in an output vector y of length n_{nodes} :

$$h_{\theta}(\mathbf{x}) = \sigma(\mathbf{W} \cdot \mathbf{x} + \mathbf{b}) = \sigma(\mathbf{z}) = \hat{\mathbf{y}} \quad (2.2)$$

In modern deep learning, neural networks commonly consist of $H + 2$ layers which are stacked to process information sequentially, with the first layer, called the input layer, transforming the input, the following H layers, typed hidden layers, transforming each previous layer's outputs and the $H + 2^{\text{th}}$ layer, called output layer, producing the final result \hat{y} . In a Multi Layer Perceptron (MLP), the previous layer's output nodes M_{h_i} are the same number as the subsequent layers input nodes. Neural networks where information flows directionally through the network from the input layer to the output layer are also called feed-forward neural networks (see figure 2.1 on the facing page).

Thus, formally, a feed-forward NN can be seen as a computational Directed Acyclic Graph (DAG) consisting of many individual, sequentially chained operations defining a function g parameterized by weights θ of an input \mathbf{x} producing an output \hat{y} :

$$g_{\theta}(\mathbf{x}) = \hat{y} \quad (2.3)$$

$$g_{\theta}(\mathbf{x}) = h_{n_{\phi_n}}(h_{n-1_{\phi_{n-1}}}(\dots h_{1_{\phi_1}}(\mathbf{x}))) = \hat{y} \quad (2.4)$$

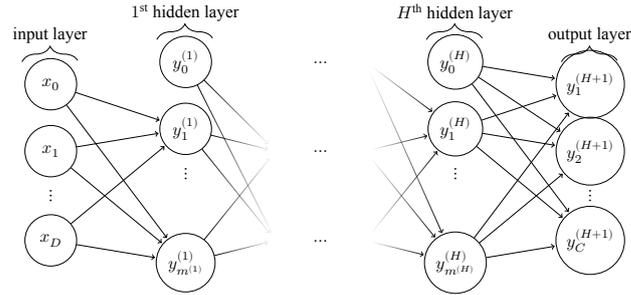


Figure 2.1: DAG of a multilayer perceptron. Displayed is a pictogram of the connectivity in a $(H + 2)$ -layer MLP with $N_x = D$ input units and $N_y = C$ output units. The H^{th} hidden layer contains $m^{(H)}$ hidden units. Figure adapted from (Stutz 2022).

Here, g_θ may consist of $H + 2$ individual layers h_H parameterized by the individual layer's weights ϕ , which process the previous layer's outputs as described previously. As typical in supervised machine learning, Neural Network are trained by minimizing a loss function $\mathcal{L}_\theta(\hat{y}, y)$ that quantifies the error between the prediction \hat{y} and the observed label y .

In order to train a neural network and optimize its weights, gradient descent through *backpropagation* is applied (Rumelhart, Hinton, and Williams 1986). In backpropagation, the gradients of the loss function w.r.t. the model's weights are passed backward through the network, and parameters at step θ^s are updated following equation 2.5 for the next step $s + 1$,

$$\theta^{s+1} = \theta^s - \alpha \frac{\delta \mathcal{L}_\theta(\hat{y}, y)}{\delta \theta} \quad (2.5)$$

where α is called the learning rate. The update rule is computed individually for each layer's set of weights and biases. Hence, the backpropagation step is often referred to as the backward pass. Typically, a neural network is not trained on the entire dataset at once, computing one single parameter update, but rather by randomly sampling so called *batches* of samples from the dataset and computing iterative parameter updates on each *batch* until the loss function has converged. Thereby a full pass over the dataset is referred to as *epoch*, and the method is known as Stochastic Gradient Descent (SGD). Over the years, several extensions to SGD have been proposed, controlling the parameter updates more carefully by tracking averages and variances of the past steps' updates (Kingma and Ba 2015).

Importantly, neural networks and MLPs are subject to limitations: Firstly, standard neural networks are not interpretable and are considered black box models as they neither inform on the importance of an input nor allow the interpretation of their internal processes. Due to the extreme number of parameters, dedicated algorithms for post-hoc interpretability have

2 Methodology

to be applied in order to attribute importance measures to the input space and explain predictions. Although many interpretability methods have been developed (Lapuschkin et al. 2019; Lundberg and S.-I. Lee 2017; Ribeiro, Singh, and Guestrin 2016; Shrikumar, Greenside, and Kundaje 2017; Sixt, Granz, and Landgraf 2020), limitations exist (Kumar et al. 2020; Sixt, Granz, and Landgraf 2020) and interpretation requires dedicated effort, especially in the medical domain (Tonekaboni et al. 2019). For a description of the interpretability methods used in this study, see section 2.5 on page 64. Relating to the issue of interpretability, the entirely data-driven nature of neural networks does not allow control over which information from the input space is utilized to make a given prediction. Therefore, model auditing and introspection are crucial to ensure sensible and fair predictions.

2.1.2 Convolutional neural networks

While fully-connected layers and MLPs are only able to process data in vectorized, thus tabular format, other types of layers have been developed, allowing the analysis of more complex domains with higher dimensional data such as images (Lecun et al. 1998), text (Vaswani et al. 2017) or time-series (Hochreiter and Schmidhuber 1997; Jordan 1986) data. Neural networks for image analysis commonly consist of convolutional layers, which allow the analysis of two- and three-dimensional inputs and exacerbate spatial information. Specifically, convolutional layers parameterize a trainable filter mask or *kernel* \mathbf{K} of specified dimensionality, which is applied over the input \mathbf{X} in a convolution operation producing the output feature map or *activation map* \mathbf{Y} :

$$\mathbf{Y}_{ij} = \sum_m \sum_n \mathbf{X}_{i-m, j-n} \mathbf{K}_{m,n} \quad (2.6)$$

where $\mathbf{K} \in \mathbb{R}^{M,N}$ and $\mathbf{X} \in \mathbb{R}^{I,J}$ with $I > M$ and $J > N$. The convolution operation is essentially the sum of the elements of the Hadamard product of the kernel and the matching selection of the input \mathbf{X}_i , computed in a sliding fashion over the input.

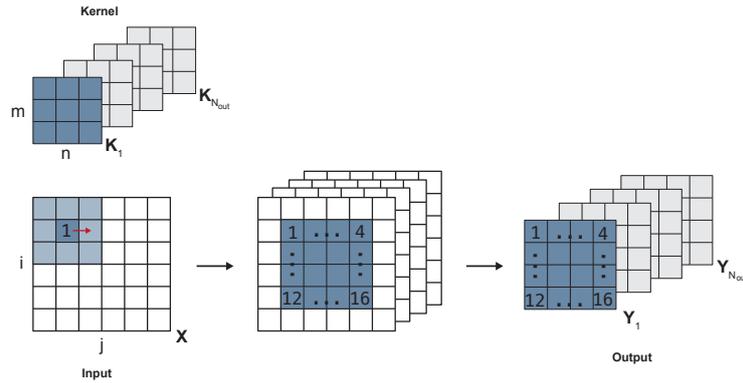


Figure 2.2: The convolution operation. Schematic depiction of the convolutional operation. An input image $\mathbf{X} \in \mathbb{R}^{i \times j}$ is convoluted with a kernel $\mathbf{K} \in \mathbb{R}^{m \times n \times N_{out}}$. Each kernel is applied individually on the input image generating a dimension of the output \mathbf{Y} .

In practice, this operation is often three-dimensional as the image has RGB channels, and a layer may have $N_{features}$ kernels, resulting in a $N_{features}$ -dimensional output per input pixel. Thus for a batch the input to a convolutional layer is a four-dimensional tensor $\mathbf{X} \in \mathbb{R}^{N_{batch}, N_{height}, N_{width}, N_{features}}$.

Convolutions possess a number of inherently desirable capabilities which have led to their dominance in image analysis and other related fields. Firstly, the sliding window approach of convolutions is conceptually close to mammal vision replicating concepts such as translation invariance of recognized shapes and objects. The nature of the convolutional operation as the sum of the Hadamard product of a given region in the input and the kernel also inherently satisfies the prior of spatial relativity, thus, the assumption that information in close proximity must be relevant. Further, convolutional layers are very parameter efficient as the same kernel is repeatedly applied over the input. A convolutional layer thus only requires $N_{inputfeatures} \times M \times N \times N_{outputfeatures} + N_{outputfeatures}$ parameters, where $N_{inputfeatures}, M, N_{outputfeatures} \ll I, J$, while a fully connected layer required one parameter for each pixel in the input space, thus $I \times J \times N_{outputfeatures}$ parameters.

Thus, convolutional layers learn filter masks to extract meaningful features from the inputs. Convolutional neural networks commonly consist of many stacked convolutional layers. Here, layers closer to the input layer extract low-level features, such as edges, while layers higher up in the network combine these features to retrieve aggregated information, such as a particular shape. In practice, convolutional layers are often applied along pooling operations. These layers essentially apply a kernel with a predefined operator, such as max,

2 Methodology

min, or average pooling, to reduce the dimensionality of the feature maps. The application of pooling layers increases the receptive field of layers higher up in the network allowing these layers to combine information over a wider spatial area of the input. At the same time, the reduced dimensionality in the feature maps results in decreased computational requirements (Tan and Le 2019). For the implementation of convolutional neural networks, this study relied on the PyTorch Python package (Paszke et al. 2017), and models were trained on NVIDIA A100 and V100 Graphics Processing Unit (GPU)s.

2.1.3 Residual neural networks and ConvNeXt

Residual neural networks or ResNets, were introduced by He et al. in 2016, are a very popular and highly capable form of neural network architecture originally proposed to circumvent the vanishing gradient problem hampering the training of large CNN architectures (He et al. 2016). Older, very large neural networks suffered from low gradients in backpropagation resulting from low and zero-valued derivatives of the applied activation functions. The ResNet architecture alleviated this problem by grouping the layers of a network (commonly convolutions) into *blocks* and introducing *residual connections* or *skip connections*, which add the input feature map of a block to its output (see Figure 2.3). Importantly, the information in this residual connection is not fed through an activation function, resulting in a higher overall derivative of the block, thus counteracting the vanishing gradients problem.

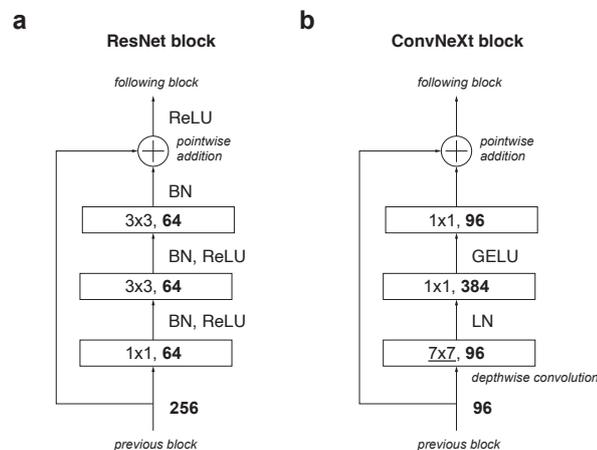


Figure 2.3: Architectures of the ResNet and ConvNeXt blocks. Displayed are flow diagrams depicting the abstract architectures of the ResNet and ConvNeXt blocks. **a**, The ResNet block relies on batch normalization and ReLU activations. **b**, Compared to the ResNet block, the ConvNeXt block utilizes layer normalization and a single GELU activation for improved gradient flow and decreased computation. Further, the ConvNeXt block utilizes a depthwise convolution in the first layer of the block and generally larger kernel sizes.

The ConvNext architecture introduced by Liu et al. is the latest evolution of the residual neural network principle, modifying the layers in the residual block to improve the performance. Specifically, Liu et al. introduce an inverse bottleneck within the block while reducing the number of channels passed between blocks from 256 to 96 and introducing a depthwise convolution with a larger kernel size of 7x7 in the first layer of the block. The ConvNeXt block generally has fewer activation functions, replaces Rectified Linear Unit (ReLU) with Gaussian Error Linear Unit (GELU) (Hendrycks and Gimpel 2016), and relies on layer normalization (Ba, Kiros, and Hinton 2016) instead of batch normalization (Ioffe and Szegedy 2015). Liu et al. introduce three different sizes of ConvNeXt networks with 29 million to 350 million parameters. This work relies on the ConvNeXt-small architecture for image inputs sized 384x384 pixels with 50 million parameters.

2.1.4 Activation functions

Activation functions or non-linearities are functions applied on the feature maps of neural networks to allow for non-linear combinations of the input features. Originally, neural networks have been trained using the sigmoid function commonly denoted as σ , which is also known from logistic regression equation (2.7). The sigmoid function is defined from $-\infty$, ∞ and thus maps any real value to the interval $]0, 1[$:

$$\sigma(x) = \frac{1}{1 + \exp^{-x}} = \frac{\exp^x}{1 + \exp^x} \quad (2.7)$$

Due to the interval of its output $]0, 1[$, the sigmoid function is often applied on the output layer, and interpreted as a probability. The applicability of the sigmoid function as an activation function for hidden layers is limited by its derivative being always lower than 1, and 0 for most real values, which has been identified as the main cause of the vanishing gradients problem (Bengio, Simard, and Frasconi 1994).

Another common activation function applied on the output layer of neural networks is the softmax function. The softmax function defines a categorical probability distribution over its inputs and is often applied for classification.

$$\text{softmax}(\mathbf{x})_i = \frac{\exp^{x_i}}{\sum_{j=1}^K \exp^{x_j}} \quad (2.8)$$

A very common activation function for hidden layers is the Rectified Linear Unit (ReLU) (Nair and Hinton 2010). The ReLU is defined as

$$\text{ReLU}(x) = \begin{cases} 0 & \text{if } x < 0 \\ x & \text{else} \end{cases} \quad (2.9)$$

For $x < 0$ the derivative of the ReLU is zero, while it is the slope for $x > 0$. While technically, the derivative at $x = 0$ cannot be computed, it is just set to zero in practice (Ian Goodfellow and Yoshua Bengio and Aaron Courville 2016). One issue encountered with ReLU activations is the dying ReLU problem, where neurons would only output zeros for any given input. The dying ReLU problem is caused by exploding gradients pushing the neurons' weights to output large negative values, thus regions where the ReLU function is zero. The neuron would not be able to recover as the derivative of the ReLU at these regions is zero as well. Despite these drawbacks, the ReLU activation function is frequently applied and has become one of the standard activation functions used in deep learning.

One of the first modifications to the ReLU activation concerned the circumvention of the dying neurons problem. The leaky ReLU (Maas, Hannun, and Ng 2013) introduces a small but positive slope α with $0 < \alpha < 1$ for negative inputs where the ReLU applies a slope of zero.

$$\text{Leaky ReLU}(x) = \begin{cases} \alpha x & \text{if } x < 0 \\ x & \text{else} \end{cases} \quad (2.10)$$

The exact value for α is commonly set prior to training and not learned but may be tuned as a hyperparameter.

Other modifications of the ReLU concern the smoothness of the curve and its derivatives. An early, smooth alternative to the ReLU activation function is the softplus activation function. The softplus activation function is defined as

$$\text{softplus}(x) = \log(1 + \exp^x) \quad (2.11)$$

and the derivative of the softplus function is the sigmoid function, providing much smoother derivatives as compared to the ReLU activation (Glorot, Bordes, and Bengio 2011).

Another smooth alternative to the ReLU activation function is the Gaussian Error Linear Unit (GELU) that relies on the cumulative distribution of a standard normal Gaussian (Hendrycks and Gimpel 2016):

$$GELU(x) = x\Phi(x) = x\frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{x}{\sqrt{2}}\right)\right] \quad (2.12)$$

where erf is the error function defined as $\operatorname{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z \exp^{-w^2} dw$. In contrast to the ReLU, the GELU weights the inputs by their percentile rank rather than gating them by whether they are larger or smaller than zero, leading to an overall smoother curve. To date, GELU are commonly applied in large transformer-based architectures, such as GPT-3 and BERT (Brown et al. 2020; Devlin et al. 2018).

In the same work that introduced GELUs, Hendrycks and Gimpel also proposed the Sigmoid Linear Unit (SiLU). The SiLU is similar in shape to the GELU but requires less computation, as it is defined as relying on the sigmoid function (Hendrycks and Gimpel 2016).

$$SiLU(x) = x\sigma(x) = x\frac{1}{1 + \exp^{-x}} \quad (2.13)$$

SiLUs have been applied in transformer-based architectures (Srinivas et al. 2021) and reinforcement learning (Elfwing, Uchibe, and Doya 2018).

Another recently proposed activation function is the Scaled Exponential Linear Unit (SELU) (Klambauer et al. 2017). The SELU is defined as

$$SELU(x) = \begin{cases} \lambda\alpha(\exp^x - 1) & \text{if } x \leq 0 \\ \lambda x & \text{else} \end{cases} \quad (2.14)$$

where λ and α are carefully determined factors defined as $\lambda \approx 1.6733$ and $\alpha \approx 1.0507$. The SELU was designed to improve internal normalization in neural networks as the choice of factors ensures that activations with approximately zero mean and unit variance will converge towards zero mean and unit variance when propagated through the network layers with SELU activations (Klambauer et al. 2017).

2.1.5 Regularization

Neural networks are inherently overparameterized, as they often contain vastly more parameters than training samples. While neural networks are thought to possess a form of internal regularization, many regularization methods have been proposed to enhance the generalizability, increase training stability and prevent overfitting. One of the most fundamental methods is dropout (Srivastava et al. 2014), where neurons are disabled or “dropped” from the computation with a given dropout probability. The selection of the disabled neuron is performed on each step individually and neurons are removed with both their incoming and

2 Methodology

outgoing connection from the forward and backward passes during training. By enforcing redundancy in model weights, dropout increases the robustness of the model prediction and thereby helps neural networks to generalize. During test time, the dropout probability is set to zero, and all neurons are activated. Further, dropout implicitly learns ensemble models, which may be sampled during test time to approximate model uncertainty (Gal and Ghahramani 2015). Another common form of regularization, commonly known as weight decay, is the application of the l^2 norm on the parameter values (equation (2.15)). The addition of the regularization term to the objective function and its minimization forces the parameter values closer to the origin. This form of regularization is similar to the Ridge regression (Hoerl and Kennard 1970). The l^2 norm is defined as

$$l^2(\mathbf{x}) = \sqrt{\sum_i^N x_i^2} = \|\mathbf{x}\|_2 \quad (2.15)$$

Other implicit forms of regularization include the application of early stopping, which involves tracking a metric, commonly the objective function, on the validation set and stopping the training once no significant improvement is detected over a defined number of epochs. This number of epochs is referred to as the grace period.

Neural networks have also been shown to train more stability if multiple objectives are minimized simultaneously. In multi-task learning, otherwise exploding gradients for one sub-objective may be regularized by the influence of the other training objectives (Ian Goodfellow and Yoshua Bengio and Aaron Courville 2016). Another common and very effective technique applied to improve the robustness and generalization of neural networks is data augmentation. In data augmentation, fake input samples are generated by distorting and manipulating existing samples to introduce more data variance. Dataset augmentation is straightforward for classification tasks and very commonly applied in the imaging domain. Here, augmentation methods include rotations, crops, noise induction, manipulations of the color and brightness as well as more advanced methods accounting for label information, such as cut-mix (Yun et al. 2019). A detailed overview of image augmentation methods is provided in section 2.3.1 on page 55. For non-image domains, such as tabular data augmentation might not be applicable as beneficial variances are often unknown, and methods for their introduction may not exist.

2.1.6 Normalization

As activation functions are optimized for receiving inputs in ranges close to the origin, internal normalization of neural networks is very important. While normalization would ideally be applied over the entire dataset, this is incompatible with the batch mode in which stochastic gradient optimization operates. Thus, solutions have been proposed to apply normalization over either subparts of data or subparts of the neural network. A very common form of normalization applied in deep learning is batch normalization (Ioffe and Szegedy 2015), where each batch of data is standardized to zero mean and unit variance by calculating the mean $\mu_B(\mathbf{x})$ and variance $\hat{s}_B(\mathbf{x})^2$ over each dimension of the data \mathbf{x} in a given batch B .

$$BN(\mathbf{x}) = \frac{\mathbf{x} - \mu_B(\mathbf{x})}{\sqrt{\hat{s}_B(\mathbf{x})^2 + \epsilon}} \boldsymbol{\gamma} + \boldsymbol{\beta} \quad (2.16)$$

where ϵ is added for numerical stability and $\boldsymbol{\gamma}$ and $\boldsymbol{\beta}$ are learnable parameters of the same dimension as the input \mathbf{x} . During training, a batch normalization layer commonly keeps running estimates of the mean $\mu_B(\mathbf{x})$ and variance $\hat{s}_B(\mathbf{x})^2$ vectors, which are updated with a momentum parameter m according to the update rule (equation (2.17)) during training,

$$\hat{s}_t = \hat{s}_{t-1}(1 - m) + m s_t \quad (2.17)$$

and utilized for normalization during inference at test time.

Another common form of normalization is layer normalization (Ba, Kiros, and Hinton 2016). In contrast to batch normalization, where the mean and variance over each dimension of the data in a given batch, layer norm computes these metrics per sample S in the batch, $\mu_S(\mathbf{x})$ and $\hat{s}_S(\mathbf{x})^2$. Layer norm is commonly used in transformer architectures and recently proposed convolutional architectures such as the ConvNeXt architecture (see section 2.1.3 on page 44) (Z. Liu, Mao, et al. 2022).

2.2 Survival models

2.2.1 Kaplan-Meier Estimator

The Kaplan-Meier estimator is one of the most fundamental tools of survival analysis. Proposed in 1958 by Kaplan and Meier, the estimator allows a non-parametric assessment of the survival function from observed data (E.L. Kaplan and P. Meier 1958). The estimator $\hat{S}(t)$ is calculated as a product of the fractions of the number of events d_i and the number of individuals at risk n_i over all observed event times t_i up to the time point of interest t :

$$\hat{S}(t) = \prod_{t_i \leq t}^N 1 - \frac{d_i}{n_i} \quad (2.18)$$

Thereby, the Kaplan-Meier estimator at the time point 0 is set to 1, $\hat{S}(0) = 1$ to reflect that the entire population is initially alive. Kaplan-Meier Estimators are commonly used to estimate and compare survival functions, often for the comparison of survival functions among different groups of individuals, for instance, between treated and untreated or high and low-risk individuals.

2.2.2 Cox Proportional Hazards model

The Cox Proportional Hazards (CPH) model is a linear regression model for time-to-event analysis first proposed by Cox et al. in 1972 (Cox 1972). The model is specified by a time-dependent baseline hazard function $h_0(t)$, which is modified by the *partial hazard* h_i to calculate the hazard for an individual i at a given time point t :

$$h_i(t|\mathbf{x}) = h_0(t)h_i = h_0(t)\exp(\boldsymbol{\beta}^T \mathbf{x}_i) \quad (2.19)$$

where \mathbf{x}_i is the vector of covariates for individual i and $\boldsymbol{\beta}$ is the vector of weights specifying the model. Taking the logarithm of the CPH hazard function, the linear relationship between the covariates and weights becomes evident:

$$\log h_i(t|\mathbf{x}) = \log h_0(t) + \beta_1 x_{i_1} + \beta_2 x_{i_2} + \dots + \beta_n x_{i_n} \quad (2.20)$$

Importantly, the central assumption of the CPH model is that the partial hazard h_i is constant and does not depend on time. Thus, the baseline hazard h_0 is scaled with a constant to yield the hazard for an individual. This assumption gives way to the possibility of computing constant hazard ratios between groups of samples, which are also constant over time:

$$HR = \frac{h_{group1}}{h_{group2}} = constant \quad (2.21)$$

Following equation (2.19) the survival function $S(t|X_i)$ is estimated via the $H(t|X_i)$,

$$H(t|\mathbf{x}_i) = \int_0^t h_0(u)\exp(\boldsymbol{\beta}^T \mathbf{x}_i) du = \exp(\boldsymbol{\beta}^T \mathbf{x}_i)H_0(t) \quad (2.22)$$

and the basic relationship between cumulative hazard function and survival function,

$$S(t|\mathbf{x}_i) = \exp(-H(t|\mathbf{x}_i)) = \exp(-H_0(t))\exp(\boldsymbol{\beta}^T \mathbf{x}_i) = S_0(t)\exp(\boldsymbol{\beta}^T \mathbf{x}_i) \quad (2.23)$$

The likelihood of observing an event at time point t for individual i is given by the cox likelihood function as the ratio of the individual hazard at the given time point and the sum of the hazards of all other individuals at risk at the given time point:

$$L_i(\boldsymbol{\beta}) = \frac{h(t|\mathbf{x}_i)}{\sum_{j:t_j \geq t} h(t|\mathbf{x}_j)} = \frac{h_0(t)\exp(\boldsymbol{\beta}^T \mathbf{x}_i)}{\sum_{j:t_j \geq t} h_0(t)\exp(\boldsymbol{\beta}^T \mathbf{x}_j)} = \frac{\exp(\boldsymbol{\beta}^T \mathbf{x}_i)}{\sum_{j:t_j \geq t} \exp(\boldsymbol{\beta}^T \mathbf{x}_j)} \quad (2.24)$$

As the baseline hazard is constant at a given time point for all individuals, the Cox likelihood is only dependent on the partial hazards. The overall likelihood of the CPH model is the product of the individual likelihoods as samples are assumed to be independent:

$$L(\boldsymbol{\beta}) = \prod_i L_i(\boldsymbol{\beta}) \quad (2.25)$$

Following this assumption, the log partial likelihood of the observed events is given by

$$l(\boldsymbol{\beta}) = \log \left[\prod_{i=1;\delta_i=1} \frac{h_0(t)\exp(\boldsymbol{\beta}^T \mathbf{x}_i)}{\sum_{j:t_j \geq t} h_0(t)\exp(\boldsymbol{\beta}^T \mathbf{x}_j)} \right] = \sum_{i=1;\delta_i=1} \left[\boldsymbol{\beta}^T \mathbf{x}_i - \log \sum_{j:t_j \geq t} h_0(t)\exp(\boldsymbol{\beta}^T \mathbf{x}_j) \right] \quad (2.26)$$

and the model is optimized by minimizing the negative log partial likelihood $-l(\boldsymbol{\beta})$.

Importantly, the CPH model can be optimized without specifying the baseline hazard function. However, the baseline hazard is commonly implemented by Weibull distributions, and spline-based methods have been proposed as well (Etezadi-Amoli and Ciampi 1987). Further, the CPH model may be implemented with regularization methods commonly applied to linear regression models (Tibshirani 1997). Although the CPH model is fairly easy to interpret thanks to the linear relationship between the covariates and training the model is feasible even with low sample sizes due to the low number of parameters, there are several limitations. Firstly, the CPH model natively does not account for tied events. Extensions have been proposed, relying on averaging partial hazards for tied events (Breslow 1974; Efron 1977). Second, the proportional hazards assumption ignores time-dependent effects on covariate importance, thus, a covariate is constrained to have the same effect regardless of the time. This assumption is violated in many real-world scenarios, and methods for assessing its validity have been proposed (Hess 1995). Third, the linear relationship between

covariates may fail to pick up higher-order interactions encountered in real-world settings.

All CPH models were implemented with the lifelines Python package (Davidson-Pilon et al. 2021). For the CPH model with Polygenic Score (PGS) and age interaction, an additional interaction term (PGS*age) was added for every polygenic score predictor variable section 4 on page 87.

2.2.3 Deep Surv

In order to overcome the linearity constraint of the CPH model, Faraggi and Simon proposed a neural network-based extension of the proportional hazards model (Faraggi and Simon 1995). Specifically, they suggested replacing the linear relationship between the weights β and the covariates \mathbf{X} with a function $g(\mathbf{X})$, parameterized by a neural network:

$$h_i(t|\mathbf{x}_i) = h_0(t)h_i = h_0(t)\exp(g(\mathbf{x}_i)) \quad (2.27)$$

While computational resources significantly hampered the application of neural networks at the time of the proposal, Katzmann et al. revisited the idea in the deep learning millennia (Katzman et al. 2017). Applying the neural network $g(\mathbf{X})$, the model minimizes the negative log-likelihood of the CPH model (equation (2.26) on the previous page) as follows:

$$l(\theta) = - \sum_{i=1; \delta_i=1} \left[g_\theta(\mathbf{x}_i) - \log \sum_{j:t_j \geq t} \exp^{g_\theta(\mathbf{x}_j)} \right] \quad (2.28)$$

In practice, the loss function is implemented to minimize the cumulative sum of the partial log-likelihood for all observed events in a given batch. DeepSurv has been applied to scalar and categorical feature sets (D.W. Kim et al. 2019; Liang et al. 2020; Matsuo et al. 2019; Yousefi et al. 2017) as well as in combination with complex convolutional architectures extracting features from images (Mobadersany et al. 2018) and videos (G.A. Bello et al. 2019). As DeepSurv may be combined with arbitrary feature extractors, it allows the utilization of high-dimensional inputs for risk modeling. For this study, the DeepSurv model was implemented in a custom Python implementation, following the specifications in the original DeepSurv repository by Katzman et al., and the reimplemention of the loss by Kvamme et al.. Details on the metabolomic state model and the retina risk model, both relying on the DeepSurv objective are provided in section 2.6 on page 67.

2.2.4 Deep Survival Machines

Deep Survival Machines proposed by Nagpal et al. are a fully parametric model to estimate survival and incidence from covariates with time-varying effects (Nagpal, X.R. Li, and Dubrawski 2021). The model is specified by a weighted mixture of *primitive* distributions, where both the weights of the mixture as well as the *primitive* distribution are parameterized by neural networks.

Thus the survival function for an individual is given by

$$S_i(t|\mathbf{x}_i, \theta, \phi, \psi) = \sum_{k=1}^K \sigma_{\theta}^{[k]}(\mathbf{x}_i) S_k(t|\mathbf{x}_i, \phi, \psi) = \sum_{k=1}^K \sigma_{\theta}^{[k]}(\mathbf{x}_i) (1 - F_k(t|\mathbf{x}_i, \phi, \psi)) \quad (2.29)$$

where K is the dimensionality of the mixture, θ, ϕ, ψ are weights of three neural networks $n_{\theta}, g_{\phi}, h_{\psi}$, and $\sigma_{\theta}^{[k]}$ is the value of a softmax function over the K dimensional output of n_{θ} at the k -th index. The cumulative incidence function $F_k(t|\mathbf{x}_i, \phi, \psi)$ is the cumulative density function of the utilized *primitive* distribution parameterized by two neural networks g_{ϕ} and h_{ψ} . Analogous to the individual survival function, the incidence function of an individual may be estimated by

$$f_i(t|\mathbf{x}_i, \theta, \phi, \psi) = \sum_{k=1}^K \sigma_{\theta}^{[k]}(\mathbf{x}_i) f_k(t|\mathbf{x}_i, \phi, \psi) \quad (2.30)$$

where the incidence function $f_k(t|\mathbf{x}_i, \theta, \phi, \psi)$ is the probability density function of the *primitive* distribution parameterized by the same two neural networks g_{ϕ} and h_{ψ} , and $\sigma_{\theta}^{[k]}$ is the same softmax output.

Thus, the total likelihood of the Deep Survival Machines model for the censored data is given by:

$$L_{total}(\theta, \phi, \psi) = L_{censored}(\theta, \phi, \psi) \times L_{uncensored}(\theta, \phi, \psi) \quad (2.31)$$

$$= \prod_{i=1}^{D_{\delta_i=0}} S_i(t|\mathbf{x}_i, \theta, \phi, \psi) \times \prod_{i=1}^{D_{\delta_i=1}} f_i(t|\mathbf{x}_i, \theta, \phi, \psi) \quad (2.32)$$

While the model theoretically allows the utilization of any function with a definition in the basic reals and a closed-form solution of the cumulative density function as *primitive*

2 Methodology

distribution, Nagpal et al. explored the application of the Lognormal and Weibull distributions. The model is optimized by maximizing the evidence lower bound of the total model log likelihood (equation (2.31) on the previous page) tuned with a regularization parameter $\alpha \in [0, 1]$ to adjust the loss for the degree of censoring in the dataset. The censored loss is computed for the censored samples, while the uncensored loss is only computed for the samples with observed events, thus where $\Delta = 1$:

$$\mathcal{L} = \alpha L_{censored} + L_{uncensored} \quad (2.33)$$

The parameters for the *primitive* distributions, thus the weights of the neural networks g_ϕ and h_ψ are learned jointly with the weights of the mixture, i.e., the parameters of the neural network n_θ . However, in the original implementation, the parameters of the *primitive* distributions are computed by

$$a_{k_i} = \tilde{a}_k + \varphi(g_\phi(\mathbf{x}_i)) \quad (2.34)$$

and

$$b_{k_i} = \tilde{b}_k + \varphi(h_\psi(\mathbf{x}_i)) \quad (2.35)$$

where φ is an activation function, h_ψ is a neural network and \tilde{b}_k is a pre-trained base-parameter. Nagpal et al. utilize the *SELU* and *tanh* activation functions for the Weibull and Log-Normal *primitive* distributions respectively. The base-parameters are estimated by fitting a singular (i.e. $K = 1$) *primitive* distribution to the train dataset. The parameters of the *primitive* distributions are required to be tightly controlled as they are subject to positivity constraints, which break the model when violated. Importantly, in this work, a different implementation is proposed: Instead of relying on the additional pre-training step, a *softplus* activation function is applied on the output of the neural networks g_ϕ and h_ψ . This simplifies the workflow, as no pre-training is necessary.

So far, Deep Survival Machines have been applied on tabular medical datasets for the estimation of breast cancer risk and competing risks of breast cancer (Nagpal, X.R. Li, and Dubrawski 2021) as well as on time series data from intensive care units to predict the duration of stay and the short term life expectancy (Nagpal, Jeanselme, and Dubrawski 2021). For this study, DeepSurvivalMachines have been implemented in Python in a custom implementation, details are provided in section 2.6 on page 67.

2.2.5 Model calibration

In survival modeling, calibration describes the agreement between a model's predicted probabilities for the occurrence of an event in a specific timeframe and the actually observed probabilities of the event in a given population. As risk models are commonly developed on observational data originating from a specific population sample, calibration is not always guaranteed when the model is to be applied to unseen data. In fact, when models are transferred between populations, risk models are often prone to over- or underestimate the actual event probabilities. Once miscalibration has been detected, recalibration of the model's predictions may help the model's predictions to match the actual event probabilities in the new population. One possible means of recalibration is through the application of an isotonic regression model to reweigh the miss-calibrated model's predictions to match the observed event probabilities more closely (Davidson-Pilon et al. 2021). Isotonic regression describes a curve-fitting approach with monotonically increasing (or decreasing) values in the output space (Chakravarti 1989; De Leeuw 1977). To apply isotonic regression for calibration, the regression model is fit on the predicted probabilities to regress on the observed probabilities. It is important to note that the set of observed event probabilities is discrete, as observed event probabilities are derived from population subgroups with similar predicted risks (see section 2.4.5 on page 63). Therefore, interpolating the observed risks with cubic splines may be helpful. As fitting the isotonic regression model requires information on the observed event time and indicator, this must not be done on the test dataset to prevent information leakage.

In this study, models were recalibrated after the completion of the training on the training dataset. First, model predictions were sorted by magnitude, and restricted cubic splines with 3 knots were applied to interpolate the observed probabilities from the binary event indicators. Subsequently, an isotonic regression model was fit on the training set using the `sklearn.isotonic.IsotonicRegression` method of the scikit-learn Python package (Pedregosa et al. 2011). At test time, the regression model was then treated as an additional model layer to output the final calibrated cumulative risks.

2.3 Image analysis

2.3.1 Image augmentation methods

In deep learning for image analysis, dataset augmentation is a crucial element of regularization. Over the years, many data augmentation methods have been developed to introduce specific variances into the data and generate invariances beneficial for generalization.

2 Methodology

Thereby, dataset augmentation methods are often tailored specifically towards the task at hand, for instance, introducing rotational variances in the data in order for the network to learn rotation-invariant patterns. One of the fundamental augmentation methods includes rotating the image by a specified, often random number of degrees or flipping the image across a given axis. Other methods involve image degradation through noise. Often Gaussian blurs are applied, or pixels are randomly set to maximum or minimum intensities. Images may also be altered on their color channels (i.e., the RGB values) or their brightness and intensity (i.e., in the HSV space) by color jittering, solarization, or posterization. The image may also be permuted in its dimensions and information content by cutting out and resizing parts of the image. Random crops, cutting at a random location in the image, the center crop, and cropping and recombining parts of the image to generate new unseen samples are also common augmentation methods. In practice, the construction of augmentation pipelines is often empirical and follows the researcher’s intuitions on inductive biases. Image manipulations may be done on the CPU prior to feeding the data onto the GPU as implemented in PyTorch’s torchvision package (Paszke et al. 2017) or directly on the GPU as implemented in the kornia package (Riba et al. 2020). In this study, image augmentation is applied for training of the retina risk model section 7 on page 141. table 2.1 provides an overview on the augmentation methods applied in this study and their specific functions. More details such as the order of augmentation methods and the specific parameters applied are provided in section 2.6 on page 67.

Table 2.1: Augmentation methods applied in augmentation of the retinal fundus photographs.

Augmentation method	Function
Random resize crop	Crops the image at a random location, based on a given aspect ratio to the original image, then resized the crop to a given size.
Random horizontal flip	Flips the image horizontally with random probability.
Random vertical flip	Flips the image vertically with random probability.
Random greyscale	Applies greyscale conversion with random probability.
Random posterization	Applies posterization with random probability.
Random sharpness adjustment	Applies sharpness adjustment with a specified factor with a random probability.
Random autocontrast	Applied contrast adjustment with a specified factor under a random probability.

2.3.2 Canny edge detection operator

The Canny edge detection algorithm is a multi-stage computational approach for edge detection in images (Canny 1986). It is defined by first converting the image to Grayscale and then applying a 5×5 Gaussian filter to smooth the image. Subsequently, an edge detection operator, such as the Sobel-Feldmann operator (equation (2.36)) (Sobel 1968) is applied to derive image gradients in the horizontal and vertical directions, S_x , S_y .

$$S_x = \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix} S_y = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{pmatrix} \quad (2.36)$$

Based on the direction-specific gradients, the direction ∇ and intensity $|S|$ of the edge can be computed, as

$$|S| = \sqrt{S_x^2 + S_y^2} \quad (2.37)$$

and

$$\nabla(x, y) = \text{atan2}\left(\frac{S_y}{S_x}\right) \quad (2.38)$$

where atan2 is defined as

$$\text{atan2}(x, y) = \begin{cases} \arctan\left(\frac{y}{x}\right) & \text{if } x > 0 \\ \arctan\left(\frac{y}{x}\right) + \pi & \text{if } x < 0, y \geq 0 \\ \arctan\left(\frac{y}{x}\right) - \pi & \text{if } x < 0, y < 0 \\ \frac{\pi}{2} & \text{if } x = 0, y > 0 \\ -\frac{\pi}{2} & \text{if } x = 0, y < 0 \end{cases} \quad (2.39)$$

After edge computation, the Canny algorithm applies non-maximum suppression to thin out the derived edges, followed by double thresholding categorizing pixels in strong, weak, and non-relevant. Strong pixels have high intensities and are considered edge-relevant, while non-relevant pixels are those below both thresholds. Those pixels in between thresholds are categorized as weak pixels and are further analyzed by hysteresis. In hysteresis, a 3×3 mask is applied on each weak pixel, setting it to a strong pixel if another strong pixel is the 3×3 neighborhood. The canny edge detection algorithm is readily implemented in the scikit-image Python package (Walt et al. 2014).

2.4 Evaluation metrics

2.4.1 Concordance index

The concordance index or Harrell’s C-index is a common metric employed to scrutinize the goodness of fit for risk models in survival analysis (Harrell et al. 1982). The C-index is a measure for ranking accuracy considering the time of an individual event, censoring, and the predicted risk. The underlying thought is that individuals at high predicted risk should be experiencing events earlier than individuals at low predicted risk. Pairs for which this ranking is correct are considered *concordant* pairs, while pairs where this is not the case, are considered *discordant*. Thus for computation, the C-index requires the assessment of all eligible pairs of samples i and j , where $i \neq j$ with predicted risks ρ_i and ρ_j and times T_i and T_j , where $T = \min C^*, T^*$ with censoring time C^* and event time T^* . Eligible pairs are pairs of samples where either both samples have had an observed event (thus $\Delta = 1$ for i and j) or one of the two samples has had an event *before* the censoring time of the other sample. In this case the pair is considered concordant if $\rho_i > \rho_j$ and $T_i < T_j$ and discordant if otherwise. The C-index is then computed by calculating the ratio of the number of concordant pairs over the number of all possible pairs:

$$c = \frac{N_{concordant}}{N_{concordant} + N_{discordant}} = \frac{\sum_{i \neq j} \mathbb{1}\{\rho_i > \rho_j\} \mathbb{1}\{\delta_i = 1\} \mathbb{1}\{T_i < T_j\}}{\sum_{i \neq j} \mathbb{1}\{\delta_i = 1\} \mathbb{1}\{T_i < T_j\}} \quad (2.40)$$

The C-index has values between 0 and 1, with 1 being a perfect ranking. Values close to 0.5 indicate random performance, while values close to 0 indicate perfectly inverse ranking. Further, from the notation in 2.40, it becomes apparent that the C-index strongly depends on the number of eligible pairs, which in turn depend on the number of observed events. This has several consequences limiting the capabilities of the C-index. First, C-indices are dataset and population-dependent, rendering inter-dataset comparisons futile. Furthermore, in datasets with high degrees of censoring, the C-index depends largely on the few samples with observed events and might not provide an accurate picture of the ranking over the entire population. As a consequence, the implications of the C-index are not linear over its scale, and a 0.1 increase from $C_1 = 0.79$ to $C_2 = 0.80$ is much more difficult to achieve than a C-index increase from $C_1 = 0.65$ to $C_2 = 0.66$. In conclusion, C-index values should only be compared on the same datasets to avoid misinterpretation. For all experiments in this study, C-indices are computed using the lifelines Python package (Davidson-Pilon et al. 2021).

2.4.2 Net reclassification improvement

The Net Reclassification Improvement (NRI) is a measure assessing how well a model categorizes individuals into risk groups compared to another model or a given standard. The net reclassification improvement was proposed in 2007 by Pencina et al., with the intention to allow for a fine-grained assessment of risk-factor utility in risk prediction models (Pencina, D’Agostino, D’Agostino, et al. 2008). Pencina et al. argue that while statistical significance is a prerequisite for the inclusion of risk factors in prediction models, statistical significance does not necessarily imply a better risk prediction. Further, actually significant and even causal risk factors might only marginally improve the performance of risk models when measured via metrics such as the C-index or the Area under the receiver operating characteristic. Thus, the net reclassification improvement is meant to provide means to assess the effect on the predictions of a risk model upon risk factor or biomarker addition. The underlying thought behind the net reclassification improvement is that clinical practice often relies on risk categories to guide interventions. Thus, any model improving on correctly assigning these risk categories is to be preferred. If, for example, in cardiovascular prevention, statins are recommended based on a certain risk threshold $\tau_{20} = 0.2$, dividing the population into high-risk $\rho_i \geq \tau_{20}$ and low-risk $\rho_i < \tau_{20}$, model performance may be evaluated by how many individuals with events are correctly assigned into this risk category, and how many individuals without events are falsely assigned into this risk category. This allows the assessment of the number of individuals that a proposed model B correctly up classifies and down classifies over an existing model A :

$$NRI_{events}(A, B) = \frac{\#correctly\ up - \#incorrectly\ down}{\#events} = \quad (2.41)$$

$$\frac{\sum_i^N \mathbb{1}\{\rho_{A_i} < \tau\} \mathbb{1}\{\rho_{B_i} \geq \tau\} \mathbb{1}\{\delta_i = 1\} - \sum_i^N \mathbb{1}\{\rho_{A_i} \geq \tau\} \mathbb{1}\{\rho_{B_i} < \tau\} \mathbb{1}\{\delta_i = 1\}}{\sum_i^N \mathbb{1}\{\delta_i = 1\}} \quad (2.42)$$

2 Methodology

$$NRI_{non-events}(A, B) = \frac{\#correctly\ down - \#incorrectly\ up}{\#non-events} = \quad (2.43)$$

$$\frac{\sum_i^N \mathbb{1}\{\rho_{A_i} \geq \tau\} \mathbb{1}\{\rho_{B_i} < \tau\} \mathbb{1}\{\delta_i = 0\} - \sum_i^N \mathbb{1}\{\rho_{A_i} < \tau\} \mathbb{1}\{\rho_{B_i} \geq \tau\} \mathbb{1}\{\delta_i = 0\}}{\sum_i^N \mathbb{1}\{\delta_i = 0\}} \quad (2.44)$$

Consequently, the net reclassification improvement of model A over model B is computed as the sum of the NRIs for events and non-events:

$$NRI(A, B) = NRI_{events}(A, B) + NRI_{non-events}(A, B) \quad (2.45)$$

The NRI has been extended to account for censored data by reformulating the 2.41 and 2.43 as probabilistic measures and applying Bayes' rule:

$$NRI = P(up | event) - P(down | event) \quad (2.46)$$

$$+ P(down | non-event) - P(up | non-event) \quad (2.47)$$

$$= \frac{P(event | up) \times N_{up} - P(event | down) \times N_{down}}{N_{total} \times P(event)} \quad (2.48)$$

$$+ \frac{P(non-event | down) \times N_{down} - P(non-event | up) \times N_{up}}{N_{total} \times P(non-event)} \quad (2.49)$$

The probabilities for events and non-events are then estimated using Kaplan-Meier Estimators (Pencina, D'Agostino, and Steyerberg 2011). Ultimately, while the NRI is a strongly relied-on metric in the medical statistical literature, it has known weaknesses. For instance, the categorical NRI is very sensitive to event rates, and the NRI generally has been shown to be overconfident, stating significant improvements even for the addition of uninformative biomarkers (Pepe et al. 2015). The categorical net reclassification index was calculated with the `nricens` R package (Inoue 2018).

2.4.3 Net benefit

The net benefit is a holistic measure allowing for assessing a recommendation's effects over a given population. A recommendation based on a risk model's prediction may entail measures such as interventional medication or treatments. As these interventions and treatments

commonly have side effects, adverse or not, the aim is to recommend interventions only to those individuals in need, thus to those at risk. In order to stratify a population based on a continuous risk score, a threshold τ is applied. The choice of the threshold depends on the disease and the benefit B obtained by treating an individual and the harm H caused by treating an individual without the disease:

$$\tau = \frac{B}{B + H} \quad (2.50)$$

Individuals with predicted risk $\rho_i \geq \tau$ would be recommended for intervention, while individuals with predicted risk below the threshold would be left untreated. This allows the assessment of true positives and false positives of the threshold-based model recommendation against the observed events:

$$N_{TP} = \sum_i^N \mathbb{1}\{\rho_i \geq \tau\} \mathbb{1}\{\delta_i = 1\} \quad (2.51)$$

$$N_{FP} = \sum_i^N \mathbb{1}\{\rho_i \geq \tau\} \mathbb{1}\{\delta_i = 0\} \quad (2.52)$$

Relying on these numbers, the net benefit for the treated population is calculated as follows

$$\mathcal{NB}_{treated} = \frac{N_{TP}}{N} - \frac{N_{FP}}{N} \times \frac{\tau}{1 - \tau} \quad (2.53)$$

While models may be compared by their net benefit of the treated, it is recommended to additionally assess the benefit of the untreated (Rousson and Zumbo 2011). Similar to the benefit of the treated, the benefit of the untreated is calculated relying on the estimated true negatives and false negatives,

$$N_{TN} = \sum_i^N \mathbb{1}\{\rho_i < \tau\} \mathbb{1}\{\delta_i = 0\} \quad (2.54)$$

and

$$N_{FN} = \sum_i^N \mathbb{1}\{\rho_i < \tau\} \mathbb{1}\{\delta_i = 1\} \quad (2.55)$$

by using a similar relationship with the threshold probability τ as for the net benefit of the treated,

$$\mathcal{NB}_{untreated} = \frac{N_{TN}}{N} - \frac{N_{FN}}{N} \times \frac{1 - \tau}{\tau} \quad (2.56)$$

While a model maximizing the net benefit for the treated will, per definition, also maximize the net benefit for the untreated, models differ in their effects for the treated and

untreated populations. Therefore Rousson et al. proposed the *overall net benefit*, a sum of the net benefit of the treated at threshold probability τ and the net benefit of the untreated at threshold probability τ (Rousson and Zumbrunn 2011):

$$\mathcal{NB}_{overall} = \frac{N_{TP}}{N} + \frac{N_{TN}}{N} - \frac{N_{FN}}{N} \times \frac{1-\tau}{\tau} - \frac{N_{FP}}{N} \times \frac{\tau}{1-\tau} \quad (2.57)$$

Similar to the individual net benefits of the treated and the untreated, the measure of the overall net benefit is equal to zero if benefit and losses compensate each other. While the benefit of the treated may be negative but is upper bound by the disease prevalence d , the benefit of the untreated is bound between $[0, 1 - d]$. Thus, the overall net benefit as the composite measure of the benefit of the treated and untreated may reach negative values and is upper bound by 1. An overall net benefit of 1 requires all individuals to benefit, thus a perfect prediction model. For all experiments, the net benefit was calculated with the `dca` R package (Sjoberg 2021).

2.4.4 Decision curve analysis

The composite measure of the overall net benefit accounts for the effect of an intervention recommended by, for instance, a prediction model, the treated and untreated in a population. However, the overall net benefit strongly depends on the choice of decision threshold τ , and the choice of the decision threshold may not always be clearly defined. Therefore, Vickers and Elkin et al. proposed the concept of decision curve analysis (Vickers and Elkin 2006). Decision curve analysis allows for assessing the benefit of an intervention over a range of possible decision thresholds. Decision curves plot the benefit (y-axis) against decision thresholds (x-axis) and facilitate interpretation by providing two extreme case scenarios per default. In the first edge case, none of the individuals are treated, resulting in zero true and false positives $TP = FP = 0$. In contrast, in the second edge case, all individuals are treated, resulting in as many true positives as prevalent cases $TP = d$ and accordingly, $FP = 1 - d$ false positives. In the first case, the “treat none” case, the benefit of the treated is always zero, irrespective of the threshold probability. In the second case, the “treat all” case, the benefit of the treated is $\mathcal{NB}_{treated} = d - (1 - d) \times \frac{\tau}{1-\tau}$, thus, a monotonically decreasing function of the decision threshold, with a maximum at the prevalence d . Between these two extreme cases, the benefit of the model can be plotted to determine at which decision thresholds d_i the benefit of the model is higher than the benefit of one of the two baseline scenarios. In a model achieving perfect prediction, the false positives and false negatives will equal zero, resulting in a net benefit of the treated of d , irrespective of the threshold probability τ . In practice, the net benefit of the treated will be a - not necessarily monoton-

ically - decreasing function of the threshold probability τ . By visualizing these functions, different models may be compared by their net benefits at different decision thresholds of interest. While extensions to the decision curve analysis have been proposed to consider the overall net benefit in place of only the benefit of the treated (Rousson and Zumbrunn 2011), common practice is still to consider the benefit of the treated (Vickers, Calster, and Steyerberg 2019; Vickers and Elkin 2006). Decision curve analysis was performed with the `dca` R package (Sjoberg 2021).

2.4.5 Calibration curve analysis

Model calibration refers to the agreement between a model's predicted probabilities and the observed probabilities in a given dataset. In a time-to-event setting, calibration is assessed at specific points in time, i.e., as the agreement between a model's predicted cumulative risk at a given time point and the observed event probability over the same timeframe. Importantly, a model may be well-calibrated on one dataset while over or under-estimating the actual probability in another dataset (see section 1.2.5 on page 14). Therefore it is vital to assess model calibration before relying on a model's absolute estimated risk. Most commonly, calibration in a population is assessed by visual examination of calibration curves (Austin, Harrell Jr, and Klaveren 2020). To calculate a calibration curve, samples in a population are binned into n subgroups based on their predicted event probability ρ_t at timepoint t , where t is the end of the observation window. Next, the observed event probability p_t over the observation window for each subgroup is calculated as the ratio of the observed events and the number of samples in the subgroup i :

$$p_t^{(i)} = \frac{N_{events}^{(i)}}{N_{samples}^{(i)}} \quad (2.58)$$

In the calibration curve, for each subgroup i , the observed event probability $p_t^{(i)}$ is then plotted against the average predicted probability $\rho_t^{(i)}$. Thus, ideal calibration results in a diagonal line through the origin with a slope of 1. A model is said to overestimate risk if the predicted probabilities are consistently higher than the observed probabilities and, vice versa, to underestimate risk. In this study, calibration curves are calculated based on a custom implementation of the code provided in (Austin, Harrell Jr, and Klaveren 2020).

2.4.6 Relative risk differences

Relative risk differences describe a measure to assess the normalized impact on individual risk predictions between two prediction models. For instance, relative risk differences allow

the assessment of the impact on the predicted risk resulting from the addition of a new variable (or multiple variables) \mathbf{y} to a given set of variables \mathbf{x} on the predicted risk of an individual i . Relative risk differences are defined as

$$RRD_i = \frac{f_z(\mathbf{z}_i) - f_x(\mathbf{x}_i)}{f_x(\mathbf{x}_i) + \epsilon} \quad (2.59)$$

where \mathbf{x} is a set of clinical covariates, $\mathbf{z} = \mathbf{x} \cup \mathbf{y}$ is another set of clinical covariates that is the union of \mathbf{x} and \mathbf{y} , f_x and f_z are two models trained on the covariate sets respectively and ϵ represents a small constant added for numerical stability. Relative risk differences are positive if the added information \mathbf{y} results in a higher risk estimate and negative if the added information leads to a lower risk estimate. Commonly, \mathbf{x} represents established clinical information, while \mathbf{y} are novel markers that are to be investigated. Relative risk differences for individuals with clinical risk ($f_x(\mathbf{x}_i)$) below 1 % are prone to numerical instabilities in calibration and should therefore be excluded. In this study, relative risk differences are applied to assess the impact of the addition of Polygenic Score (PGS) predictors to commonly assessed cardiovascular predictors on individual predicted risk trajectories (see section 4 on page 87).

2.5 Attribution methods

2.5.1 Shapley additive explanations

SHapley Additive exPlanations (SHAP) is an attribution method to explain individual predictions in neural networks (Lundberg and S.-I. Lee 2017). SHAP builds on the concept of Shapley values to quantify the impact of single features on an individual prediction and may be used to explain both local and global feature importances. Importantly, SHAP values fulfill three important conditions: local accuracy, missingness, and consistency (Molnar 2022). A single attribution under the explanation model $p_{\mathbf{w}}$ in SHAP is defined by

$$p(\mathbf{w}) = \alpha_0 + \sum_{j=1}^C \alpha_j w_j \quad (2.60)$$

where α_j is the attribution value for the feature w_j and $\alpha_0 = \mathcal{E}_{\mathbf{x}}(f(\mathbf{x}))$, the expected average prediction of $f(\mathbf{x})$, the model to be explained. The coalition vector $\mathbf{w} \in \{0, 1\}$ of length C , indicates if a feature is present 1 or missing 0. Local accuracy states, that the individual attribution values α_j for each member of a coalition w_j explain a given prediction $f(\mathbf{x})$

$$p(\mathbf{x}) = f(\mathbf{x}) = \alpha_0 + \sum_{j=1}^C \alpha_j \quad (2.61)$$

as all features in \mathbf{x} are present, all values in \mathbf{w} are 1.

The missingness property states that if a coalition member is missing ($w_j = 0$), the explanation value α_j is 0 as well. Further, the consistency property states that the attribution value α_j changes according to the marginal contribution of a feature, reflecting changes independently of other features. The computation time of exact Shapley values grows exponentially with an increasing number of features; therefore, numerous approximation methods have been proposed to improve computational efficiency. In this study, SHAP values were estimated using DeepSHAP, an adaptation of the DeepLIFT method (Shrikumar, Greenside, and Kundaje 2017). Calculations were performed using the DeepExplainer method implemented in version 0.39 of the shap package (Lundberg 2022). In order to derive the metabolite importances for the metabolomic state model (see section 5 on page 101, per-sample SHAP attributions were calculated for each metabolite and endpoint. Subsequently, attributions were aggregated per endpoint to derive a global metabolite-specific set of attributions. Importantly, individual attributions were considered to be important if the absolute SHAP value was in the top and bottom 1 % percentiles of the SHAP value distribution over all attributions.

2.5.2 Information bottleneck attribution

Information Bottleneck Attribution (IBA) is an image attribution method for deep neural networks proposed by Schulz and Sixt et al. (Sixt, Granz, and Landgraf 2020). IBA introduces an information bottleneck into a trained neural network to suppress unimportant information. Specifically, the information bottleneck learns parameters to replace information irrelevant for a prediction with noise while retaining the predictive, thus important information. The objective for the parameters of the information bottleneck is described by

$$\mathcal{L} = \mathcal{L}_{pred} + \beta \mathcal{L}_I \quad (2.62)$$

where \mathcal{L}_{pred} is the original prediction task and \mathcal{L}_I is a quantification of the mutual information in the inputs and the noise added by the bottleneck. Specifically, the bottleneck replaces intermediate features with noise of identical mean and variance following a linear interpolation:

$$\mathbf{Z} = \lambda(\mathbf{X})\mathbf{R} + (1 - \lambda(\mathbf{X})) \quad (2.63)$$

2 Methodology

where \mathbf{R} is the tensor of intermediate features and $\lambda(\mathbf{X})$ is a neural network predicting the parameters of the information bottleneck. The parameters of $\lambda(\mathbf{X})$ can be learned on an individual sample or the entire dataset. Ultimately, the importance measure is derived by evaluating the shared information between the intermediate features and the output of the bottleneck (i.e., the \mathcal{L}_I):

$$\mathcal{L}_I = \mathcal{D}_{KL}[P(\mathbf{Z} | \mathbf{R}) || Q(\mathbf{Z})] \quad (2.64)$$

where $P(\mathbf{Z}|\mathbf{R})$ is defined by the neural network of the information bottleneck and $Q(\mathbf{Z})$ is a variational approximation of the prior $P(\mathbf{Z})$, the added noise. In practice, the KL-divergence is evaluated for each dimension of \mathbf{R} independently and summed up to derive a heatmap of important features. Then the output is scaled up to the resolution of the original image. In this study, the original implementation of the IBA is applied, as defined in version 0.0.1 of the IBA Python package (Karl Schulz 2020): Specifically, IBA was calculated for a selection of six endpoints with significant added predictive information over basic demographic predictors. In order to generate attributions for an image, IBA was applied for 64 random crops (see figure 2.7 on page 71) with crop ratio 0.66 and images were subsequently aggregated to retrieve a full retinal fundus photograph. Repeating this procedure, attributions were calculated to maximize information, both, increasing and decreasing the predicted retinal state (i.e., the log partial hazards quantified by the adapted CPH (Faraggi and Simon 1995; Katzman et al. 2017)), receiving two independent attribution maps. For visualization, the direction of the aggregated attribution (represented by the colormap) was determined as the difference of the positive (risk contributing) and negative (risk reducing) attributions, and the strength of the attribution was determined as the maximum of the absolutes of both attribution maps (represented by the alpha variable in the visualization). To reduce border attribution effects and highlight all non-marginal features even more, attribution maps were masked circularly by the 32 outmost pixels in each direction. Colormaps were scaled to the 95 % quantile borders of the full-image attribution maps calculated against the retinal state adjusted for age and biological sex.

2.6 Model implementation

Based on the work presented in the results chapters section 4 on page 87 and section 5 on page 101 two manuscript have been prepared, submitted and published (Buerger, Steinfeldt, Looek, et al. 2022; Buerger, Steinfeldt, Ruyoga, et al. 2022). Along the respective results both manuscripts contain descriptions of the model implementations and methods. Detailed information on the contributors is presented at the beginning of the respective results sections and the published manuscripts.

2.6.1 NeuralCVD

The NeuralCVD model was implemented in Python 3.7 using PyTorch 1.7 (Paszke et al. 2017) and PyTorch-lightning 1.2. A Multi Layer Perceptron (MLP) with two fully connected linear layers with 256 and 128 nodes, batch normalization, dropout (Srivastava et al. 2014) of 0.1, Scaled Exponential Linear Unit (SELU) activations (Klambauer et al. 2017) is applied to learn a representation of the patient state. For details on the activation functions and MLPs see section 2.1.1 on page 40 and section 2.1.4 on page 45. In short, the combination of several fully connected linear layers each followed by a non-linear activation function, allows learning higher-order interactions as each weight in the upstream layer is exposed to each (non-linear) output of the previous layer. This representation is passed to an adapted implementation of the Deep Survival Machines (DSM) to parametrize a weighted mixture of four Weibull distributions with additional three linear layers for each of the parameters (Nagpal, Jeanselme, and Dubrawski 2021; Nagpal, X.R. Li, and Dubrawski 2021).

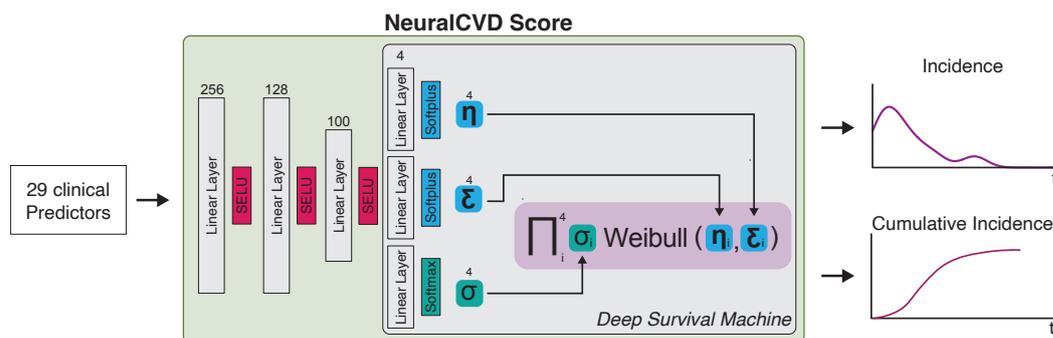


Figure 2.4: Architecture of the NeuralCVD score. Displayed is a pictogram of the architecture of the NeuralCVD score. NeuralCVD score builds on the architecture of Deep Survival Machines (Nagpal, X.R. Li, and Dubrawski 2021) to learn a patient representation from the input features to parametrize a mixture of Weibull distributions to model the incidence function over a continuous time scale. The NeuralCVD score outputs both cumulative incidence and survival functions, as well as the estimated risk.

2 Methodology

Importantly, the implementation deviates from the original DSM setup in discarding the pretraining and instead adding Softplus activations to the linear layers for the shape and scale parameters to enforce the Weibull distribution parameters' positivity constraint (see methods subsection section 2.2.4 on page 53). This setup proved more reliable in the face of high censoring rates. Like the original DSM setup, gradient descent is applied to minimize two negative log-likelihood losses, one for the censored events maximizing the survival probability at the censoring time, and another for the uncensored observations, maximizing the event probability at the observation time. In order to balance the losses, the loss for censored observations was weighted down by a factor of 0.5. The NeuralCVD model was developed in 22-fold spatial cross-validation as described in section 3.2.6 on page 81. After architecture development, hyperparameter search was run on train and validation splits of partition zero as random search over a constrained parameter space tuning batch size, initial learning rate, number of nodes in the layers of the feature extractor, size of its output vector and number of dimensions of the Weibull distribution mixture. Final models were trained for a maximum of 60 epochs using the Adam optimizer (Kingma and Ba 2015) with default parameters, stochastic weight averaging, a learning rate of 0.0005, and early stopping tracking the performance on the validation set. In the NeuralCVD model, risk at a specific time point ($f(t)$) and the cumulative risk ($F(t)$) is predicted by parameterizing the Weibull function with a forward pass through the entire model.

2.6.2 Metabolomic state model

The Metabolomic State Model is a residual neural network simultaneously predicting metabolomic states for each of the 24 endpoints. The metabolomic state is a scalar summary of an individual's risk for future onset of that specific endpoint. The architecture of the metabolomic state model consists of two integral parts, a shared network and an array of smaller endpoint-specific head networks figure 2.5 on the facing page. The shared neural network comprises three fully connected linear layers, each with batch normalization, dropout (Srivastava et al. 2014) of 0.3, and Sigmoid Linear Unit (SiLU) (Elfwing, Uchibe, and Doya 2017) activations with 256, 256, and 512 nodes. It outputs a representation of size 512, which is passed on to the endpoint-specific residual head networks. Thereby, each of the 24 residual head networks takes two inputs: the shared representation learned by the shared network and the original 168 metabolomic markers. Each residual head network consists again of two parts: First, a small Multi Layer Perceptron (MLP) with three layers of 256, 128, and 32 nodes with a dropout of 0.6, batch normalization and SiLU activations, which transforms the shared representation, and second, a skip-connection (He et al. 2016) network of three layers with 128, 128, and 32 nodes transforming the 168 metabolomic markers. The

outputs of both networks are subsequently added in a skip-connection and fed through another two-layer fully connected network of 128 and 128 nodes with a dropout of 0.6, batch normalization, and SiLU activations. For details on skip-connections, see section 2.1.3 on page 44. Then, the scalar metabolomic state is computed through a final single output linear layer with identity activation. For each endpoint, and thus for each metabolomic state, an adapted proportional hazards loss (Faraggi and Simon 1995; Katzman et al. 2017) is calculated individually, excluding prevalent events endpoint-specifically. The individual losses are averaged and then summed up to derive the final loss of the Metabolomic State Model.

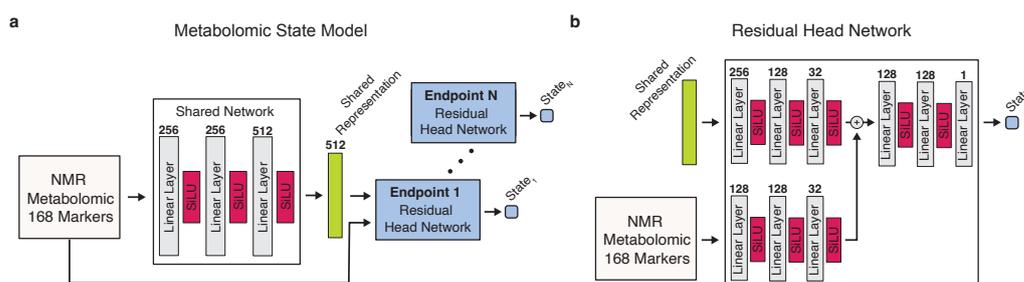


Figure 2.5: Architecture of the Metabolomic State Model. **a**, Overview on the residual architecture of the metabolomic state model. 168 circulating metabolomic markers are fed to the trunk network to learn a common shared representation. Endpoint-specific head networks then predict the metabolomic state for each endpoint from the shared representation and the input using a residual connection. **b**, Details of the individual endpoint-specific head network featuring residual connections.

After architecture development, hyperparameter search was run on train and validation splits of partition zero as random search over a constrained parameter space tuning batch size, initial learning rate, number of nodes in the layers of the endpoint heads, and size of the output vector of the shared network. The final models were trained with batch size 1,024 for a maximum of 100 epochs using the Adam optimizer (Kingma and Ba 2015) with default parameters, stochastic weight averaging, a learning rate of 0.001, and early stopping tracking of the performance on each partition's validation set. Further, a multi-step learning rate schedule with gamma 0.1 and steps at 20, 30, and 40 epochs was applied. The Metabolomic State Model was implemented in Python 3.7 using PyTorch 1.7 (Paszke et al. 2017) and PyTorch-lightning 1.4.

2.6.3 Medical history model

The medical history model is a multi-task neural network that relies on a binary representation of an individual's medical history to simultaneously predict log partial hazards for a set of 1,883 endpoints. The binary representation of an individual's prior medical history

2 Methodology

is a binary vector indicating whether each of a selection of 15,595 medical concepts was observed in this individual's medical records.

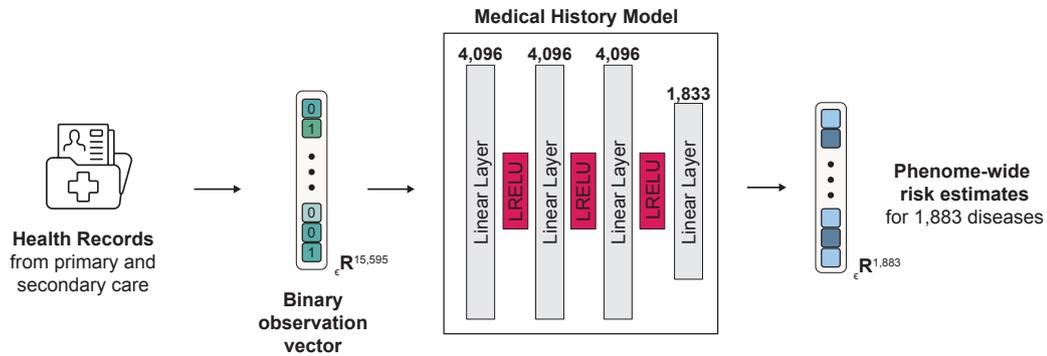


Figure 2.6: Architecture of the Medical History Model. Overview on the residual architecture of the medical history model. Medical concepts observed prior to baseline are mapped to OMOP standard concepts, deriving a binary observation vector of length 15,595. The vector is fed into the medical history model to predict the risk for incident disease onset for 1,833 endpoint simultaneously.

The medical history model consists of three fully connected linear layers, each with 4,096 hidden units, layer normalization (Ba, Kiros, and Hinton 2016), dropout (Srivastava et al. 2014), and leaky ReLU activations. The last layer outputs a 1,883-dimensional vector representing the endpoint-specific log partial hazard (the risk state). Similar to the metabolomic state model section 2.6.2 on page 68, for each endpoint, and thus for each risk state, an adapted proportional hazards loss (Faraggi and Simon 1995; Katzman et al. 2017) is calculated individually. The individual losses are averaged and then summed up to derive the final loss of the medical history model. After architecture development, hyperparameters were tuned using bayesian optimization on train and validation splits over a constrained parameter space, tuning batch size, learning rate, weight decay, number of nodes in the layers of the endpoint heads, number of hidden layers, dropout rates, and size of the output vector of the shared network. The final models were trained with a batch size of 512 using the Adam optimizer (Kingma and Ba 2015) with a learning rate of 0.0006 and weight decay of 0.3, and early stopping tracking of the performance on the validation set. We implemented the model in Python 3.9 using PyTorch 1.11 (Paszke et al. 2017) and PyTorch-lightning 1.5.5.

2.6.4 Retinal risk model

The retinal risk model is a deep convolutional neural network simultaneously predicting a scalar summary of the risk for future incident disease onset, i.e., a retinal state, for 1,171 endpoints. The retinal risk model consists of two major parts, a large convolutional neural network for feature extraction and a smaller head network to reweigh the features and

predict the retinal states. The convolutional feature extractor is the trunk of a ConvNext-small network (Z. Liu, Mao, et al. 2022) consisting of 53 million parameters pre-trained on the ImageNet dataset (Deng et al. 2009). Instead of the original fully connected prediction layer, a custom head network is added. The head network consists of three fully connected linear layers, the first two with 512 nodes, batch normalization, dropout (Srivastava et al. 2014) of 0.5, and leaky rectified linear units activations and the last one with 1,171 nodes and identity activation as out layer. For each endpoint, and thus for each retinal state, an adapted proportional hazards loss (Faraggi and Simon 1995; Katzman et al. 2017) is calculated individually, excluding prevalent events endpoint-specifically. The individual losses are averaged and then summed up to derive the final loss of the retinal risk model.

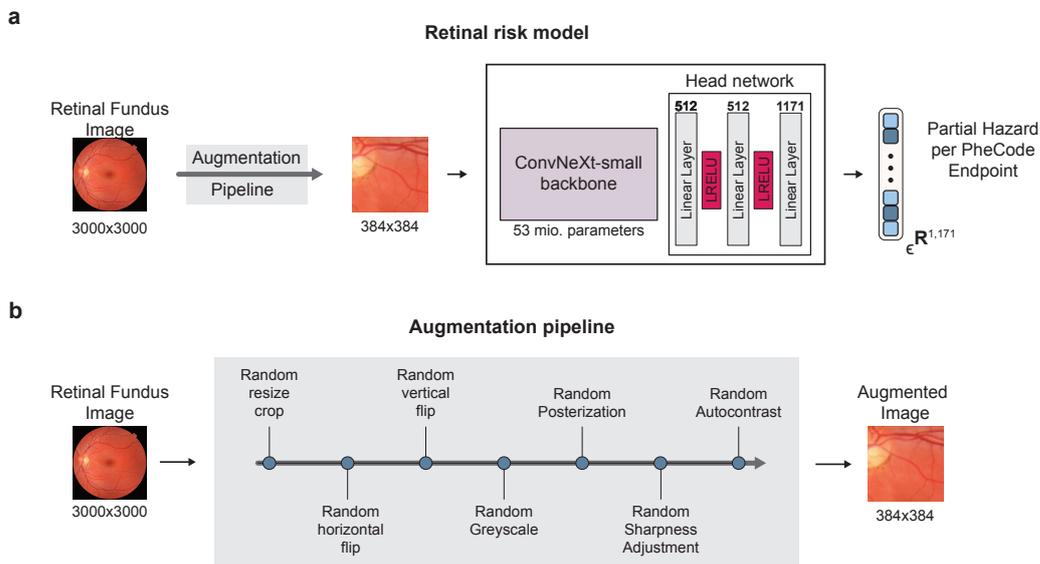


Figure 2.7: Architecture of the retinal risk model and the augmentation pipeline. **a**, The retinal risk model is a convolutional architecture based on the ConvNext-small model, with a custom head network predicting log partial hazards for 1,171 endpoints simultaneously from a 384x384 sized image input. **b**, For training, retinal fundus images are pushed through the augmentation pipeline sequentially, applying each augmentation with a random probability before being fed to the retinal risk model.

After architecture development, hyperparameter search was run on train and validation splits of partition zero as a random search over a constrained parameter space tuning batch size, initial learning rate, number of nodes in the head network, number of warmup epochs, and the weight decay. The final models were trained with a batch size of 256 for a maximum of 100 epochs using the Adam optimizer (Kingma and Ba 2015) with a weight decay of 0.001, a learning rate of 0.0001, a warm up period of 8 epochs, and early stopping tracking the performance on each partition’s validation set with a grace period of 40 epochs. The

2 Methodology

retinal state model was implemented in Python 3.9 using PyTorch 1.7 (Paszke et al. 2017) and PyTorch-lightning 1.4.

During training the model receives resized random crops of the original retinal fundus photographs to present the model with varying feature-resolutions. A random resized crop is a percentage-based random crop followed by a resize to a given pixel shape. During training, the model receives random resized crops with crop ratios drawn for each sample from a uniform distribution over $\{0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0\}$ resized to 420×420 pixels. The cropped images are subsequently fed through an extensive augmentation pipeline consisting of a random rotation for up to 25 degrees, a center crop to 384×384 pixels, a random horizontal flip with probability 0.5, a random vertical flip with probability 0.5, random application of greyscaling with probability 0.1, random posterization for 4 bits with a probability of 0.2, random sharpness adjustment with factor 2 and probability of 0.3, and random application of auto contrast with a probability of 0.3. All augmentations were implemented with the respective augmentation methods offered in `torchvision.transforms` as part of the PyTorch Python package (Paszke et al. 2017). During the validation epochs, the crop ratio was fixed at 0.8, and instead of applying the augmentation pipeline, images were only center cropped to 384×384 pixels.

Finally, at test time, the predictions on the test sets were generated using test-time augmentation of 100 samples per image with a crop ratio of 0.66 followed by a center crop of 384×384 pixels. Thus, at test-time, each image was fed 100 times through the augmentation pipeline and the network, retrieving 100 predictions. For each endpoint and image, the arithmetic mean of these 100 predictions was used as the final retinal state for this endpoint. More details on the foundations of the augmentations applied are provided in subsection 2.3.1.

3 Data sources and preprocessing

3.1 Data sources

3.1.1 UK Biobank

The UK biobank cohort, a sample of the UK's general population, is a large-scale population cohort with longitudinal follow-up (Bycroft et al. 2018). Participants were enrolled from 2006 to 2010 in 22 recruitment centers across the United Kingdom; the follow-up is ongoing, and records until the 24th of September 2021 are included in this analysis. The UK Biobank cohort comprises 273,353 women and 229,107 men aged between 37-73 years at the time of their assessment visit. Participants are linked to routinely collected records from primary care (GP), hospital episode statistics (HES), and death registries (ONS), providing longitudinal information on diagnosis, procedures, and prescriptions for the entire cohort from Scotland, Wales, and England. The UK Biobank contains demographic information, standard laboratory measures, questionnaires on lifestyle, and genotyping for all study participants. Further, the UK Biobank offers many complex measurements for sub-selections of participants. Contained are $^1\text{H-NMR}$ metabolomics, measured at recruitment for a random subset of individuals, 63,903 women and 54,078 men aged between 37 and 73 years at the time of their baseline assessment. The cohort also contains retinal fundus photographs taken during an eye examination at baseline for 85,223 participants. All models considered throughout this work were developed on UK Biobank data. Depending on the model, subsets of the UK Biobank cohort were used, and detailed descriptions of the study populations are provided in the respective results sections. A detailed description of the data preprocessing is provided in section 3.2 on page 76, and a description of the partitioning applied is provided in section 3.2.6 on page 81. Data from the UK Biobank cohort is utilized in all experiments performed in this study and details of the respective study populations, are presented in each results subsection individually.

3.1.2 Whitehall II

The Whitehall II cohort is an ongoing prospective cohort study of adults, consisting of 10,308 people (3,413 women and 6,895 men) recruited at ages 35–55 (Marmot and Brunner 2005). At the time of recruitment 1985–88, all study participants were employed at the offices of the 20 London Whitehall departments of the UK public service. Participants have been followed up regularly throughout the years, with questionnaire and self-examination conducted in 5 year intervals. Additionally, participants have been genotyped, and Proton Nuclear Magnetic Resonance ($^1\text{H-NMR}$) profiling was performed from serum samples between 1997-1999 for a random subset of the population. The Whitehall II cohort was utilized

in the validation process of the Metabolomic State model and a detailed description of the cohort population, as well as a comparison with the UK Biobank study population, can be found in table 8 on page 303.

3.1.3 PROspective Study of Pravastatin in the Elderly at Risk

The PROspective Study of Pravastatin in the Elderly at Risk (PROSPER) trial aimed to investigate the effects pravastatin (40 mg/day) in elderly individuals who are at risk of cvd (Shepherd, Blauw, Murphy, Cobbe, et al. 1999; Shepherd, Blauw, Murphy, and Gaw 2002). The trial was conducted as a double-blind, randomized, placebo-controlled trial including a total of 5,804 participants between the ages of 70 and 82. Participants were recruited in the primary care setting between December 1997 and May 1999 from 3 centers: Glasgow, Scotland; Cork, Ireland and Leiden, the Netherlands. The mean follow-up period was 3.2 years. All included patients either had evidence of vascular disease (physician-diagnosed stable angina, stroke, Transient Ischaemic Attack (TIA) or Myocardial Infarction (MI)), or high risk of vascular disease as determined by hypertension, diabetes, or smoking status. Fasting venous blood samples were collected at baseline and at 3-month intervals and biobanked at -80 °C. For the present study, all individuals recruited at the Leiden recruitment center, with Proton Nuclear Magnetic Resonance (1H-NMR) metabolomics data available through the BBMRI-NL consortium (in total 960 individuals), were included, employing the study as a cohort study. 1H-NMR metabolomics were quantified from previously unfrozen 6-month post-randomization samples. The PROSPER cohort was utilized in the validation process of the Metabolomic State model, and a detailed description of the cohort population, as well as a comparison with the UK Biobank study population, can be found in table 8 on page 303.

3.1.4 Leiden Longevity Study

The Leiden Longevity Study (LLS) includes data on 421 long-lived families of European descent. Families were included in the study if at least two long-lived siblings were alive and fulfilled the age criterion of 89 years or older for males and 91 years or older for females, representing < 0.5 % of the Dutch population in 2001 (Schoenmaker et al. 2006). In total, 944 long-lived proband siblings (mean age = 94 years, range = 89-104), 1,671 offsprings (mean age = 61 years, range = 39-81) and 744 spouses/partners thereof (mean age = 60 years, range = 36-79) were included. For all included participants, follow-up is available until the 27th of October 2016. 1H-NMR Metabolomics assays were conducted for 843 long-lived probands, 1,157 of their offsprings, and 684 controls using non-fasted EDTA plasma samples. For the experiments presented in this study, we included all 1,655 samples of the

3 Data sources and preprocessing

offspring and spouse population with complete baseline covariates and NMR metabolomics available in the BBMRI-NL platform. The LLS cohort was utilized in the validation process of the Metabolomic State model, and a detailed description of the cohort population, as well as a comparison with the UK Biobank study population, can be found in table 8 on page 303.

3.1.5 Rotterdam Study

The Rotterdam Study is a prospective, population-based cohort study (M.A. Ikram et al. 2017) that was set up with the aim of investigating the occurrence of common diseases in older adults. The study's baseline examination took place in 1990 with approximately 7,983 persons aged 55 years and older undergoing a home interview and extensive physical examination. Participants were re-visited every 3-4 years (RS-I) to collect follow-up information. The study extended in two stages as of 2008 contained 14,926 subjects in total. All participants provided written informed consent and the Medical Ethics Committee of the Erasmus Medical Center, Rotterdam, approved the study. ¹H-NMR metabolomics assays were conducted in fasted EDTA plasma samples. For this study, we included all 2,949 samples with complete baseline covariates and ¹H-NMR metabolomics that were available in the BBMRI-NL platform. The Rotterdam study cohort was utilized in the validation process of the Metabolomic State model, and a detailed description of the cohort population, as well as a comparison with the UK Biobank study population, can be found in table 8 on page 303.

3.2 Data preprocessing

3.2.1 Extraction of covariates and clinical predictors

Basic demographic predictors and clinical covariates were utilized in all experiments in this study. For the UK Biobank cohort, basic demographic information was extracted from primary care records and matched with data collected at the study's recruitment interview. Lifestyle information was extracted from the questionnaire completed at recruitment. Physical measurements and laboratory measures were taken at recruitment. Pre-existing medical conditions were extracted from the questionnaire, interview at recruitment, primary care records, and hospital episode statistics. Medications were extracted from the recruitment interview. Cardiovascular predictors were selected based on the ESC and AHA recommended cardiovascular risk scores for primary prevention, the AHA-ASCVD score (Goff et al. 2014), and the ESC-SCORE2 (SCORE2 working group and ESC Cardiovascular risk collaboration 2021) as well as the QRISK3 risk calculator (Hippisley-Cox, C. Coupland,

and Brindle 2017).

Table 3.1: Clinical predictors utilized in this study. A = Assessment, L = Blood Laboratory Assessment, Q = Questionnaire, I = Interview, P = Physical Assessment, HES = Hospital Episode Statistics, GP = Primary Care Records.

Predictor	Source	Coding	ASCVD	SCORE	QRISK3	PANEL	NeuralCVD	CAIDE	FINDRISC
Age at Recruitment	Q/I	Field ID: 21022	x	x	x	x	x	x	x
Sex	Q/I	Field ID: 31	x	x	x	x	x	x	x
Education	Q/I	Field ID: 21001				x		x	
Ethnicity	Q/I	Field ID: 21000	x		x	x	x		
Townsend Deprivation Index	Q/I	Field ID: 189			x		x		
Overall Health Rating	Q/I	Field ID: 2178					x		
Smoking Status	Q/I	Field ID: 20116	x	x	x	x	x		
Body-Mass-Index	P	Field ID: 21001				x	x	x	x
Weight	P	Field ID: 21002			x	x	x		
Standing Height	P	Field ID: 50			x	x	x		
Waist Circumference	P	Field ID: 48				x			x
Waist-Hip Ratio	P	Field ID: 48/49				x			
Physical Activity	Q/I	Field ID: 20116				x		x	x
Alcohol Intake	Q/I	Field ID: 1558				x			
Glucose	L	Field ID: 30740				x			
HbA1c	L	Field ID: 30750				x			
AST	L	Field ID: 30650				x			
ALT	L	Field ID: 30620				x			
AP	L	Field ID: 30610				x			
Albumin	L	Field ID: 30600				x			
Creatinine	L	Field ID: 30700				x			
Cystatine C	L	Field ID: 30720				x			
Urea	L	Field ID: 30670				x			
Urate	L	Field ID: 30880				x			
C-reactive Protein	L	Field ID: 30710				x			
Erythrocyte Count	L	Field ID: 30010				x			
Leukocyte Count	L	Field ID: 30000				x			
Platelet Count	L	Field ID: 30080				x			
Hemoglobin	L	Field ID: 30020				x			
Hematocrit	L	Field ID: 30030				x			
MCH	L	Field ID: 30050				x			
MCV	L	Field ID: 30040				x			
MCHC	L	Field ID: 30060				x			
Systolic Blood Pressure	A	Field ID: 4080	x	x	x	x	x		
Diastolic Blood Pressure	A	Field ID: 4079					x		
Total Cholesterol	A	Field ID: 30690	x	x	x	x	x	x	
HDL Cholesterol	A	Field ID: 30760	x		x	x	x		
LDL Cholesterol	A	Field ID: 30780				x	x		
Triglycerides	A	Field ID: 30870				x	x		

3 Data sources and preprocessing

Table 3.1 continued from previous page

Predictor	Source	Coding	ASCVD	SCORE	QRISK3	PANEL	NeuralCVD	CAIDE	FINDRISC
FH Heart Disease	Q/I	Field ID: 20107, 20110			x		x		
FH Type-2 Diabetes	Q/I	Field ID: 20107, 20110, 20111				x			x
Antihypertensive Treatment	Q/I	ATC: C02	x		x	x	x		x
Aspirin	Q/I	ATC: B01					x		
Atypical Antipsychotics	Q/I	ATC: N05			x		x		
Glucocorticoids	Q/I	ATC: H02			x		x		
Type-1 Diabetes	Q/I/GP/HES	ICD10: E10			x		x		
Type-2 Diabetes	Q/I/GP/HES	ICD10: E11, E12, E13, E14	x		x		x		
Chronic Kidney Disease	Q/I/GP/HES	ICD10: I12, N18, N19			x		x		
Atrial Fibrillation	Q/I/GP/HES	ICD10: I47, I48			x		x		
Migraine	Q/I/GP/HES	ICD10: G43, G44			x		x		
Rheumatoid Arthritis	Q/I/GP/HES	ICD10: J99, M05, M06, M08, M12, M13			x		x		
Systemic Lupus Erythematosus	Q/I/GP/HES	ICD10: M32			x		x		
Severe Mental Illness	Q/I/GP/HES	ICD10: F20, F25, F30, F31, F32, F33, F44			x		x		
Erectile Dysfunction	Q/I/GP/HES	ICD10: F52, N48			x		x		

For the NeuralCVD score, the QRISK3 variable set was extended with available information on the overall health status and the BMI. For the experiments involving the metabolomic state model, predictors were grouped into three sets of clinical predictors, Age+Sex, ASCVD, and the PANEL set (see section 5 on page 101). The PANEL predictor set contains, additional predictors from the CAIDE (Kivipelto et al. 2006) and FINDRISC (Lindström and Tuomilehto 2003) scores as well as comprehensive information on lifestyle, demographics, physical measurements, and lab values available in the primary care setting (see table 3.1 on the preceding page). The Age+Sex and ASCVD predictor sets were replicated on the PROSPER, LLS, RS, and Whitehall II cohorts for external validation. In each of these cohorts, data was already obtained in tabular format, and predictors did not require a dedicated extraction from primary care.

3.2.2 Calculation of PGS and the PGSMETA Score

Individual Polygenic Score (PGS) were calculated on the imputed genotyping data for all participants in the UK Biobank cohort and utilized in the experiments investigating the in-

tegration of PGS with clinical covariates (see section 4 on page 87). Predefined Polygenic Score (PGS), five for coronary artery disease (PGS000011 (Tada et al. 2016), PGS000018 (Inouye et al. 2018), PGS000057 (Natarajan et al. 2017), PGS000058 (Morieri et al. 2018) and PGS000059 (Hajek Catherine et al. 2018)) and one for stroke (PGS000039 (Abraham et al. 2019)) were selected from the PGS catalog (Lambert, Gil, et al. 2021) and calculated for all participants. The PGS were developed on multiple external cohorts and covered a diverse set of patients. The polygenic scores were developed on the Malmö Diet and Cancer Study (MDC, PGS000011), the West of Scotland Coronary Prevention Study (WOSCOPS, PGS000057), the Coronary Artery Risk Development in Young Adults (CARDIA, PGS000057) and BioImage cohorts, the Action to Control Cardiovascular Risk in Diabetes (ACCORD, PGS000058), the Outcome Reduction With Initial Glargine Intervention (ORIGIN, PGS000058), the Multi-Ethnic Study of Atherosclerosis (MESA, PGS000059) and the CARDIoGRAMplusC4D (PGS000018) and MEGASTROKE (PGS000039) Genome Wide Association Study (GWAS) results. All PGSs were calculated with the published weights from the PGS catalog (Lambert, Gil, et al. 2021), the imputed genotype information from the UK Biobank, and the R package PRSice-2 (S.W. Choi, Mak, and O'Reilly 2020). In order to be able to provide a holistic analysis of the overall genetic risk, the individual PGS are aggregated to a composite measure to construct a polygenic meta score (PGSMETA). PGSMETA is derived by summing up the individual percentile ranks for each of the six PGS scores and calculating a new aggregated percentile rank over the sum. All models are trained under the utilization of the six individual PGS scores, not the PGSMETA score.

3.2.3 Processing of NMR Metabolomics data

All experiments in this study involving Proton Nuclear Magnetic Resonance (¹H-NMR) metabolomic profiling rely on the same standardized assay offered by Nightingale Health, Inc. The assay covers 168 metabolites from multiple amino acids over lipids, lipoproteins, cholesterol subtypes, and inflammation markers (Soininen et al. 2015; Würtz, A.J. Kangas, et al. 2017). Additionally, the assay offers 81 precomputed percentage ratios derived from combinations of the 168 original measures; as these were not primary metabolite measurements, they were not included in the analysis. A full list of the metabolomic predictors applied in this study is presented in table 2 on page 245.

Prior to model development and training, the corpus of available ¹H-NMR metabolomics data in the UK Biobank was split into spatially decoupled train, test, and validation sets following the procedure described in section 3.2.6 on page 81. Subsequently, for each partition, the ¹H-NMR metabolomics data was imputed along the demographic and clinical

3 Data sources and preprocessing

information, then log-transformed and standardized to zero mean and unit variance. For external validation on the PROSPER, LLS, RS, and Whitehall II cohorts, 1H-NMR data was imputed and normalized utilizing the imputers and standardization procedures of the 22 models developed in the partitions of the UK Biobank cohort.

3.2.4 Processing of EHR data

In the UK Biobank cohort, participants' data has been linked to NHS health records, composed of primary care data, hospital episode statistics, and mortality information Bycroft et al. 2018; Sudlow et al. 2015. The linked primary care data originates from two different healthcare provider systems, EMIS and TPP, coded predominantly in CTV3, SNOMED CT, and dm+d. The data contains coded clinical events (including diagnoses, history, symptoms, lab results & procedures), prescriptions (i.e., medications that are prescribed but not necessarily dispensed), and a range of administrative codes (e.g., referrals to specialist hospital clinics). The hospital inpatient data originates from the Data Access Request Service (DARS) for England, the Secure Anonymised Information Linkage (SAIL) Databank for Wales, and the electronic Data Research and Innovation Service (eDRIS) for Scotland with records predominantly coded in ICD-10 and OPCS-4. The mortality data originates from the NHS Digital for participants in England & Wales and the NHS Central Register (NHSCR) for participants in Scotland, with causes of death coded in ICD-10.

For preprocessing, all these data sources were mapped to standard vocabularies of the OMOP CDM (SNOMED CT, RxNorm and CVX) using the mapping tables provided by NHS Digital, SNOMED International and the UK Biobank. The data was restricted to the OMOP domains of "Observation", "Condition", "Procedure", "Drug" and "Device" with 71,036 unique concepts in total. Subsequently, a threshold was applied on the observation frequency of each concept, retaining only concepts observed in more than 50 individuals. Thresholding yielded a subset of 15,595 unique concepts which was utilized for analysis. For each individual, these 15,595 unique concepts were aggregated in a binary observation vector informing on whether or not a concept is part of the individual's medical history. A full list of the 15,595 unique concepts and their definitions is provided in table 3 on page 248.

3.2.5 Retinal Fundus Photographs

Prior to model development and training, the corpus of available retinal fundus photographs was cleaned and preprocessed to exclude corrupted and faulty samples. Retinal images usually consist of a central, circular image structure representing the retinal disc, surrounded

by dark areas. However, as the retinal fundus images in the UK biobank were automatically taken during optical coherence tomography and not manually curated, the dataset contains faulty images, which are heavily underexposed, heavily overexposed, or may even show the eye from the outside. Therefore, the preprocessing pipeline included the removal of images where the retinal disc was not detectable, either because of over- or under-exposition, over- or under-saturation, lack of focus or frame, or due to strong shifts in either of the three color channels. In order to remove such faulty samples, a set of 1,000 valid images was first manually curated with the aid of ophthalmological domain experts. Next, each image of the dataset, including the manually curated samples, was first tested for a detectable retinal disc by first converting the image to grayscale and applying the Canny edge detection operator. If the retinal disc was detectable in the grayscale converted image, all areas outside the retinal disc in the original image were removed (i.e., set to NaN values), and the image was trimmed such that the retinal disc touched all four image borders. Next, distributions (mean and standard deviations) for the three RGB color channels and the three HSV channels were computed on the curated set. Subsequently, all images of the remaining dataset were compared in the means of their RGB and HSV channels to the distribution of the curated samples. All images with a mean outside the 2σ interval in any of the six assessed channels defined by the curated reference dataset, were excluded from the analysis. In total, the UK Biobank cohort contains 175,000 retinal fundus photographs for 85,000 participants from both the left and right eyes. Of these images, 136,000 items for 68,000 participants were measured at baseline, while the remainder of the images was measured 5 years after baseline. In order to relate the images to the observed outcomes, the analysis was restricted to images recorded at baseline. After the application of the preprocessing pipeline, the cleaned dataset contained 113,122 images for 61,264 participants measured at baseline (see table 3.2).

Table 3.2: Retinal fundus photographs available in the UK Biobank.

	Left Eye		Right Eye	
	Images	Participants	Images	Participants
UK Biobank	87,546	84,744	88,242	85,223
At Baseline	68,207	67,685	68,798	68,114
After Preprocessing	55,832	55,582	57,290	56,985

3.2.6 Dataset partitions and imputation

All experiments, thus, the development and testing of the NeuralCVD model, the Metabolomic State Model, and the Retinal Risk model, were conducted in spatial cross-validation in the UK Biobank cohort. For each experiment, the respective UK Biobank study population was

3 Data sources and preprocessing

split into 22 spatially separated partitions based on the location of the assessment center at recruitment (see figure 3.1). The data was analyzed in 22-fold nested cross-validation, setting aside one of the spatially separated partitions as a test set, aggregating the remaining partitions, and randomly selecting 10 % of the aggregated data for the validation set. Within each of the 22 cross-validation loops, the individual test set (i.e., the spatially separated partition) remained untouched throughout model development, and the validation set was used to validate the fitting progress and checkpoint selection. All 22 obtained models were then evaluated on their respective test sets, and model predictions were aggregated for downstream analysis.

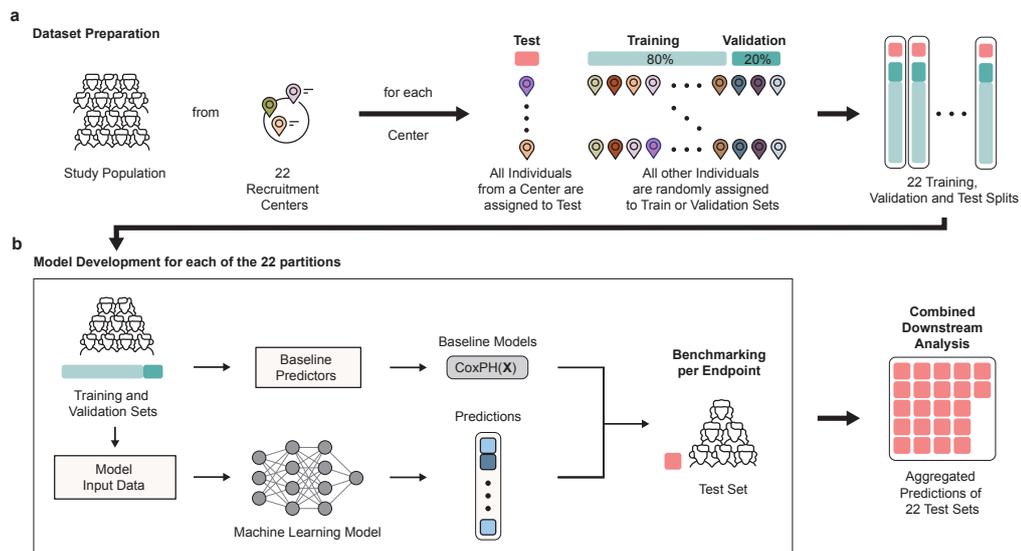


Figure 3.1: Spatial cross-validation in the UK Biobank cohort. **a**, To develop machine learning models, the eligible study population, a subset of the UK Biobank population, is split into train, validation and test sets in 22-fold nested cross-validation based on the assigned UK Biobank assessment center. **b**, For each of the 22 partitions, the machine learning model is trained on the input variables, and predictions are gathered for each partitions test set. Additionally, baseline models are trained on baseline predictors. Ultimately, predictions of the machine learning model and the baseline models are aggregated for benchmarking and downstream analysis.

Assuming random occurrence of missingness, data was imputed relying on multiple imputations using chained equations with random forests (Stekhoven and Bühlmann 2012). Continuous variables were standardized; Categorical variables were one-hot encoded. Imputation models were fitted on the training sets and applied to the respective validation and test sets. This procedure was independently repeated for each of the four described experiments.

3.3 Endpoint definitions

3.3.1 MACE endpoint

The cardiovascular endpoint of Major Adverse Cardiac Events (MACE) was defined based on International Classification of Diseases (ICD) codes. Specifically, by the earliest recorded event of fatal or non-fatal myocardial infarction (ICD10 codes I21, I22, I23, I24, I25) or fatal or non-fatal transient ischaemic attack or ischaemic stroke (ICD10 codes G45, I63, I64) either in the primary care records, the hospital episode statistics, or death records. The MACE endpoint was utilized to develop and validate a neural network-based risk model (NeuralCVD) integrating polygenic and clinical information to predict the 10-year cardiovascular disease risk in disease-free participants from the UK Biobank cohort (see section 4 on page 87).

3.3.2 Common disease endpoints

In order to investigate the added information and usability of 1H-NMR metabolomics in risk modeling of common diseases, a set of 24 endpoints was adapted from an earlier study: Pietzner et al. defined a set of common disease endpoints from ICD10 codes (Pietzner et al. 2021). The set was replicated in the UK Biobank cohort, defining each of the 24 endpoints by the earliest occurrence of one of its defining ICD10 codes in the primary care records, the hospital episode statistics, or death records of the UK Biobank cohort. A full list of the endpoint definitions can be found in table 3.3. For each endpoint, only disease-free individuals, thus individuals without a prior record of the endpoint, were considered eligible. Additionally, individuals with lipid-lowering therapy records were excluded in the case of cardiovascular endpoints. In the case of predominantly sex-specific diseases such as prostate and breast cancer, only men or women were analyzed respectively.

Table 3.3: Definitions, exclusions, and incident events of the common diseases endpoints in the UK Biobank cohort. All endpoints were defined based on ICD-10 code records in primary care, HES and death records. Exclusion criteria were applied to each endpoint specifically, and excluded individuals were not considered in the endpoint loss or any downstream analyses.

Abbreviation	Description	ICD 10 Codes	Exclusion Criteria	Eligible participants	Incident Event Rate (%)
Dementia	All-Cause-Dementia	F00, F01, F02, F03, G30, G31	Prior disease records	117245	2.06%
MACE	Major adverse cardiac event	G45, I21, I22, I23, I24, I25, I63, I64	Prior disease records, self reported lipid lowering therapy	92629	8.73%

3 Data sources and preprocessing

Table 3.3 continued from previous page

Abbreviation	Description	ICD 10 Codes	Exclusion Criteria	Eligible participants	Incident Event Rate (%)
T2 Diabetes	Type-2 Diabetes	E10, E11, E12, E13, E14	Prior disease records	111745	4.77%
Liver Disease	Liver Disease	B15, B16, B17, B18, B19, C22, E83, E88, I85, K70, K72, K73, K74, K75, K76, R18, Z94	Prior disease records	115468	5.16%
Renal Disease	Renal Disease	N00, N01, N02, N03, N04, N05, N06, N07, N08, N09, N10, N11, N12, N13, N14, N15, N16, N17, N18, N19, N25, N26, N27, N28, N29	Prior disease records	111704	10.70%
Atrial Fibrillation	Atrial Fibrillation	I48	Prior disease records	116026	5.68%
Heart Failure	Heart Failure	I50	Prior disease records	113636	3.53%
CHD	Coronary Heart Disease	I20, I21, I22, I23, I24, I25	Prior disease records	108551	7.86%
Venous Thrombosis	Venous Thrombosis	I80, I81, I82	Prior disease records	114986	1.70%
Cerebral Stroke	Cerebral Stroke	I63, I65, I66	Prior disease records	117342	1.85%
AAA	Abdominal Aortic Aneurysm	I71	Prior disease records	117826	0.77%
PAD	Peripheral Artery Disease	I70, I71, I72, I73, I74, I75, I76, I77, I78, I79	Prior disease records	115132	3.51%
Asthma	Asthma	J45, J46	Prior disease records	102392	3.79%
COPD	Chronic Obstructive Pulmonary Disease	J40, J41, J42, J43, J44, J47	Prior disease records	110465	7.90%
Lung Cancer	Lung Cancer	C33, C34	Prior disease records	117896	1.08%
Skin Cancer	Non-melanoma Skin Cancer	C44	Prior disease records	115996	4.24%
Colon Cancer	Colon Cancer	C18	Prior disease records	117617	1.04%
Rectal Cancer	Rectal Cancer	C19, C20	Prior disease records	117699	0.65%
Prostate Cancer	Prostate Cancer	C61	Prior disease records, females	53393	2.21%
Breast Cancer	Breast Cancer	C50	Prior disease records, males	62009	2.15%
Parkinson's	Parkinson's Disease	G20, G21, G22	Prior disease records	117764	0.61%

Table 3.3 continued from previous page

Abbreviation	Description	ICD 10 Codes	Exclusion Criteria	Eligible participants	Incident Event Rate (%)
Fractures	Fractures	S02, S12, S22, S32, S42, S52, S62, S72, S82, S92, T02, T08, T10	Prior disease records	106392	7.43%
Cataracts	Cataracts	H25, H26	Prior disease records	114345	9.67%
Glaucoma	Glaucoma	H40	Prior disease records	115934	2.31%

3.3.3 Phecode endpoints

While the MACE and common disease endpoints were defined by custom and domain knowledge specific remappings of individual ICD-10 codes, phecodes were applied for the assessment of risks on a phenome-wide scale. Phecodes aim to facilitate associative analyses with medical phenotypes by providing a mapping from ICD-10 and ICD-10-CM codes into clinically meaningful phenotypes (Wei et al. 2017; P. Wu et al. 2019). Thus, ICD-10 codes in the primary care records (GP), the hospital episode statistics (HES), or death records of individuals were mapped to PheCodes X by either mapping directly from ICD-10 (HES, death records) or mapping from SNOMED to ICD-10 (using the official SNOMED mapping table) and then to Phecodes X. All-cause death was defined as an additional, custom endpoint, due to the lack of an appropriate phecode. Additionally, congenital, developmental, and neonatal endpoints were excluded, as all UK Biobank participants are adults. For each endpoint, time-to-event outcomes were extracted, defined by the first occurrence after recruitment in primary care, the hospital episode statistics, or death records. Subsequently, for both experiments, the development of the medical history model and the retinal risk model, a population-specific frequency cutoff was applied to exclude endpoints recorded in less than 100 individuals in the respective study population. In the case of the medical history models, this procedure yielded 1,833 endpoints, on the subset of the UK Biobank population with retinal fundus photographs 1,171 endpoints were derived. A full table of all utilized endpoints is presented in the appendix: see table 4 on page 265 for both the endpoints used in the retinal risk model as well as for the endpoints used in the medical history model.

4 A Neural Survival model integrates polygenic information for cardiovascular risk prediction

Based on the work presented in this chapter a manuscript has been prepared, submitted and published in *The Lancet Digital Health* (Buerger, Steinfeldt, Loock, et al. 2022). The published manuscript was written and prepared by Thore Buerger, Dr. Jakob Steinfeldt, Professor John Deanfield, Professor Roland Eils and Professor Ulf Landmesser. Details on the individual author’s contributions to the publication are presented in the published manuscript.

4.1 Experimental Setup

This section presents the development and validation of a novel neural network-based cardiovascular disease risk model in the UK Biobank cohort. NeuralCVD (see figure 2.4 on page 67), based on Deep Survival Machines (Nagpal, X.R. Li, and Dubrawski 2021) relies on a set of 29 cardiovascular risk factors utilized in well-established clinical cardiovascular risk scores, the ESC score (Conroy 2003), the AHA/ASCVD score (Goff et al. 2014) and the QRISK3 score (Hippisley-Cox, C. Coupland, and Brindle 2017) to predict risk for incident Major Adverse Cardiac Events (MACE). A list of all utilized predictors is presented in table 3.1 on page 77, the endpoint definition is presented in section 3.3.1 on page 83 and a detailed description of the NeuralCVD score is provided in section 2.6 on page 67. In order to assess the composite of clinical and genetic predictors, an additional version of the NeuralCVD score is developed on an extended covariate set including six well-established Polygenic Score (PGS) against coronary artery disease and stroke (see section 3.2.2 on page 78). To benchmark the NeuralCVD approach and to evaluate the impact of the PGS-contained information on the predictive performance of cardiovascular risk, five additional models are trained: replications of the ESC-SCORE, AHA/ASCVD, and QRISK3 score, respectively, as well as two CPH models developed on the same predictor sets as the two NeuralCVD scores. All models are developed in 22-fold spatial cross-validation as described in section 3.2.6 on page 81 and evaluated against the occurrence of the first recorded MACE in the observation window. After training all models, the NeuralCVD score is scrutinized against the established cardiovascular risk scores. Further, the PGS’s added benefit is compared between the linear CPH model and the nonlinear NeuralCVD.

4.2 Characteristics of the study population

Information on demographics, clinical records, and outcomes were extracted from the entire UK Biobank cohort (Bycroft et al. 2018; Sudlow et al. 2015) (see section 3.1.1 on page 74).

4.2 Characteristics of the study population

From the 502,505 individuals, sixteen participants who withdrew their participation agreement, one participant without information on sex, and 106,775 with prior records of myocardial infarction, stroke, or lipid-lowering treatment were excluded (see figure 4.1).

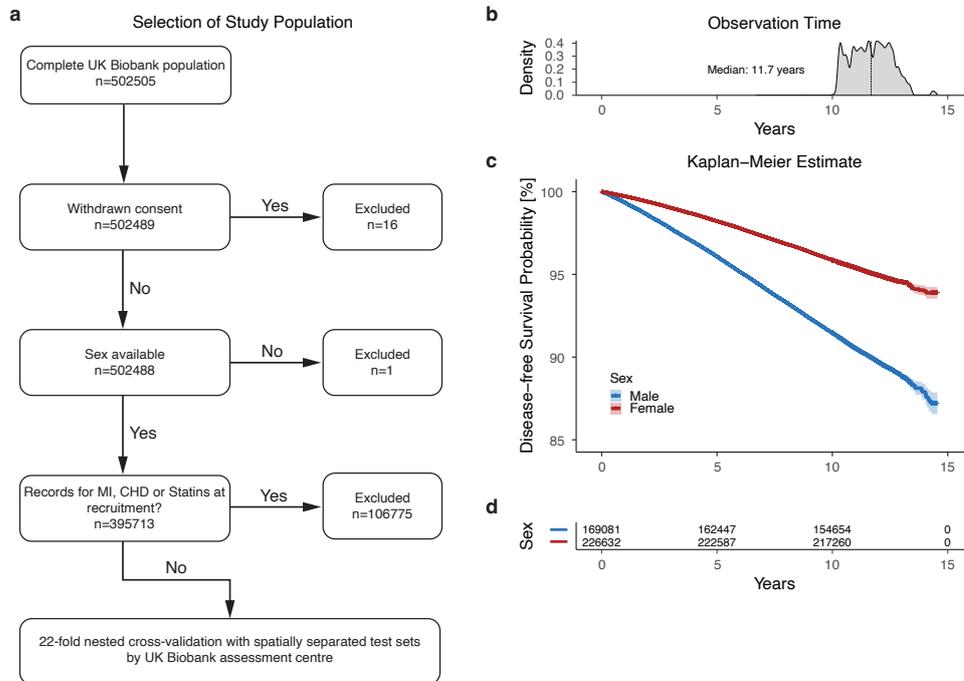


Figure 4.1: Selection and composition of the study population. **a**, Starting with the entire UK Biobank population, individuals who withdrew consent had missing information about their sex or had earlier records of incident myocardial infarction or stroke or lipid-lowering treatment at baseline were excluded. The remaining set was split into training, validation, and test sets in 22-fold nested cross-validation based on the UK Biobank assessment center. **b**, Distribution of observation times for the derived study population. The median observation time was 11.7 years (IQR 11.0, 12.3). **c**, Kaplan-Meier estimates for the disease-free survival functions stratified by sex. **d**, Numbers at risk in 5-year intervals stratified by sex.

The remaining 395,713 participants had a median age of 56 years (IQR 49, 62), 57.3 % were female, the majority being of White/British ethnicity, relatively affluent, and “good” self-reported overall health (table 4.1 on the following page). Of the study population, 11 % were current smokers, and 1.5 % were diagnosed with type-2 diabetes (table 4.1 on the next page). The median Body Mass Index (BMI) was 26.4 (IQR 23.8, 29.4), systolic blood pressure was 135 (IQR 124, 148) mmHg and total cholesterol was 5.84 (IQR 5.16, 6.56) mmol/l. Participants were observed for a median follow-up time of 11.7 years, and 28,083 (7.1 %) participants experienced a Major Adverse Cardiac Events (MACE), defined as Fatal + Non-fatal Myocardial Infarction (MI), Fatal + Nonfatal Transient Ischaemic Attack (TIA)/Stroke

4 A Neural Survival model integrates polygenic information for cardiovascular risk prediction

or Cardiovascular Death within the observation period (figure 4.1 on the preceding page).

Table 4.1: Summary statistics of the study population. Reported are medians (IQR) and absolute counts (%).

Characteristic	Male, N = 169,081	Female, N = 226,632	Overall, N = 395,713
Age at Recruitment	56 (48, 62)	56 (49, 62)	56 (49, 62)
Ethnicity			
Asian	3,523 (2.1%)	3,588 (1.6%)	7,111 (1.8%)
Black	2,782 (1.7%)	3,871 (1.7%)	6,653 (1.7%)
Chinese	491 (0.3%)	866 (0.4%)	1,357 (0.3%)
Mixed	892 (0.5%)	1,624 (0.7%)	2,516 (0.6%)
White	158,761 (95%)	213,464 (96%)	372,225 (95%)
Missing	2,632	3,219	5,851
Townsend Deprivation Index	-2.16 (-3.67, 0.53)	-2.19 (-3.67, 0.37)	-2.18 (-3.67, 0.44)
Missing	230	276	506
Overall Health Rating			
Excellent	30,892 (18%)	42,310 (19%)	73,202 (19%)
Good	98,290 (59%)	136,866 (61%)	235,156 (60%)
Fair	32,890 (20%)	39,057 (17%)	71,947 (18%)
Poor	5,813 (3.5%)	6,970 (3.1%)	12,783 (3.3%)
Missing	1,196	1,429	2,625
Smoking Status			
Current	21,454 (13%)	20,012 (8.9%)	41,466 (11%)
Previous	58,717 (35%)	69,074 (31%)	127,791 (32%)
Never	87,924 (52%)	136,326 (60%)	224,250 (57%)
Missing	986	1,220	2,206
Body-Mass-Index	26.9 (24.7, 29.5)	25.8 (23.2, 29.3)	26.4 (23.8, 29.4)
Missing	1,161	1,180	2,341
Weight (kg)	84 (76, 93)	68 (61, 78)	75 (66, 86)
Missing	1,026	1,105	2,131
Standing Height (cm)	176 (172, 181)	163 (158, 167)	168 (162, 175)
Missing	1,017	959	1,976
Systolic Blood Pressure (mmHg)	138 (128, 150)	132 (120, 146)	135 (124, 148)
Missing	10,193	13,679	23,872
Diastolic Blood Pressure (mmHg)	84 (78, 91)	80 (74, 87)	82 (75, 89)
Missing	10,192	13,678	23,870
Total Cholesterol (mmol/l)	5.72 (5.07, 6.41)	5.93 (5.24, 6.68)	5.84 (5.16, 6.56)
Missing	10,480	15,514	25,994
HDL Cholesterol (mmol/l)	1.26 (1.08, 1.47)	1.57 (1.34, 1.83)	1.43 (1.20, 1.70)
Missing	22,622	34,937	57,559
LDL Cholesterol (mmol/l)	3.67 (3.17, 4.20)	3.66 (3.13, 4.25)	3.67 (3.15, 4.23)
Missing	10,835	15,858	26,693
Triglycerides (mmol/l)	1.68 (1.16, 2.43)	1.30 (0.94, 1.84)	1.44 (1.02, 2.09)
Missing	10,659	15,634	26,293
FH Heart Disease	57,025 (34%)	90,847 (40%)	147,872 (37%)
Antihypertensive Treatment	1,697 (1.0%)	1,610 (0.7%)	3,307 (0.8%)
Aspirin	1,456 (0.9%)	935 (0.4%)	2,391 (0.6%)
Atypical Antipsychotics	2,086 (1.2%)	3,865 (1.7%)	5,951 (1.5%)
Glucocorticoids	122 (0.1%)	233 (0.1%)	355 (0.1%)

4.3 NeuralCVD outperforms established CVD-Risk models

Table 4.1 continued from previous page

Characteristic	Male, N = 169,081	Female, N = 226,632	Overall, N = 395,713
Type 1 Diabetes	795 (0.5%)	586 (0.3%)	1,381 (0.3%)
Type 2 Diabetes	3,379 (2.0%)	2,440 (1.1%)	5,819 (1.5%)
Chronic Kidney Disease	6,052 (3.6%)	8,253 (3.6%)	14,305 (3.6%)
Atrial Fibrillation	2,687 (1.6%)	2,303 (1.0%)	4,990 (1.3%)
Migraine	7,027 (4.2%)	20,879 (9.2%)	27,906 (7.1%)
Rheumatoid Arthritis	6,052 (3.6%)	16,916 (7.5%)	22,968 (5.8%)
Systemic Lupus Erythematosus	186 (0.1%)	725 (0.3%)	911 (0.2%)
Severe Mental Illness	14,303 (8.5%)	28,902 (13%)	43,205 (11%)
Erectile Dysfunction	7,731 (4.6%)	0 (0%)	7,731 (2.0%)

4.3 NeuralCVD outperforms established CVD-Risk models

In order to determine whether neural networks improved risk discrimination over conventional approaches, the NeuralCVD was first compared against established clinical baselines and a linear CPH model trained on the same 29 cardiovascular risk factors. All scores were evaluated independently on all 22 assessment centers of the UK Biobank cohort with the Concordance Index (see section 2.4.1 on page 58) and the categorical net-reclassification-improvement (see section 2.4.2 on page 59) at the 10 % threshold (as recommended in the NICE guidelines (National Institute for health and Care Excellence (NICE) 2014)) as metrics for the risk discrimination.

NeuralCVD score outperformed the SCORE, ASCVD, and QRISK3 models with a difference in C-Index of +0.037 (0.034-0.039), +0.024 (0.023-0.026), +0.01 (0.009-0.011) and Net Reclassification Improvement (NRI) of +0.1043 (0.0981-0.1103), +0.0704 (0.0648-0.0765), +0.0488 (0.0442-0.0534) respectively (see figure 4.2 on the next page, table 4.2 on the following page, table 2 on page 285). Furthermore, NeuralCVD improved over a linear CPH model fitted on the same set of covariates with a difference in C-Index of +0.003 (0.002-0.004) and NRI of +0.0469 (0.0429-0.0511) (see table 2 on page 285). The discrimination performance was stable over the 22 spatially distinct assessment centers (see figure 1 on page 202, table 4 on page 287), and all models were well-calibrated over the observed risk spectrum (see figure 4.2 on the next page and figure 1 on page 202).

4 A Neural Survival model integrates polygenic information for cardiovascular risk prediction

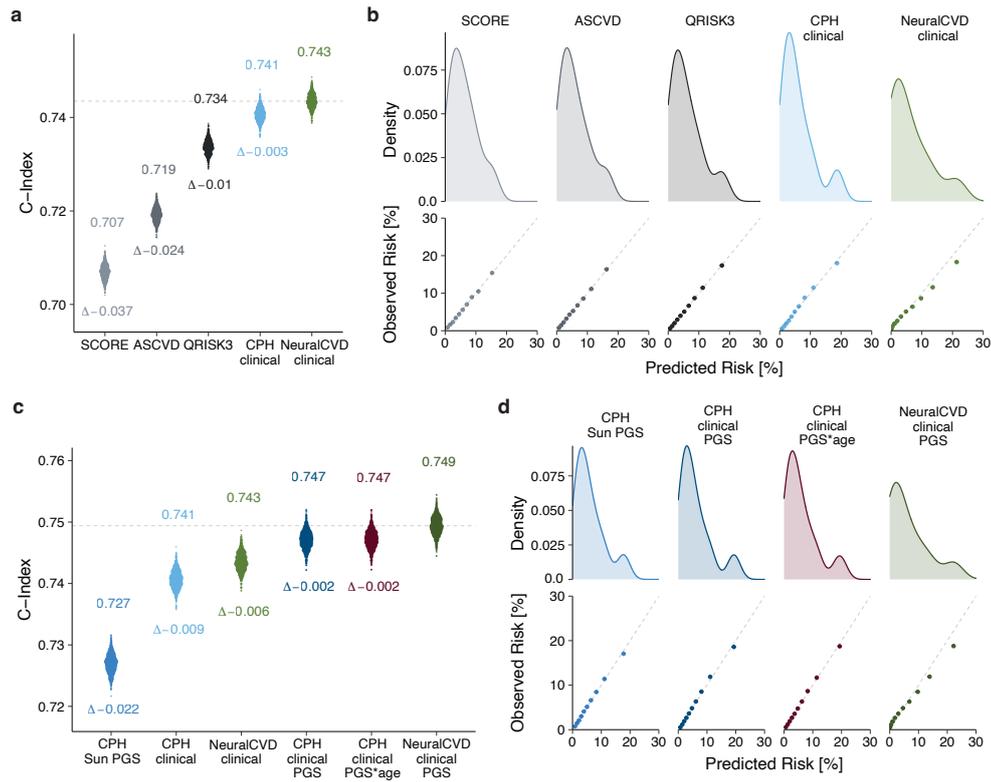


Figure 4.2: Performance and calibration of the NeuralCVD score and established CVD risk models. **a**, Displayed are discriminatory performances of the NeuralCVD score, trained Cox Proportional Hazards (CPH) models, and established CVD risk scores in terms of C-index. The NeuralCVD score outperformed existing approaches in discrimination of Major Adverse Cardiac Events (MACE) at 10 years. Evaluated over the entire study population, this corresponded to an increment of 0.01 compared with the best-performing baseline model, the QRISK3 score table 2 on page 285. **b-d**, Calibration curves and distributions of the predicted risks for each model at 10 years after baseline. PGS = Polygenic risk score.

Table 4.2: Categorical Net Reclassification Improvement of NeuralCVD in comparison to other scores at the 10 % threshold

	SCORE	ASCVD	QRISK3	CPH
NRI	0.1043 (0.0981, 0.1103)	0.0704 (0.0648, 0.0765)	0.0488 (0.0442, 0.0534)	0.0469 (0.0429, 0.0511)
Cases (n=23,786)	15.65% (15.06%, 16.23%)	11.41% (10.86%, 12.01%)	9.62% (9.17%, 10.06%)	10.75% (10.35%, 11.17%)
Non-Cases (n=371,889)	-5.22% (-5.33%, -5.11%)	-4.38% (-4.47%, -4.27%)	-4.74% (-4.83%, -4.65%)	-6.06% (-6.14%, -5.98%)

4.4 Assessing the ability of NeuralCVD to integrate PGS information

Prior to integrating the PGS into the variable set for the development of the composite NeuralCVD and CPH models, the six PGS were assessed for validity in the study population. All six PGS, as well as the derived PGSMETA (see section 3.2.2 on page 78), were found to be associated with incident MACE onset in the study population (figure 2 on page 203).

Subsequently, to investigate the potential of the non-linear NeuralCVD model to integrate polygenic information with established clinical covariates, the six PGS against coronary artery disease and stroke were added to the covariate set, and both the NeuralCVD model as well as a linear CPH model were retrained. Furthermore, to allow the CPH model to assess potential non-linear effects between age and genetic information, interaction terms between age and the polygenic scores were additionally evaluated. Beyond that, all models were scrutinized against a previously published ASCVD-based model involving polygenic information (L. Sun et al. 2021).

Table 4.3: Categorical net reclassification improvement of the PGS addition to NeuralCVD (NeuralCVD+PGS) over other scores at the 10 % threshold.

	CPH Sun PGS	CPH PGS	CPH PGS*age	CPH	NeuralCVD
NRI	0.074 (0.0678, 0.0795)	0.0424 (0.0383, 0.0464)	0.0359 (0.0321, 0.0394)	0.0585 (0.0538, 0.0625)	0.0116 (0.0066, 0.0159)
Cases (n=23,790)	12.92% (12.39%, 13.47%)	10.34% (9.89%, 10.76%)	9% (8.65%, 9.34%)	11.87% (11.42%, 12.27%)	1.12% (0.62%, 1.54%)
Non-Cases (n=371,909)	-5.52% (-5.64%, -5.42%)	-6.1% (-6.16%, -6%)	-5.41% (-5.48%, -5.34%)	-6.01% (-6.11%, -5.92%)	0.05% (-0.03%, 0.12%)

The addition of polygenic scores to the predictor set improved the discrimination performance for the NeuralCVD model in terms of C-index by +0.006 (0.005-0.007) and in terms of NRI by +0.0116 (0.0066-0.0159) when compared to the NeuralCVD model trained on covariates only (figure 4.2 on the preceding page, table 4.3, table 3 on page 286). While the addition of the polygenic scores resulted in improvements in discriminative performance for the CPH model as well, the NeuralCVD model remained superior in C-Index (CPH+PGS: +0.002 (0.002-0.003), CPH+PGS*age: +0.002 (0.002-0.003)) and NRI (CPH+PGS: +0.0424 (0.0383-0.0464), CPH+PGS*age: +0.0359 (0.0321-0.0394)). Compared to the model proposed by Sun et al., substantial improvements in the C-Index of +0.022 (0.021-0.024) and of +0.0740 (0.0687-0.0790) in terms of NRI became apparent (see table 3 on page 286). All models were well-calibrated over the full spectrum of risk (see figure 4.2 on the preceding page), and the differences were consistent in the spatially separated assessment centers (fig-

ure 1 on page 202, table 4 on page 287).

4.5 Individual differences in predicted risk after PGS addition

In order to further characterize the impact of the additional genetic information on predictions for single individuals in both the CPH and the NeuralCVD models, relative risk differences between each model's predictions with and without genetic information were calculated (see section 2.4.6 on page 63). Relative risk differences are positive if the polygenic scores resulted in a higher risk estimate and negative if they led to a lower risk estimate.

The NeuralCVD model predicted relative risk differences in the range of +805 % and -84 % compared to values in the range of +152 % and -63 % (+249 % and -72 % with PGS*age interaction) for the linear CPH model (see figure 4.3 on page 96).

In order to examine this wide interval of relative risk differences obtained for the NeuralCVD model, associations between the information added by the PGS (i.e., the relative risk difference) and the observed clinical phenotype were examined. Individual conventional risk factors were not substantially associated with relative risk differences at ten years (figure 3 on page 204). However, associations with the overall clinical risk (see figure 4.3 on page 96) and with age (see figure 4.3 on page 96) were notable. In the subpopulation of individuals with high genetic risk (i.e., individuals in the top 5 % PGSMETA), pronounced risk differences in the predicted risk were evident in younger individuals with low to intermediate clinical risk: the relative risk difference diminished with increasing clinical risk, and age. This effect was most notable for individuals with high genetic risk and flattened out with lower genetic predisposition (figure 4 on page 205). These differences were non-existent in the linear CPH model without interaction terms but observable in the CPH model with the PGS*age interaction terms.

Further, the observed effect reflects the predicted cardiovascular risk trajectories stratified by clinical risk and age (figure 4.3 on page 96). Young and low-risk individuals were predicted to have the highest relative risk increase from high genetic predisposition ($RR_{t=10} : 2.64(2.52 - 2.76)$). Individuals between 50 and 60 years at intermediate clinical risk were predicted to have a lower impact ($RR_{t=10} : 1.81(1.78 - 1.85)$), and individuals > 60 years at already high clinical risk witnessed the smallest effect on their overall risk with

4.5 Individual differences in predicted risk after PGS addition

the additional high genetic risk information with $RR_{t=10} : 1.40(1.37 - 1.42)$.

To further substantiate these findings, the number of events stratified by clinical risk and age at the end of the observation window for different genetic risk strata was examined (figure 5 on page 206). The relative risk for high genetic risk (Top 5 % PGSMETA) subpopulation was $RR_{t=10} : 1.96(1.59-2.44)$ for the young and low risk, $RR_{t=10} : 1.64(1.40-1.86)$ for the middle age and intermediate-risk and $RR_{t=10} : 1.49(1.32 - 1.60)$ for the older and high-risk subgroup at the end of the observation window (see table 5 on page 288:).

These findings suggest that the additional predictive polygenic information depends on the clinical phenotype (i.e., clinical risk) and that the NeuralCVD model is able to reflect this residual contribution and assess the residual predictive information correctly (see figure 4.4 on page 97). High genetic risk barely affected the overall risk in older individuals when their clinical risk was already high. In contrast, the risk in young individuals at low to intermediate clinical risk was sharply increased.

4 A Neural Survival model integrates polygenic information for cardiovascular risk prediction

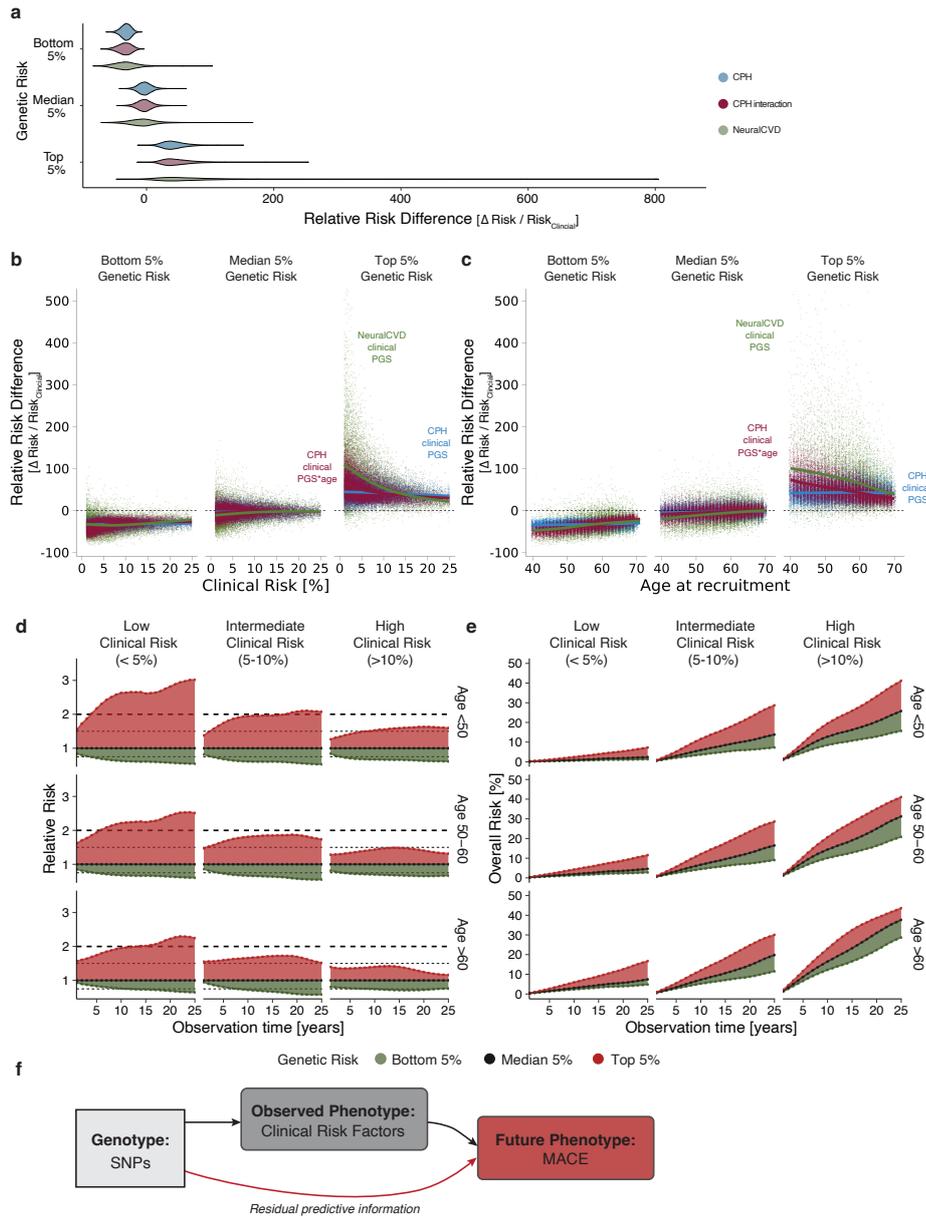


Figure 4.3: The NeuralCVD score learned the residual contribution of the polygenic information over the clinical phenotype. **a**, Relative Risk Difference (RRD) for three genetic strata (bottom, median, and top 5% PGSMETA). Higher genetic risk increased the RRDs for all models. The distributions of RRDs for the NeuralCVD score were wide, with RRDs of up to +805% for the top 5% genetic stratum compared to the predicted risk based on the clinical factors. **b**, RRDs within the two CPH models and the NeuralCVD score upon the PGS addition for the bottom, median, and top 5% of PGSMETA. Increasing genetic risk yielded positive RRDs for both models. RRDs for the CPH model were constant over the spectrum of clinical risk. In contrast, NeuralCVD learned a residual contribution of the polygenic risk over the clinical risk. In the high genetic risk group, RRDs were the highest for the low to the intermediate clinical risk group and declined with clinical risk > 15%. **c**, A similar, albeit weaker trend, is observable. **d**, 25-year risk ratios stratified by genetic risk (bottom, median, and top 5% PGSMETA), age, and clinical risk. Additional genetic information increases risk most in individuals with low to intermediate clinical risk and age <50. **e**, 25-year cumulative risk stratified by genetic risk (bottom, median and top 5% PGSMETA), age, and clinical risk. Risk ratios from **d** are reflected in the CVD risk trajectories and the proportion of polygenic risk in the overall risk. The difference in trajectories was most pronounced in individuals with low to intermediate clinical risk and age <50.

4.5 Individual differences in predicted risk after PGS addition

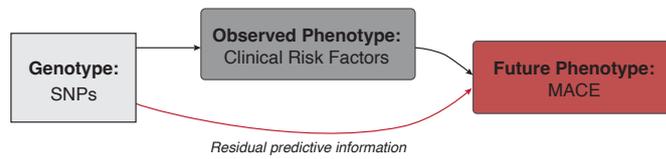


Figure 4.4: The information transition hypothesis. A schematic depiction of the proposed mechanism for the impact of polygenic information on overall risk, adapted from Janssens, 2019. Parts of the Single Nucleotide Polymorphism (SNP)s included in PGS mediate through the manifestation of a clinical phenotype. As conventional risk factors contain this information, the information gained by PGS addition is the residual information.

4.6 Discussion

In recent years, Polygenic Score (PGS) predicting the lifetime risk for many common diseases have been developed. Especially in cardiovascular prevention, an application of PGS in primary prevention could help to improve risk predictions and thus better guide interventions. However, as polygenic risk scores commonly model an individual's lifetime risk, it is mandatory to understand the relationship between phenotypic information contained in commonly assessed cardiovascular risk factors and genetic information in a given PGS. Further, despite neural networks representing the state-of-the-art performance in survival modeling, few medical studies have exploited this potential so far.

In this experiment, NeuralCVD, a novel neural network based cardiovascular risk model predicting 10-year risk for incident Major Adverse Cardiac Events (MACE), was developed. On data from the UK Biobank cohort, the NeuralCVD score improved discrimination and reclassification at the 10 % risk threshold over the latest currently available clinical scores and a Cox Proportional Hazards (CPH) baseline trained on the same phenotypic data. Notably, the improvements of NeuralCVD over established models do not require additional predictor collection and thus encourage the application of neural survival models in cardiovascular prevention. Retraining NeuralCVD on a predictor set, including genetic information, its performance in discrimination, and categorical reclassification at the 10 % risk threshold improved. These observations align with previous studies reporting the identification of more high-risk individuals upon addition of PGS information to common cardiovascular predictors (Elliott et al. 2020; L. Sun et al. 2021). Previous studies integrated PGS with clinical predictors in an additive relationship relying on established linear models, irrespective of biological mechanisms of action and potential mediatory effects on the observed phenotype (Riveros-Mckay Fernando et al. 2021; L. Sun et al. 2021). While the impact of PGS on risk discrimination are known to be small on the population level, its impact is not equally distributed over individuals, and for some individuals, the impact is considerable. The Relative Risk Difference (RRD) calculated in this experiment revealed that the NeuralCVD score inherently accounts for the transition of predictive information from the genotype to the composite clinical phenotype by learning higher-order interactions between clinical factors and a common representation of the patient state. The transition is accounted for in a learned attenuating effect of observed phenotypes, summarized in the learned patient representation, with increasing clinical risk on information gained by the PGS addition in the high-genetic risk strata. A similar but weaker interaction was reproduced by modeling an interaction between age and the added PGS in the CPH model. The observation of the

attenuating relationship implies that substantial parts of the genetic lifetime risk captured in PGS act through phenotypic manifestation and that age represents a lossy proxy on the information transitioning to the patient state. Ultimately, however, it is the added information of PGS information over the clinical risk factors (see figure 4.4 on page 97) which matters in an applied clinical setting (Janssens 2019). This hypothesis of information transitioning from the genotype to the phenotype was first postulated by Janssens et al.: while genetic information is independent at birth, the effects of SNPs are mediated through the manifestation of clinical risk factors (e.g. LDL, elevated blood pressure, obesity). Thus, when assessed later in life and in combination with risk factors, the genotype only provides information on the residual genetic risk; alas, the information not yet manifested in the phenotype. The experiments performed in the UK Biobank, confirm this heterogeneity in the importance of genetic information on future disease risk (figure 5 on page 206, table 5 on page 288). The findings confirming the transitioning hypothesis have two separate implications: First, PGS are most relevant for identifying individuals at high risk before they have a manifested severe clinical phenotype. Second, once the predictive information transitioned from the genotype to the phenotype (i.e., the clinical risk is considerable), the PGS contain just modest information on future risk trajectories.

Despite the large sample size, the presented experiments are subject to limitations. First, as demonstrated previously, individuals in the UK Biobank study cohort have a lower risk for cardiovascular events and are generally healthier than the general population in the UK (Fry et al. 2017; L. Sun et al. 2021). Thus, recalibration with a relevant data source, e.g., the UK Clinical Practice Research Datalink (CPRD), should be performed before a clinical application of the NeuralCVD score. Second, despite the extensive spatial cross-validation over 22 recruitment centers, and the absence of overfitting, the NeuralCVD score has not been validated on an external cohort. This is of particular importance for every model incorporating polygenic scores, as a generalization to ancestrally distant populations is controversial (Schultz et al. 2022). Third, while discrimination, reclassification, and calibration are critical criteria for evaluating predictive models and allow a comparison with the established baselines, these metrics cannot assess clinical benefit. In order to quantify relevant metrics, such as life years saved by identifying the correct individuals for early intervention, and demonstrate clinical utility, prospective studies are required. This is especially relevant, as most individuals are not expected to experience strong risk modifications after consideration of polygenic scores.

4 A Neural Survival model integrates polygenic information for cardiovascular risk prediction

In summary, NeuralCVD is a highly performant neural-network-based risk model for primary cardiovascular prevention. The score outperformed conventional scores and learned the residual genetic contribution to identifying individuals at the highest risk of cardiovascular events. This demonstrates the capability of neural survival models and opens up new opportunities for targeted primary CVD prevention, integrating both clinical and genetic risk factors.

5 Metabolomic profiles predict individual multi-disease outcomes

Based on the results presented in this chapter a manuscript has been prepared, submitted and published in *Nature Medicine* (Buergele, Steinfeldt, Ruyoga, et al. 2022). The published manuscript was written and prepared by Thore Buergele, Dr. Jakob Steinfeldt, Professor Roland Eils and Professor Ulf Landmesser. Details on the individual author’s contributions to the publication are presented in the published manuscript.

5.1 Experimental Setup

This experiment presents a novel neural network-based risk model to exploit the potential of Proton Nuclear Magnetic Resonance (¹H-NMR)-based blood profiling as a single-domain assay to simultaneously predict multi-disease onset. Specifically, a deep residual multi-task neural network, the Metabolomic State Model (see section 2.6.2 on page 68), is developed, trained, and validated to simultaneously learn disease-specific metabolomic states for 24 conditions, including common metabolic, vascular, respiratory, musculoskeletal and neurological disorders and cancers (figure 5.1 on the facing page). The scalar metabolomic states, contained in a 24-dimensional vector, were derived from 168 circulating metabolomic markers measured in 120,000 individuals in the UK Biobank population cohort (Bycroft et al. 2018). The learned metabolomic states were subsequently investigated by integration in Cox Proportional Hazards (CPH) models, assessing the risk for individual endpoints and demonstrating that information gained through ¹H-NMR metabolomic profiling is additive to known clinical predictors. This analysis was replicated in four independent cohorts, the Whitehall II cohort (Marmot and Brunner 2005), the Rotterdam Study (M.K. Ikram et al. 2006), the Leiden Longevity Study (Schoenmaker et al. 2006), and the PROspective Study of Pravastatin in the Elderly at Risk (Shepherd, Blauw, Murphy, Cobbe, et al. 1999; Shepherd, Blauw, Murphy, and Gaw 2002) (figure 5.1 on the next page, section 3 on page 73), and the metabolomic states were investigated for their clinical utility. Similarly to the experiment described in section 4 on page 87, the metabolomic state model was developed and trained in 22-fold spatial cross-validation over the recruitment centers of the UK Biobank cohort (see section 3.2.6 on page 81). Further relying on the spatial separation, the predictions of the metabolomic state model were subsequently integrated in endpoint-specific survival models to scrutinize the information in the metabolomic states against clinical predictors. Subsequently, predictions of the survival models were aggregated for downstream analysis.

In summary, the experimental approach involved four steps: first, the information of the ¹H-NMR metabolomic profile on disease risk was investigated by assessing the stratifica-

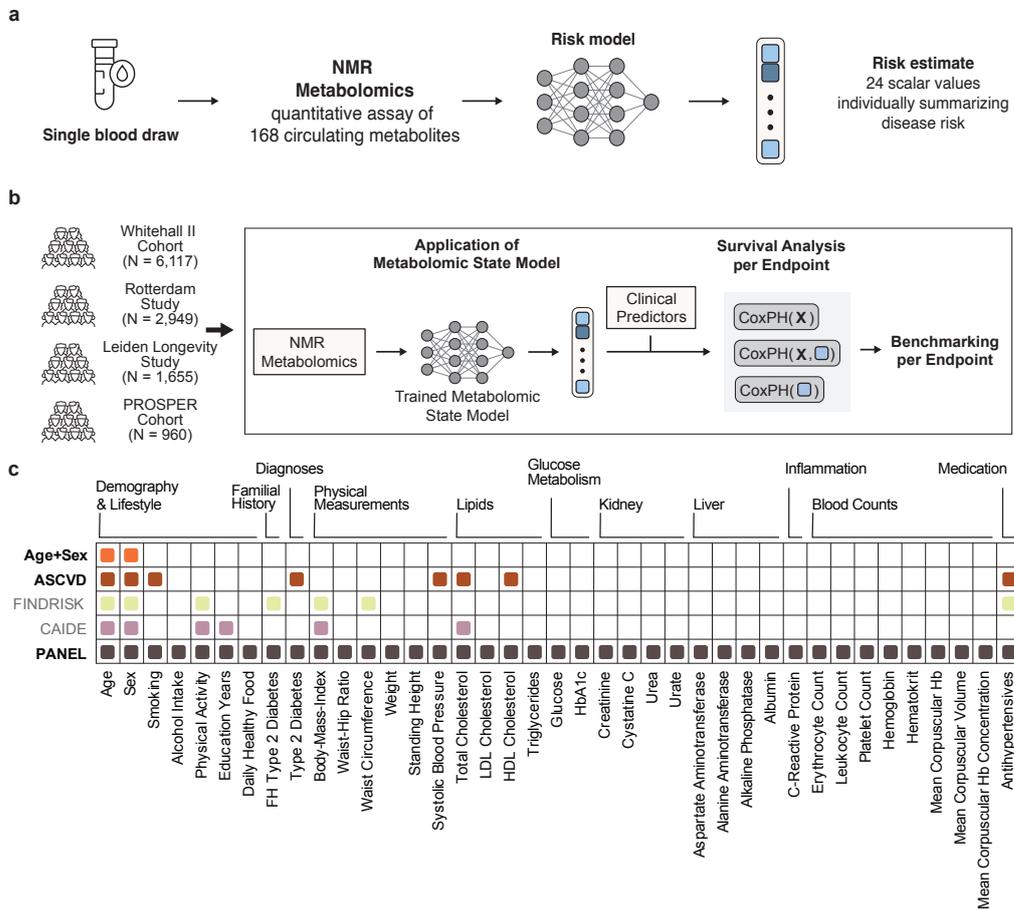


Figure 5.1: Overview of the experimental setup. **a**, Based on a single blood draw and 1H-NMR Metabolomics profiling, the metabolomic state model estimates *metabolomic states* summarizing disease risks for 24 common disease endpoints. For model development the eligible UK Biobank population (with 1H-NMR blood metabolomics and valid consent) was split into train, validation, and test sets in 22-fold nested cross-validation based on the assigned UK Biobank assessment center as described in section 3.2.6 on page 81. **b**, The Metabolomic State Model was externally validated in four independent cohorts, the Whitehall II Cohort and three cohorts from the BBMRI-NL consortium, the Rotterdam Study, the Leiden Longevity Study, and the PROSPER cohort. **c**, Overview of the clinical predictor sets considered in this experiment. Predictors originate from scores commonly applied in primary prevention. Additionally, variables are integrated into a comprehensive predictor set (PANEL) to investigate overlapping information with the metabolomic state. A full list of clinical predictors and their utilization in the different predictor sets is provided in table 3.1 on page 77.

5 Metabolomic profiles predict individual multi-disease outcomes

tion of event rates and ten-year discriminative performances of the metabolomic states; next, the additive information of the metabolomic states over clinical predictors was explored, by scrutinizing the models against conventional clinical predictors and calculating adjusted risk trajectories in both the UK Biobank and the external validation cohorts; third, followed an assessment of the clinical utility of metabolomic state additions over clinical predictors, and fourth, metabolomic profiles associated with high risk for specific endpoints were closely investigated, both on the population as well as on the individual level.

5.2 Characteristics of the study population

The metabolomic state model was developed on a subset of the UK Biobank cohort with Proton Nuclear Magnetic Resonance (¹H-NMR) metabolomic measurements, representing a sample of the UK general population (Bycroft et al. 2018) (see section 3.1.1 on page 74). In total, clinical predictors and disease endpoints for 117,981 individuals with serum ¹H-NMR profiling at the time of cohort recruitment (see section 3.2.3 on page 79, table 3.3 on page 83, table 3.1 on page 77, and table 2 on page 245) were extracted. The study population had a median age of 58 (IQR 50, 63) years; 54.2 % were female, 11 % current smokers, and 5.2 % were diagnosed with type-2 diabetes (table 5.1). The median BMI was 26.8 (IQR 24.2, 29.9), systolic blood pressure was 136 (IQR 124, 149) mmHg, total cholesterol was 5.65 (IQR 4.90, 6.42) mmol/l and glucose was 4.93 (IQR 4.60, 5.32) mmol/l. The median follow-up time was 12.2 years with ~1,435,340 overall patient-years.

Table 5.1: Summary statistics of the study population. Reported are medians (IQR) and absolute counts (%).

	Male N = 54,078	Female N = 63,903	Overall N = 117,981
Age at Recruitment	58 (50, 64)	57 (50, 63)	58 (50, 63)
Education years	15.00 (11.00, 15.00)	13.00 (11.00, 15.00)	13.00 (11.00, 15.00)
Current Smoker	6,724 (12%)	5,747 (9.0%)	12,471 (11%)
Daily Alcohol Intake	13,651 (25%)	10,191 (16%)	23,842 (20%)
Daily Moderate to Vigorous Physical Activity	50 (15, 105)	45 (10, 90)	45 (10, 90)
Daily Healthy Food	52,974 (98%)	63,290 (99%)	116,264 (99%)
Family History Diabetes	8,827 (16%)	11,266 (18%)	20,093 (17%)
Type 2 Diabetes	3,882 (7.2%)	2,295 (3.6%)	6,177 (5.2%)
Weight (kg)	84 (76, 94)	69 (62, 79)	76 (66, 88)
Standing Height (cm)	176 (171, 180)	162 (158, 167)	168 (162, 175)
BMI	27.3 (25.0, 30.1)	26.1 (23.5, 29.7)	26.8 (24.2, 29.9)
Waist-Hip-Ratio	0.93 (0.89, 0.98)	0.81 (0.77, 0.86)	0.87 (0.80, 0.94)
Waist Circumference (cm)	96 (89, 103)	83 (76, 92)	90 (80, 99)

5.2 Characteristics of the study population

Table 5.1 continued from previous page

	Male N = 54,078	Female N = 63,903	Overall N = 117,981
Systolic Blood Pressure (mmHg)	139 (128, 152)	133 (121, 147)	136 (124, 149)
Total Cholesterol (mmol/L)	5.45 (4.70, 6.21)	5.80 (5.07, 6.58)	5.65 (4.90, 6.42)
LDL Cholesterol (mmol/L)	3.46 (2.87, 4.05)	3.56 (3.00, 4.17)	3.52 (2.94, 4.12)
HDL Cholesterol (mmol/L)	1.24 (1.06, 1.45)	1.55 (1.32, 1.82)	1.40 (1.17, 1.67)
Triglycerides (mmol/L)	1.69 (1.18, 2.44)	1.33 (0.96, 1.89)	1.48 (1.04, 2.14)
Glucose (mmol/L)	4.96 (4.61, 5.37)	4.91 (4.59, 5.28)	4.93 (4.60, 5.32)
Glycated Hemoglobin (%)	35.3 (32.8, 38.1)	35.2 (32.7, 37.7)	35.2 (32.8, 37.9)
Creatinine (umol/L)	80 (72, 88)	63 (57, 70)	70 (61, 81)
Cystatin C (mg/L)	0.92 (0.84, 1.01)	0.86 (0.78, 0.95)	0.88 (0.80, 0.98)
Urea (mmol/L)	5.45 (4.68, 6.33)	5.10 (4.33, 5.95)	5.26 (4.49, 6.13)
Urate (umol/L)	350 (305, 399)	264 (225, 309)	303 (250, 361)
Aspartate Aminotransferase (U/L)	26 (23, 31)	23 (20, 27)	24 (21, 29)
Alanine Aminotransferase (U/L)	24 (18, 32)	18 (14, 23)	20 (15, 27)
Alkaline Phosphatase (U/L)	79 (67, 93)	82 (67, 98)	80 (67, 96)
Albumin (g/L)	45.52 (43.80, 47.24)	44.91 (43.21, 46.63)	45.20 (43.47, 46.93)
C-Reactive Protein (mg/L)	1.29 (0.67, 2.55)	1.38 (0.65, 2.95)	1.33 (0.66, 2.76)
Erythrocytes (1012 cells/L)	4.74 (4.51, 4.98)	4.32 (4.10, 4.54)	4.50 (4.23, 4.79)
Leucocytes (109 cells/L)	6.68 (5.66, 7.89)	6.61 (5.60, 7.81)	6.64 (5.62, 7.85)
Platelets (109 cells/L)	234 (202, 269)	261 (226, 301)	248 (214, 287)
Haemoglobin (g/dL)	15.00 (14.37, 15.64)	13.50 (12.90, 14.10)	14.15 (13.31, 15.02)
Haematocrit (%)	43.3 (41.4, 45.2)	39.2 (37.5, 41.0)	41.0 (38.7, 43.5)
Mean Corpuscular Volume (fl)	91.4 (88.8, 94.1)	91.1 (88.4, 93.7)	91.2 (88.6, 93.9)
Mean Corpuscular Haemoglobin (pg)	31.69 (30.70, 32.70)	31.37 (30.33, 32.37)	31.50 (30.50, 32.50)
Mean Corpuscular Haemoglobin (g/dL)	34.60 (34.00, 35.22)	34.36 (33.80, 35.00)	34.48 (33.90, 35.10)
Antihypertensives	1,090 (2.0%)	680 (1.1%)	1,770 (1.5%)

For validation, four independent cohorts with the same Proton Nuclear Magnetic Resonance (1H-NMR) metabolomics assay were analyzed, the Whitehall II cohort (Marmot and Brunner 2005), the Rotterdam Study (M.K. Ikram et al. 2006), the Leiden Longevity Study (Schoenmaker et al. 2006), and the PROspective Study of Pravastatin in the Elderly at Risk (Shepherd, Blauw, Murphy, Cobbe, et al. 1999; Shepherd, Blauw, Murphy, and Gaw 2002). The Whitehall II cohort is an ongoing prospective cohort study, including 1H-NMR metabolomic measurements for 6,197 participants aged between 44 and 69 years (Marmot and Brunner 2005). The Rotterdam Study is a prospective, population-based cohort study including individuals recruited in the Ommoord district in Rotterdam (Netherlands). The study offers 1H-NMR metabolomic measurements for 2,949 participants with a median age of 74 years (IQR 70-79) (M.K. Ikram et al. 2006). The Leiden Longevity PAROFFS Study (LLS) comprises offsprings and spouses of long-lived individuals, with 1H-NMR

metabolomics measurements for 1,655 individuals with a mean age of 59 years (IQR 54-63) (Schoenmaker et al. 2006). Finally, the PROspective Study of Pravastatin in the Elderly at Risk (PROSPER) was set up as a clinical trial to investigate pravastatin effects (Shepherd, Blauw, Murphy, Cobbe, et al. 1999; Shepherd, Blauw, Murphy, and Gaw 2002). Of the recruited population, 960 samples with a median age of 76 years (IQR 73-78) are included in the BBMRI-NL platform. Detailed characteristics of the four replication cohorts are given in section 3 on page 73, and a comparison between the cohorts is presented in table 1 on page 236.

5.3 The metabolomic state is associated with event rates and stratifies the population according to disease risk

A critical component of prevention is identifying individuals at high risk of developing a disease, often at an early subclinical stage. The first step was, therefore, to investigate the ability of the 1H-NMR-derived metabolomic states to stratify the population by observed incident event rates in the observation period (figure 5.3). Further, to allow a comparison between the endpoints despite the large differences in observed event rates (e.g., Parkinson's disease: 0.6 %, Major Adverse Cardiac Events (MACE): 8.7 %), the observed event rate ratio (OR) between the individuals in the top and bottom 10 % of metabolomic states (figure 5.3, table 5 on page 290) was additionally calculated including 95 % confidence intervals (table 5 on page 290).

All endpoints, except breast cancer, displayed increasing event rates over the metabolomic state percentiles indicating stratification performance. For 15 of the 24 endpoints (62.5 %), the event rate in the top 10 % was >5-times higher than in the bottom 10 % of the metabolomic state (see Figure 2). For a few conditions, such as type-2 diabetes (Top 10 %: 21.87 %, Bottom 10 %: 0.36 %, OR: 61.45, 95 % CI: 47, 86.12), abdominal aortic aneurysm (Top 10 %: 2.46 %, Bottom 10 %: 0.18 %, OR: 14.1 95 % CI: 9.93, 24.45) and heart failure (Top 10 %: 10.80 %, Bottom 10 %: 0.96 %, OR: 11.27, 95 % CI: 9.43, 13.5) the ratio was more than 10. However, the ratios for most other diseases were lower, e.g., cerebral stroke: 9.66 (95 % CI: 7.64, 12.14), MACE 9.25 (95 % CI: 8.12, 10.53), atrial fibrillation: 8.13 (95 % CI: 6.95, 9.37), all-cause dementia 6.39 (95 % CI: 5.4, 8.09) or COPD 4.98 (95 % CI: 4.37, 5.62). On the other end, for a few diseases e.g., glaucoma (Top 10 %: 3.47 %, Bottom 10 %: 1.57 %, OR: 2.19, 95 % CI: 1.91, 2.62) or asthma (Top 10 %: 5.52 %, Bottom 10 %: 2.48 %, OR: 2.22, 95 % CI: 2.01, 2.57), ratios were much smaller suggesting the metabolomic state to

5.3 The metabolomic state is associated with event rates and stratifies the population according to disease risk

provide less information on disease onset.

These observations were further reflected in the cumulative event trajectories, where the metabolomic state separated trajectories for the bottom, median, and top 10 % of metabolomic states for all endpoints except breast cancer (see figure 5.3 on the preceding page). Cumulative event rates were separated most notably for type-2 diabetes, renal disease, and heart failure but also to a much lesser extent for glaucoma or asthma. In summary, the metabolomic state contained stratifying information for 23 of 24 diseases, most notably for conditions with a known metabolic or cardiovascular contribution.

5 Metabolomic profiles predict individual multi-disease outcomes

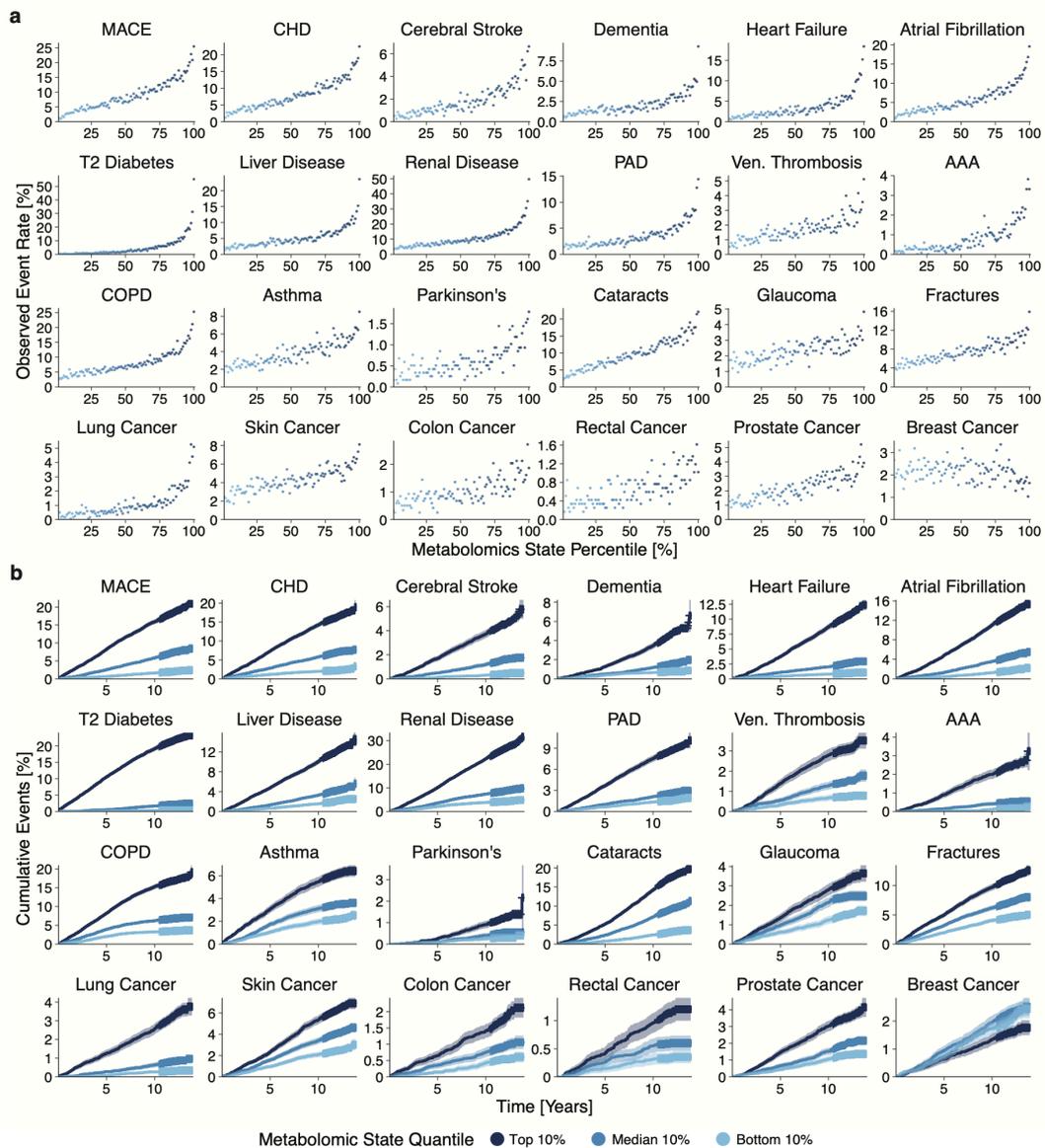


Figure 5.2: The metabolomic state is associated with observed event rates and stratifies survival. **a**, Displayed are observed event frequencies for incident disease onset plotted against the metabolomic state percentiles over the entire study population for all 24 endpoints. A higher metabolomic state is associated with increased observed event rates across all investigated endpoints but breast cancer. **b**, Displayed are cumulative event rates over the observation time for all assessed endpoints, stratified by metabolomic state quantiles (light blue: bottom 10 %, blue: median 10 %, dark blue: top 10 %). Stratification by metabolomic state yields separated risk trajectories for all endpoints but breast cancer. Major Adverse Cardiac Events (MACE), Coronary Heart Disease (CHD), Peripheral Artery Disease (PAD), Abdominal Aortic Aneurysm (AAA), Chronic Obstructive Pulmonary Disease (COPD).

5.4 The metabolomic state shares information with common clinical predictors

Many clinical predictors are readily available in primary care and commonly used to stratify the risk of common diseases, such as cardiovascular disease (Conroy 2003; Goff et al. 2014), kidney disease (Velde et al. 2009), dementia (Kivipelto et al. 2006) or diabetes (Lindström and Tuomilehto 2003). While risk scores involving more complex predictors have been proposed (Khera, Chaffin, Aragam, et al. 2018; Marsh, Janes, and Pepe 2020), the trade-off between the added predictive information and the resources in time and cost required to collect the new data has limited clinical adoption (C. Huang et al. 2021). Thus, the predictive information of the relatively affordable and standardized NMR metabolomics assay was scrutinized against common clinical predictors in both the UK Biobank and four independent validation cohorts.

First, disease risk was modeled for each endpoint in the UK Biobank using Cox Proportional Hazards (CPH) models for three clinical predictor sets with increasing complexity: Age+Sex, highly predictive and available ahead of any test, ASCVD, a set of readily available cardiovascular predictors, and PANEL a comprehensive selection of clinical predictors including in-depth blood measurements (see figure 5.1 on page 103) exceeding what is typically available in primary care. For all sets, the performance of the CPH models was benchmarked against CPH models based on the sets' combinations with the Metabolomic State (MET). The discriminative performances in terms of Harrell's C-Index of all models ten years after baseline are displayed in figure 5.3 on the next page. Subsequently, to externally validate the metabolomic states, the trained metabolomic state model was transferred and applied to the external validation cohorts. This was followed by a replication of the CPH models with and without metabolomic state addition for the Age+Sex predictor set for all endpoints available. The external validation results are displayed in figure 6 on page 208. The discriminative performance of the metabolomic state (MET) was highly disease dependent: while the metabolomic state contained significantly less predictive information than the clinical predictors for cataracts, glaucomas, skin, colon, rectal, and prostate cancers, this was not the case for renal disease, liver disease, and type-2 diabetes. Here, the metabolomic state contained a greater predictive value than Age+Sex and even ASCVD. Generally, the discriminative performance improved upon the addition of more comprehensive clinical predictors systematically across all endpoints. Further, the model performances were stable over different age groups, biological sexes, and ethnic backgrounds (see figure 7 on page 209). Details on the absolute discriminatory performances for all models trained and

5 Metabolomic profiles predict individual multi-disease outcomes

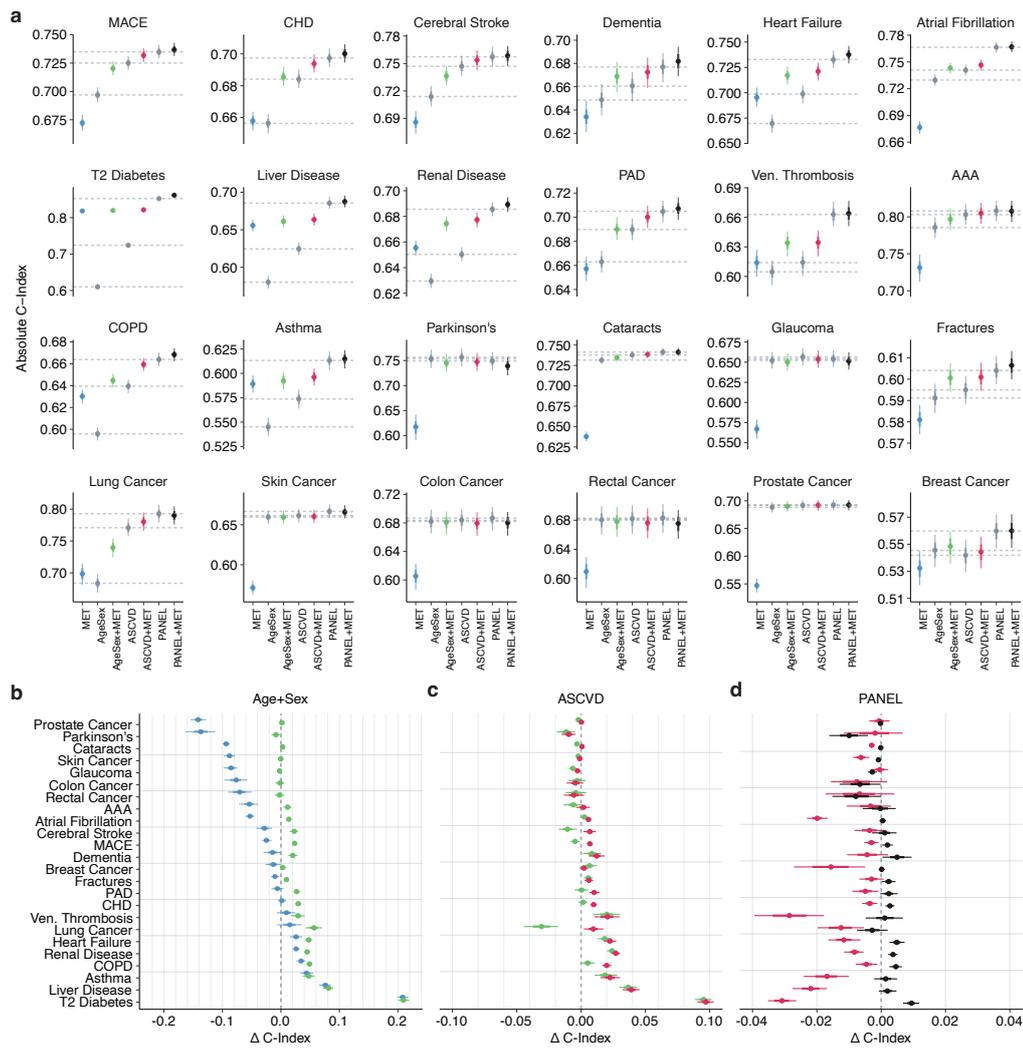


Figure 5.3: Predictive value of the metabolomic state is endpoint dependent. **a**, Displayed are discriminative performances in terms of C-index of the Cox Proportional Hazards (CPH) models trained on the metabolomic state only (MET), the three clinical predictor sets (Age+Sex, ASCVD, and PANEL) and the sets' combinations with the metabolomic state. The horizontal dashed lines indicate the median performance of the three clinical predictor sets. Whiskers indicate the 95 % Confidence Interval (CI) determined by bootstrapping over 1000 iterations. **b**, Displayed are pair-wise differences in discriminative performance between the Age+Sex baseline (gray dashed line), the metabolomic state only (blue), and the combination of Age+Sex and the metabolomic state (orange). The metabolomic state in isolation outperforms the Age+Sex baseline for 6 of 24 endpoints, while its addition to the baseline set improves discriminative performance for 18 of 24 endpoints. **c**, Displayed are differences in discriminative performance between the ASCVD predictors (grey dashed line), the combination of Age+Sex and the metabolomic state (orange), and the combination of the metabolomic state and the ASCVD predictors (red). Age+Sex+MET performs better than or equal to the ASCVD set for 15 endpoints. The addition of the metabolomic state to ASCVD improves the performance of type-2 diabetes dramatically. **d**, Discriminative performance differences between the comprehensive PANEL predictors (dashed line), the ASCVD+MET (red), and the PANEL+MET (black). E.g., all-cause dementia and type-2 diabetes addition of the metabolomic state results in a clear benefit in discrimination over the PANEL. **b-d**, All performances are reported as medians (dots), and whiskers extend to the 95 % CI as estimated by bootstrapping over 1000 iterations. Note that the x-axis range differs across panels; the vertical grid lines indicate differences of 0.02 C-Index. Major Adverse Cardiac Events (MACE), Coronary Heart Disease (CHD), Peripheral Artery Disease (PAD), Abdominal Aortic Aneurysm (AAA), Chronic Obstructive Pulmonary Disease (COPD).

5.4 *The metabolomic state shares information with common clinical predictors*

the direct model comparisons in the UK Biobank cohort are provided in table 6 on page 291 and table 7 on page 292.

Next, the predictive value of the metabolomic state and the combination of the metabolomic state with clinical predictors was compared directly by calculating C-index deltas between pairs of models for the same endpoint (figure 5.3 on the facing page). In this direct comparison, CPH models fit on solely the metabolomic state (MET) performed competitively or better than Age+Sex for 10 of the 24 endpoints, including type-2 diabetes and COPD, but also heart failure, liver disease, and renal disease (figure 5.3 on the preceding page). This competitive performance compared to Age+Sex was validated in external cohorts for type-2 diabetes, COPD, heart failure, CHD, and all-cause dementia (see figure 6 on page 208, table 8 on page 303).

Interestingly, CPH models fit on the combination of the metabolomic state with Age+Sex (Age+Sex+MET) performed comparably to or better than the ASCVD predictors for 15 of the 24 endpoints, including type-2 diabetes, liver disease, renal disease, heart failure, venous thrombosis, and dementia (figure 5.3 on the facing page). Further, the comprehensive PANEL score generally contained the most predictive information. Nevertheless, performance gains over the combinations of ASCVD and the metabolomic state and Age+Sex and the metabolomic state were only modest (figure 5.3 on the preceding page). To exclude differences in model complexity as a determinant for performance, the complex metabolomic state model architecture was additionally trained on the predictors of the PANEL set (figure 8 on page 210). Here, no systematic performance improvements were noticeable, indicating the performance differences to be driven by the information in the NMR metabolomic profiles. Further, the metabolomic state model outperformed endpoint-specific linear models and linear models trained on principal components, demonstrating its ability to retrieve information from shared metabolite profiles while retaining flexibility in fitting endpoint-specific variations (see figure 9 on page 211).

In summary, these findings indicate the metabolomic state to contain substantial predictive information, shared in part with comprehensive clinical predictors. This suggests the potential of metabolomic states to augment simple demographic predictors such as age and biological sex and add clinically relevant information for many common diseases simultaneously.

5.5 The metabolomic state adds predictive value over clinical variables

Besides investigating the shared information, quantifying the additive predictive value over readily available clinical variables is crucial to determining the clinical utility of a novel predictor. Therefore, the NMR metabolomics assay variables were first directly compared with the predictors of the PANEL set. Apart from shared measures such as glucose, albumin, or creatinine, lipids, and creatinine/cystatin c, no strong correlations ($|r| > 0.5$) between predictors in the PANEL set and variables from the NMR metabolomics assay were noticeable (see table ?? on page ??). Further, an assessment of the aggregated coefficients of the CPH models trained on the PANEL predictor set revealed basic demographic information, medical history, and physical measurements to provide the most predictive information over all endpoints (see table 11 on page 313).

Consequently, the performance differences between the CPH models' fit on clinical predictors and those with the added metabolomic state (see table 8 on page 303) were assessed next. In the UK Biobank, the metabolomic state significantly added predictive information over age and sex for 18 of the 24 endpoints; 6 endpoints, those with a comparably low predictive value of the metabolomics state, such as Parkinson's disease, skin cancer, colon cancer, rectal cancer, glaucoma, and cataracts did not benefit from metabolomic state addition. A replication of the analysis in the four external validation cohorts confirmed significant discriminative improvements over Age+Sex for coronary heart disease, heart failure, atrial fibrillation, type-2 diabetes, and COPD. Detailed results are presented in figure 6 on page 208. Specifics on the characteristics of the external validation cohorts are presented in table 1 on page 236. The addition of the metabolomic state to the cardiovascular predictors from the ASCVD set further significantly improved discriminative performance for 15 of the 24 endpoints (figure 5.3 on page 110). Even when added to the comprehensive PANEL set, the metabolomic state provided significant additional discriminatory value for 8 of the 24 endpoints (figure 5.3 on page 110), in terms of the C-index, including type-2 diabetes (+0.009, 95 % CI: +0.007, +0.012), dementia (+0.005, 95 % CI: +0.000, +0.009), heart failure (+0.005, 95 % CI: +0.003, +0.007), COPD (+0.005, 95 % CI: +0.003, +0.006), renal disease (+0.004, 95 % CI: +0.002, +0.005), coronary heart disease (+0.003, 95 % CI: +0.001, +0.004) and MACE (+0.002, 95 % CI: +0.000, +0.004).

The development of the CPH models allowed for a calculation of adjusted hazard ratios and partial effects of the metabolomic state under consideration of established clinical pre-

5.5 The metabolomic state adds predictive value over clinical variables

dictors. Partial effects and hazard ratios (HR, per SD metabolomic state, with 95 % CI) for combinations of the metabolomic state with the predictor sets Age+Sex, ASCVD, and PANEL for the 18 endpoints with discrimination benefits over Age+Sex set are presented in figure 5.4 on the next page. A separation between the top, median, and bottom 10 % of metabolomic states was notable for 14 of the 18 endpoints when adjusted for more comprehensive clinical predictors. A change of one standard deviation in the metabolomic state for type-2 diabetes resulted in substantial adjusted HRs ($HR_{\text{Age+Sex}}$: 3.83 (95 % CI: 3.71-4.01), HR_{PANEL} : 2.5 (95 % CI: 2.34-2.67)), which were replicated with adjustment for Age+Sex in the independent cohorts (see figure 6 on page 208). Other investigated endpoints such as all-cause dementia ($HR_{\text{Age+Sex}}$: 1.56 (95 % CI: 1.54-1.72), HR_{PANEL} : 1.46 (95 % CI: 1.43-1.47)), heart failure ($HR_{\text{Age+Sex}}$: 1.8 (95 % CI: 1.74-1.86), HR_{PANEL} : 1.45 (95 % CI: 1.38-1.52)), COPD ($HR_{\text{Age+Sex}}$: 1.56 (95 % CI: 1.53-1.6), HR_{PANEL} : 1.35 (95 % CI: 1.31-1.39)) or MACE ($HR_{\text{Age+Sex}}$: 1.63 (95 % CI: 1.58-1.69), HR_{PANEL} : 1.4 (95 % CI: 1.33-1.46)) show less pronounced, yet clear separation of the risk trajectories. As for type-2 diabetes, the HRs of the metabolomic states were externally validated with adjustment for Age+Sex for all-cause dementia, heart failure, atrial fibrillation, CHD, and COPD (see figure 6 on page 208). In contrast, the metabolomic state only marginally modified the risk trajectories for diseases such as asthma ($HR_{\text{Age+Sex}}$: 1.37 (95 % CI: 1.3-1.44), HR_{PANEL} : 1.09 (95 % CI: 1.03-1.16)) and cataract ($HR_{\text{Age+Sex}}$: 1.22 (95 % CI: 1.18-1.25), HR_{PANEL} : 1.08 (95 % CI: 1.05-1.11)).

The findings indicate that the metabolomic state adds significant predictive information over basic and comprehensive clinical predictors for many common diseases. A replication of the findings in four independent cohorts confirms the substantial detrimental role of adverse metabolomic profiles on the risk of incident disease onset.

5 Metabolomic profiles predict individual multi-disease outcomes

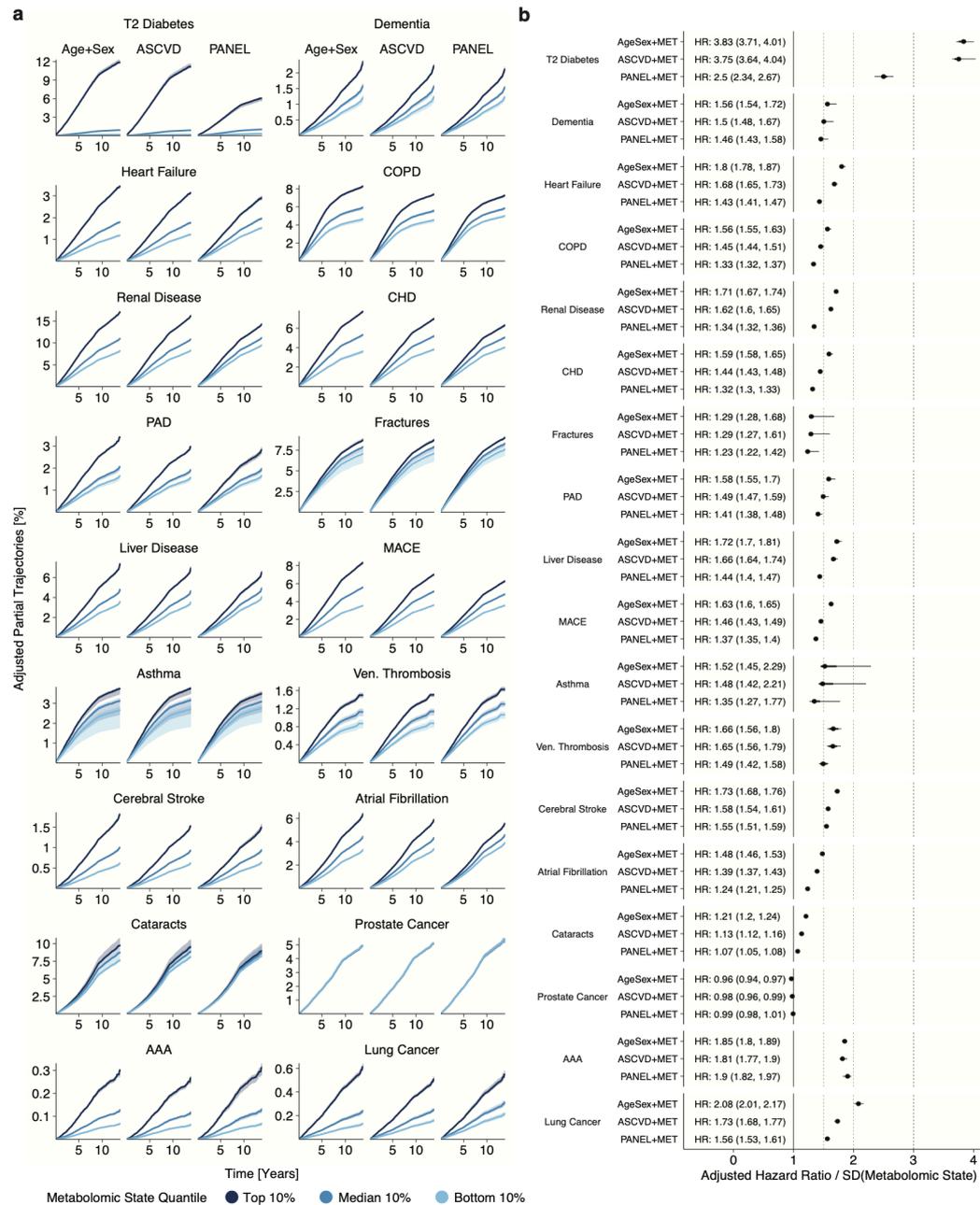


Figure 5.4: Adjusted effect of the metabolomic state is endpoint dependent. **a**, Adjusted trajectories representing the partial cumulative risk dependent on the metabolomic state over time for the endpoints where the metabolomic state added information to the Age+Sex baseline (see figure 5.3 on page 110) for the bottom (light blue), median (blue) and top (dark blue) 10 % metabolomic state quantiles. The shaded area indicates the 95 % confidence interval as estimated by bootstrapping over 1000 iterations. **b**, Adjusted Hazard Ratio (HR) for the metabolomic state in combination with the three clinical predictor sets. A unit standard deviation increase in the metabolomic state corresponds to a HR increase in predicted risk. Adjustment for increasing complex clinical predictors generally decreases the HRs, however, for endpoints where the metabolomic state has a high predictive value (see figure 5.3 on page 110), e.g. type-2 diabetes the HRs remain high even after exhaustive adjustment with comprehensive clinical predictors. The 95 % confidence interval is indicated and estimated by bootstrapping over 1000 iterations.

5.6 Discriminative performance translates into potential clinical utility

While discrimination and additive performance are critical, the clinical utility of any risk model depends on calibration and the choice of adequate thresholds for interventions. All models were found to be well-calibrated in the UK Biobank cohort (see figure 5.5 on page 117 and figure 10 on page 218 for details on all endpoints). The large sample size and close follow-up allowed an assessment of the clinical utility over a wide range of clinically reasonable intervention thresholds. However, adequate clinical decision thresholds depend on interventions' benefits and harms and disease prevalence. Therefore, decision curve analysis (Vickers, Calster, and Steyerberg 2019; Vickers, Van Calster, and Steyerberg 2016) (see section 2.4.4 on page 62) was performed to estimate the benefit of including the NMR metabolomic assay into the predictor set for clinical risk modeling (again, see figure 5.5 on page 117 and figure 10 on page 218 for details on all endpoints). Further, clinically-relevant metrics, such as the sensitivity, positive predictive value, and the positive likelihood ratio over multiple false-positive rates, were calculated additionally (see table 12 on page 314).

Specifically, the metabolomic state was assessed in two scenarios: First, in combination with Age+Sex and with the less resource-intensive, non-laboratory predictors of the PANEL set as a potentially economical and practical option; and second, in combination with the entire PANEL set (including all laboratory predictors), to assess whether the possibility of a net benefit even beyond comprehensive predictors. Generally, previously observed discriminative gains (figure 5.3 on page 110) translated into utility gains (figure 5.5 on page 117 and figure 10 on page 218 for details on all endpoints). For most endpoints, the metabolomic state substantially added to age and sex, and the inclusion of additional non-laboratory predictors of the PANEL set either closed (12 of the 24 endpoints, including type-2 diabetes, stroke, heart failure, and lung cancer) or narrowed the gap (an additional 4 of the 24 endpoints, including dementia, atrial fibrillation, and renal disease) to the complete comprehensive set of PANEL predictors. A direct combination of the metabolomic state and the PANEL predictors further improved the utility for 11 of the 24 endpoints (most notably type-2 diabetes, heart failure, and dementia, see figure 10 on page 218 for details on all endpoints, figure 11 on page 219 for additional analyses investigating *APOE4* carrier status for dementia). Contrariwise, if there were no discriminatory value improvements, no relevant clinical utility improvements could be found. These observations were directly reflected in the positive predictive values and the positive likelihood ratios (table 12 on page 314).

5 Metabolomic profiles predict individual multi-disease outcomes

In summary, these findings demonstrate that combinations of the metabolomic state with both, easily obtainable and extensive clinical predictors, bear the potential of delivering not only discriminative improvements but also clinical utility over clinically relevant decision thresholds.

5.6 Discriminative performance translates into potential clinical utility

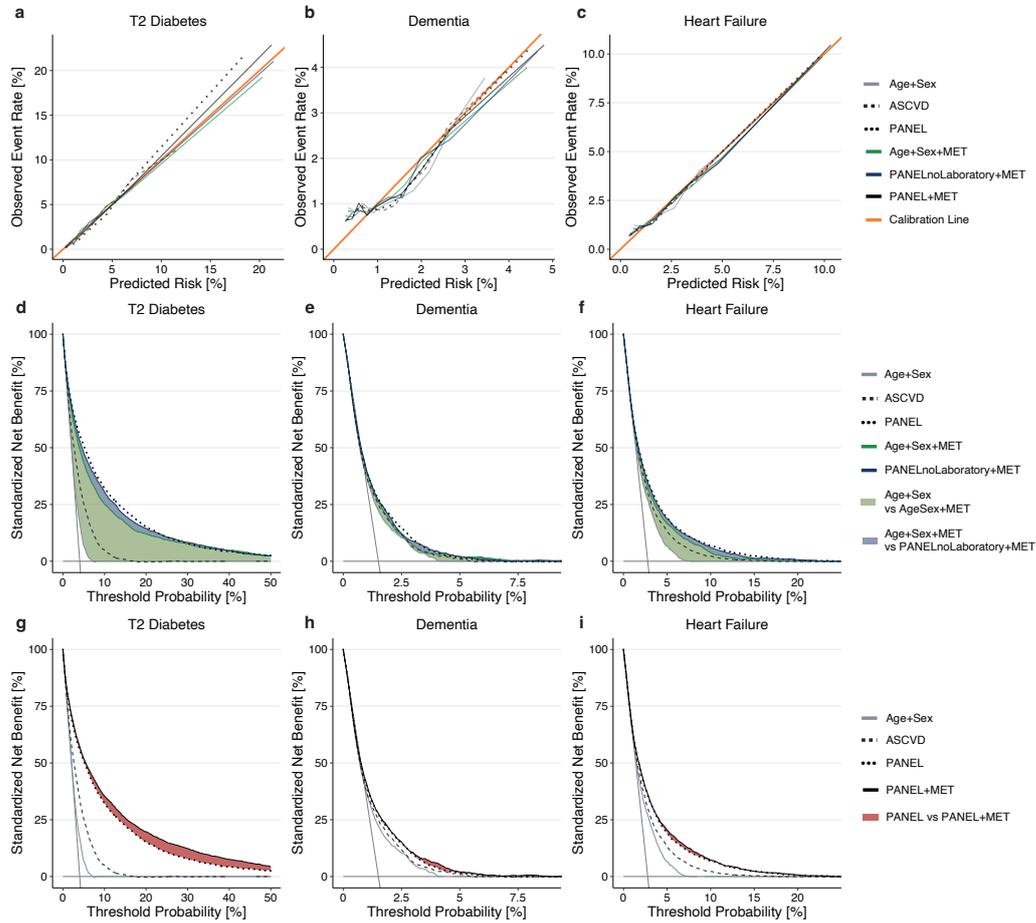


Figure 5.5: Model calibration and additive predictive value of the metabolomic state translate to potential clinical utility. **a-c**, Displayed are calibration curves for Cox Proportional Hazards (CPH) models trained on the three clinical predictor sets AgeSex, ASCVD, and PANEL as well as their extensions with the metabolomic state (i.e., AgeSex+MET) for three endpoints, **(a)** type-2 diabetes, **(b)** dementia and **(c)** heart failure. All models are well-calibrated. **d-f**, Displayed are net benefit curves for the same three endpoints, standardized by the endpoint's prevalence. The horizontal solid gray line indicates *treat none*, and the vertical solid gray line indicates *treat all*. The standardized net benefit of AgeSex (gray line), ASCVD (dashed dark gray line), and the PANEL set (dotted black line) are compared to AgeSex+MET (green line) and additional non-laboratory predictors of the PANEL (PANELnoLaboratory, blue line). The green colored area indicates the added benefit of the metabolomic state addition over AgeSex, while the blue alpha indicates added performance over PANELnoLaboratory, respectively. **g-i**, The same standardized net benefit curves comparing the performance of PANEL+MET (solid black line) against all baselines: AgeSex (gray line), ASCVD (dashed dark gray line), and the PANEL set (dotted black line).

5.7 Identification of disease-specific metabolite profiles

A requirement for the adoption of neural networks in medicine is explainability (Gilpin et al. 2018; Molnar 2022). SHapley Additive exPlanations (SHAP) is a post-hoc attribution method informing on the importance of individual input variables for a given prediction (Lundberg 2022; Lundberg and S.-I. Lee 2017). Thus, applied to the metabolomic state model, informs on which metabolites most affect disease risk. Generally, the larger the absolute SHAP value, the more important a metabolite is for an individual prediction. Based on the direction of the effect of a metabolite's contribution, increasing or decreasing the predicted risk, the SHAP value can take positive or negative values (see section 2.5 on page 64).

First, to understand the role of individual metabolites over the disease spectrum, absolute SHAP values were calculated per metabolite and disease (figure 5.6 on the next page, figure 12 on page 220). Most of the high-impact metabolites were linked to multiple diseases: Plasma levels of metabolites with consistently high contribution included the amino acids glutamine, glycine, and tyrosine, metabolites related to carbohydrate metabolism, albumin, the kidney function marker creatinine, Glycoprotein Acetylation (GlycA), and the ketone bodies acetone and acetoacetate. An important role was further implied for fatty acids such as Arachidonic Acid (AA) and multiple lipoprotein components, including free cholesterol in very large HDL, triglycerides in large LDL, phospholipids in small LDL, and sphingomyelins. Beyond the shared metabolite profiles between diseases, marked associations were notable between creatinine and abdominal aortic aneurysm, glucose and type-2 diabetes, as well as between Glycoprotein Acetylation (GlycA) and both lung cancer and chronic obstructive pulmonary disease (COPD). Notably, for diseases with a strong discriminatory value of the metabolomic state (e.g., type-2 diabetes), the predicted metabolite contributions were considerably higher than for diseases with less discriminatory metabolomic information (figure 5.6 on the next page).

The two diseases with the strongest metabolic contributions over the comprehensive clinical predictors and indications of clinical utility (see above), type-2 diabetes (figure 5.6 on the facing page) and all-cause dementia (figure 5.6 on the next page), were further investigated for individual metabolomic profiles. For the prediction of metabolomic state for type-2 diabetes, metabolites related to carbohydrate metabolisms, such as glucose and lactate, showed the highest importance for the metabolomic state model (figure 5.6 on the facing page). Further, metabolites such as amino acids, ketone bodies, lipids, fatty acids, and creatinine and albumin displayed high relevance, all in line with earlier findings (Mahendran, Cederberg,

5 Metabolomic profiles predict individual multi-disease outcomes

Figure 5.6: Introspection of the metabolomic state informs on metabolite profiles associated with disease risk. **a**, Displayed is a heatmap of metabolite importances, quantified by the absolute global SHAP value estimates per endpoint for the 75 globally most important metabolites. Important metabolites are often relevant for multiple diseases. The endpoints are sorted by the discriminative performance of the metabolomic state (left to right, see figure 5.3 on page 110). Generally, the more predictive the metabolomic state, the higher the absolute attributed metabolite importances. **b**, Displayed is the global metabolite attribution profile for type-2 diabetes. Individual attributions are aggregated by percentiles, where each dot represents one percentile. The more distant a dot from the circular baseline, the stronger the absolute attribution for the percentile. Deviations towards the center represent negative (protective), and deviations towards the outside represent positive (risk-increasing) contributions to the metabolomic state. The color indicates the metabolite's normalized mean plasma value. Glycolysis-related metabolites are estimated to have the highest impact, with many other metabolites contributing, such as various amino acids, fatty acids, and lipoprotein components. **c**, Global metabolite attributions for all-cause dementia. Amino acids, fluid balance metabolites, fatty acids, glycolysis-related metabolites, and ketone bodies contribute most to the predicted metabolomic state.

et al. 2013; Mahendran, Vangipurapu, et al. 2013; Würtz, Tiainen, et al. 2012). Additionally confirmed were the association of higher plasma levels of the fatty acids DocosaHexaenoic Acid (DHA) and LA with lower risk (Virtanen et al. 2014; J.H. Wu et al. 2017). The observed contribution of the lipid content across the whole density gradient of lipoproteins, including a high TriGlyceride (TG) content in large and medium-sized LDL or free cholesterol content in very small VLDL and HDL, was very distinct with notable associations with lower risk, as well as with increased risk. For all-cause dementia, creatinine, albumin, and the amino acids glutamine, leucine, and tyrosine were found to be predominant contributors to predicted risk (figure 5.6 on the previous page). Similarly to type-2 diabetes, observed associations were in line with earlier findings: fatty acids such as LA, Monounsaturated Fatty Acid (MUFA), and Saturated Fatty Acid (SFA) were associated with increased risk, while branched-chain amino acids displayed protective effects. The findings further implicated associations of glucose, the ketone bodies acetate, acetoacetate, acetone, and beta-hydroxybutyrate. Finally, several lipoproteins were strongly associated, most notably free cholesterol in very large HDL and cholesteryl ester in extremely large VLDL. Comprehensive data for all investigated endpoints, including the most important metabolites, can be found in the supplements (figure 12 on page 220, table 14 on page 321). Disease-specific attribution profiles for all assessed endpoints are provided in the figure 13 on page 233.

5.8 High-risk individuals for type-2 diabetes differ in their personal metabolomic profiles

The computation of SHAP values enables derivation of attribution profiles for each individual and disease, informing on the specific contribution of metabolites to individual risk. Further, the computation of SHAP values allows the assessment of individual risk attribution

5.8 *High-risk individuals for type-2 diabetes differ in their personal metabolomic profiles*

profiles for each participant and each disease, informing on the impact of single metabolites on a given prediction. To explore the landscape of risk attribution profiles, attribution profiles for type-2 diabetes were embedded in 2-dimensional Uniform Manifold Approximation and Projection (UMAP) space and visualized along the UMAP coordinates (figure 5.7 on the next page) (McInnes, Healy, and Melville 2018).

Interestingly, the UMAP space is resolved by the estimated importance of glucose. Thus the SHAP values assigned to glucose regarding the predicted risk for type-2 diabetes (figure 5.7 on the following page). This point is further illustrated in the example of three high-risk individuals whose attribution profiles are presented with highlighted high-impact metabolites ($\text{SHAP} \in [-0.2, 0.2]$) corresponding to the top and bottom 1 % percentiles). While most high-risk individuals (top 1 % metabolomic state) were located at coordinates with strong glucose attribution, high-risk individuals were scattered over the entire attribution space (figure 5.7 on the next page). Most notably, the attribution profiles of high-risk individuals were not consistently dominated by glucose but by, for instance, low levels of albumin, LA and DHA, histidine, and glycine (figure 5.7 on the following page). This observation is further reflected in the 1H-NMR metabolite concentrations: here, substantial differences between concentrations of glucose, LA, fatty acids, and triglycerides became apparent when comparing individuals in the area with the most substantial glucose attribution with individuals in two spatially distinct, high-risk UMAP regions (figure 14 on page 234).

In summary, these findings demonstrate the metabolomic state model to leverage established metabolite-disease associations and illustrate the potential of attribution methods to inform on the impact of single metabolites on an individual's predicted risk.

5 Metabolic profiles predict individual multi-disease outcomes

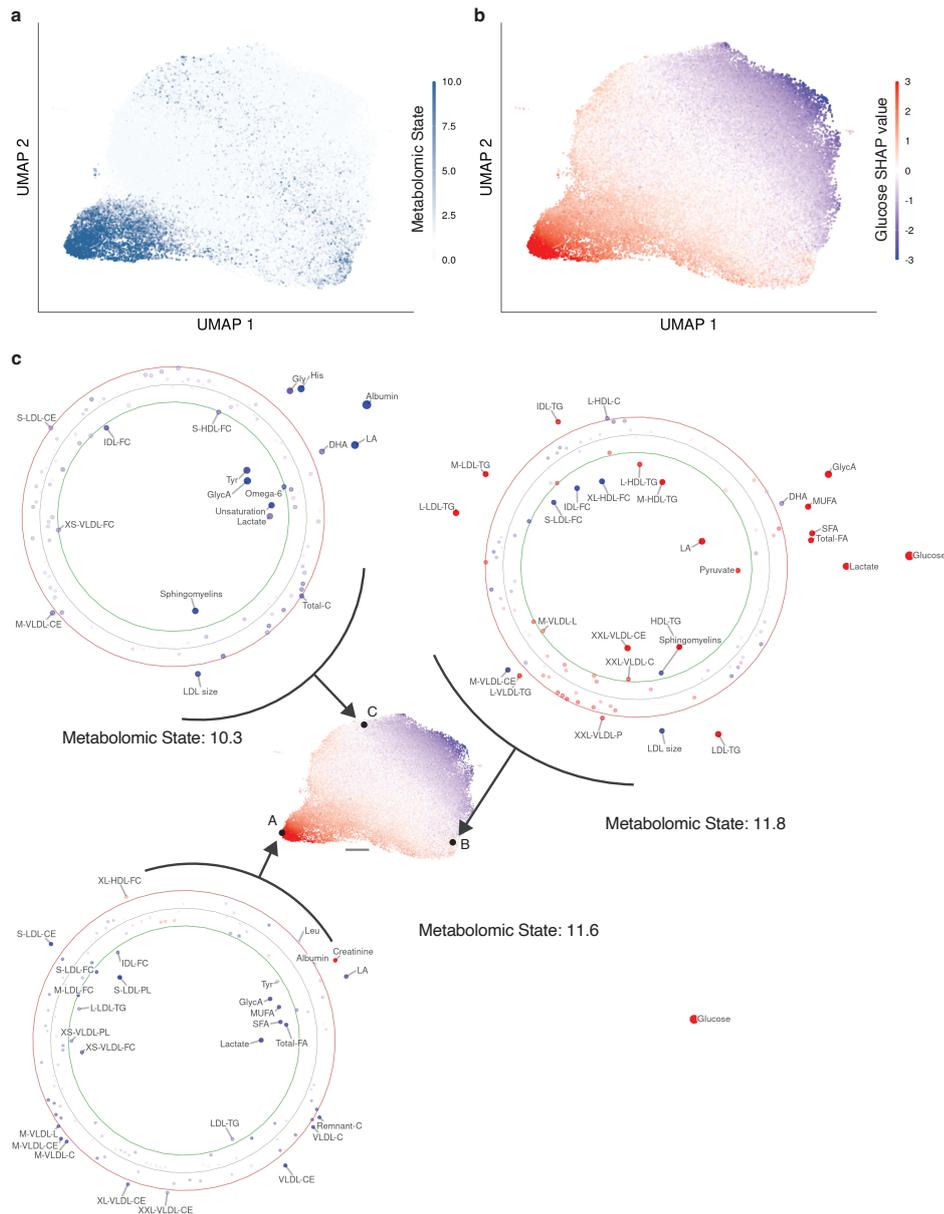


Figure 5.7: Individual attribution profiles diverge for high-risk individuals in type-2 diabetes. The Uniform Manifold Approximation and Projection (UMAP) projection allows an assessment of the complex, high-dimensional manifold of attribution values in 2-dimensional space. For visualization, 41 unconnected outliers of 117,981 total observations were excluded. **a**, Displayed is the high-dimensional manifold of attribution values of the SHAP metabolite attributions for type-2 diabetes in 2-dimensional UMAP space. Visualized is the entire study population. Each dot represents an individual and is colored by the individual's predicted metabolomic state for type-2 diabetes. **b**, The same 2-dimensional UMAP colored by the Glucose SHAP value. **c**, Displayed are three individual attribution profiles for high-risk (metabolomic state > 10, top 1% metabolomic state percentile) individuals, indicated by the letters A, B, C in the central UMAP. The three individual attribution profiles are dominated by different metabolites. The scale bar represents a unit in the UMAP space. The individual attribution profiles are set up equivalently to figure 5.6 on page 119: Each point in an individual attribution profile indicates one metabolite; the position, size, and color of the point indicate the magnitude and direction of the attributed contribution to the predicted risk. The green and red circles represent the bounds of the top and bottom percentile of the global SHAP distribution, respectively, indicating outliers in the SHAP global distribution.

5.9 Discussion

A robust and accurate assessment of future disease risk is a critical component of disease prevention. However, comprehensive risk assessment often requires the careful acquisition of predictors, one disease at a time. In consequence, the resources (time and cost) required for the collection can severely limit the adoption and utility (Steyerberg, Moons, et al. 2013) for each disease-specific risk score individually. Interestingly, many common diseases involve metabolic alterations, and human blood metabolomic patterns contain rich systemic information on the underlying physiology (Ahadi et al. 2020; Deelen et al. 2019; Julkunen et al. 2021; Nicholson et al. 2012; Pietzner et al. 2021; Schüssler-Fiorenza Rose et al. 2019). While individual metabolites have long been linked to disease risk, systemic information from blood metabolomics could inform on multiple diseases at once. Importantly, in recent years, assays such as Proton Nuclear Magnetic Resonance ($^1\text{H-NMR}$) spectroscopy have matured and allow the retrieval of information on serum metabolites very robustly and at comparatively low costs (Soininen et al. 2015; Würtz, A.J. Kangas, et al. 2017). However, the potential of the metabolomic profile as a single-domain multi-disease assay in primary care has not been investigated thus far.

This experiment explored the potential of NMR-derived metabolomic profiles as a tool for individualized outcome prediction across 24 common diseases. Over 1.4 ~million person-years of follow-up allowed a robust exploitation of the systemic information in metabolomic profiles to derive integrative metabolomic states for many diseases simultaneously. The metabolomic states were found to be predictive for all but one of the investigated diseases, and findings were externally validated in four independent cohorts for available endpoints. Further, the predictive value beyond clinical variables was assessed, and a subset of endpoints with potential clinical utility was identified. Finally, the experiment examined metabolite attributions confirming a multitude of disease-associated metabolites as well as a shared metabolomic background of common diseases.

Importantly, the predictive information of the metabolomic state matched established clinical variables for multiple of the investigated endpoints. In line with previous reports on NMR-metabolite associations, this confirms the predictivity of metabolomic profiles for e.g., type-2 diabetes (Ahola-Olli et al. 2019), dementia (Tynkkynen et al. 2018), and cardiovascular diseases (Deelen et al. 2019; Holmes et al. 2018; Würtz, A.S. Havulinna, et al. 2015), such as coronary heart disease and heart failure (Delles et al. 2018). Generally, the additional predictive information decreased when combined with increasingly compre-

hensive clinical predictors. This indicated that substantial parts of the metabolomic state's discriminatory information were shared with established clinical predictors. Despite that, the metabolomic state contained complementary information that added predictive value even over comprehensive laboratory measurements for multiple endpoints, including type-2 diabetes, all-cause dementia, and heart failure. These findings largely translate into the potential clinical utility of NMR-based metabolomic profiling as a replacement for comprehensive laboratory examinations and as an additional source of discriminatory information to refine comprehensive risk assessments.

The attributions for each individual allowed an assessment of how differences in the metabolomic profiles affect disease risk. Here, the impact of metabolites like albumin and creatinine was confirmed. Albumin and creatinine are two metabolites that have previously been associated with all-cause and disease-specific mortality (Deelen et al. 2019; Fischer et al. 2014) and are already measured in routine care (Ronit et al. 2020; Wannamethee, Shaper, and Perry 1997). Further, the experiment confirmed multi-disease associations for Ainoic Acid (LA), tyrosine, glycine, and cholesteryl esters in extremely large VLDL, reinforcing the hypothesis of systemic metabolomic information (Pietzner et al. 2021). Dissecting the disease-specific attribution profiles for type-2 diabetes and dementia, multiple metabolite-disease associations previously reported in the literature were observed. In the case of type-2 diabetes, associations included metabolites beyond glucose: a positive association between high levels of glycoprotein acetyls, branched-chain amino acids, lactate, and fatty acids (Monounsaturated Fatty Acid (MUFA) and Saturated Fatty Acid (SFA)) and the predicted disease risk, as well as protective associations with lower disease risk such as linoleic acid or glycine (Ahola-Olli et al. 2019; Mahendran, Cederberg, et al. 2013) were discovered. For dementia, associations with branched-chain amino acids, including leucine and valine, and fatty acids, most notably LA, were replicated (S.J. v.d. Lee et al. 2018; Tynkkynen et al. 2018). Further, on a broader disease spectrum, associations of Glycoprotein Acetylation (GlycA) with cardiovascular disease, type-2 diabetes, COPD, and lung cancer were reflected in the attributions (Kettunen et al. 2018; Komaromy et al. 2020). In consequence, these findings suggest the metabolomic state capture systemic information in NMR-derived metabolomic profiles based on established shared and highly specific metabolite-disease associations.

Further, the findings render Proton Nuclear Magnetic Resonance ($^1\text{H-NMR}$) metabolomics profiling an attractive candidate for a single-domain multi-disease assay to identify individuals at risk and to guide interventions potentially. Since many countries already recommend regular check-ups entailing blood tests for adults in the prevention and early detection of se-

lected common diseases (Bundesgesundheitsministerium 2022; NHS 2022), the results of this experiment demonstrate the predictive potential of NMR-metabolomic profiling in combination with easy-to-obtain clinical predictors, including age and biological sex, but also to refine risk trajectories in combination with comprehensive laboratory assays. Additionally, the estimation of metabolomic risk profiles over many common diseases at once could be of potential value for the guidance of not just pharmacological but also lifestyle interventions. This is especially relevant for diseases such as type-2 diabetes, where interventions on modifiable risk factors have been shown to delay disease onset (Pronk, Remington, and Force* 2015) and prevent subsequent comorbidities (Haffner et al. 1998; Yusuf et al. 2004). Similarly, the Lancet 2020 commission suggested that up to 40 % of worldwide dementia may be preventable by interventions on modifiable risk factors (Livingston et al. 2017). Here, an easily accessible parameter – below the level of, e.g., cerebrospinal fluid analysis – could ideally serve as an argument to motivate affected persons for preventive lifestyle adjustments (Rosenberg et al. 2018). This is particularly compelling, as today’s pharmacological treatment options for dementia are scarce and of only insufficient symptomatic benefit (S.-M. Wang et al. 2021). However, the efficacy of various lifestyle interventions is disputed (Ngandu et al. 2022; Silarova et al. 2019), calling for further experimental investigation.

Before application in routine care, substantial challenges remain to be solved. While the Proton Nuclear Magnetic Resonance ($^1\text{H-NMR}$) assay is very robust and orders of magnitude cheaper than Mass-Spectrometry (MS) based alternatives, the sensitivity is lower. Also, the current metabolite coverage is relatively narrow and lipid-focused (Soininen et al. 2015; Würtz, A.J. Kangas, et al. 2017). Although the metabolite coverage is expected to be expanded in the future, it presents a limitation for the clinical utility to date. Further, while the NMR assay is by design very robust, downstream quantification from the raw NMR spectra requires harmonization of individual measurements for the reliable application of multivariable prediction models. Additionally, even though the applied study population for model development comprised almost 120 thousand individuals, the population is more healthy and less deprived than the general UK population (Fry et al. 2017). While the external validation in the four independent cohorts indicated a general transferability of the findings, the validation was limited by the available endpoint information. Thus, validation could only be performed for a subset of 7 endpoints. In light of the limitations, an application of the metabolomic state model to predict disease risk beyond the validated conditions and in specific populations outside the research context is not feasible. Ultimately, a broad roll-out of NMR metabolomics for clinical care requires multiple logistical questions to be addressed,

5 Metabolomic profiles predict individual multi-disease outcomes

including both sample processing and transport.

Taken together, the integrated analysis of NMR-derived metabolomic profiles demonstrates the potential of metabolomic profiles for predicting and preventing diseases such as type-2 diabetes, heart failure, and dementia. In addition, the experiment illustrates the potential of metabolomic profiling as a multi-disease assay for the personalized prevention of many common diseases at once.

6 Medical history predicts future health trajectories over the human phenome

6 Medical history predicts future health trajectories over the human phenome

The findings presented in this chapter include contributions by Dr. Jakob Steinfeldt and Dr. Benjamin Wild. Specifically, Dr. Jakob Steinfeldt contributed with the mapping of the Electronic Health Record (EHR) data to the OMOP CDM standard, and jointly with Dr. Benjamin Wild to the implementation of the medical history model.

6.1 Experimental setup

This experiment explores the predictive potential of an individual’s medical history and proposes a systematic approach for phenome-wide risk stratification. The experiment introduces the medical history model, a deep neural network, simultaneously estimating disease risk for 1,883 endpoints, including common and actionable metabolic, cardiovascular, respiratory, musculoskeletal, and neurological disorders (figure 6.1). For this experiment, the prior medical history is defined as the entirety of routine health records before recruitment.

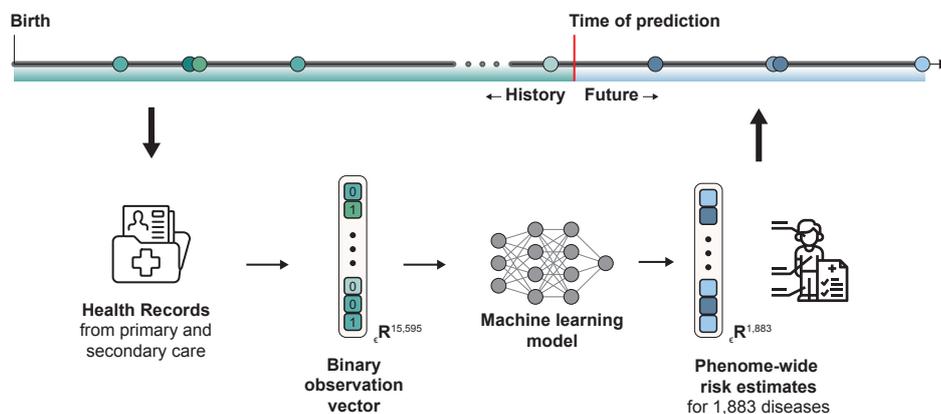


Figure 6.1: Utilizing individual medical histories for phenome-wide disease onset prediction. The medical history captures primary and secondary care encounters, including diagnoses, medications, and procedures (ideally) from birth. The data is aggregated into a binary observation vector indicating if a medical code was present prior to baseline (i.e., the time point of prediction). A multi-layer perceptron is trained on the observational data to simultaneously predict phenome-wide incident disease onset for 1,883 endpoints.

The medical history model was developed, trained, and validated on aggregated electronic health records from the UK Biobank cohort (Bycroft et al. 2018; Sudlow et al. 2015), and details on the model’s architecture are presented in section 2.6.3 on page 69. Details on the data preprocessing and aggregation are presented in section 3.2.4 on page 80. Similarly to the experiments previously described section 4 on page 87, the metabolomic state model was developed and trained in 22-fold spatial cross-validation over the recruitment centers of the UK Biobank cohort (see section 3.2.6 on page 81). Subsequently, the predicted risk states were integrated with Cox Proportional Hazards (CPH) models (Cox 1972) to scrutinize the

6.2 Characteristics of the study population

phenome-wide predictive potential over simple demographic predictors (age and biological sex), and for cardiovascular diseases over guideline-recommended scores for primary prevention (SCORE2, ASCVD, and QRISK3) (Goff et al. 2014; Hippisley-Cox, C. Coupland, and Brindle 2017; SCORE2 working group and ESC Cardiovascular risk collaboration 2021). For an overview of the utilized predictors, see table 3.1 on page 77. Ultimately, the learned phenome-wide risk states were investigated for their capability to react to emerging health threats in the example of COVID-19. In summary, the results demonstrate the potential of routine health records for the simultaneous prediction of phenome-wide disease risk.

6.2 Characteristics of the study population

The medical history model was developed on Electronic Health Record (EHR) data of individuals of the UK Biobank cohort, representing a sample of the UK general population (Bycroft et al. 2018) (see section 3.1.1 on page 74). In total, EHR data, clinical predictors, and disease endpoints for 502,460 individuals were extracted. The UK Biobank cohort is a longitudinal population cohort with relatively healthy individuals of mostly British descent, with a median age of 58 (IQR 50, 63) years, 54.4 % biological females, 11 % current smokers, and a median BMI of 26.7 (IQR 24.1, 29.9) at recruitment (see table 6.1). Individuals were followed for a median of 12.6 years resulting in ~6.2M overall person-years on 1,883 phenome-wide endpoints (P. Wu et al. 2019) with more than 100 incident events (see section 3.3.3 on page 85).

Table 6.1: Summary statistics of the study population. Reported are medians (IQR) and absolute counts (%).

		Male, N = 229,107	Female, N = 273,353	Overall, N = 502,460
Age (years)		58 (50, 64)	57 (50, 63)	58 (50, 63)
Ethnicity	Asian	5,878 (2.6%)	5,575 (2.0%)	11,453 (2.3%)
	Black	3,407 (1.5%)	4,653 (1.7%)	8,060 (1.6%)
	Mixed	1,105 (0.5%)	1,851 (0.7%)	2,956 (0.6%)
	White	215,244 (95%)	257,413 (96%)	472,657 (95%)
	Unknown	3,473 (1.5%)	3,861 (1.4%)	7,334 (1.5%)
Smoking status	Current	28,607 (13%)	24,364 (9.0%)	52,971 (11%)
	Previous	87,604 (38%)	85,440 (31%)	173,044 (35%)
	Never	111,460 (49%)	162,037 (60%)	273,497 (55%)
	Unknown	1,436 (0.6%)	1,512 (0.7%)	2,948 (0.5%)
Body Mass Index		27.3 (25.0, 30.1)	26.1 (23.5, 29.7)	26.7 (24.1, 29.9)
	Unknown	1,646 (0.7%)	1,458 (0.6%)	3,104 (0.6%)
Systolic Blood Pressure (mmHg)		139 (128, 152)	133 (121, 147)	136 (124, 150)
	Unknown	13,579 (5.9%)	16,536 (6.0%)	30,115 (6.0%)

6.3 The medical history is associated with future disease onset

The central prerequisite of clinical utility is a predictor's ability to stratify a population by risk for disease onset. The more accurately low- and high-risk individuals can be distinguished, the more effective targeted interventions, prevention, and disease diagnoses become. Thus, to answer the question of whether health records contain information to identify high-risk individuals, the relationship between the estimated endpoint-specific risk states and the risk of future disease was assessed over the entire human phenome (figure 6.2 on page 132). First, the incident events were counted over percentile ranks of the predicted risk states for each endpoint, and subsequently, event rate ratios between the top and bottom 10 % of risk states were calculated. Importantly, the risk states predicted from health records alone resulted in a substantial separation of event rates in the top and bottom 10 % of risk states, reflecting a strong stratification of high and low-risk individuals for almost all endpoints covering a broad range of disease categories and etiologies. For 1,404 of 1,883 endpoints (74.6 %), stratification by risk states resulted in >10-times more events for individuals in the top 10 % of the predicted risk states compared to the bottom 10 %. For instance, these endpoints included rheumatoid arthritis (Ratio ~12.5), coronary heart disease (Ratio ~23), or chronic obstructive pulmonary disease (Ratio ~63). For 286 (15.1 %) of the 1,883 conditions, including abdominal aortic aneurysm (Ratio ~212), and all-cause death (Ratio ~107), individuals in the top 10 % suffered more than 100 times the number of incident events compared to the bottom 10 %. In contrast, the separation between high and low-risk individuals was smaller (Ratio less than 10) for 479 (25.4 %) endpoints, which included hypertension (Ratio ~5.5) and anemia (Ratio ~6.9), both diagnoses often made earlier in life or precursors for future comorbidities. Importantly, the ratios were greater than 1 for all 1,883 investigated endpoints, even though all models were developed in spatially segregated assessment centers. The complete list of incident event counts over risk state deciles is presented in table 15 on page 324, and the full list of event rate ratios for all endpoints can be found in table 16 on page 340.

In addition to the phenome-wide distribution of event rate ratios, detailed associations between the risk percentiles and incident event ratios, as well as cumulative event rates for up to 15 years of follow-up for the top, median, and bottom percentiles, were calculated for a subset of 24 selected endpoints (see figure 6.2 on page 132). These endpoints include actionable and common diseases with significant societal burdens, specific cardiovascular conditions with pharmacological and surgical interventions, as well as additional interesting

6.3 The medical history is associated with future disease onset

endpoints without any preexisting tools to stratify risk to date.

In summary, the results demonstrated that disease-specific risk states predicted from individual medical histories (i.e., EHR data) are able to capture the risk of onset for all 1,883 investigated endpoints over a wide range of disease categories and etiologies. Consequently, routine health records contain a vast, currently unused potential for the phenome-wide risk estimation of disease onset in the general population.

6 Medical history predicts future health trajectories over the human phenome

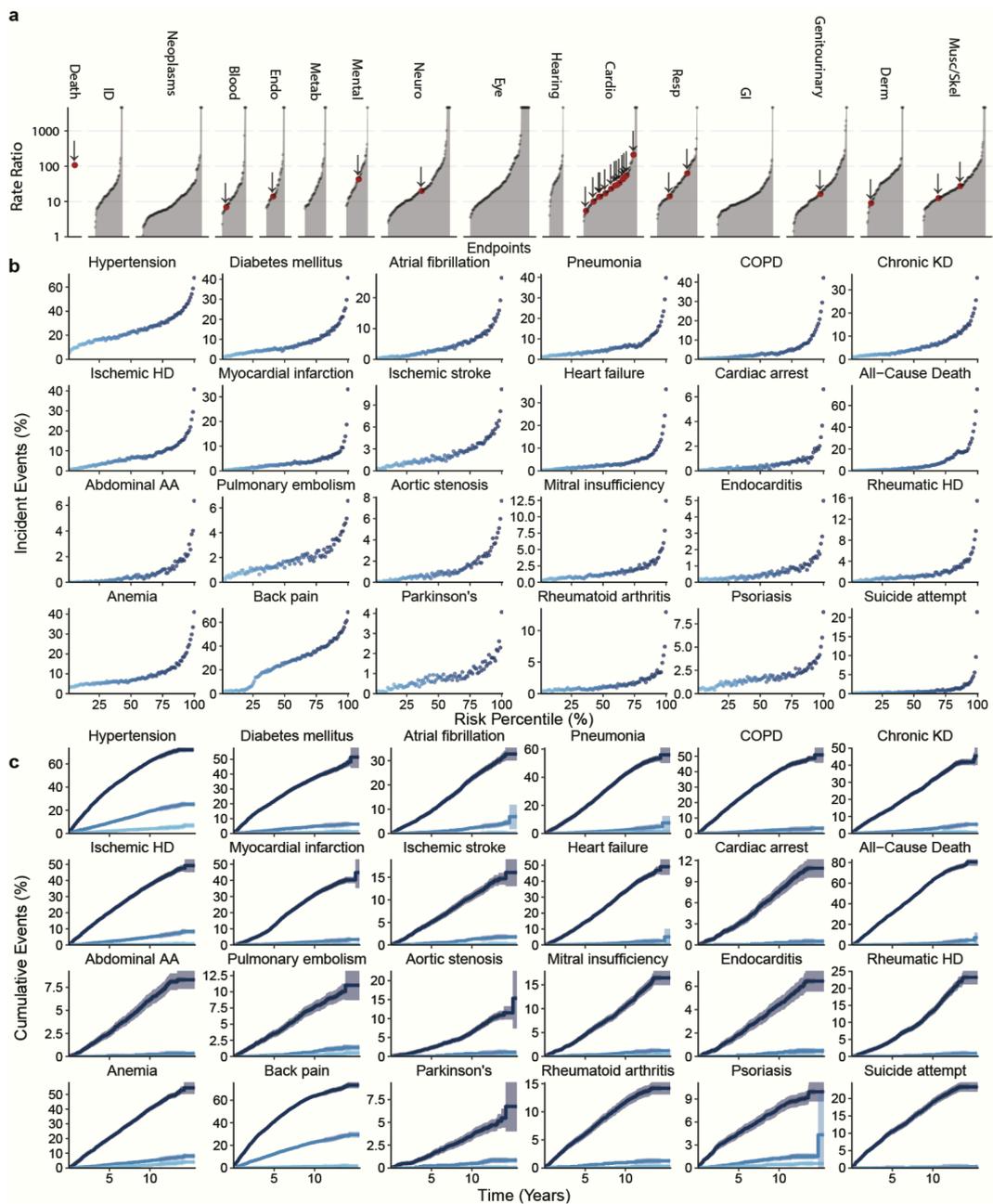


Figure 6.2: Routine health records stratify phenome-wide disease onset. **a**, Event rate ratios of incident events between the top 10 % and the bottom 10 % of the estimated risk states for each endpoint. Event rate ratios are greater than 1 for all of the 1,883 investigated endpoints indicating the presence of relevant information in Electronic Health Record (EHR) to predict future disease onset. **b**, Observed incident event frequencies plotted against risk state percentiles over the entire study population for a selection of 24 endpoints. **c**, Cumulative event rates over the observation time for selected endpoints, stratified by risk state quantiles (light blue: bottom 1 %, blue: median 1 %, dark blue: top 1 %). Stratification by risk state yields separated risk trajectories. Statistical measures were derived from 502,460 individuals. Individuals with prevalent diseases were excluded from the endpoint-specific analyses.

6.4 Medical history yields discriminative improvements over basic predictors

While routinely collected health records contain information that is able to stratify the population by disease risk, this does not prove clinical utility. Towards an application, the extracted information should ideally inform beyond ubiquitously available predictors. Thus, the predictive information in EHR data was scrutinized against age and biological sex to investigate clinical utility further. Specifically, the risk of disease onset was modeled for all 1,883 endpoints using CPH models trained on age and biological sex (Age+Sex) and age, biological sex, and the learned risk state. This allowed a direct performance comparison between the baseline model (Age+Sex) and the model augmented with the medical risk state (Age+Sex+MedicalHistory), and further the calculation of adjusted hazard ratios (see table 18 on page 394) and 10-year discriminative improvements (denoted as Delta C-index, see figure 6.3 on the following page).

The addition of information from the medical history resulted in significant improvements over the baseline model (Age+Sex) for 1,800 (95.6 %) of the 1,883 investigated endpoints (figure 6.3 on the next page, table 17 on page 356) and for a wide array of these endpoints, the discriminative improvements were considerable (Delta C-Index Q_{25} : 0.099, Q_{50} : 0.119, Q_{75} : 0.139). For the highlighted subset of 24 endpoints (defined in figure 6.2 on the preceding page), the most remarkable discriminative improvements were observed for the prediction of suicide attempts (C-Index: 0.608 (CI 0.602, 0.615) \rightarrow 0.831 (CI 0.826, 0.837)), back pain (C-Index: 0.519 (CI 0.517, 0.521) \rightarrow 0.72 (CI 0.719, 0.722)), all-cause mortality (C-Index 0.701 (CI 0.699, 0.703) \rightarrow 0.878 (CI 0.877, 0.88)) and Chronic Obstructive Pulmonary Disease (COPD) (C-Index 0.662 (0.66, 0.666) \rightarrow 0.818 (CI 0.815, 0.82)). In contrast, other endpoints, such as Parkinson's disease (C-Index: 0.738 (CI 0.732, 0.745) \rightarrow 0.737 (CI 0.731, 0.743)) did not benefit from the information contained in the medical history.

These findings indicate that health records contain substantial predictive information beyond fundamental demographic predictors for a wide range of endpoints from diverse disease categories, indicating potential utility for risk stratification. Thus, individual phenome-wide risk estimations have the potential to inform on actionable disease endpoints systematically.

6 Medical history predicts future health trajectories over the human phenome

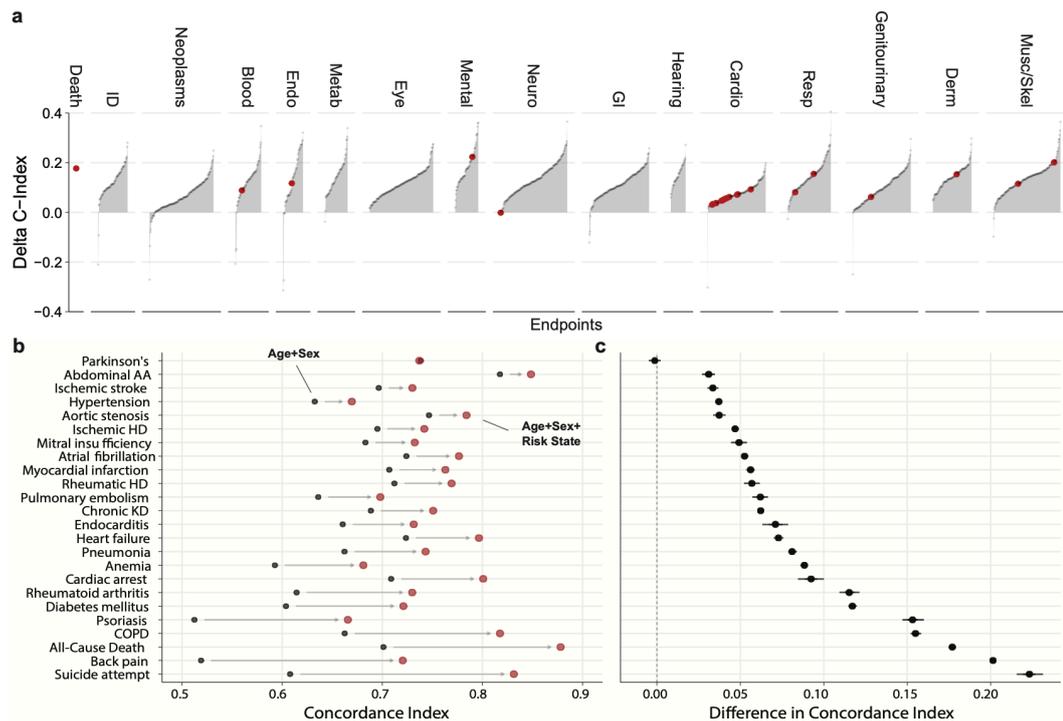


Figure 6.3: Discriminative performance of the medical history addition over basic demographic predictors. **a**, C-Index deltas between CPH models trained on Age+Sex (baseline) and Age+Sex+MedicalHistory for all 1,883 endpoints. 1,800 (95.6%) of 1,883 investigated endpoints show significant improvements over the baseline model. **b**, Absolute C-Index values for the baseline (Age+Sex, black point) and the model including the estimated risk state (Age+Sex+MedicalHistory, red point) for a selection of 24 endpoints. **c**, Direct C-index differences between the baseline (Age+Sex) and the model including the estimated risk state (Age+Sex+MedicalHistory) for a selection of 24 endpoints. All performances are reported as medians (dots), and whiskers extend to the 95% CI as estimated by bootstrapping over 100 iterations. Individuals with prevalent diseases were excluded from the endpoint-specific analyses.

6.5 Learned medical risk states are applicable for cardiovascular prevention and public health reactions

In contrast to the vast majority of endpoints, in the primary prevention of cardiovascular diseases, the utilization of risk scores has long been established. In fact, the use of risk scores is recommended by major international guidelines for the primary prevention of cardiovascular events to guide preventive lipid-lowering interventions (National Institute for health and Care Excellence (NICE) 2014). While the predictors for these primary prevention scores are accessible at a low cost, they require physical and laboratory measurements and must be manually updated by healthcare providers. If the information derived from the medical history was competitive to these predictors, this could substantially reduce friction and cost in primary care. Thus, the risk of cardiovascular disease onset was modeled using CPH models and predictors from well-validated primary prevention scores, the European Society of

6.5 *Learned medical risk states are applicable for cardiovascular prevention and public health reactions*

Cardiology SCORE2 (SCORE2 working group and ESC Cardiovascular risk collaboration 2021), the American Heart Association ASCVD (Goff et al. 2014) and the British QRISK3 (Hippisley-Cox, C. Coupland, and Brindle 2017) score, recommended by the NHS Health Check (NHS 2022).

Interestingly, the discriminative performance of a simplified risk model, including only age, sex, and the predicted risk state (Age+Sex+MedicalHistory), was competitive or better than the established cardiovascular risk scores for all investigated cardiovascular endpoints (figure 6.4 on page 137, table 19 on page 414), with differences in the C-Index of +0.002 (CI -0.001, 0.005) for ischemic stroke, +0.003 (CI 0.001, 0.005) for ischemic heart disease and +0.007 (CI 0.004, 0.009) for myocardial infarction compared with the comprehensive QRISK3 score. In addition, the discriminative improvements for later stage outcomes, including heart failure (+0.02 (CI 0.017, 0.022)), cardiac arrest (+0.055 (CI 0.049, 0.062)), and all-cause death (+0.136 (CI 0.134, 0.138)), were dramatically better when the prior health records are considered.

Thus far, this experiment has shown the potential utility of routine health records for the onset prediction of a broad spectrum of 1,883 established disease endpoints over the human phenome and the direct applicability of the information extracted from medical history for the primary prevention of cardiovascular diseases. However, novel diseases like COVID-19 do not have defined medical codes. Nevertheless, these emerging health threats may require fast and reliable risk stratification in light of scarce or missing information on outcomes. Here, pre-trained medical risk states estimated from routinely collected EHRs may be repurposed to stratify and prioritize patients.

This possible application was examined at the recent example of the COVID-19 pandemic. COVID-19 is a respiratory infection with pneumonia and sepsis as common, life-threatening complications of severe cases. Thus, the risk states (i.e., the log partial likelihoods estimated from the routine health records) for pneumonia, sepsis, and all-cause death were repurposed to calculate a severity score under strict limitation to information available at the end of 2019 (the time before COVID-19 spread globally). Interestingly, a combination of age and the three risk states improved the discriminative performance for both severe and fatal COVID-19 outcomes over age alone (Severe: C-Index (Age) 0.62 → C-Index (Age+RiskStates) 0.65 and Fatal: C-Index (Age) 0.70 → C-Index (Age+RiskStates) 0.78). In an applied setting, these discriminative improvements would have translated into higher cumulative incidence in the Top 5 % of the population compared to age alone (figure 6.4 on page 137). In the top 5 % of the age-based risk score (~79 (IQR 77, 81) years old), 0.42 %

6 Medical history predicts future health trajectories over the human phenome

(CI 0.34 %, 0.5 %, n=105) would have been hospitalized, and 0.26 % (CI 0.2 %, 0.33 %, n=66) would have died by the end of the first wave. By the end of the second wave, around 0.96 % (CI 0.83 %, 1.08 %, n=240) would have been hospitalized and 0.44 % (0.36 %, 0.52 %, n=111) would have died. When stratified based on the estimated risk states, much higher rates of hospitalization and death were found when selecting the same number of individuals: By the end of the first wave, around 0.61 % (CI 0.51 %, 0.71 %, n=153) would have been hospitalized, and 0.54 % (0.45 %, 0.63 %, n=136) would have died. By the end of the second wave, 1.24 % (CI 1.1 %, 1.38 %, n=312) would have been hospitalized and 0.83 % (0.72 %, 0.94 %, n=208) would have died.

Thus, in summary, these findings showed that the predictive information in routine health records could be competitive in identifying high-risk individuals with established guideline-recommended risk scores for the primary prevention of cardiovascular diseases. Further, the extracted predictive information captured in risk states may be leveraged to respond to emerging health threats. Notably, both applications, in primary prevention of cardiovascular disease and public health, require no additional data or dedicated measurements.

6.5 Learned medical risk states are applicable for cardiovascular prevention and public health reactions

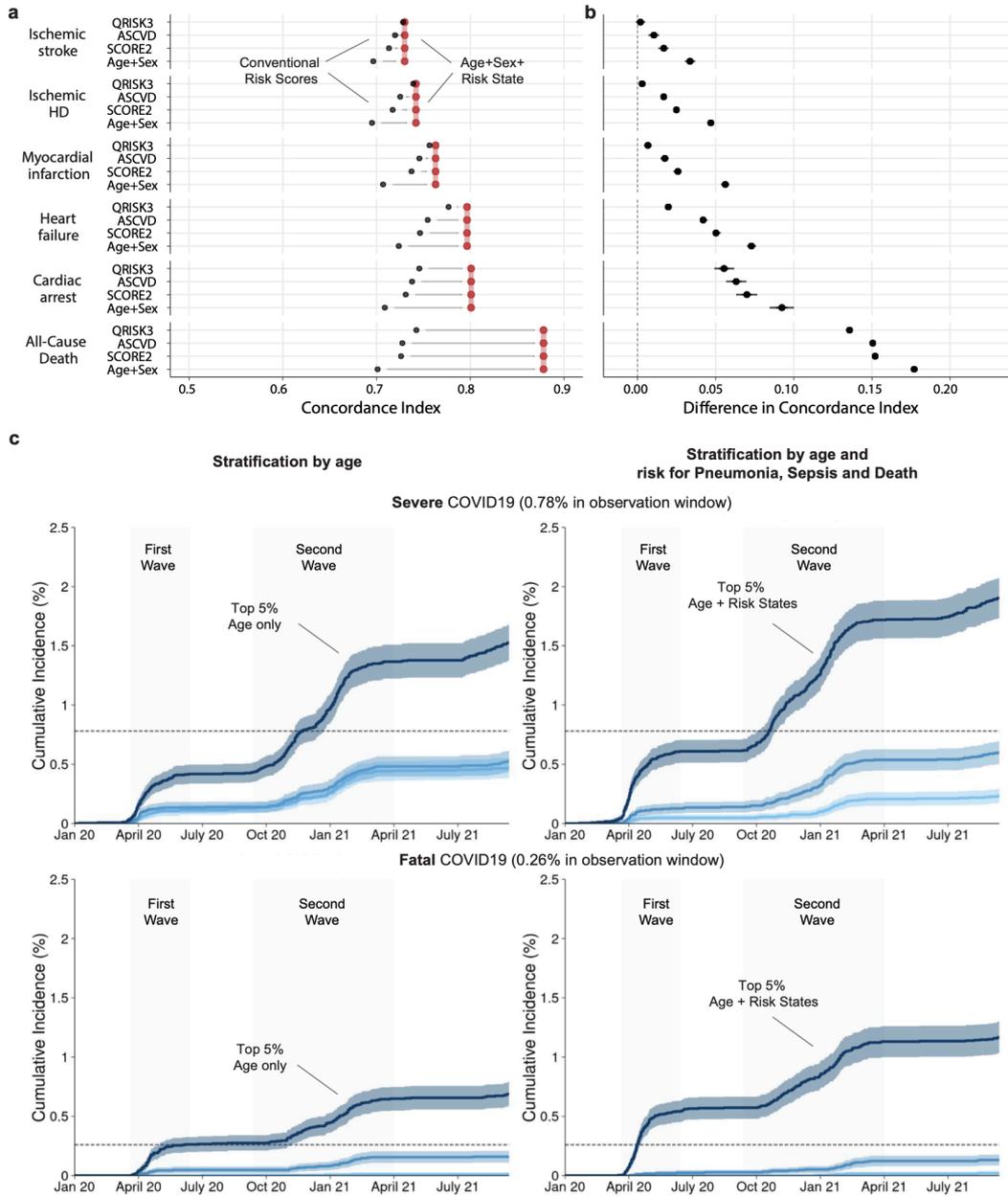


Figure 6.4: Showcases of applications for learned risk states. **a**, Absolute C-Index comparing risk scores (age and sex, SCORE2, ASCVD, and QRISK3 as indicated, black point) with the simplified model (Age+Sex+MedicalHistory, red segment). **b**, Direct C-Index comparisons matching the models compared in **a**. All performances are reported as medians (dots), and whiskers extend to the 95 % CI as estimated by bootstrapping over 100 iterations. Individuals with prevalent diseases were excluded from the endpoint-specific analyses. **c**, Estimated cumulative event trajectories of severe (individual was hospitalized) and fatal COVID-19 outcomes stratified by the top, median, and bottom 5 % based on age (left) or risk states of pneumonia, sepsis, and all-cause mortality as estimated by Kaplan-Meier analysis.

6.6 Discussion

Individual risk assessment is the foundation for personalized prevention and efficient early diagnostics through screening. In the current medical practice, however, a systematic risk assessment is severely limited by the spectrum of available risk scores and the necessity to acquire the predictors for their calculation. As existing risk models are centered around the prediction of a single disease at a time, the time and effort for dedicated measurements severely limit the clinical adoption, utility, and disease coverage of risk scores in primary prevention settings (Steyerberg, Moons, et al. 2013). At the same time, over recent years, longitudinal EHR have become available on the population level. These health records represent an individual's medical history, informing on prior diagnoses, treatments, medications, and a multitude of related procedural measures. Thus, EHR presents a potentially very informative data source readily available without dedicated measurement. However, the potential of longitudinal health records for systematic phenome-wide risk estimation has not been investigated thus far.

This experiment presented the medical history model, a neural network-based survival model ingesting information on an individual's medical history to simultaneously predict risk states for 1,833 endpoints. The medical history model allowed for a dedicated investigation of the information contained in EHR data to predict risk for future disease onset to enable primary prevention on a phenome-wide scale. The predicted risk states were evaluated in over 500,000 individuals and six million person-years of follow-up in the UK Biobank cohort. The individual medical history was found to be highly predictive for most investigated endpoints, generalizing across 22 spatially segregated recruitment centers and three healthcare provider systems from England, Scotland, and Wales. The results indicate that the information in EHR data provides potential utility beyond conventional predictors, even for diseases without existing risk stratification tools. Towards applicability, a simple combination of basic demographics and the predicted risk states was found to be competitive with guideline-recommended scores for the primary prevention of cardiovascular diseases. Further, investigating the example of the COVID-19 pandemic, risk states were found to be applicable as a quick-reaction tool to respond to emerging disease threats.

Importantly, the presented approach provides the first-ever means to assess the risk for future disease onset for many diseases at once. Three main scenarios of potential utility are imaginable: First, in the best case, the high-risk disease is preventable and effective interventions exist, e.g., lipid-lowering medication in the primary prevention of cardiovascular

disease. Second, while most diseases are not preventable, early detection can substantially slow down the progression and development of adverse events for other diseases, e.g., type-2 diabetes (Pozzilli, Stollo, and Bonora 2014), or systolic heart failure (Fonarow et al. 2011). Ultimately, even if an outcome is not preventable or treatable, estimates of prospective risk can be highly actionable, e.g., for personal decisions or advanced care planning. For illustration, a high risk of cardiac arrest could trigger discussions of the preferences for future medical care, including cardiopulmonary resuscitation, ventilatory support, or enteral feeding. Similarly, a high risk of death could enable prospective palliative consultations and shared decisions for transitioning from curative to palliative care in the right individuals (Adelson et al. 2017; Weissman and D.E. Meier 2011).

Further investigations are necessary before this approach can be deployed in routine care outside the UK Biobank. First, health records fundamentally differ from modalities commonly used in conventional risk scores (Vayena 2021) as they are recorded as a consequence of interactions with the medical system and are thus subject to biological, procedural, and socioeconomic biases. In addition, recorded information is not static but conditional on the ever-changing nature of medical knowledge and policy changes reflected in the applied coding systems. Therefore, differences in data providers and population characteristics can substantially complicate the learning of robust features and hamper transferability between healthcare systems. Another challenge is that large study populations and aggressive regularization are required to avoid overfitting due to the vast complexity of the records. Thus, despite the large cohort of over 500,000 individuals and over six million person-years of follow-up, about 55,000 of 71,000 original concepts had to be excluded due to low observation frequencies. Likewise, only the 1,883 most common endpoints with more than 100 incident events were included in the analysis. The third limitation arises from model development in the UK Biobank cohort and its restriction to the British healthcare system. While the findings indicated generalization between different locations in the UK (England, Scotland, and Wales), applications critically depend on the right choice of decision thresholds and, therefore, the absolute risk estimates. As the UK Biobank cohort is healthier and has lower disease prevalences than the UK general population (Fry et al. 2017), absolute risks are expected to be underestimated. With many national initiatives emerging to curate routine health records for millions of individuals, future studies will allow a better understanding of the challenges and limitations inherent to the information contained in EHR data.

In summary, this experiment presented the first systematic approach to simultaneously perform risk stratification for thousands of diseases across many clinical specialties. It is

6 Medical history predicts future health trajectories over the human phenome

based on currently unused information, which is theoretically available in real-time at no additional cost. The findings of this experiment suggest a considerable potential to disrupt medical practice by performing data-driven risk assessments to inform and guide early diagnosis, treatments, and preventive interventions.

7 Phenome-wide prediction of disease onset from retinal fundus photographs

The findings presented in this chapter include contributions by Dr. Jakob Steinfeldt and Lukas Loock. Specifically, Dr. Jakob Steinfeldt contributed with the mapping of the Electronic Health Record (EHR) data to the OMOP CDM standard, and jointly with Lukas Loock to the implementation of the retinal risk model.

7.1 Experimental setup

In this experiment, information predictive for incident disease onset in retinal fundus photographs was explored on a phenome-wide scale (see section 1.3.5 on page 28). To this end, a novel neural network-based risk model, the Retinal Risk Model (section 2.6.4 on page 70), was developed to simultaneously extract meaningful information from retinal fundus photographs across all clinical specialties, including metabolic, vascular, respiratory, musculoskeletal and neurological disorders as well as cancers (figure 7.1 on the next page). The Retinal Risk Model is a multi-task convolutional neural network based on the ConvNext architecture (Z. Liu, Mao, et al. 2022) fine-tuned on retinal fundus photographs to model the integrative retinal state over the entire human phenome for 1,171 endpoints simultaneously (figure 7.1 on the facing page, details provided in section 2.6.4 on page 70). The broad supervision, in combination with the individual objectives, allows the model to leverage the systemically relevant information in retinal fundus photographs while retaining flexibility in fitting endpoint-specific variations. The scalar risk states were derived from retinal fundus photographs recorded in 61,256 individuals in the UK Biobank population cohort (Bycroft et al. 2018; Sudlow et al. 2015). Similarly to the experiments previously described in this work, the Retinal Risk Model was developed and trained in 22-fold spatial cross-validation over the recruitment centers of the UK Biobank cohort (see section 3.2.6 on page 81). After development, training, and validation, the learned retinal states were extensively investigated for their information content by integrating them in Cox Proportional Hazards (CPH) models (Cox 1972) to model the risk for individual endpoints and investigate if the information gained through retinal fundus photographs is additive to basic demographic predictors. While the integration into survival models also relied on spatial separation, subsequent predictions of the survival models were aggregated for downstream analysis.

In summary, this experiment was conducted in four steps: first, the information extracted from the retinal fundus photograph was examined for its ability to stratify the study population by event rates, and the ten-year discriminative performances of the risk states were assessed; second, the additive information of the retinal risk states over basic demographic predictors was explored by comparing the CPH models trained on age and biological sex

7.2 Characteristics of the study population

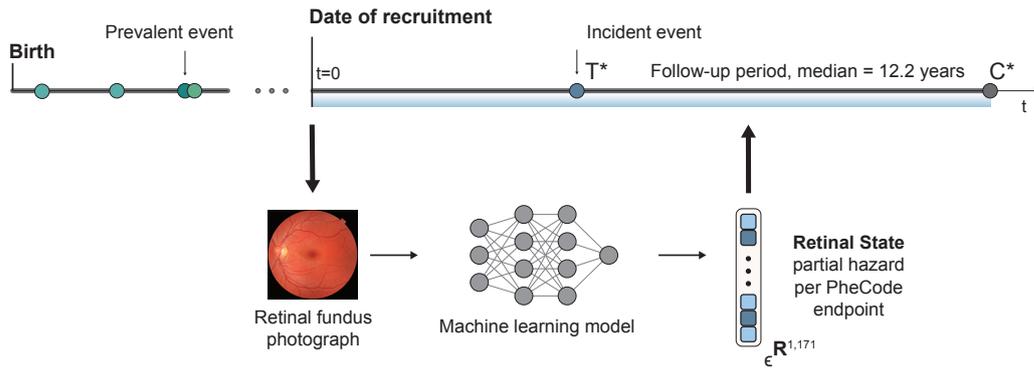


Figure 7.1: Overview on the experiment and the retinal risk model. The Retinal Risk Model is a neural network-based survival model ingesting retinal fundus photographs to simultaneously predict retinal states capturing information on incident disease onset for 1,171 diseases simultaneously.

against models extended with the retinal risk states; third, the applicability of the extracted retinal information for risk modeling in cardiovascular primary prevention was investigated by scrutinizing the retinal risk states against established cardiovascular predictors. Fourth and last, image attributions were calculated to pinpoint regions of the retina contributing most to disease risk.

7.2 Characteristics of the study population

The Retinal Risk Model was developed and trained on retinal fundus photographs contained in the UK Biobank cohort (Bycroft et al. 2018; Sudlow et al. 2015), a sample of the general population. Participants were enrolled from 2006 to 2010 in 22 recruitment centers across the United Kingdom; the follow-up is ongoing. The UK Biobank provides retinal fundus photographs recorded at recruitment for a subset of ~68,000 individuals. Data filtering, cleaning, and preprocessing (see section 3.2.5 on page 80) derived the study population of 61,256 individuals (see table 7.1. Clinical predictors and disease endpoints were extracted as described in section 3.2.1 on page 76 and section 3.3.3 on page 85.

Table 7.1: Summary statistics of the study population. Reported are medians (IQR) and absolute counts (%).

	Male, N = 27,971	Female, N = 33,285	Overall, N = 61,256
Age at Recruitment	59 (50, 64)	58 (50, 63)	58 (50, 63)
Ethnicity			
Asian	794 (2.9%)	712 (2.2%)	1,506 (2.5%)
Black	651 (2.4%)	895 (2.7%)	1,546 (2.6%)
Chinese	81 (0.3%)	162 (0.5%)	243 (0.4%)
Mixed	157 (0.6%)	300 (0.9%)	457 (0.8%)
White	25,753 (94%)	30,539 (94%)	56,292 (94%)

Table 7.1 continued from previous page

	Male, N = 27,971	Female, N = 33,285	Overall, N = 61,256
Current Smoker	3,232 (12%)	2,608 (7.8%)	5,840 (9.5%)
Type-2 Diabetes	2,024 (7.2%)	1,148 (3.4%)	3,172 (5.2%)
Weight (kg)	85 (76, 94)	69 (62, 79)	77 (66, 88)
Standing Height (cm)	176 (171, 180)	163 (159, 167)	168 (162, 176)
BMI	27.3 (25.0, 30.0)	26.0 (23.3, 29.6)	26.7 (24.1, 29.8)
Systolic Blood Pressure (mmHg)	138 (128, 150)	133 (120, 146)	136 (124, 148)
Total Cholesterol (mmol/L)	5.44 (4.70, 6.20)	5.82 (5.10, 6.58)	5.65 (4.90, 6.42)
HDL Cholesterol (mmol/L)	1.26 (1.09, 1.48)	1.59 (1.36, 1.86)	1.43 (1.20, 1.71)

The study population had a median age of 58 (IQR 50, 63) years; 54.3 % were female, 9.5 % current smokers, and 5.2 % were diagnosed with type 2 diabetes. The median BMI was 26.7 (IQR 24.1, 29.8), systolic blood pressure was 136 (IQR 124, 148) mmHg, total cholesterol was 5.65 (IQR 4.90, 6.42) mmol/l, and HDL cholesterol was 1.43 (IQR 1.20, 1.71) mmol/l. The median follow-up time was 11.4 years, with 684,968 overall patient years.

7.3 Phenome-wide association of retinal states with future disease onset

Biological features in the human retina have long been associated with cardiovascular, neurological, and metabolic diseases, and the retina has been discussed as a potential source of systemic information. While retinal fundus photographs have been leveraged for the onset prediction of individual disease endpoints, so far, the predictive potential of retinal fundus photographs has not been investigated systematically over a broader range of endpoints. To identify endpoints for which retinal fundus photographs contain predictive information, the relationship between the learned endpoint-specific risk states and the risk of future disease (figure 7.2 on page 146) was assessed for all 1,171 endpoints across the entire human phenome. Specifically, incident events were counted over the percentiles of the risk states, and subsequently observed event rates were assessed between the top and bottom 10 % of risk states over the entire phenome (figure 7.2 on page 146). Details on the observed event rates are provided in the example of all-cause mortality with a Ratio of 13.3. Importantly, event rates increased gradually, reflecting a stratification of high and low-risk individuals for the majority of endpoints covering a broad range of disease categories and etiologies (figure 7.2 on page 146): For 338 of 1,171 endpoints (28.8 %), >5-times as many events for individuals in the top 10 % of the predicted risk states compared to the bottom 10 % were observed and for 338 endpoints (28.8 %) even >5-times as many. These endpoints included, for instance, type-2 diabetes (Ratio ~7.1), chronic kidney disease (Ratio ~9.6),

7.3 Phenome-wide association of retinal states with future disease onset

coronary heart disease (Ratio ~13.2), chronic obstructive pulmonary disease (Ratio ~14.8), and Parkinson's disease (Ratio ~20.00). For 27 (2.3 %) of the 1,171 conditions, including lung cancer (Ratio ~52.7), cataracts (Ratio ~58.6), age-related macular degeneration (Ratio ~116.7), and Alzheimer's disease (Ratio ~133.0), individuals in the top 10 % encountered more than 50 times the number of incident events compared to the bottom 10 %. In contrast, for 164 (14.0 %) endpoints, including, for instance, hepatitis (Ratio ~1.40), major depressive disorders (Ratio ~1.89), and psoriasis (Ratio ~1.51), no separation between high and low-risk individuals (Ratio equal to less than 2) was observed. Notably, the ratios were greater than 3 for 585 (49.95 %) of the 1,171 endpoints, even though all models were developed in spatially segregated assessment centers. The complete list of all endpoints and corresponding statistics can be found in table 21 on page 417 and table 22 on page 438.

In addition to the phenome-wide analysis of 1,171 endpoints, cumulative event rates for up to 15 years (figure 7.2 on the following page) of follow-up for the top, median, and bottom percentiles were calculated for a subset of 24 selected endpoints. These selected endpoints cover common diseases across clinical specialties and diseases with established retinal associations. Retinal information was found to stratify the risk trajectories well for cardiovascular conditions such as myocardial infarction and heart failure, metabolic conditions such as type-2 diabetes, neurological conditions such as all-cause dementia and Parkinson's, neoplastic diseases such as skin cancer, colon cancer, and lung cancer, as well as ocular conditions such as age-related macular degeneration and cataracts.

In summary, information from retinal fundus photographs stratified the risk of onset for the majority of the 1,171 investigated endpoints across clinical specialties. These findings indicate that retinal fundus photographs contain relevant information for risk stratification in the general population across the human phenome.

7 Phenome-wide prediction of disease onset from retinal fundus photographs

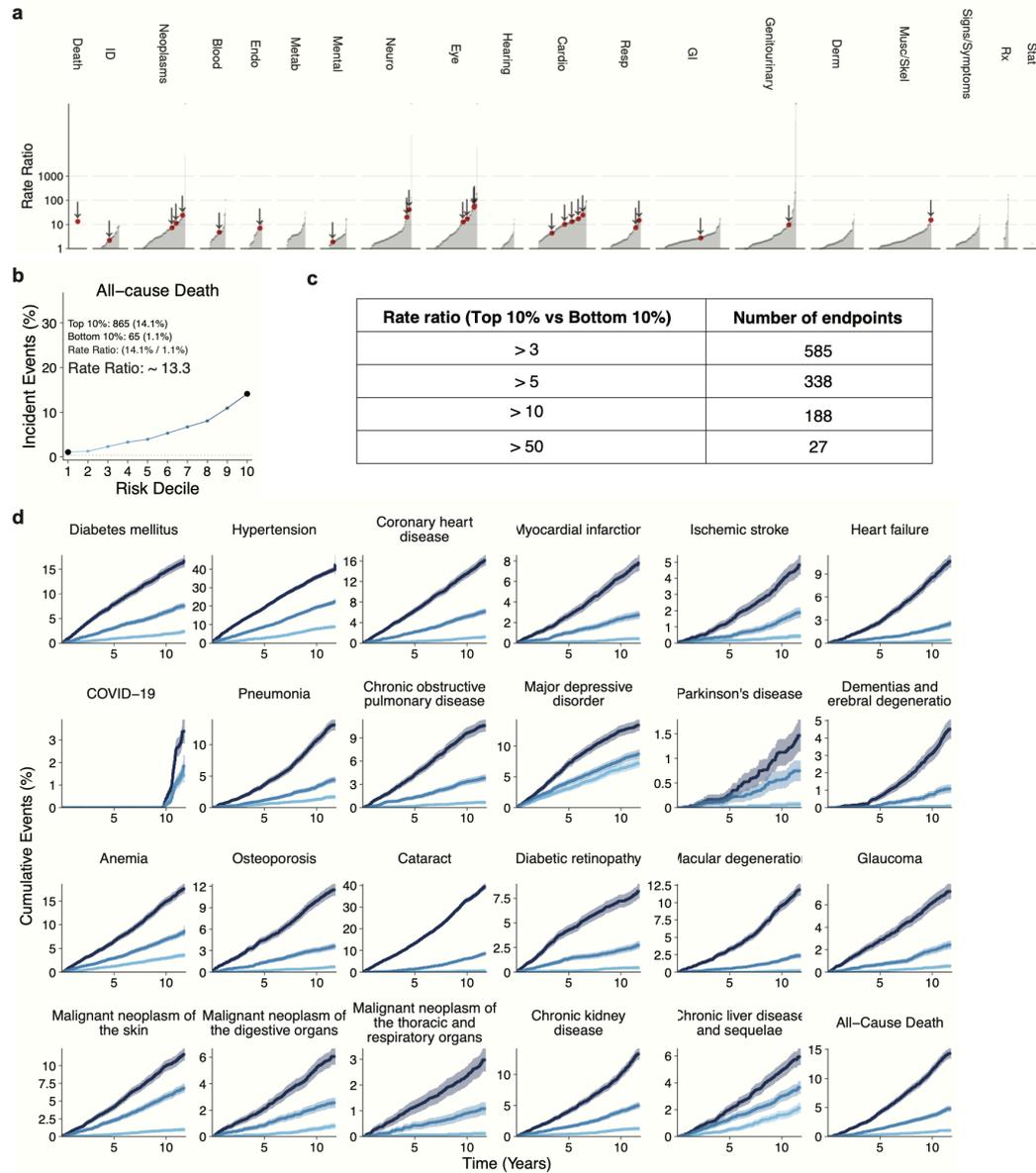


Figure 7.2: Information in retinal fundus photographs stratifies incident disease onset. a, Ratio of incident events in the Top 10 % compared with the Bottom 10 % of the estimated risk states. Event rates in the Top 10 % are greater than double the event rate in the Bottom 10 % for 1,007 of the 1,171 investigated endpoints (85.9 %). Red dots indicate the 24 selected endpoints displayed in subfigure d. b, To illustrate, 865 (14.1 %) individuals in the top risk decile for all-cause death experienced an event compared with only 65 (1.1 %) in the bottom decile, with a risk ratio of 13.3. c, Overview of the number of endpoints surpassing a given rate ratio. d, Cumulative event rates for a selection of 24 endpoints for the Top 1 %, median, and Bottom 1 % of risk percentiles over 15 years. Statistical measures were derived from 62,256 individuals. Individuals with prevalent diseases were excluded from the endpoints-specific analysis.

7.4 Retinal information yields discriminative improvements over basic predictors

While retinal fundus photographs contained stratifying information about disease onset over a broad spectrum of endpoints, clinical utility requires information to be additive to ubiquitously available predictors. Therefore, in the following experimental step, the risk of disease onset was modeled using CPH models for all 1,171 endpoints to test whether retinal fundus photographs contained predictive information beyond age and biological sex. Specifically, the performance of the baseline model (based on age and biological sex) was compared with the model augmented with the risk learned from retinal fundus photographs (based on age, biological sex, and the retinal state). Additionally, the modeling allowed the calculation of adjusted hazard ratios (denoted as HR in table 24 on page 485) and 10-year discriminative improvements (denoted as Delta C-index in figure 7.3 on the following page).

Significant improvements over the baseline model (age and biological sex only) were observed for 456 (38.94 %) of the 1,171 investigated endpoints (figure 7.3 on the next page, Supplementary Table 4). Interestingly 79.1 % of all assessed metabolic diseases, 64.5 % of all assessed eye-related diseases, and 18.7 % of all assessed cardiovascular diseases were found to gain significant discriminatory performance from the addition of retinal information (table 23 on page 460). For the highlighted subset of 24 endpoints (indicated in figure 7.3 on the next page), significant discriminative improvements were noted for the prediction of cataracts (C-Index: 0.732 (95 % CI 0.728, 0.736) → 0.784 (95 % CI 0.049, 0.055)), diabetic retinopathy (C-Index: 0.63 (95 % CI 0.621, 0.637) → 0.708 (95 % CI 0.697, 0.714)), heart failure (C-Index: 0.737 (95 % CI 0.731, 0.748) → 0.747 (95 % CI 0.739, 0.756)), myocardial infarction (C-Index: 0.707 (95 % CI 0.699, 0.715) → 0.715 (95 % CI 0.707, 0.723)), all-cause death (C-Index: 0.707 (95 % CI 0.702, 0.714) → 0.718 (95 % CI 0.712, 0.723)) and COVID-19 (C-Index: 0.504 (95 % CI 0.468, 0.542) → 0.61 (95 % CI 0.57, 0.651)). Improvements were suggested for ischaemic stroke (C-Index: 0.704 (95 % CI 0.692, 0.715) → 0.711 (95 % CI 0.7, 0.721)) and chronic kidney disease (C-Index: 0.684 (95 % CI 0.677, 0.692) → 0.688 (95 % CI 0.681, 0.693)), albeit not significant in the investigated population. In contrast, the retinal information did not improve discrimination beyond age and biological sex for endpoints such as dementia (C-Index: 0.788 (95 % CI 0.776, 0.802) → 0.783 (95 % CI 0.772, 0.796)), Parkinson's disease (C-Index: 0.724 (95 % CI 0.709, 0.74) → 0.716 (95 % CI 0.698, 0.731)) and cancers of the digestive organs (C-Index: 0.672 (95 % CI 0.663, 0.683) → 0.669 (95 % CI 0.659, 0.68)).

7 Phenome-wide prediction of disease onset from retinal fundus photographs

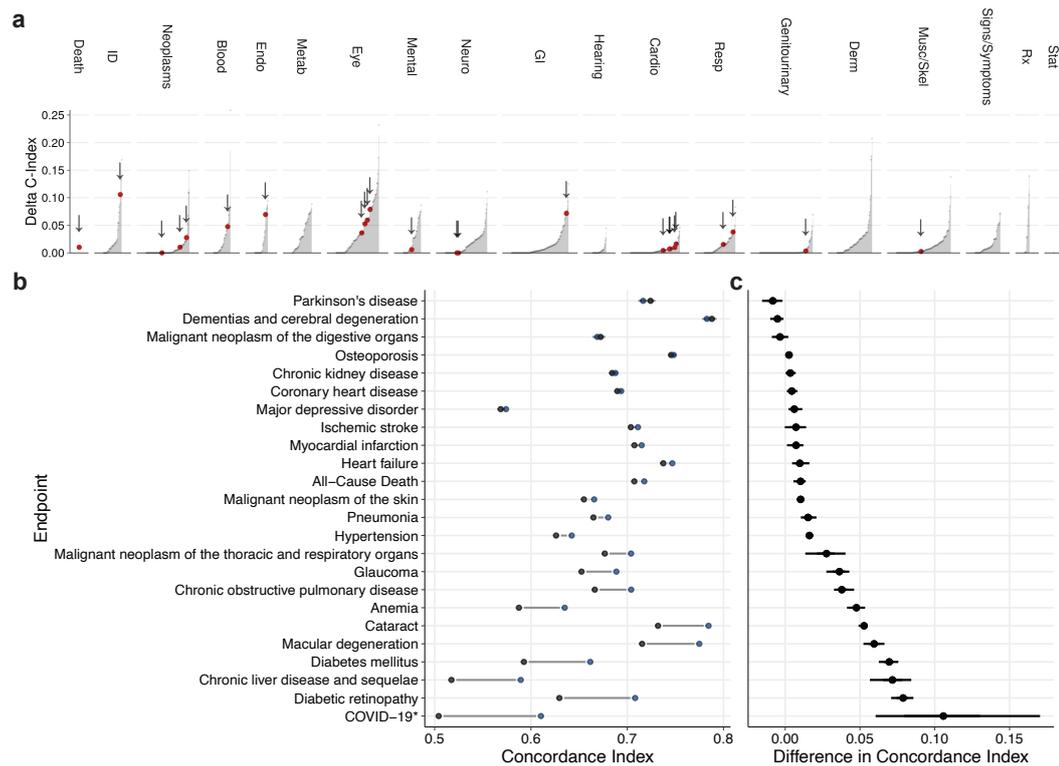


Figure 7.3: Retinal information adds significant gains in discriminatory performance. **a**, Differences in discriminatory performance quantified by the C-Index between Cox Proportional Hazards (CPH) models trained on Age+Sex and Age+Sex+Retina for all 1,171 endpoints. Significant improvements over the baseline model (Age+Sex, age, and biological sex only) are found for 456 (38.9 %) of the 1,171 investigated endpoints. Red dots indicate selected endpoints in subfigure **b**. **b**, Absolute discriminatory performance in terms of C-Index comparing the baseline (Age+Sex, black point) with the added retinal information (Age+Sex+Retina, red point) for selected endpoints. **c**, Relative discriminatory performance in terms of C-Index deltas comparing the baseline (Age+Sex) with the added retinal information (Age+Sex+Retina) for selected endpoints. Dots indicate medians and whiskers extend to the 95 % confidence interval for distributions bootstrapped with 100 iterations.

The adjusted hazard ratios (HR, per standard deviation retinal state, with 95 % CI) of the CPH models trained on the combinations of the retinal information and the predictors age and biological sex for all endpoints are presented in table 24 on page 485. A change of one standard deviation in the retinal risk state was estimated to substantially alter the risk trajectory as indicated by the adjusted HRs for diabetic retinopathy (adjusted HR: 3.22 (95 % CI: 2.87, 3.47)), Glaucoma (adjusted HR: 3.02 (95 % CI: 2.78, 3.38)) or All-Cause Death (adjusted HR: 3.25 (95 % CI: 3, 3.53)). Even for diseases where discriminatory benefits over age and biological sex were not significant, such as Dementia (HR: 3.68 (95 % CI: 3.24, 4.15); adjusted HR: 2.5 (95 % CI: 2.25, 2.66)), Parkinson's disease (HR: 4.26 (95 % CI: 3.5, 4.73); adjusted HR: 2.23 (95 % CI: 1.95, 2.61)) or cancer of the digestive organs (HR: 3.56 (95 % CI: 3.2, 3.83); adjusted HR: 2.87 (95 % CI: 2.56, 3.18)) adjusted HRs were substantial.

7.5 Extracted retinal information is similar to cardiovascular predictors

In the primary prevention of cardiovascular diseases, risk scores are well established to guide preventive lipid-lowering interventions. While the utilized predictors are accessible at a low cost, they require dedicated physical and laboratory measurements by healthcare providers. Vascular information in the human retina has been linked to the risk of the onset of many cardiovascular diseases (C.Y. Cheung, D. Xu, et al. 2020; Poplin et al. 2018), and retinal fundus photographs have been suggested as a drop-in replacement for cardiovascular risk factors (Rudnicka et al. 2022).

In order to explicitly investigate the utility of retinal information and its ability to augment or simplify cardiovascular risk scores, the risk for cardiovascular disease onset was modeled using Cox Proportional Hazards (CPH) models with predictors from well-validated primary prevention scores: the European Society of Cardiology SCORE2 (SCORE2 working group and ESC Cardiovascular risk collaboration 2021), the American Heart Association ASCVD (Goff et al. 2014) and the British QRISK3 score (Hippisley-Cox, C. Coupland, and Brindle 2017), recommended by the NHS Health Check (NHS 2022). These cardiovascular risk scores were found to consistently outperform a combination of the retinal state with age and biological sex for all investigated cardiovascular endpoints (figure 7.4 on the next page). The addition of the retinal state to the established cardiovascular predictor sets revealed information from retinal fundus photographs to not significantly improve discriminatory performance for any of the investigated cardiovascular endpoints, including myocardial infarction (+0.001 C-Index (95 % CI -0.003, 0.005) over SCORE2) and -0.001 C-Index (95 % CI -0.005, 0.002) over ASCVD), coronary heart disease (+0 C-Index (95 % CI -0.002, 0.003) over SCORE2 and -0.001 C-Index (95 % CI -0.004, 0.001) over ASCVD) and heart failure (+0.003 C-Index (95 % CI -0.001, 0.007) over SCORE2 and -0.001 C-Index (95 % CI -0.005, 0.003) over ASCVD) (see figure 7.4 on the following page and table 25 on page 509).

Therefore, these findings indicate that information extracted from retinal fundus photographs with the methodology applied in this study is not additive over current guideline-recommended cardiovascular risk models.

7 Phenome-wide prediction of disease onset from retinal fundus photographs

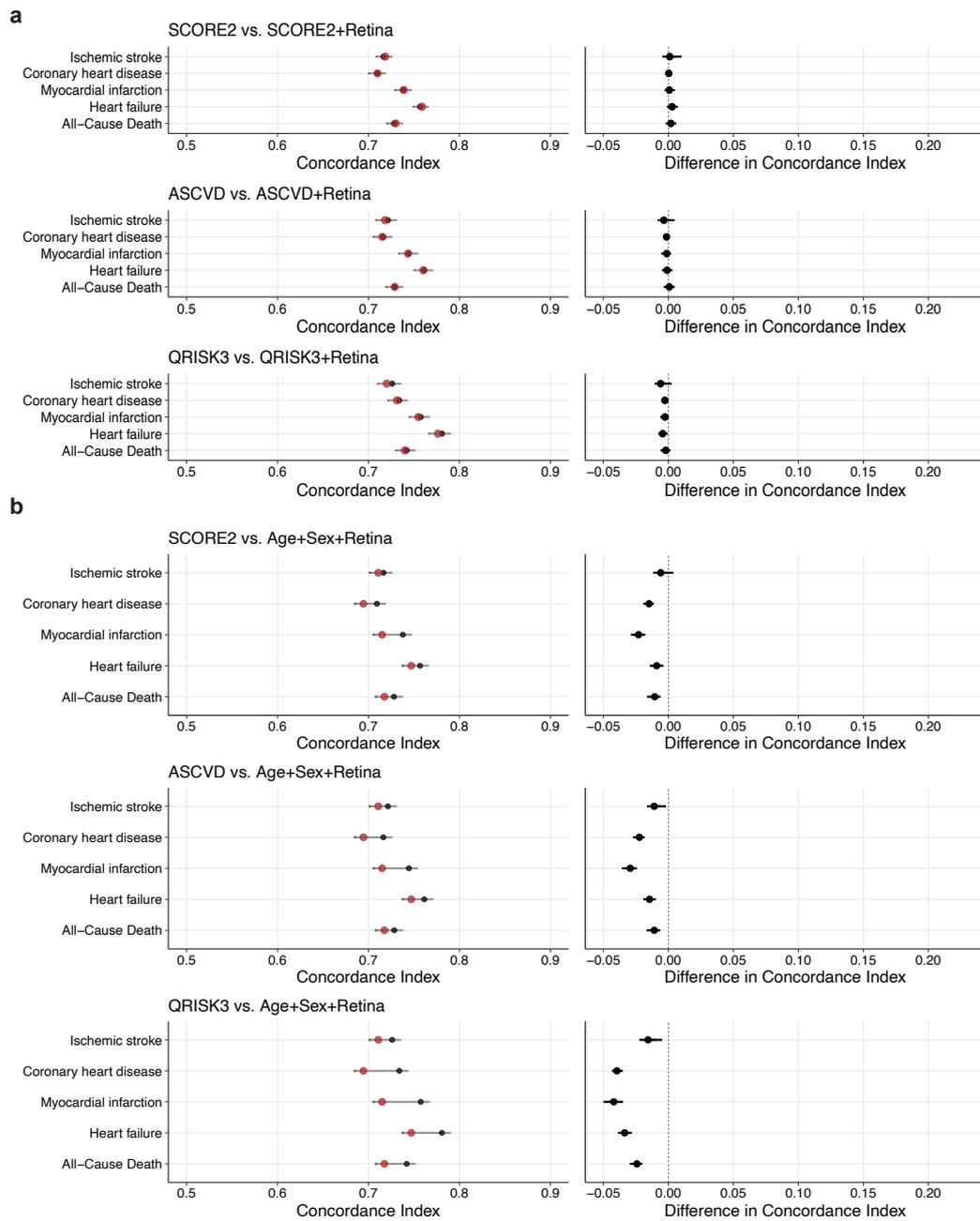


Figure 7.4: Retinal information is not additive to cardiovascular risk factors. **a**, Discriminatory performances (left) and direct differences (right) in terms of C-index between CPH models trained on sets of established cardiovascular predictors (SCORE2, ASCVD, and QRISK as black dots) and their extensions with retinal information (red dots). **b**, Discriminatory performances (left) and direct differences (right) in terms of C-index between CPH models trained on sets of established cardiovascular predictors (SCORE2, ASCVD, and QRISK as black dots) and a simplified risk model based on Age+Sex+Retina. Dots indicate medians and whiskers extend to the 95 % confidence interval for distributions bootstrapped over 100 iterations.

7.6 Image attributions reveal risk-related regions in retinal fundus photographs

While neural networks are not inherently interpretable, methods have been developed to overcome this challenge (Schulz et al. 2020). While information has been attributed to retinal fundus photographs in previous studies, the additive information has not been explicitly investigated. Therefore, in the following experimental step, Information Bottleneck Attribution (IBA) (Schulz et al. 2020) were calculated for six diseases with high relevance and incidence to identify which physiological regions of the human retina inform on future disease onset (figure 7.5 on the following page). Importantly, attributions were computed against the retinal state alone and against the additive information over age and biological sex.

Comparing the attributions against the retinal state, predictions were found to be driven by distributed regions of known importance, such as patterns in the vascularization and the regions of the fovea or the macula (figure 7.5 on the next page). Notably, the predictions for an individual were influenced by the same consistent regions for multiple endpoints, albeit with endpoint-dependent directionality. Assessing the adjusted attributions revealed the origins and directional effects of the additive information over the basic demographic predictors age and biological sex. For instance, for the individual in figure 7.5 on the following page, the retinal state was dominated by the young age and female sex of the individual, resulting in predominantly protective attributions, while the adjusted attributions revealed the origins of the information modifying the predicted health trajectory. Generally, information corresponding to age and biological sex centered around the fovea and the vascular regions. In contrast, additive information was found to be distributed over the retinal fundus photograph, with particular features highlighted.

In summary, the image attributions revealed the additive predictive information to originate from well-described retinal regions, the macula, the fovea, and the vascularization. With these regions being consistent across individuals and endpoints, our model attributions might help guide further molecular research and help ophthalmologists focus their efforts on associating retinal features with disease onset.

7 Phenome-wide prediction of disease onset from retinal fundus photographs

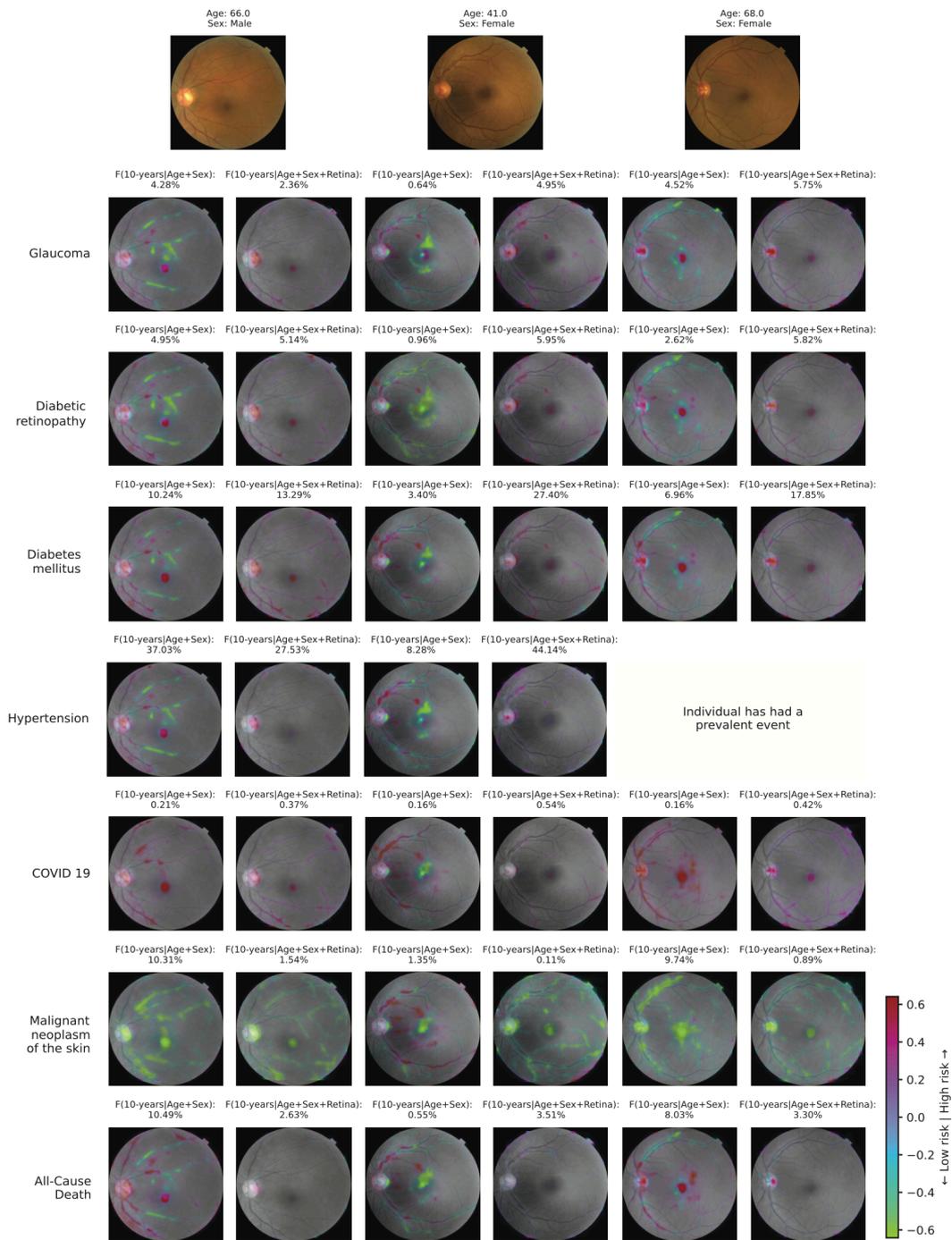


Figure 7.5: Attribution reveals risk-related regions in retinal fundus photographs. The image attributions for six selected endpoints and three individuals were derived by applying the Information Bottleneck Attribution (IBA)-method against the retinal state as well as against its adjustment with age and biological sex. The color map indicates the directionality of the impact (risk increasing, red, or decreasing, green), while the color intensity indicates the absolute scale of the attributions. The bounds of the colormap represent the minimum and maximum values observed in the adjusted attribution maps.

7.7 Discussion

The retina has long been postulated as a broad source of relevant clinical information for ophthalmological, cardiovascular, neurological, and metabolic diseases (MacGillivray et al. 2014). Retinal funduscopy is a non-invasive method to image the human retina, commonly conducted in ophthalmological examinations. While the information in retinal fundus photographs has already been leveraged for disease diagnosis (Ahn and Shah 2021; Mitani, Hammel, and Y. Liu 2021; Mitani, A. Huang, et al. 2020; Sabanayagam et al. 2020; W. Xiao et al. 2021), risk factor estimation (Poplin et al. 2018; Rim, C.J. Lee, et al. 2021; Rim, G. Lee, et al. 2020) and outcome prediction (Arcadu et al. 2019; C.Y. Cheung, D. Xu, et al. 2020; Rudnicka et al. 2022; K. Zhang et al. 2021) for individual disease endpoints, the extent, and relevance of predictive information contained in retinal fundus photographs is currently unclear. In this experiment, the predictive information for predicting individual disease onset contained in retinal fundus photographs was mapped out for the first time over the entire human phenome. With ~685,000 person-years of follow-up, information in retinal fundus photographs was robustly leveraged to derive integrative retinal states for 1,171 disease endpoints simultaneously. The systematic analysis allowed for the identification of diseases where information from retinal fundus photographs has the potential to augment individual risk assessment and impact clinical decision-making. The learned retinal states were found to contain enough information to stratify the population by incident disease risk for 625 of the 1,171 investigated diseases. Further, by investigating the predictive value beyond common demographic predictors, a subset of endpoints where retinal information provides complementary information was identified. A specific examination of a possible application of retinal information for risk modeling in cardiovascular prevention yielded no evidence for added information over established cardiovascular predictors. Finally, image attributions confirmed the importance of known retinal features and a shared physiological background of diseases with additive retinal information.

Importantly, the predictive information contained in retinal fundus photographs was found to add significant discriminative improvements over the basic demographic predictors age and biological sex for 456 endpoints across all categories of the human phenome. These endpoints include diseases with established associations to features in retinal fundus photographs: cardiovascular diseases such as coronary heart disease and myocardial infarction (Poplin et al. 2018; Rudnicka et al. 2022); metabolic diseases like type-2 diabetes (K. Zhang et al. 2021) and anemia (Mitani, A. Huang, et al. 2020), ocular diseases such as age-related macular degeneration (Peng, Dharssi, et al. 2019; Peng, Keenan, et al. 2020; Yan et al. 2020)

and diabetic retinopathy (Arcadu et al. 2019; Bora et al. 2021) as well as all-cause mortality (Nusinovici et al. 2022; Zhu et al. 2020). However, a multitude of additional diseases where retinal information significantly improved the discriminatory performance over basic demographic predictors was identified, across all disease categories of the human phenome, indicating retinal fundus photographs to capture systemic information relevant over a much wider disease spectrum. Examples include inflammatory diseases such as chronic obstructive pulmonary disease, infectious diseases such as pneumonia and COVID-19, additional cardiovascular endpoints such as hypertension and heart failure, cancers such as malignant neoplasms of the skin and respiratory organs, and mental disorders such as major depression. For these endpoints, the information contained in retinal fundus photographs may be leveraged in combination with primary demographic predictors as a low-cost tool for population stratification and identification of high-risk individuals. In contrast, for other endpoints such as chronic kidney disease (Sabanayagam et al. 2020; K. Zhang et al. 2021) or all-cause dementia (C.Y. Cheung, Ran, et al. 2022), the results of this experiment reproduced previous studies identifying predictive information, however, suggesting this information to be a proxy for basic demographic predictors.

The retina is considered a non-invasive window into the human vascular condition, and links between retinal vascularization and cardiovascular diseases have long been established (MacGillivray et al. 2014; McGeechan et al. 2008; Owen et al. 2019; T.Y. Wong, Islam, et al. 2006; T.Y. Wong, Klein, et al. 2003). In line with recent work, the results of this experiment demonstrate the importance of retinal information for predicting cardiovascular disease onset (Poplin et al. 2018; Rudnicka et al. 2022) able to stratify the population and provide significant discriminative performance gains over basic demographic predictors. However, a detailed investigation of the possibility of augmenting the primary prevention of cardiovascular diseases yielded no evidence for retinal information providing discriminatory improvements over comprehensive sets of cardiovascular risk factors. Previous studies have shown that retinal fundus photographs contain (lossy) information on cardiovascular risk factors (Gerrits et al. 2020; Poplin et al. 2018; Rim, C.J. Lee, et al. 2021), and although studies have extracted more complex risk factors, such as the retinal vessel caliber (C.Y. Cheung, D. Xu, et al. 2020; McGeechan et al. 2008) or the coronary artery calcium score (Rim, C.J. Lee, et al. 2021), this information could not be retrieved with the applied methodology, suggesting the need of explicit supervision for its extraction. Nevertheless, retinal fundus photographs have the potential to augment no-cost predictors like age and biological sex as an information proxy on cardiovascular risk factors to enable automated background examinations in ophthalmological care to identify individuals for further screening procedures at

no additional cost.

Calculating attributions for six endpoints with high relevance and incidence allowed an assessment of which and how regions of the retinal fundus photographs contributed to risk. Comparing the attribution maps between cardiovascular, metabolic, and neurological diseases, the importance of the retinal vascularization patterns, the macula, and the fovea, was confirmed. Importantly, these regions are commonly examined by ophthalmologists and long suspected to contain relevant features for neurological and cardiovascular conditions (MacGillivray et al. 2014; Poplin et al. 2018; T.Y. Wong, Islam, et al. 2006; T.Y. Wong, Klein, et al. 2003). Interestingly, individual features were generally important across multiple disease endpoints, indicating that these regions of the human retina contain systemic information on the human condition. Further, information on age and biological sex appeared to be enriched in consistently attributed regions, while added information was found to be distributed more selectively over the image.

Before, information from retinal fundus photographs may be leveraged for risk-modeling in clinical practice; however, substantial challenges remain. Even though the study population comprised over 61 thousand individuals, the population is healthier and less deprived than the general UK population (Fry et al. 2017). Further, while the model was validated in spatially separated samples from the individual assessment centers, and no signs of overfitting were observed, the developed retinal risk model has yet to be evaluated in an entirely independent cohort. External validation is of particular importance as retinal fundus photographs are known to be sensitive to batch effects as well as procedure and equipment-dependent sources of variance (Beede et al. 2020). In light of these limitations, careful scrutinization is imperative prior to applying the retinal risk model outside the research context. Ultimately, broad utilization of retinal fundus photography for preventive care requires not only a further examination of the connection between information retrievable from images and disease onset but also questions on infrastructure and logistics to be addressed.

Taken together, this study systematically assessed predictive information in retinal fundus photographs on a phenome-wide scale. By mapping out the predictive information and the potential of retinal fundus photographs for disease onset prediction, the experiment confirmed the potential of retinal fundus photographs to inform non-invasively on the human condition.

8 Conclusion

8 Conclusion

Modern medicine is centered around treatment, not around prevention and health maintenance. In light of rising healthcare costs worldwide, an adaptation of prevention as a central dogma is urgently needed. However, a fundamental requirement for personalized prevention is the ability to predict disease risks early and accurately in order to stratify populations and deliver the right interventions to the right individuals at the right time. However, a comprehensive, systematic assessment of disease risk is currently not feasible due to methodological limitations. On the one hand, there exist no risk stratification tools for a large number of diseases, and on the other hand, newly proposed risk models are hardly adapted into clinical practice due to the severe resource limitations constraining the collection of required predictors in the primary care setting. At the same time, in modern medicine, more and more data on patients is available in the clinical and research context that could be exploited for risk modeling. Especially complex data modalities, such as -omics measurements or images, bear supposedly valuable information on an individual's future health trajectory. This information could potentially disrupt medical practice and decision-making by allowing the prediction for many endpoints simultaneously from a single assay and optimizing resource requirements by rendering the collection of many individual predictors obsolete. However, this potential so far has remained unassessed due to a lack of dedicated methods allowing for the extraction of clinically relevant information from complex data modalities and the integration of this information for risk modeling.

Thus, the aim of this study was to build on recent advances in machine learning to investigate the applicability of complex data modalities for risk modeling in primary prevention. Towards this goal, four novel neural network-based survival models were developed, trained, and validated in four independent experiments involving four distinct data modalities not yet leveraged for risk modeling in primary prevention to date: polygenic risk scores for cardiovascular prevention, ¹H-NMR metabolomics as a drop-in replacement for the measurement of costly laboratory assays in the primary prevention of common diseases, electronic health records as a general source of predictive information relevant over a wide spectrum of diseases and retinal fundus photographs as a systemic information source across many endpoints over the human phenome. For each model and modality, the predictions were examined to scrutinize the extracted predictive information against commonly collected predictors and pinpoint clinical utility. Further, where appropriate, interpretability methods were applied to assess the model predictions and identify the origin of predictive information. Throughout the entire analysis, special emphasis was placed on quantifying the clinical benefit of the information added by complex data modalities in the primary care

setting.

Importantly, this study demonstrated the applicability of neural survival models on real-world clinical data for risk modeling in primary prevention and their superior performance in comparison to linear baselines. While performance comparisons have been presented before (Katzman et al. 2017; Nagpal, X.R. Li, and Dubrawski 2021), this study adds an extensive evaluation in real-world populations of sufficient size and a benchmarking against clinically adopted baselines under consideration of metrics beyond pure discrimination. While the added discriminatory performance of a fully-parametric neural survival model over a Cox Proportional Hazards (CPH) model was modest when trained on a selection of scalar, almost uncorrelated parameters (see the NeuralCVD results, figure 4.2 on page 92), the neural survival model still generated a positive Net Reclassification Improvement (NRI), thus benefiting the reclassification of individuals into risk groups. Further, the results of the Polygenic Score (PGS) integration show that neural survival models are inherently capable of modeling medically relevant nonlinear interactions between the input covariates. While the nonlinear effect was partly recoverable with an explicit interaction, the fully parametric model picked up the nonlinearity inherently and more accurately. Consequently, the utilization of neural survival models may be beneficial whenever there are nonlinear interactions in the feature space to model these effects correctly and inherently. Incremental benefits in terms of discrimination were also found when comparing a CPH model trained on the PANEL predictor set with the complex metabolomic state model architecture (see figure 8 on page 210). Here, the benefit may be caused both by the increased model complexity allowing higher expressivity, as well as by the better regularization through supervision with multiple endpoints simultaneously.

Besides these benefits, this study strongly underlines the versatility of neural survival models in their ability to extract meaningful features from complex data modalities. Even large architectures like the ConvNeXt feature extractor utilized to extract information from the retinal fundus photographs with 53 million parameters or the complex metabolomic state model architecture with residual head networks were efficiently trainable in a neural survival setup. Importantly, the versatility of neural survival models is also reflected in the output space: Historically, clinical risk modeling involved the collection of predictors to calculate a risk score in a single endpoint setting one disease at a time. While neural survival models have been shown to be capable of multi-disease risk modeling, (C. Lee et al. 2018; Nagpal, X.R. Li, and Dubrawski 2021), this study explicitly investigates multi-disease risk modeling in a clinically applicable context. For instance, the metabolomic state model proposed in this

8 Conclusion

study is able to reliably predict disease risk for 24 common disease endpoints from a single input. Extending this approach, the retinal risk model generates risk estimates for 1,171 diseases and the medical history model even for 1,833 diseases across all categories of the human phenome. While these three approaches were developed to assess the information content of their respective input data and not *per se* for risk modeling on an individual level, the architectures and the results of their evaluation provide a glimpse into a neural survival model-enabled future of systematic risk assessment.

Exploiting the superiority and increased flexibility of neural survival models, this study presented four novel neural network-based risk models, each tailored to extract predictive information from a specific data modality. The successful development, training, and validation of each model allowed a scrutinization of the learned information and quantification of the added predictive information over established clinical predictors. Importantly, for each of these data modalities, this study builds on previous work by explicitly characterizing the extent of contained predictive information against established clinical predictors with special emphasis on clinical utility. For instance, PGS had been combined with clinical covariates in previous studies (Mosley et al. 2020; L. Sun et al. 2021), with reports of limited added discriminatory performance. However, the utilization of advanced methodology in NeuralCVD and the examination of metrics beyond discrimination, such as the Net Reclassification Improvement (NRI), allowed for identifying subpopulations with the clear clinical benefit of PGS addition. In the case of the 1H-NMR metabolomics data, Electronic Health Record (EHR) data, and the retinal fundus photographs, previous studies had exploited the data modalities to estimate risk for individual endpoints (Bagheri et al. 2020; Bora et al. 2021; C.Y. Cheung, Ran, et al. 2022; Deelen et al. 2019; Julkunen et al. 2021; Langham et al. 2021; Rudnicka et al. 2022; Solares et al. 2021). This study complements these prior investigation by explicitly characterizing the predictive information and quantifying the added predictive performance over conventional clinical predictors available in primary care. A quantification of the added predictive performance is essential to decide whether there is truly more (unique) information provided by the data modality or whether the contained information is merely a proxy for information already accessible through conventional clinical covariates.

Importantly, the results demonstrate that while a data modality may contain information *per se* able to stratify a population for a given endpoint, that information must not necessarily be additive to conventionally assessed covariates. In the case of the retinal fundus photographs, extracted information stratified the population for 625 of 1,171 diseases (53.7%),

while significant additional information over Age+Sex was observed for a much smaller set of 456 endpoints (38.9 %). Medical history showed significant benefits over age and sex for 1,800 of 1,833 diseases (95.6 %), thus relevance over most parts of the human phenome, although endpoints with no benefit were observed as well. For cardiovascular endpoints, information from the medical history was additive to cardiovascular predictors from the SCORE2, ASCVD, and QRISK3 scores. In the case of the nmr) metabolomics data, the contained information was predictive for 23 out of 24 endpoints, and significant additions over Age+Sex were observed for 16 and additions over the ASCVD predictors for eight endpoints. Thus, the information extracted from 1H-NMR metabolomics data overlapped with information on age and biological sex and, to a lesser extent, with information on more complex predictors, i.e., from the ASCVD set.

On the other hand, information extracted from retinal fundus photographs did not significantly add discriminative information over cardiovascular predictors from the SCORE2, ASCVD, and QRISK3 sets, indicating a substantial overlap in predictive information. Previous studies have demonstrated that cardiovascular predictors may be estimated from retinal fundus photographs, albeit as lossy reconstructions (Poplin et al. 2018; Rim, G. Lee, et al. 2020). While other studies indicated the possibility of extracting additional information with specific supervision, an evaluation of clinical utility is pending (Rim, C.J. Lee, et al. 2021; Rudnicka et al. 2022). Depending on the overlap of the extracted information, various applications of complex data modalities can be envisioned. Therefore, prior to considering an application of a complex data modality for risk modeling in primary care, it is imperative to characterize the information contained explicitly. In the trivial case, the extracted information is additive even to complex predictor sets and provides knowledge on an individual's risk trajectory that could otherwise not be gained or only at substantially higher costs. In this case, the extracted information may be leveraged to refine and augment future health trajectories. One example is the consideration of 1H-NMR metabolomic data along the ASCVD predictor set for the prediction of type-2 diabetes, dementia, or heart failure, where significant performance increases translating into increased clinical utility were found (see figure 5.5 on page 117). Other examples are the consideration of an individual's medical history along cardiovascular predictors of the SCORE2, ASCVD, or QRISK3 sets (see figure 6.4 on page 137) or the addition of PGS to the already very comprehensive NeuralCVD predictor set (see table 4.3 on page 93). In the case of the 1H-NMR metabolomics and the PGS, the added information comes at the expense of additional measurements, calling for a precise definition of the benefitting subpopulation that should receive the assay. However, even if the information in the complex data modality is not exactly additive to complex vari-

8 Conclusion

able sets, there is still vast potential for clinical utility through optimizing workflows and reducing resource requirements. The experiments in this study have demonstrated that by utilizing the information extracted from a complex data modality in combination with easily accessible predictors such as age and biological sex, the risk for multiple diseases may be estimated quite reliably at low costs. Such a risk model would promise to reduce friction in the primary care setting by reducing the time and effort required for predictor collection while simultaneously rendering a systematic risk assessment feasible. Importantly, the appeal of such an approach increases with every predictable endpoint, even if the performance of the extracted information in combination with age and biological sex is only on par with established, specialized risk models. One example is the use of $^1\text{H-NMR}$ metabolomics in combination with Age+Sex as a drop-in replacement for multiple individual risk factor measurements requiring laboratory assays to derive a risk model able to assess 10-year disease onset reliably for 15 common diseases, performing on par with ASCVD predictors. Another example is the combination of information from medical history with Age+Sex to allow systemic risk assessment for a vast number of endpoints. Importantly, this system could run as a *guardian-angel* running in the background of care platforms. Additionally, this combination performed comparably or better than SCORE2, ASCVD, and QRISK3, indicating direct applicability with the potential to simplify cardiovascular risk modeling in primary prevention.

Nevertheless, while a significant addition of predictive information is a prerequisite of clinical utility, it is insufficient as proof. In the case of the $^1\text{H-NMR}$ metabolomics data, decision curve analysis was performed for all endpoints with significant additions over age and biological sex, leading to the identification of explicit decision thresholds at which the added discriminatory performance would translate into an overall net benefit (i.e., into more correctly identified high-risk individuals, see results for heart failure, dementia, and type-2 diabetes, figure 5.5 on page 117). However, as decision curve analysis requires manual inspection, in the case of the retinal fundus photographs and the medical history, the systematic tracing of clinical utility over decision thresholds was impeded by the vast number of endpoints with a significant benefit over age and biological sex. Consequently, future studies must investigate which endpoints the discriminatory benefits translate into clinical benefit and quantify the clinical benefit over the spectrum of possible decision thresholds.

Further, while decision curve analysis is the state-of-the-art method to holistically evaluate the clinical benefit given established interventions and decision thresholds, it is no replacement for a real-world demonstration. Thus, towards applying the developed models in primary prevention, the added value of intervening in identified high-risk individuals has

to be demonstrated explicitly in randomized controlled trials. While the administration of interventional medication has been demonstrated to be effective, for instance, in the example of statins in cardiovascular prevention (Armitage et al. 2019; Collins et al. 2016; Tonelli et al. 2011), and interventions on modifiable risk factors have been proven to delay the onset of type-2 diabetes (Pronk, Remington, and Force* 2015), the benefit of risk communication for the induction of lifestyle changes is debated (Ngandu et al. 2022; Silarova et al. 2019). This is especially critical, as the majority of the endpoints modeled by, e.g., the metabolomic state model or the medical history model, lack established medicinal interventions. Thus, trials are required to investigate the benefits of lifestyle advice or further diagnostic screening, especially for diseases without available medicative interventions. Another strong call for real-world evidence emerges from the very homogeneous study population, composed predominantly of middle-aged caucasians. While external validation has been performed for the metabolomic state model, the validation cohorts were limited to the European continent. Therefore, additional investigations are needed to assess the transferability of the findings to genetically distant populations or different non-European healthcare systems.

In summary, this study has demonstrated the capability of neural survival models as a versatile tool to unlock the potential of complex data modalities for primary prevention. Exemplified in four experiments for four independent data modalities, this study has showcased the application of neural survival models for the extraction of risk-relevant information from complex data modalities and its scrutinization against established risk models and predictors. Many more interesting and information-rich data sources worth examining exist, and the experimentative framework solidified in this study may serve as a blueprint for further investigations. While the models developed in this study will likely not enter clinical practice, their demonstrated capabilities and proposed benefits make a strong case to lay focus on exploiting underused data modalities and multi-disease risk modeling to improve discrimination, reduce resource requirements and increase clinical benefit. Collectively, this study illustrates the potential of machine learning models to disrupt medical practice and enter an age of data-driven prevention-centered medicine on a population scale.

Acronyms

1H-NMR	Proton Nuclear Magnetic Resonance
AAA	Abdominal Aortic Aneurysm
AMD	Age-related Macula Degeneration
ANN	Artificial Neural Network
ASCVD	Atherosclerotic Cardiovascular Disease
BCRAT	Breast Cancer Risk Assessment Tool
BMI	Body Mass Index
CAP	Coronary Artery Disease
CHD	Coronary Heart Disease
CI	Confidence Interval
CKD	Chronic Kidney Disease
CNN	Convolutional Neural Network
COPD	Chronic Obstructive Pulmonary Disease
CPH	Cox Proportional Hazards
CVD	Cardiovascular Disease
DAG	Directed Acyclic Graph
DHA	DocosaHexaenoic Acid
DSM	Deep Survival Machines
ECG	Electronic Cardiogram
EHR	Electronic Health Record
FP	Family Physician
GELU	Gaussian Error Linear Unit
GlycA	Glycoprotein Acetylation
GP	General Practitioner
GPU	Graphics Processing Unit
GWAS	Genome Wide Association Study
HDL	High-Density Lipoprotein
HR	Hazard Ratio

Acronyms

IBA	Information Bottleneck Attribution
ICD	International Classification of Diseases
IFG	Impaired Fasting Glucose
IQR	Inter Quartile Range
LA	Linoleic Acid
LDL	Low-Density Lipoprotein
LDL-C	Low-Density Lipoprotein Cholesterol
Leaky ReLU	Leaky Rectified Linear Unit
LR	Logistic Regression
MACE	Major Adverse Cardiac Events
MET	Metabolomic State
MI	Myocardial Infarction
MLP	Multi Layer Perceptron
MS	Mass-Spectrometry
MUFA	Monounsaturated Fatty Acid
NaN	Not a Number
NGS	Next-Generation Sequencing
NHS	National Health Service
NN	Neural Network
NRI	Net Reclassification Improvement
PAD	Peripheral Artery Disease
PCA	Principal Component Analysis
PCE	Pooled Cohort Equations
PCP	Primary Care Provider
PGS	Polygenic Score
PPV	Positive Predictive Value
PRS	Polygenic Risk Score
RCT	Randomized Controlled Trial
ReLU	Rectified Linear Unit
ResNet	Residual Neural Network
RRD	Relative Risk Difference
SCORE	Systematic COronary Risk Evaluation
SELU	Scaled Exponential Linear Unit
SFA	Saturated Fatty Acid
SGD	Stochastic Gradient Descent
SHAP	SHapley Additive exPlanations

SiLU	Sigmoid Linear Unit
SNP	Single Nucleotide Polymorphism
TG	TriGlyceride
TIA	Transient Ischaemic Attack
TNR	True Negative Rate/Specificity
TPR	True Positive Rate/Recall or Sensitivity
TPU	Tensor Processing Unit
UMAP	Uniform Manifold Approximation and Projection

Bibliography

- Abbasi, A. et al. (Sept. 2012). Prediction models for risk of developing type 2 diabetes: systematic literature search and independent external validation study. en. *BMJ* 345, e5900.
- Abraham, G. et al. (Dec. 2019). Genomic risk score offers predictive performance comparable to clinical risk factors for ischaemic stroke. en. *Nat. Commun.* 10.1, 5819.
- Adelson, K. et al. (2017). Standardized criteria for palliative care consultation on a solid tumor oncology service reduces downstream health care use. *Journal of Oncology Practice* 13.5, e431–e440.
- Aderemi, A.V. et al. (June 2021). Metabolomics: A Scoping Review of Its Role as a Tool for Disease Biomarker Discovery in Selected Non-Communicable Diseases. en. *Metabolites* 11.7.
- Ahadi, S. et al. (Jan. 2020). Personal aging markers and ageotypes revealed by deep longitudinal profiling. en. *Nat. Med.* 26.1, 83–90.
- Ahmed, I. et al. (Jan. 2014). Developing and validating risk prediction models in an individual participant data meta-analysis. en. *BMC Med. Res. Methodol.* 14, 3.
- Ahn, J.C. and Shah, V.H. (Feb. 2021). Deep learning-based detection of hepatobiliary disorders in ophthalmic imaging. en. *The Lancet. Digital health* 3.2, e68–e69. issn: 2589-7500. doi: 10.1016/S2589-7500(20)30319-8. url: [http://dx.doi.org/10.1016/S2589-7500\(20\)30319-8](http://dx.doi.org/10.1016/S2589-7500(20)30319-8).
- Ahola-Olli, A.V. et al. (Dec. 2019). Circulating metabolites and the risk of type 2 diabetes: a prospective study of 11,896 young adults from four Finnish cohorts. en. *Diabetologia* 62.12, 2298–2309.
- Ala-Korpela, M. et al. (Aug. 2021). Apt interpretation of comprehensive lipoprotein data in large-scale epidemiology: disclosure of fundamental structural and metabolic relationships. en. *Int. J. Epidemiol.*
- Alba, A.C. et al. (Oct. 2017). Discrimination and Calibration of Clinical Prediction Models: Users' Guides to the Medical Literature. en. *JAMA* 318.14, 1377–1384.
- Alderwick, H. and Dixon, J. (Jan. 2019). The NHS long term plan. en. *BMJ* 364, 184.
- Alpert, A. et al. (Mar. 2019). A clinically meaningful metric of immune age derived from high-dimensional longitudinal monitoring. en. *Nat. Med.* 25.3, 487–495.

Bibliography

- Appelbaum, L. et al. (Jan. 2021). Development and validation of a pancreatic cancer risk model for the general population using electronic health records: An observational study. en. *Eur. J. Cancer* 143, 19–30.
- Arcadu, F. et al. (2019). Deep learning algorithm predicts diabetic retinopathy progression in individual patients. *npj Digital Medicine* 2.1, 92.
- Armitage, J. et al. (Feb. 2019). Efficacy and safety of statin therapy in older people: a meta-analysis of individual participant data from 28 randomised controlled trials. *Lancet* 393.10170, 407–415.
- Attia, Z.I. et al. (Sept. 2019). An artificial intelligence-enabled ECG algorithm for the identification of patients with atrial fibrillation during sinus rhythm: a retrospective analysis of outcome prediction. *Lancet* 394.10201, 861–867.
- Austin, P.C., Harrell Jr, F.E., and Klaveren, D. van (2020). Graphical calibration curves and the integrated calibration index (ICI) for survival models. *Statistics in Medicine* 39.21, 2714–2742. doi: <https://doi.org/10.1002/sim.8570>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/sim.8570>. url: <https://onlinelibrary.wiley.com/doi/abs/10.1002/sim.8570>.
- Ayala Solares, J.R. et al. (Jan. 2020). Deep learning for electronic health records: A comparative review of multiple deep neural architectures. en. *J. Biomed. Inform.* 101, 103337.
- Ba, J.L., Kiros, J.R., and Hinton, G.E. (July 2016). Layer Normalization. arXiv: 1607.06450 [stat.ML].
- Babenko, B. et al. (Mar. 2022). Detection of signs of disease in external photographs of the eyes via deep learning. en. *Nat Biomed Eng.*
- Bagheri, A. et al. (Aug. 2020). Multimodal Learning for Cardiovascular Risk Prediction using EHR Data. arXiv: 2008.11979 [cs.LG].
- Baker, M.L. et al. (Apr. 2008). Retinal signs and stroke: revisiting the link between the eye and brain. en. *Stroke* 39.4, 1371–1379.
- Beaulieu-Jones, B.K., Greene, C.S., and Pooled Resource Open-Access ALS Clinical Trials Consortium (Dec. 2016). Semi-supervised learning of the electronic health record for phenotype stratification. *J. Biomed. Inform.* 64, 168–178.
- Beede, E. et al. (Apr. 2020). A Human-Centered Evaluation of a Deep Learning System Deployed in Clinics for the Detection of Diabetic Retinopathy. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, CHI '20*. (Honolulu, HI, USA: Association for Computing Machinery), pp. 1–12. isbn: 9781450367080. doi: 10.1145/3313831.3376718. url: <https://doi.org/10.1145/3313831.3376718>.
- Bello, G.A. et al. (2019). Deep-learning cardiac motion analysis for human survival prediction. *Nature Machine Intelligence* 1.2, 95–104.

- Bello, I. et al. (Mar. 2021). Revisiting ResNets: Improved Training and Scaling Strategies. arXiv: 2103.07579 [cs.CV].
- Ben Miled, Z. et al. (Jan. 2020). Predicting dementia with routine care EMR data. en. *Artif. Intell. Med.* 102, 101771.
- Bengio, Y., Simard, P., and Frasconi, P. (Mar. 1994). Learning long-term dependencies with gradient descent is difficult. *IEEE Trans. Neural Netw.* 5.2, 157–166.
- Bennett, C.C. (Aug. 2012). Utilizing RxNorm to support practical computing applications: capturing medication history in live electronic health records. en. *J. Biomed. Inform.* 45.4, 634–641.
- Berry, D.A. et al. (June 2002). BRCAPRO validation, sensitivity of genetic testing of BRCA1/BRCA2, and prevalence of other breast cancer susceptibility genes. en. *J. Clin. Oncol.* 20.11, 2701–2712.
- Bolnick, H.J. et al. (Oct. 2020). Health-care spending attributable to modifiable risk factors in the USA: an economic attribution analysis. en. *Lancet Public Health* 5.10, e525–e535.
- Bora, A. et al. (Jan. 2021). Predicting the risk of developing diabetic retinopathy using deep learning. en. *Lancet Digit Health* 3.1, e10–e19.
- Boyle, E.A., Li, Y.I., and Pritchard, J.K. (June 2017). An Expanded View of Complex Traits: From Polygenic to Omnigenic. en. *Cell* 169.7, 1177–1186.
- Bratt, O. et al. (Oct. 2016). Family History and Probability of Prostate Cancer, Differentiated by Risk Category: A Nationwide Population-Based Study. en. *J. Natl. Cancer Inst.* 108.10.
- Bray, F. et al. (Nov. 2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. en. *CA Cancer J. Clin.* 68.6, 394–424.
- Breslow, N. (Mar. 1974). Covariance analysis of censored survival data. en. *Biometrics* 30.1, 89–99.
- Brown, T. et al. (2020). Language models are few-shot learners. *Advances in neural information processing systems* 33, 1877–1901.
- Buergel, T., Steinfeldt, J., Looock, L., et al. (2022). Neural network-based integration of polygenic and clinical information: development and validation of a prediction model for 10-year risk of major adverse cardiac events in the UK Biobank cohort. *The Lancet Digital Health* 4.2, e84–e94.
- Buergel, T., Steinfeldt, J., Ruyoga, G., et al. (2022). Metabolomic profiles predict individual multidisease outcomes. *Nature Medicine*, 1–12.
- Bundesgesundheitsministerium (2022). *Gesundheits-Check-up*. de. www.bmg.de/checkup. Accessed: 2022-10-11.

Bibliography

- Bycroft, C. et al. (Oct. 2018). The UK Biobank resource with deep phenotyping and genomic data. en. *Nature* 562.7726, 203–209.
- Caldwell, M. et al. (May 2019). Prospects for the Primary Prevention of Myocardial Infarction and Stroke. en. *J. Cardiovasc. Pharmacol. Ther.* 24.3, 207–214.
- Canny, J. (June 1986). A computational approach to edge detection. en. *IEEE Trans. Pattern Anal. Mach. Intell.* 8.6, 679–698.
- Cavestro, G.M. et al. (June 2018). Early onset sporadic colorectal cancer: Worrisome trends and oncogenic features. en. *Dig. Liver Dis.* 50.6, 521–532.
- Chakravarti, N. (1989). Isotonic median regression: a linear programming approach. *Mathematics of operations research* 14.2, 303–308.
- Chan, A.T. and Giovannucci, E.L. (June 2010). Primary prevention of colorectal cancer. en. *Gastroenterology* 138.6, 2029–2043.e10.
- Chan, J.M., Gann, P.H., and Giovannucci, E.L. (Nov. 2005). Role of diet in prostate cancer development and progression. en. *J. Clin. Oncol.* 23.32, 8152–8160.
- Chapfuwa, P. et al. (Apr. 2018). Adversarial Time-to-Event Modeling. 35th International Conference on Machine Learning, ICML 2018 2, 1143–1156.
- Chatterjee, S., Khunti, K., and Davies, M.J. (June 2017). Type 2 diabetes. en. *Lancet* 389.10085, 2239–2251.
- Cheung, C.Y., Ran, A.R., et al. (Nov. 2022). A deep learning model for detection of Alzheimer’s disease based on retinal photographs: a retrospective, multicentre case-control study. en. *The Lancet. Digital health* 4.11, e806–e815. issn: 2589-7500. doi: 10.1016/S2589-7500(22)00169-8. url: [http://dx.doi.org/10.1016/S2589-7500\(22\)00169-8](http://dx.doi.org/10.1016/S2589-7500(22)00169-8).
- Cheung, C.Y., Xu, D., et al. (Oct. 2020). A deep-learning system for the assessment of cardiovascular disease risk via the measurement of retinal-vessel calibre. *Nature Biomedical Engineering*, 1–11. issn: 2157-846X. doi: 10.1038/s41551-020-00626-4. url: <https://doi.org/10.1038/s41551-020-00626-4>.
- Cheung, C.Y.-L. et al. (May 2011). Retinal vascular tortuosity, blood pressure, and cardiovascular risk factors. en. *Ophthalmology* 118.5, 812–818.
- Cheung, N., Mitchell, P., and Wong, T.Y. (July 2010). Diabetic retinopathy. en. *Lancet* 376.9735, 124–136.
- Choi, E. et al. (Oct. 2018). MiME: Multilevel Medical Embedding of Electronic Health Records for Predictive Healthcare. arXiv: 1810.09593 [cs.LG].
- Choi, S.W., Mak, T.S.-H., and O’Reilly, P.F. (Sept. 2020). Tutorial: a guide to performing polygenic risk score analyses. en. *Nat. Protoc.* 15.9, 2759–2772.

- Cintolo-Gonzalez, J.A. et al. (July 2017). Breast cancer risk models: a comprehensive overview of existing models, validation, and clinical applications. en. *Breast Cancer Res. Treat.* *164.2*, 263–284.
- Clalit Institute (2022). *Data*. en. <http://clalitresearch.org/about-us/our-data/>. Accessed: 2022-8-1.
- Collins, R. et al. (Nov. 2016). Interpretation of the evidence for the efficacy and safety of statin therapy. *Lancet* *388.10059*, 2532–2561.
- Conroy, R. (June 2003). Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project. *Eur. Heart J.* *24.11*, 987–1003.
- Cook, N.R. and Ridker, P.M. (June 2009). Advances in measuring the effect of individual predictors of cardiovascular risk: the role of reclassification measures. en. *Ann. Intern. Med.* *150.11*, 795–802.
- Cornelis, M.C. et al. (Apr. 2009). Joint effects of common genetic variants on the risk for type 2 diabetes in U.S. men and women of European ancestry. en. *Ann. Intern. Med.* *150.8*, 541–550.
- Costantino, J.P. et al. (Sept. 1999). Validation studies for models projecting the risk of invasive and total breast cancer incidence. en. *J. Natl. Cancer Inst.* *91.18*, 1541–1548.
- Cox, D.R. (Jan. 1972). Regression Models and Life-Tables. *J. R. Stat. Soc. Series B Stat. Methodol.* *34.2*, 187–202.
- D’Agostino, R.B. and Nam, B.-H. (Jan. 2003). “Evaluation of the Performance of Survival Analysis Models: Discrimination and Calibration Measures”. *Handbook of Statistics*, vol. 23. (Elsevier), pp. 1–25.
- Damen, J.A.A.G. et al. (May 2016). Prediction models for cardiovascular disease risk in the general population: systematic review. en. *BMJ* *353*, i2416.
- Davidson-Pilon, C. et al. (Jan. 2021). *CamDavidsonPilon/lifelines: v0.25.8*.
- De Leeuw, J. (1977). Correctness of Kruskal’s algorithms for monotone regression with ties. *Psychometrika* *42.1*, 141–144.
- Deelen, J. et al. (Aug. 2019). A metabolic profile of all-cause mortality risk identified in an observational study of 44,168 individuals. en. *Nat. Commun.* *10.1*, 3346.
- Delles, C. et al. (2018). Nuclear magnetic resonance-based metabolomics identifies phenylalanine as a novel predictor of incident heart failure hospitalisation: results from PROSPER and FINRISK 1997. *European journal of heart failure* *20.4*, 663–673.
- Dencker et al. (2022). Deep Learning for Long-Term Post-Discharge Mortality Prediction on ICU Patient Data. cs.cmu.edu.
- Deng, J. et al. (June 2009). ImageNet: A large-scale hierarchical image database. In 2009 IEEE Conference on Computer Vision and Pattern Recognition, (), pp. 248–255.

Bibliography

- Devlin, J. et al. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805.
- Diaz-Pinto, A. et al. (Jan. 2022). Predicting myocardial infarction through retinal scans and minimal personal information. en. *Nature Machine Intelligence* 4.1, 55–61.
- Dieleman, J.L. et al. (Nov. 2017). Factors Associated With Increases in US Health Care Spending, 1996-2013. en. *JAMA* 318.17, 1668–1678.
- Dk, S. (June 2016). *Danish eHealth Portal*. en. <https://www.sundhed.dk/borger/service/om-sundheddk/om-organisationen/ehealth-in-denmark/background/>. Accessed: 2022-8-1.
- Donnelly, K. (2006). SNOMED-CT: The advanced terminology and coding system for eHealth. en. *Stud. Health Technol. Inform.* 121, 279–290.
- Efron, B. (1977). The Efficiency of Cox’s Likelihood Function for Censored Data. *J. Am. Stat. Assoc.* 72.359, 557–565.
- Elfving, S., Uchibe, E., and Doya, K. (Feb. 2017). Sigmoid-Weighted Linear Units for Neural Network Function Approximation in Reinforcement Learning. arXiv: 1702.03118 [cs.LG].
- (Nov. 2018). Sigmoid-weighted linear units for neural network function approximation in reinforcement learning. *Neural Netw.* 107, 3–11.
- Elliott, J. et al. (Feb. 2020). Predictive Accuracy of a Polygenic Risk Score-Enhanced Prediction Model vs a Clinical Risk Score for Coronary Artery Disease. en. *JAMA* 323.7, 636–645.
- Erdmann, M. et al. (2021). *Deep Learning for Physics Research (WORLD SCIENTIFIC)*.
- Esteva, A. et al. (Feb. 2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 542.7639, 115–118.
- Etezadi-Amoli, J. and Ciampi, A. (1987). Extended Hazard Regression for Censored Survival Data with Covariates: A Spline Approximation for the Baseline Hazard Function. *Biometrics* 43.1, 181–192.
- Faraggi, D. and Simon, R. (Jan. 1995). A neural network model for survival data. *Stat. Med.* 14.1, 73–82.
- Ferdian, E. et al. (Feb. 2020). Fully Automated Myocardial Strain Estimation from Cardiovascular MRI-tagged Images Using a Deep Learning Framework in the UK Biobank. en. *Radiol Cardiothorac Imaging* 2.1, e190032.
- Fischer, K. et al. (Feb. 2014). Biomarker profiling by nuclear magnetic resonance spectroscopy for the prediction of all-cause mortality: an observational study of 17,345 persons. en. *PLoS Med.* 11.2, e1001606.

- Fizelova, M. et al. (May 2015). Associations of multiple lipoprotein and apolipoprotein measures with worsening of glycemia and incident type 2 diabetes in 6607 non-diabetic Finnish men. en. *Atherosclerosis* 240.1, 272–277.
- Fletcher, G.S. (Feb. 2022). “Evidence-based approach to prevention”. UpToDate, J.G. Elmore and J. Givens, ed. (Waltham, MA: UpToDate).
- Flores, A.M. et al. (Dec. 2021). Unsupervised Learning for Automated Detection of Coronary Artery Disease Subgroups. en. *J. Am. Heart Assoc.* 10.23, e021976.
- Fonarow, G.C. et al. (2011). Potential impact of optimal implementation of evidence-based heart failure therapies on mortality. *American heart journal* 161.6, 1024–1030.
- Frampton, M. and Houlston, R.S. (Mar. 2017). Modeling the prevention of colorectal cancer from the combined impact of host and behavioral risk factors. en. *Genet. Med.* 19.3, 314–321.
- Freedman, A.N. et al. (Feb. 2009). Colorectal cancer risk prediction tool for white men and women without known susceptibility. en. *J. Clin. Oncol.* 27.5, 686–693.
- Fry, A. et al. (Nov. 2017). Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants With Those of the General Population. en. *American journal of epidemiology* 186.9, 1026–1034. issn: 0002-9262, 1476-6256. doi: 10.1093/aje/kwx246. url: <http://dx.doi.org/10.1093/aje/kwx246>.
- Fu, Y. et al. (Aug. 2020). Pan-cancer computational histopathology reveals mutations, tumor composition and prognosis. en. *Nat Cancer* 1.8, 800–810.
- Furlow, B. (2020). Information overload and unsustainable workloads in the era of electronic health records. *The Lancet Respiratory Medicine* 8.3, 243–244.
- Gail, M.H. et al. (Dec. 1989). Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. en. *J. Natl. Cancer Inst.* 81.24, 1879–1886.
- Gal, Y. and Ghahramani, Z. (June 2015). Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning. 33rd International Conference on Machine Learning, ICML 2016 3, 1651–1660.
- Galea, S. and Maani, N. (Oct. 2020). The cost of preventable disease in the USA. en. *Lancet Public Health* 5.10, e513–e514.
- Ganna, A. et al. (Aug. 2019). Large-scale GWAS reveals insights into the genetic architecture of same-sex sexual behavior. en. *Science* 365.6456.
- Gensheimer, M.F. and Narasimhan, B. (2018). *A Scalable Discrete-Time Survival Model for Neural Networks*. Tech. rep.
- Gerrits, N. et al. (June 2020). Age and sex affect deep learning prediction of cardiometabolic risk factors from retinal images. en. *Scientific reports* 10.1, 9432. issn: 2045-2322. doi:

Bibliography

- 10.1038/s41598-020-65794-4. url: <http://dx.doi.org/10.1038/s41598-020-65794-4>.
- Gerrits, N. et al. (Jan. 2021). Publisher Correction: Age and sex affect deep learning prediction of cardiometabolic risk factors from retinal images. en. *Sci. Rep.* *11.1*, 1198.
- Gilpin, L.H. et al. (2018). Explaining Explanations: An Overview of Interpretability of Machine Learning. 2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA), 80–89.
- Glorot, X., Bordes, A., and Bengio, Y. (2011). Deep Sparse Rectifier Neural Networks. In *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, G. Gordon, D. Dunson, and M. Dudík, ed. Vol. 15. *Proceedings of Machine Learning Research*. (Fort Lauderdale, FL, USA: PMLR), pp. 315–323.
- Goff, D.C. et al. (July 2014). 2013 ACC/AHA guideline on the assessment of cardiovascular risk: A report of the American college of cardiology/American heart association task force on practice guidelines. *J. Am. Coll. Cardiol.* *63.25 PART B*, 2935–2959.
- Goldstein, B.A. et al. (Jan. 2017). Opportunities and challenges in developing risk prediction models with electronic health records data: A systematic review. *J. Am. Med. Inform. Assoc.* *24.1*, 198–208.
- Gottesman, R.F. et al. (Oct. 2017). Associations Between Midlife Vascular Risk Factors and 25-Year Incident Dementia in the Atherosclerosis Risk in Communities (ARIC) Cohort. en. *JAMA Neurol.* *74.10*, 1246–1254.
- Groenendyk, J.W., Greenland, P., and Khan, S.S. (Oct. 2022). Incremental Value of Polygenic Risk Scores in Primary Prevention of Coronary Heart Disease: A Review. en. *JAMA Intern. Med.* *182.10*, 1082–1088.
- Guo, S. and Yang, Z. (Jan. 2018). Multi-Channel-ResNet: An integration framework towards skin lesion analysis. *Informatics in Medicine Unlocked* *12*, 67–74.
- Haffner, S.M. et al. (1998). Mortality from coronary heart disease in subjects with type 2 diabetes and in nondiabetic subjects with and without prior myocardial infarction. *New England journal of medicine* *339.4*, 229–234.
- Hajek Catherine et al. (Oct. 2018). Coronary Heart Disease Genetic Risk Score Predicts Cardiovascular Disease Risk in Men, Not Women. *Circulation: Genomic and Precision Medicine* *11.10*, e002324.
- Hampton, J.R. et al. (May 1975). Relative contributions of history-taking, physical examination, and laboratory investigation to diagnosis and management of medical outpatients. en. *Br. Med. J.* *2.5969*, 486–489.
- Harrell Jr, F.E. et al. (May 1982). Evaluating the yield of medical tests. en. *JAMA* *247.18*, 2543–2546.

- He, K. et al. (June 2016). Deep Residual Learning for Image Recognition. In 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), (), pp. 770–778.
- Health, E.M. of (Oct. 2021). *e-Health Record*. en. <https://e-estonia.com/solutions/healthcare/e-health-records/>. Accessed: 2022-8-1. url: <https://e-estonia.com/solutions/healthcare/e-health-records/>.
- Hendrycks, D. and Gimpel, K. (June 2016). Gaussian Error Linear Units (GELUs). arXiv: 1606.08415 [cs.LG].
- Herold, K.C. et al. (Aug. 2019). An Anti-CD3 Antibody, Teplizumab, in Relatives at Risk for Type 1 Diabetes. *N. Engl. J. Med.* 381.7, 603–613.
- Hess, K.R. (Aug. 1995). Graphical methods for assessing violations of the proportional hazards assumption in Cox regression. en. *Stat. Med.* 14.15, 1707–1723.
- Hill, N.R. et al. (Nov. 2019). Predicting atrial fibrillation in primary care using machine learning. en. *PLoS One* 14.11, e0224582.
- Hippisley-Cox, J. and Coupland, C. (June 2010). Predicting the risk of chronic Kidney Disease in men and women in England and Wales: prospective derivation and external validation of the QKidney Scores. en. *BMC Fam. Pract.* 11, 49.
- (Aug. 2013). Predicting risk of emergency admission to hospital using primary care data: derivation and validation of QAdmissions score. en. *BMJ Open* 3.8, e003482.
 - (Mar. 2015). Development and validation of risk prediction algorithms to estimate future risk of common cancers in men and women: prospective cohort study. en. *BMJ Open* 5.3, e007825.
 - (May 2021). Predicting the risk of prostate cancer in asymptomatic men: a cohort study to develop and validate a novel algorithm. en. *Br. J. Gen. Pract.* 71.706, e364–e371.
- Hippisley-Cox, J., Coupland, C., and Brindle, P. (May 2017). Development and validation of QRISK3 risk prediction algorithms to estimate future risk of cardiovascular disease: Prospective cohort study. *BMJ* 357.
- Hippisley-Cox, J., Coupland, C.A., et al. (Sept. 2021). Risk prediction of covid-19 related death and hospital admission in adults after covid-19 vaccination: national prospective cohort study. en. *BMJ* 374, n2244.
- Hochreiter, S. and Schmidhuber, J. (Nov. 1997). Long Short-Term Memory. *Neural Comput.* 9.8, 1735–1780.
- Hoerl, A.E. and Kennard, R.W. (Feb. 1970). Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Technometrics* 12.1, 55–67.
- Holmes, M.V. et al. (Feb. 2018). Lipids, Lipoproteins, and Metabolites and Risk of Myocardial Infarction and Stroke. en. *J. Am. Coll. Cardiol.* 71.6, 620–632.

Bibliography

- Howard, A.G. et al. (Apr. 2017). MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. arXiv: 1704.04861 [cs.CV].
- Huang, C. et al. (2021). Performance metrics for the comparative analysis of clinical risk prediction models employing machine learning. *Circulation: Cardiovascular Quality and Outcomes* 14.10, e007526.
- Huff, S.M. et al. (May 1998). Development of the Logical Observation Identifier Names and Codes (LOINC) vocabulary. en. *J. Am. Med. Inform. Assoc.* 5.3, 276–292.
- Hung, R.J. et al. (Mar. 2021). Assessing Lung Cancer Absolute Risk Trajectory Based on a Polygenic Risk Model. en. *Cancer Res.* 81.6, 1607–1615.
- Ian Goodfellow and Yoshua Bengio and Aaron Courville (2016). *Deep Learning* (MIT Press).
- Ikram, M.A. et al. (Sept. 2017). The Rotterdam Study: 2018 update on objectives, design and main results. en. *Eur. J. Epidemiol.* 32.9, 807–850.
- Ikram, M.K. et al. (Feb. 2006). Retinal vessel diameters and risk of hypertension: the Rotterdam Study. en. *Hypertension* 47.2, 189–194.
- Inoue, E. (May 2018). *nricens: R Functions to calculate NRI*. en. <https://rdrr.io/cran/nricens/man/nricens-package.html>. Accessed: 2022-10-14.
- Inouye, M. et al. (Oct. 2018). Genomic Risk Prediction of Coronary Artery Disease in 480,000 Adults: Implications for Primary Prevention. en. *J. Am. Coll. Cardiol.* 72.16, 1883–1893.
- International Consortium for Blood Pressure Genome-Wide Association Studies et al. (Sept. 2011). Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. en. *Nature* 478.7367, 103–109.
- Ioffe, S. and Szegedy, C. (2015). Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. In *Proceedings of the 32nd International Conference on Machine Learning*, F. Bach and D. Blei, ed. Vol. 37. *Proceedings of Machine Learning Research*. (Lille, France: PMLR), pp. 448–456.
- Janssens, A.C.J.W. (Nov. 2019). *Validity of polygenic risk scores: are we measuring what we think we are?*
- Johnson, C.M. et al. (June 2013). Meta-analyses of colorectal cancer risk factors. en. *Cancer Causes Control* 24.6, 1207–1222.
- Joint OECD, EUROSTAT and WHO Health Accounts SHA Questionnaires (JHAQ) (2022). *Per capita health expenditure and financing in OECD countries, measured in 2010 international dollars*. en. <https://stats.oecd.org/>. Accessed: 2022-8-7.
- Jordan, M.I. (May 1986). *Serial order: a parallel distributed processing approach. Technical report, June 1985-March 1986*. en. Tech. rep. AD-A-173989/5/XAB; ICS-8604. California Univ., San Diego, La Jolla (USA). Inst. for Cognitive Science.

- Julkunen, H. et al. (May 2021). Metabolic biomarker profiling for identification of susceptibility to severe pneumonia and COVID-19 in the general population. en. *Elife* 10.
- Kaplan, E.L. and Meier, P. (1958). Nonparametric Estimation from Incomplete Observations. *J. Am. Stat. Assoc.* 53.282, 457–481.
- Karl Schulz, L.S. (2020). *IBA: Information Bottlenecks for Attribution*. en.
- Katzman, J.L. et al. (2017). *DeepSurv: Personalized Treatment Recommender System Using A Cox Proportional Hazards Deep Neural Network*. Tech. rep.
- Kettunen, J. et al. (2018). Biomarker glycoprotein acetyls is associated with the risk of a wide spectrum of incident diseases and stratifies mortality risk in angiography patients. *Circulation: Genomic and Precision Medicine* 11.11, e002234.
- Khan, A., Choi, Y., et al. (Mar. 2020). High-resolution metabolomics study revealing l-homocysteine sulfinic acid, cysteic acid, and carnitine as novel biomarkers for high acute myocardial infarction risk. en. *Metabolism* 104, 154051.
- Khan, A., Turchin, M.C., et al. (July 2022). Genome-wide polygenic score to predict chronic kidney disease across ancestries. en. *Nat. Med.* 28.7, 1412–1420.
- Khera, A.V., Chaffin, M., Aragam, K.G., et al. (Sept. 2018). Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nat. Genet.* 50.9, 1219–1224.
- Khera, A.V., Chaffin, M., Wade, K.H., et al. (Apr. 2019). Polygenic Prediction of Weight and Obesity Trajectories from Birth to Adulthood. en. *Cell* 177.3, 587–596.e9.
- Khoury MJ, M.G.A. (June 2019). *Is it Time to Integrate Polygenic Risk Scores into Clinical Practice? Let's Do the Science First and Follow the Evidence Wherever it Takes Us!* <https://blogs.cdc.gov/genomics/2019/06/03/is-it-time/>. Accessed: 2021-2-5.
- Kim, D.W. et al. (May 2019). Deep learning-based survival prediction of oral cancer patients. en. *Sci. Rep.* 9.1, 6994.
- Kim, Y.D. et al. (Mar. 2020). Effects of Hypertension, Diabetes, and Smoking on Age and Sex Prediction from Retinal Fundus Images. en. *Sci. Rep.* 10.1, 4623.
- Kingma, D.P. and Ba, J.L. (2015). Adam: a Method for Stochastic Optimization. *International Conference on Learning Representations 2015*.
- Kivipelto, M. et al. (Sept. 2006). Risk score for the prediction of dementia risk in 20 years among middle aged people: a longitudinal, population-based study. en. *Lancet Neurol.* 5.9, 735–741.
- Klambauer, G. et al. (June 2017). Self-Normalizing Neural Networks. arXiv: 1706.02515 [cs.LG].

Bibliography

- Kolb, H., Fernandez, E., and Nelson, R. (1995). *Webvision: the organization of the retina and visual system [Internet]*. <https://pubmed.ncbi.nlm.nih.gov/NBK11530>. Accessed: 2022-6-9.
- Komaromy, A. et al. (2020). Glycoprotein biomarkers and analysis in chronic obstructive pulmonary disease and lung cancer with special focus on serum immunoglobulin G. *Clinica Chimica Acta* 506, 204–213.
- Krizhevsky, A., Sutskever, I., and Hinton, G.E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. In *Advances in Neural Information Processing Systems*, F. Pereira et al., ed. Vol. 25. (Curran Associates, Inc.).
- Kumar, I.E. et al. (2020). Problems with Shapley-value-based explanations as feature importance measures. *Proceedings of Machine Learning Research* 119. H.D. Iii and A. Singh, ed., 5491–5500.
- Kumuthini, J. et al. (Nov. 2022). The clinical utility of polygenic risk scores in genomic medicine practices: a systematic review. *Hum. Genet.* 141.11, 1697–1704.
- Kvamme, H., Borgan, Ø., and Scheel, I. (2019). *Time-to-Event Prediction with Neural Networks and Cox Regression*. Tech. rep., pp. 1–30.
- Lambert, S.A., Abraham, G., and Inouye, M. (Nov. 2019). Towards clinical utility of polygenic risk scores. *Hum. Mol. Genet.* 28.R2, R133–R142.
- Lambert, S.A., Gil, L., et al. (Apr. 2021). The Polygenic Score Catalog as an open database for reproducibility and systematic evaluation. *Nat. Genet.* 53.4, 420–425.
- Langham, J. et al. (Dec. 2021). Predicting risk of dementia with machine learning and survival models using routine primary care records. In *2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, (), pp. 3036–3042.
- Lapuschkin, S. et al. (Mar. 2019). Unmasking Clever Hans predictors and assessing what machines really learn. *Nat. Commun.* 10.1, 1–8.
- Larson, E.B. (Nov. 2019). “Risk factors for cognitive decline and dementia”. UpToDate, S.T. DeKosky and J.L. Wilterdink, ed. (Waltham, MA: UpToDate).
- Leavell et al. (1953). *Textbook of preventive medicine* (New York: McGraw-Hill).
- Lecun, Y. et al. (Nov. 1998). Gradient-based learning applied to document recognition. *Proc. IEEE* 86.11, 2278–2324.
- Lécuyer, L. et al. (Apr. 2018). NMR metabolomic signatures reveal predictive plasma metabolites associated with long-term risk of developing breast cancer. *Int. J. Epidemiol.* 47.2, 484–494.
- Lee, A. et al. (Aug. 2019). BOADICEA: a comprehensive breast cancer risk prediction model incorporating genetic and nongenetic risk factors. *Genet. Med.* 21.8, 1708–1718.

- Lee, C. et al. (Apr. 2018). DeepHit: A Deep Learning Approach to Survival Analysis With Competing Risks. en. *AAAI* 32.1.
- Lee, S.J. van der et al. (May 2018). The effect of APOE and other common genetic variants on the onset of Alzheimer’s disease and dementia: a community-based cohort study. en. *Lancet Neurol.* 17.5, 434–444.
- Li, B. et al. (Mar. 2021). Development and evaluation of a deep learning model for the detection of multiple fundus diseases based on colour fundus photography. en. *Br. J. Ophthalmol.*
- Li, F. et al. (July 2021). Deep learning-based automated detection for diabetic retinopathy and diabetic macular oedema in retinal fundus photographs. en. *Eye.*
- Li, Y. et al. (Dec. 2020). BEHRT: Transformer for Electronic Health Records. *Sci. Rep.* 10.1, 1–12.
- Liang, W. et al. (July 2020). Early triage of critically ill COVID-19 patients using deep learning. en. *Nat. Commun.* 11.1, 3543.
- Liew, G. et al. (Feb. 2011). Fractal analysis of retinal microvasculature and coronary heart disease mortality. en. *Eur. Heart J.* 32.4, 422–429.
- Lin, T.-Y. et al. (May 2014). Microsoft COCO: Common Objects in Context. arXiv: 1405.0312 [cs.CV].
- Lindström, J. and Tuomilehto, J. (Mar. 2003). The diabetes risk score: a practical tool to predict type 2 diabetes risk. en. *Diabetes Care* 26.3, 725–731.
- Liu, Z., Lin, Y., et al. (Mar. 2021). Swin Transformer: Hierarchical Vision Transformer using Shifted Windows. arXiv: 2103.14030 [cs.CV].
- Liu, Z., Mao, H., et al. (Jan. 2022). A ConvNet for the 2020s. arXiv: 2201.03545 [cs.CV].
- Livingston, G. et al. (Dec. 2017). Dementia prevention, intervention, and care. en. *Lancet* 390.10113, 2673–2734.
- Lombardo, E. et al. (Mar. 2021). Distant metastasis time to event analysis with CNNs in independent head and neck cancer cohorts. en. *Sci. Rep.* 11.1, 6418.
- Longato, E., Vettoretti, M., and Di Camillo, B. (Aug. 2020). A practical perspective on the concordance index for the evaluation and selection of prognostic time-to-event models. en. *J. Biomed. Inform.* 108, 103496.
- Lorenzoni, L. et al. (May 2019). *Health Spending Projections to 2030*. Tech. rep.
- Luchsinger, J.A. et al. (Aug. 2005). Aggregation of vascular risk factors and risk of incident Alzheimer disease. en. *Neurology* 65.4, 545–551.
- Luck, M. et al. (2017). *Deep Learning for Patient-Specific Kidney Graft Survival Analysis*. Tech. rep.

Bibliography

- Lundberg, S. (2022). *shap: A game theoretic approach to explain the output of any machine learning model*. en.
- Lundberg, S. and Lee, S.-I. (May 2017). A Unified Approach to Interpreting Model Predictions. arXiv: 1705.07874 [cs.AI].
- Lyssenko, V. et al. (Nov. 2008). Clinical risk factors, DNA variants, and the development of type 2 diabetes. en. *N. Engl. J. Med.* 359.21, 2220–2232.
- Ma, Y. and Zhou, X. (Nov. 2021). Genetic prediction of complex traits with polygenic scores: a statistical review. en. *Trends Genet.* 37.11, 995–1011.
- Maas, A.L., Hannun, A.Y., and Ng, A.Y. (2013). Rectifier nonlinearities improve neural network acoustic models. In *ICML Workshop on Deep Learning for Audio, Speech and Language Processing*, ().
- MacGillivray, T.J. et al. (Aug. 2014). Retinal imaging as a source of biomarkers for diagnosis, characterization and prognosis of chronic illness or long-term conditions. en. *Br. J. Radiol.* 87.1040, 20130832.
- Mach, F. et al. (Jan. 2020). *2019 ESC/EAS Guidelines for the management of dyslipidaemias: Lipid modification to reduce cardiovascular risk*.
- Mahendran, Y., Cederberg, H., et al. (Nov. 2013). Glycerol and fatty acids in serum predict the development of hyperglycemia and type 2 diabetes in Finnish men. en. *Diabetes Care* 36.11, 3732–3738.
- Mahendran, Y., Vangipurapu, J., et al. (Oct. 2013). Association of ketone body levels with hyperglycemia and type 2 diabetes in 9,398 Finnish men. en. *Diabetes* 62.10, 3618–3626.
- Markley, J.L. et al. (Feb. 2017). The future of NMR-based metabolomics. en. *Curr. Opin. Biotechnol.* 43, 34–40.
- Marmot, M. and Brunner, E. (Apr. 2005). Cohort Profile: The Whitehall II study. *Int. J. Epidemiol.* 34.2, 251–256.
- Marsh, T.L., Janes, H., and Pepe, M.S. (Sept. 2020). Statistical inference for net benefit measures in biomarker validation studies. en. *Biometrics* 76.3, 843–852.
- Matsuo, K. et al. (Apr. 2019). Survival outcome prediction in cervical cancer: Cox models vs deep-learning model. en. *Am. J. Obstet. Gynecol.* 220.4, 381.e1–381.e14.
- Mavaddat, N. et al. (Jan. 2019). Polygenic Risk Scores for Prediction of Breast Cancer and Breast Cancer Subtypes. en. *Am. J. Hum. Genet.* 104.1, 21–34.
- McGeechan, K. et al. (July 2008). Risk prediction of coronary heart disease based on retinal vascular caliber (from the Atherosclerosis Risk In Communities [ARIC] Study). en. *The American journal of cardiology* 102.1, 58–63. issn: 0002-9149. doi: 10.1016/j.amjcard.2008.02.094. url: <http://dx.doi.org/10.1016/j.amjcard.2008.02.094>.

- McInnes, L., Healy, J., and Melville, J. (2018). Umap: Uniform manifold approximation and projection for dimension reduction. arXiv preprint arXiv:1802.03426.
- Meigs, J.B. et al. (Nov. 2008). Genotype score in addition to common risk factors for prediction of type 2 diabetes. en. *N. Engl. J. Med.* 359.21, 2208–2219.
- Minaee, S. et al. (Oct. 2020). Deep-COVID: Predicting COVID-19 from chest X-ray images using deep transfer learning. en. *Med. Image Anal.* 65, 101794.
- Mitani, A., Hammel, N., and Liu, Y. (June 2021). Retinal detection of kidney disease and diabetes. en. *Nat Biomed Eng* 5.6, 487–489.
- Mitani, A., Huang, A., et al. (Jan. 2020). Detection of anaemia from retinal fundus images via deep learning. *Nature Biomedical Engineering* 4.1, 18–27.
- Mobadersany, P. et al. (Mar. 2018). Predicting cancer outcomes from histology and genomics using convolutional networks. en. *Proc. Natl. Acad. Sci. U. S. A.* 115.13, E2970–E2979.
- Molnar, C. (Feb. 2022). *Interpretable Machine Learning: A Guide For Making Black Box Models Explainable* (Independently published).
- Morey, J.R. et al. (July 2020). “Impact of Viz LVO on time-to-treatment and clinical outcomes in large vessel occlusion stroke patients presenting to Primary Stroke Centers”.
- Morieri, M.L. et al. (Nov. 2018). Genetic Tools for Coronary Risk Assessment in Type 2 Diabetes: A Cohort Study From the ACCORD Clinical Trial. en. *Diabetes Care* 41.11, 2404–2413.
- Mosley, J.D. et al. (Feb. 2020). Predictive Accuracy of a Polygenic Risk Score Compared With a Clinical Risk Score for Incident Coronary Heart Disease. en. *JAMA* 323.7, 627–635.
- My Health Record (Aug. 2022). *My Health Record*. en. <https://www.myhealthrecord.gov.au/>. Accessed: 2022-8-1.
- Nagpal, C., Jeanselme, V., and Dubrawski, A. (2021). Deep Parametric Time-to-Event Regression with Time-Varying Covariates. *Proceedings of Machine Learning Research* 146. R. Greiner et al., ed., 184–193.
- Nagpal, C., Li, X.R., and Dubrawski, A. (Jan. 2021). Deep Survival Machines: Fully Parametric Survival Regression and Representation Learning for Censored Data with Competing Risks. en. *IEEE J Biomed Health Inform PP*.
- Nair, V. and Hinton, G.E. (June 2010). Rectified linear units improve restricted boltzmann machines. In *Proceedings of the 27th International Conference on International Conference on Machine Learning, ICML’10*. (Haifa, Israel: Omnipress), pp. 807–814.
- Natarajan, P. et al. (May 2017). Polygenic Risk Score Identifies Subgroup With Higher Burden of Atherosclerosis and Greater Relative Benefit From Statin Therapy in the Primary Prevention Setting. en. *Circulation* 135.22, 2091–2101.

Bibliography

- National Institute for health and Care Excellence (NICE) (July 2014). *Cardiovascular disease: risk assessment and reduction, including lipid modification*. en. <https://www.nice.org.uk/guidance/cg181>. Accessed: 2022-9-16.
- Nelson, C.A. et al. (Jan. 2022). Embedding electronic health records onto a knowledge network recognizes prodromal features of multiple sclerosis and predicts diagnosis. en. *J. Am. Med. Inform. Assoc.* 29.3, 424–434.
- Ngandu, T. et al. (2022). The effect of adherence on cognition in a multidomain lifestyle intervention (FINGER). *Alzheimer's & Dementia* 18.7, 1325–1334.
- NHS (Oct. 2022). *NHS Health Check*. en. <https://www.nhs.uk/conditions/nhs-health-check/>. Accessed: 2022-10-6.
- NHS Connecting for Health, Great Britain: Department of Health, and Health and Social Care Information Centre (Feb. 2006). *OPCS Classifications of Interventions and Procedures* en. (The Stationery Office).
- Nicholson, J.K. et al. (2012). Metabolic phenotyping in clinical and surgical environments. *Nature* 491.7424, 384–392.
- Nikpay, M. et al. (Sept. 2015). A comprehensive 1000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat. Genet.* 47.10, 1121–1130.
- Nusinovici, S. et al. (Apr. 2022). Retinal photograph-based deep learning predicts biological age, and stratifies morbidity and mortality risk. en. *Age Ageing* 51.4.
- Owen, C.G. et al. (Jan. 2019). Retinal Vasculometry Associations with Cardiometabolic Risk Factors in the European Prospective Investigation of Cancer—Norfolk Study. *Ophthalmology* 126.1, 96–106. issn: 0161-6420. doi: 10.1016/j.ophtha.2018.07.022. url: <https://www.sciencedirect.com/science/article/pii/S0161642018312557>.
- Parekh, D.J. et al. (Dec. 2006). External validation of the Prostate Cancer Prevention Trial risk calculator in a screened population. en. *Urology* 68.6, 1152–1155.
- Paszke, A. et al. (2017). Automatic differentiation in PyTorch. *Advances in Neural Information Processing Systems* 30 Nips, 1–4.
- Patton, N. et al. (Apr. 2005). Retinal vascular image analysis as a potential screening tool for cerebrovascular disease: a rationale based on homology between cerebral and retinal microvasculatures. en. *J. Anat.* 206.4, 319–348.
- Pauker, S.G. and Kassirer, J.P. (May 1980). The threshold approach to clinical decision making. en. *N. Engl. J. Med.* 302.20, 1109–1117.
- Paul Robertson, R. (Jan. 2022). “Type 2 diabetes mellitus: Prevalence and risk factors”. UpToDate, J.E. Mulder and D.M. Nathan, ed. (Waltham, MA: UpToDate).

- Pearlman, R. et al. (Apr. 2017). Prevalence and Spectrum of Germline Cancer Susceptibility Gene Mutations Among Patients With Early-Onset Colorectal Cancer. en. *JAMA Oncol* 3.4, 464–471.
- Pedregosa, F. et al. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* 12, 2825–2830.
- Pencina, M.J., D’Agostino Sr, R.B., D’Agostino Jr, R.B., et al. (Jan. 2008). Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. en. *Stat. Med.* 27.2, 157–72, discussion 207–12.
- Pencina, M.J., D’Agostino Sr, R.B., and Steyerberg, E.W. (Jan. 2011). Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. en. *Stat. Med.* 30.1, 11–21.
- Peng, Y., Dharssi, S., et al. (Apr. 2019). DeepSeeNet: A Deep Learning Model for Automated Classification of Patient-based Age-related Macular Degeneration Severity from Color Fundus Photographs. en. *Ophthalmology* 126.4, 565–575.
- Peng, Y., Keenan, T.D., et al. (Aug. 2020). Predicting risk of late age-related macular degeneration using deep learning. en. *NPJ Digit Med* 3, 111.
- Pepe, M.S. et al. (Oct. 2015). The Net Reclassification Index (NRI): a Misleading Measure of Prediction Improvement Even with Independent Test Data Sets. en. *Stat. Biosci.* 7.2, 282–295.
- Pietzner, M. et al. (Mar. 2021). Plasma metabolites to profile pathways in noncommunicable disease multimorbidity. en. *Nat. Med.* 27.3, 471–479.
- Poplin, R. et al. (Mar. 2018). Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning. en. *Nat Biomed Eng* 2.3, 158–164.
- Porter, M.E. (Dec. 2010). What Is Value in Health Care? *N. Engl. J. Med.* 363.26, 2477–2481.
- Porter, M.E. and Kaplan, R.S. (July 2016). How to pay for health care. en. *Harv. Bus. Rev.* 94.7-8, 88–98, 100, 134.
- Pozzilli, P., Strollo, R., and Bonora, E. (2014). One size does not fit all glycemic targets for type 2 diabetes. *Journal of Diabetes Investigation* 5.2, 134–141.
- Pronk, N.P., Remington, P.L., and Force*, C.P.S.T. (2015). Combined diet and physical activity promotion programs for prevention of diabetes: community preventive services task force recommendation statement. *Annals of internal medicine* 163.6, 465–468.
- Rajkomar, A. et al. (May 2018). Scalable and accurate deep learning with electronic health records. en. *NPJ Digit Med* 1, 18.

Bibliography

- Rasmy, L. et al. (May 2021). Med-BERT: pretrained contextualized embeddings on large-scale structured electronic health records for disease prediction. en. *NPJ Digit Med* 4.1, 86.
- Reichstein, M. et al. (Feb. 2019). Deep learning and process understanding for data-driven Earth system science. en. *Nature* 566.7743, 195–204.
- Ren, K. et al. (Sept. 2019). Deep Recurrent Survival Analysis. Proc. Conf. AAAI Artif. Intell. 33, 4798–4805.
- Riba, E. et al. (Sept. 2020). A survey on Kornia: an Open Source Differentiable Computer Vision Library for PyTorch. arXiv: 2009.10521 [cs.CV].
- Ribeiro, M.T., Singh, S., and Guestrin, C. (Aug. 2016). “Why Should I Trust You?”: Explaining the Predictions of Any Classifier. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD ’16. (San Francisco, California, USA: Association for Computing Machinery), pp. 1135–1144.
- Rietschel, C., Yoon, J., and Van Der Schaar, M. (2018). *Feature Selection for Survival Analysis with Competing Risks using Deep Learning*. Tech. rep.
- Rim, T.H., Lee, C.J., et al. (May 2021). Deep-learning-based cardiovascular risk stratification using coronary artery calcium scores predicted from retinal photographs. en. *Lancet Digit Health* 3.5, e306–e316.
- Rim, T.H., Lee, G., et al. (Oct. 2020). Prediction of systemic biomarkers from retinal photographs: development and validation of deep-learning algorithms. en. *Lancet Digit Health* 2.10, e526–e536.
- Rim, T.H., Teo, A.W.J., et al. (Mar. 2020). Retinal Vascular Signs and Cerebrovascular Diseases. en. *J. Neuroophthalmol.* 40.1, 44–59.
- Riveros-Mckay Fernando et al. (2021). An Integrated Polygenic Tool Substantially Enhances Coronary Artery Disease Prediction. en. *Circulation: Genomic and Precision Medicine* 0.0.
- Roe, M.T. and Ohman, E.M. (Jan. 2012). A new era in secondary prevention after acute coronary syndrome. en. *N. Engl. J. Med.* 366.1, 85–87.
- Ronit, A. et al. (2020). Plasma albumin and incident cardiovascular disease: results from the CGPS and an updated meta-analysis. *Arteriosclerosis, thrombosis, and vascular biology* 40.2, 473–482.
- Rosenberg, A. et al. (2018). Multidomain lifestyle intervention benefits a large elderly population at risk for cognitive decline and dementia regardless of baseline characteristics: The FINGER trial. *Alzheimer’s & Dementia* 14.3, 263–270.
- Rosenblatt, F. (Nov. 1958). The perceptron: a probabilistic model for information storage and organization in the brain. en. *Psychol. Rev.* 65.6, 386–408.

- Rousson, V. and Zumbo, T. (June 2011). Decision curve analysis revisited: overall net benefit, relationships to ROC curve analysis, and application to case-control studies. en. *BMC Med. Inform. Decis. Mak.* 11, 45.
- Ruderfer, D.M. et al. (Apr. 2016). Polygenic overlap between schizophrenia risk and antipsychotic response: a genomic medicine approach. en. *Lancet Psychiatry* 3.4, 350–357.
- Rudnicka, A.R. et al. (Oct. 2022). Artificial intelligence-enabled retinal vasculometry for prediction of circulatory mortality, myocardial infarction and stroke. en. *The British journal of ophthalmology*. issn: 0007-1161, 1468-2079. doi: 10.1136/bjo-2022-321842. url: <http://dx.doi.org/10.1136/bjo-2022-321842>.
- Rumelhart, D.E., Hinton, G.E., and Williams, R.J. (Oct. 1986). Learning representations by back-propagating errors. en. *Nature* 323.6088, 533–536.
- Rush, R. (July 2019). Taking Note. en. *The New England journal of medicine* 381.1, 9. issn: 0028-4793, 1533-4406. doi: 10.1056/NEJMp1816537. url: <http://dx.doi.org/10.1056/NEJMp1816537>.
- Saaristo, T. et al. (May 2005). Cross-sectional evaluation of the Finnish Diabetes Risk Score: a tool to identify undetected type 2 diabetes, abnormal glucose tolerance and metabolic syndrome. en. *Diab. Vasc. Dis. Res.* 2.2, 67–72.
- Sabanayagam, C. et al. (June 2020). A deep learning algorithm to detect chronic kidney disease from retinal photographs in community-based populations. en. *Lancet Digit Health* 2.6, e295–e302.
- Schoenmaker, M. et al. (Jan. 2006). Evidence of genetic enrichment for exceptional survival using a family approach: the Leiden Longevity Study. en. *Eur. J. Hum. Genet.* 14.1, 79–84.
- Schultz, L.M. et al. (Apr. 2022). Stability of polygenic scores across discovery genome-wide association studies. en. *HGG advances* 3.2, 100091. issn: 2666-2477. doi: 10.1016/j.xhgg.2022.100091. url: <http://dx.doi.org/10.1016/j.xhgg.2022.100091>.
- Schulz, K. et al. (2020). Restricting the flow: Information bottlenecks for attribution. arXiv preprint arXiv:2001.00396.
- Schüssler-Fiorenza Rose, S.M. et al. (May 2019). A longitudinal big data approach for precision health. en. *Nat. Med.* 25.5, 792–804.
- Schwartz, A.G. and Cote, M.L. (2016). “Epidemiology of Lung Cancer”. *Lung Cancer and Personalized Medicine: Current Knowledge and Therapies*, A. Ahmad and S. Gadgeel, ed. (Cham: Springer International Publishing), pp. 21–41.
- Schwarz, P.E.H. et al. (Feb. 2009). Tools for predicting the risk of type 2 diabetes in daily practice. en. *Horm. Metab. Res.* 41.2, 86–97.

Bibliography

- SCORE2 working group and ESC Cardiovascular risk collaboration (July 2021). SCORE2 risk prediction algorithms: new models to estimate 10-year risk of cardiovascular disease in Europe. en. *Eur. Heart J.* 42.25, 2439–2454.
- Sekelj, S. et al. (May 2021). Detecting undiagnosed atrial fibrillation in UK primary care: Validation of a machine learning prediction algorithm in a retrospective cohort study. en. *Eur. J. Prev. Cardiol.* 28.6, 598–605.
- Shang, N. et al. (Apr. 2021). Medical records-based chronic kidney disease phenotype for clinical care and “big data” observational and genetic studies. en. *NPJ Digit Med* 4.1, 70.
- Sharma, V. et al. (Feb. 2021). Adoption of clinical risk prediction tools is limited by a lack of integration with electronic health records. en. *BMJ Health Care Inform* 28.1.
- Shepherd, J., Blauw, G.J., Murphy, M.B., Cobbe, S.M., et al. (1999). The design of a prospective study of pravastatin in the elderly at risk (PROSPER). *Am. J. Cardiol.* 84.10, 1192–1197.
- Shepherd, J., Blauw, G.J., Murphy, M.B., and Gaw, A. (2002). A prospective study of Pravastatin in the Elderly at Risk (PROSPER). *Lancet* 360, 685–696.
- Sherry, S.T. et al. (Jan. 2001). dbSNP: the NCBI database of genetic variation. en. *Nucleic Acids Res.* 29.1, 308–311.
- Shomorony, I. et al. (Jan. 2020). An unsupervised learning approach to identify novel signatures of health and disease from multimodal data. en. *Genome Med.* 12.1, 7.
- Shrikumar, A., Greenside, P., and Kundaje, A. (2017). Learning Important Features Through Propagating Activation Differences. *Proceedings of Machine Learning Research* 70. D. Precup and Y.W. Teh, ed., 3145–3153.
- Siegel, R.L. et al. (Jan. 2022). Cancer statistics, 2022. en. *CA Cancer J. Clin.* 72.1, 7–33.
- Silarova, B. et al. (2019). Effect of communicating phenotypic and genetic risk of coronary heart disease alongside web-based lifestyle advice: the INFORM Randomised Controlled Trial. *Heart* 105.13, 982–989.
- Simonyan, K. and Zisserman, A. (Sept. 2014). Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv: 1409.1556 [cs.CV]*.
- Sindi, S. et al. (Sept. 2015). The CAIDE Dementia Risk Score App: The development of an evidence-based mobile application to predict the risk of dementia. en. *Alzheimers Dement.* 1.3, 328–333.
- Sixt, L., Granz, M., and Landgraf, T. (2020). When explanations lie: Why many modified bp attributions fail. *International Conference on*.
- Sjoberg, D.D. (2021). “dcurves: decision curve analysis for model evaluation”.
- Sobel, I. (1968). “An Isotropic 3x3 Image Gradient Operator”.

- Soininen, P. et al. (Feb. 2015). Quantitative serum nuclear magnetic resonance metabolomics in cardiovascular epidemiology and genetics. en. *Circ. Cardiovasc. Genet.* 8.1, 192–206.
- Solares, J.R.A. et al. (July 2021). Transfer Learning in Electronic Health Records through Clinical Concept Embedding. arXiv: 2107.12919 [cs.LG].
- Srinivas, A. et al. (Jan. 2021). Bottleneck Transformers for Visual Recognition. arXiv: 2101.11605 [cs.CV].
- Srivastava, N. et al. (2014). Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *J. Mach. Learn. Res.* 15.56, 1929–1958.
- Stancáková, A. et al. (July 2012). Hyperglycemia and a common variant of GCKR are associated with the levels of eight amino acids in 9,369 Finnish men. en. *Diabetes* 61.7, 1895–1902.
- Stanton, A.V. et al. (Dec. 1995). Vascular network changes in the retina with age and hypertension. en. *J. Hypertens.* 13.12 Pt 2, 1724–1728.
- Stekhoven, D.J. and Bühlmann, P. (Jan. 2012). MissForest–non-parametric missing value imputation for mixed-type data. en. *Bioinformatics* 28.1, 112–118.
- Steyerberg, E.W., Roobol, M.J., et al. (Jan. 2007). Prediction of indolent prostate cancer: validation and updating of a prognostic nomogram. en. *J. Urol.* 177.1, 107–12, discussion 112.
- Steyerberg, E.W., Moons, K.G.M., et al. (Feb. 2013). Prognosis Research Strategy (PROGRESS) 3: prognostic model research. en. *PLoS Med.* 10.2, e1001381.
- Stutz, D. (2022). *Collection of LaTeX resources and examples*. <https://github.com/davidstutz/latex-resources>. Accessed on 10.08.2022.
- Sudlow, C. et al. (2015). UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. *PLoS medicine* 12.3, e1001779. issn: 1549-1277, 1549-1676. doi: 10.1371/journal.pmed.1001779. url: <https://journals.plos.org/plosmedicine/article/file?id=10.1371/journal.pmed.1001779&type=printable>.
- Sun, H. et al. (Jan. 2022). IDF Diabetes Atlas: Global, regional and country-level diabetes prevalence estimates for 2021 and projections for 2045. en. *Diabetes Res. Clin. Pract.* 183, 109119.
- Sun, L. et al. (2021). Polygenic risk scores in cardiovascular risk prediction: A cohort study and modelling analyses.
- Szegedy, C. et al. (Dec. 2015). Rethinking the Inception Architecture for Computer Vision. arXiv: 1512.00567 [cs.CV].
- Tada, H. et al. (Feb. 2016). Risk prediction by genetic risk scores for coronary heart disease is independent of self-reported family history. en. *Eur. Heart J.* 37.6, 561–567.

Bibliography

- Tan, M. and Le, Q.V. (May 2019). EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. arXiv: 1905.11946 [cs.LG].
- Tang, A.S. et al. (Feb. 2022). Deep phenotyping of Alzheimer's disease leveraging electronic medical records identifies sex-specific clinical associations. en. *Nat. Commun.* 13.1, 675.
- Tang, E.Y.H. et al. (Sept. 2015). Current Developments in Dementia Risk Prediction Modelling: An Updated Systematic Review. en. *PLoS One* 10.9, e0136181.
- Theuns, D.A.M.J. et al. (Nov. 2010). Effectiveness of prophylactic implantation of cardioverter-defibrillators without cardiac resynchronization therapy in patients with ischaemic or non-ischaemic heart disease: a systematic review and meta-analysis. en. *Europace* 12.11, 1564–1570.
- Tibshirani, R. (Feb. 1997). The lasso method for variable selection in the Cox model. en. *Stat. Med.* 16.4, 385–395.
- Tikkanen, R. and Abrams, M.K. (2020). US health care from a global perspective, 2019: higher spending, worse outcomes. The Commonwealth Fund.
- Tiwari, P. et al. (Jan. 2020). Assessment of a Machine Learning Model Applied to Harmonized Electronic Health Record Data for the Prediction of Incident Atrial Fibrillation. en. *JAMA Netw Open* 3.1, e1919396.
- Tolkach, Y. et al. (July 2020). High-accuracy prostate cancer pathology using deep learning. en. *Nature Machine Intelligence* 2.7, 411–418.
- Tomašev, N. et al. (Aug. 2019). A clinically applicable approach to continuous prediction of future acute kidney injury. en. *Nature* 572.7767, 116–119.
- Tonekaboni, S. et al. (2019). What Clinicians Want: Contextualizing Explainable Machine Learning for Clinical End Use. *Proceedings of Machine Learning Research* 106. F. Doshi-Velez et al., ed., 359–380.
- Tonelli, M. et al. (Nov. 2011). Efficacy of statins for primary prevention in people at low cardiovascular risk: a meta-analysis. en. *CMAJ* 183.16, E1189–202.
- Torkamani, A., Wineinger, N.E., and Topol, E.J. (Sept. 2018). The personal and clinical utility of polygenic risk scores. en. *Nat. Rev. Genet.* 19.9, 581–590.
- Tynkkynen, J. et al. (June 2018). Association of branched-chain amino acids and other circulating metabolites with risk of incident dementia and Alzheimer's disease: A prospective study in eight cohorts. en. *Alzheimers. Dement.* 14.6, 723–733.
- Uffelmann, E. et al. (Aug. 2021). Genome-wide association studies. en. *Nature Reviews Methods Primers* 1.1, 1–21.
- UK Department of Health and Social Care (2019). *Advancing our health: prevention in the 2020s*. <https://www.gov.uk/government/consultations/advancing-our->

- health-prevention-in-the-2020s/advancing-our-health-prevention-in-the-2020s-consultation-document. Accessed: 2021-2-5.
- US Burden of Disease Collaborators et al. (Apr. 2018). The State of US Health, 1990-2016: Burden of Diseases, Injuries, and Risk Factors Among US States. *en. JAMA* 319.14, 1444–1472.
- US Preventive Services Task Force et al. (May 2021). Screening for Colorectal Cancer: US Preventive Services Task Force Recommendation Statement. *en. JAMA* 325.19, 1965–1977.
- Vale-Silva, L.A. and Rohr, K. (June 2021). Long-term cancer survival prediction using multimodal deep learning. *en. Sci. Rep.* 11.1, 13505.
- Vaswani, A. et al. (2017). Attention is All you Need. In *Advances in Neural Information Processing Systems*, I. Guyon et al., ed. Vol. 30. (Curran Associates, Inc.).
- Vayena, E. (2021). Value from health data: European opportunity to catalyse progress in digital health. *The Lancet* 397.10275, 652–653.
- Velde, M. van der et al. (2009). Screening for albuminuria identifies individuals at increased renal risk. *Journal of the American Society of Nephrology* 20.4, 852–862.
- Vickers, A.J., Calster, B. van, and Steyerberg, E.W. (Oct. 2019). A simple, step-by-step guide to interpreting decision curve analysis. *en. Diagn Progn Res* 3, 18.
- Vickers, A.J. and Elkin, E.B. (Nov. 2006). Decision curve analysis: a novel method for evaluating prediction models. *en. Med. Decis. Making* 26.6, 565–574.
- Vickers, A.J., Van Calster, B., and Steyerberg, E.W. (Jan. 2016). Net benefit approaches to the evaluation of prediction models, molecular markers, and diagnostic tests. *en. BMJ* 352.
- Virtanen, J.K. et al. (2014). Serum omega-3 polyunsaturated fatty acids and risk of incident type 2 diabetes in men: the Kuopio Ischemic Heart Disease Risk Factor study. *Diabetes care* 37.1, 189–196.
- Vugt, H.A. van et al. (July 2012). Selecting men diagnosed with prostate cancer for active surveillance using a risk calculator: a prospective impact study. *en. BJU Int.* 110.2, 180–187.
- Walt, S. van der et al. (June 2014). scikit-image: image processing in Python. *en. PeerJ* 2, e453.
- Walters, K. et al. (Jan. 2016). Predicting dementia risk in primary care: development and validation of the Dementia Risk Score using routinely collected data. *en. BMC Med.* 14, 6.
- Wang, D. et al. (June 2016). Deep Learning for Identifying Metastatic Breast Cancer. *arXiv: 1606.05718 [q-bio.QM]*.

Bibliography

- Wang, S.-M. et al. (2021). Late-life depression, subjective cognitive decline, and their additive risk in incidence of dementia: a nationwide longitudinal study. *Plos one* 16.7, e0254639.
- Wannamethee, S.G., Shaper, A.G., and Perry, I.J. (1997). Serum creatinine concentration and risk of cardiovascular disease: a possible marker for increased risk of stroke. *Stroke* 28.3, 557–563.
- Ward, J. et al. (Sept. 2018). Polygenic risk scores for major depressive disorder and neuroticism as predictors of antidepressant response: Meta-analysis of three treatment cohorts. en. *PLoS One* 13.9, e0203896.
- Webb, S. (Feb. 2018). Deep learning for biology. en. *Nature* 554.7693, 555–557.
- Wei, W.-Q. et al. (July 2017). Evaluating phecodes, clinical classification software, and ICD-9-CM codes for phenome-wide association studies in the electronic health record. en. *PLoS One* 12.7, e0175508.
- Weinstein, M.C., Fineberg, H.V., Elstein, A.S., et al. (1980). *Clinical decision analysis* (Philadelphia: Saunders).
- Weissman, D.E. and Meier, D.E. (2011). Identifying patients in need of a palliative care assessment in the hospital setting a consensus report from the Center to Advance Palliative Care. *Journal of palliative medicine* 14.1, 17–23.
- Who (1992). *International classification of diseases*. WHO [Internet].
- Wilson, P.W. et al. (May 1998). Prediction of coronary heart disease using risk factor categories. en. *Circulation* 97.18, 1837–1847.
- Wilson, P.W.F., Castelli, W.P., and Kannel, W.B. (May 1987). Coronary risk prediction in adults (The Framingham Heart Study). *Am. J. Cardiol.* 59.14, G91–G94.
- Wilson, P.W.F., Meigs, J.B., et al. (May 2007). Prediction of incident diabetes mellitus in middle-aged adults: the Framingham Offspring Study. en. *Arch. Intern. Med.* 167.10, 1068–1074.
- Wong, Hubbard, Cruickshanks, et al. (2004). The relation of retinal vessel caliber to the incidence and progression of diabetic retinopathy: Xix: The wisconsin epidemiologic study of diabetic retinopathy. *Arch. nat. conserv. landsc. res.*
- Wong, T.Y., Islam, F.M.A., et al. (June 2006). Retinal vascular caliber, cardiovascular risk factors, and inflammation: the multi-ethnic study of atherosclerosis (MESA). en. *Invest. Ophthalmol. Vis. Sci.* 47.6, 2341–2350.
- Wong, T.Y., Klein, R., et al. (Apr. 2003). The prevalence and risk factors of retinal microvascular abnormalities in older persons: The Cardiovascular Health Study. en. *Ophthalmology* 110.4, 658–666.
- Wong, T.Y. and Mitchell, P. (Feb. 2007). The eye in hypertension. en. *Lancet* 369.9559, 425–435.

- Wood, A. et al. (Apr. 2021). Linked electronic health records for research on a nationwide cohort of more than 54 million people in England: data resource. en. *BMJ* 373.
- Wu, J.H. et al. (2017). Omega-6 fatty acid biomarkers and incident type 2 diabetes: pooled analysis of individual-level data for 39 740 adults from 20 prospective cohort studies. *The lancet Diabetes & endocrinology* 5.12, 965–974.
- Wu, P. et al. (Nov. 2019). Mapping ICD-10 and ICD-10-CM Codes to Phecodes: Workflow Development and Initial Evaluation. en. *JMIR Med Inform* 7.4, e14325.
- Würtz, P., Havulinna, A.S., et al. (Mar. 2015). Metabolite profiling and cardiovascular event risk: a prospective study of 3 population-based cohorts. en. *Circulation* 131.9, 774–785.
- Würtz, P., Kangas, A.J., et al. (Nov. 2017). Quantitative Serum Nuclear Magnetic Resonance Metabolomics in Large-Scale Epidemiology: A Primer on -Omic Technologies. en. *Am. J. Epidemiol.* 186.9, 1084–1096.
- Würtz, P., Tiainen, M., et al. (Aug. 2012). Circulating metabolite predictors of glycemia in middle-aged men and women. en. *Diabetes Care* 35.8, 1749–1756.
- Wyatt, J.C. and Altman, D.G. (Dec. 1995). Commentary: Prognostic models: clinically useful or quickly forgotten? en. *BMJ* 311.7019, 1539–1541.
- Xiao, C., Choi, E., and Sun, J. (Oct. 2018). Opportunities and challenges in developing deep learning models using electronic health records data: a systematic review. en. *J. Am. Med. Inform. Assoc.* 25.10, 1419–1428.
- Xiao, W. et al. (Feb. 2021). Screening and identifying hepatobiliary diseases through deep learning using ocular images: a prospective, multicentre study. en. *The Lancet. Digital health* 3.2, e88–e97. issn: 2589-7500. doi: 10.1016/S2589-7500(20)30288-0. url: [http://dx.doi.org/10.1016/S2589-7500\(20\)30288-0](http://dx.doi.org/10.1016/S2589-7500(20)30288-0).
- Xie, S. et al. (Nov. 2016). Aggregated Residual Transformations for Deep Neural Networks. arXiv: 1611.05431 [cs.CV].
- Xiu, Z., Tao, C., and Henao, R. (Mar. 2020). Variational Learning of Individual Survival Distributions. *ACM CHIL 2020 - Proceedings of the 2020 ACM Conference on Health, Inference, and Learning*, 10–18.
- Xu, Z. et al. (July 2022). Incremental value of risk factor variability for cardiovascular risk prediction in individuals with type 2 diabetes: results from UK primary care electronic health records. en. *Int. J. Epidemiol.*
- Yan, Q. et al. (Feb. 2020). Deep-learning-based Prediction of Late Age-Related Macular Degeneration Progression. en. *Nat Mach Intell* 2.2, 141–150.
- Yengo, L. et al. (Oct. 2022). A saturated map of common genetic variants associated with human height. en. *Nature*.

Bibliography

- Yousefi, S. et al. (Dec. 2017). Predicting clinical outcomes from large scale cancer genomic profiles with deep survival models. *Sci. Rep.* 7.1, 1–11.
- Yun, S. et al. (May 2019). CutMix: Regularization Strategy to Train Strong Classifiers with Localizable Features. arXiv: 1905.04899 [cs.CV].
- Yusuf, S. et al. (2004). Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study. *The lancet* 364.9438, 937–952.
- Zhang, J. et al. (Oct. 2018). Patient2Vec: A Personalized Interpretable Deep Representation of the Longitudinal Electronic Health Record. arXiv: 1810.04793 [q-bio.QM].
- Zhang, K. et al. (June 2021). Deep-learning models for the detection and incidence prediction of chronic kidney disease and type 2 diabetes from retinal fundus images. *en. Nat Biomed Eng* 5.6, 533–545.
- Zhang, Z. et al. (May 2019). Pathologist-level interpretable whole-slide cancer diagnosis with deep learning. *en. Nature Machine Intelligence* 1.5, 236–245.
- Zhao, J. et al. (Dec. 2019). Learning from Longitudinal Data in Electronic Health Record and Genetic Data to Improve Cardiovascular Event Prediction. *Sci. Rep.* 9.1, 717.
- Zheng et al. (2020). ... Colorectal Cancer Risk Prediction Model Incorporating Family History, Personal Characteristics, and Environmental Factors Colorectal Cancer Risk Prediction. *Cancer Epidemiol.*
- Zhu, Z. et al. (Dec. 2020). “Retinal age as a predictive biomarker for mortality risk”.

Declaration of authorship

I declare to the Freie Universität Berlin that I have completed the submitted dissertation independently and without the use of sources and aids other than those indicated. The present thesis is free of plagiarism. I have marked as such all statements that are taken literally or in content from other writings. This dissertation has not been submitted in the same or similar form in any previous doctoral procedure.

I agree to have my thesis examined by a plagiarism examination software.

Berlin,

Place, Date

Signature

Zusammenfassung

Die Kosten im Gesundheitswesen steigen systematisch und derzeitige therapieorientierte Gesundheitssysteme sind nicht nachhaltig. Angesichts vieler verhinderbarer Krankheiten stellt die Prävention ein veritables Instrument zur Verringerung von Kosten und Leiden dar. Risikostratifizierung ist die grundlegende Voraussetzung für ein präventionszentriertes Gesundheitswesen um Personen mit hohem Risiko zu identifizieren und Maßnahmen einzuleiten. Heute ist eine systematische Risikostratifizierung jedoch nur begrenzt möglich, da für die meisten Krankheiten keine Risikomodelle existieren und sich verfügbare Modelle auf einzelne Krankheiten beschränken. Weil für deren Berechnung jeweils spezielle Sets an Prädiktoren zu erheben sind werden in Praxis oft nur wenige Modelle angewandt. Gleichzeitig versprechen komplexe Datenmodalitäten, wie Bilder oder -omics-Messungen, systemische Informationen über zukünftige Gesundheitsverläufe, mit potentieller Relevanz für viele Endpunkte gleichzeitig. Da es an dedizierten Methoden zur Extraktion klinisch relevanter Informationen fehlt, sind diese Daten jedoch für die Risikomodellierung unzugänglich, und ihr Potenzial blieb bislang unbewertet. Diese Studie nutzt maschinelles Lernen, um die Anwendbarkeit von vier Datenmodalitäten in der Primärprävention zu untersuchen: polygene Risikoscores für die kardiovaskuläre Prävention, NMR Metabolomicsdaten, elektronische Gesundheitsakten und Netzhautfundusfotos. Pro Datenmodalität wurde ein neuronales Risikomodell entwickelt, um relevante Informationen zu extrahieren, additive Information gegenüber üblicherweise erfassten Kovariaten zu quantifizieren und den potenziellen klinischen Nutzen der Datenmodalität zu ermitteln. Die entwickelte Methodik konnte polygene Risikoscores für die kardiovaskuläre Prävention integrieren. Im Falle der NMR-Metabolomik erschloss die entwickelte Methodik wertvolle Informationen über den zukünftigen Ausbruch von Krankheiten. Unter Einsatz einer phänomenweiten Risikomodellierung erwiesen sich elektronische Gesundheitsakten als Quelle prädiktiver Information mit hoher systemischer Relevanz. Bei der Analyse von Fundusfotografien der Netzhaut wurden Krankheiten identifiziert für deren Vorhersage Netzhautinformationen genutzt werden könnten. Zusammengefasst zeigten die Ergebnisse das Potential neuronaler Risikomodelle die medizinische Praxis in Richtung einer datengesteuerten, präventionsorientierten Medizin zu verändern.

Acknowledgements

I would like to express my sincere gratitude to my supervisor Roland Eils, for his support and mentorship throughout my doctoral studies. His trust in my ideas and capabilities was the foundation of this research journey and is what encouraged me to grow over the past years. I will always admire him for his expertise and friendship.

I am also grateful for the continuous guidance and help of Maik Pietzner, Tim Landgraf, Ulf Landmessser, and John Deanfield. Their knowledge, constructive criticism, and enthusiasm were instrumental in this project.

Further, special thanks to my colleagues and friends Jakob Steinfeldt, Benjamin Wild, Julius Upmeier zu Belzen, and Paul Kittner for their help in every step of this journey. Their willingness to discuss ideas, readiness to help and neverending encouragement has been invaluable.

Lastly, I extend my gratitude towards the cluster administrator(s) for their patience and assistance throughout this entire process.

Acknowledgements

Table 1: Third party material utilized in this study.

Figure	Source	Description
figure 5.1 on page 103	https://www.flaticon.com/free-icon/blood-tube_2978935?related_id=2978935	Blood draw icon
figure 1.2 on page 9, figure 3.1 on page 82, figure 5.1 on page 103	https://www.flaticon.com/free-icon/crowd_3382490?term=population&related_id=3382490	Population icon
figure 3.1 on page 82	https://www.flaticon.com/free-icon/location_535239?term=location&page=1&position=1&page=1&position=1&related_id=535239&origin=style	Location icon
figure 3.1 on page 82	https://www.flaticon.com/premium-icon/locations_1916701?related_id=1916701	Location icon 2
figure 2.6 on page 70	https://www.flaticon.com/free-icon/patient_2376100?term=health%20records&page=1&position=1&page=1&position=1&related_id=2376100&origin=style	Patient record icon

Supplementary Figures

- 1 A Neural Survival model integrates polygenic information for cardiovascular risk prediction

Supplementary Figures

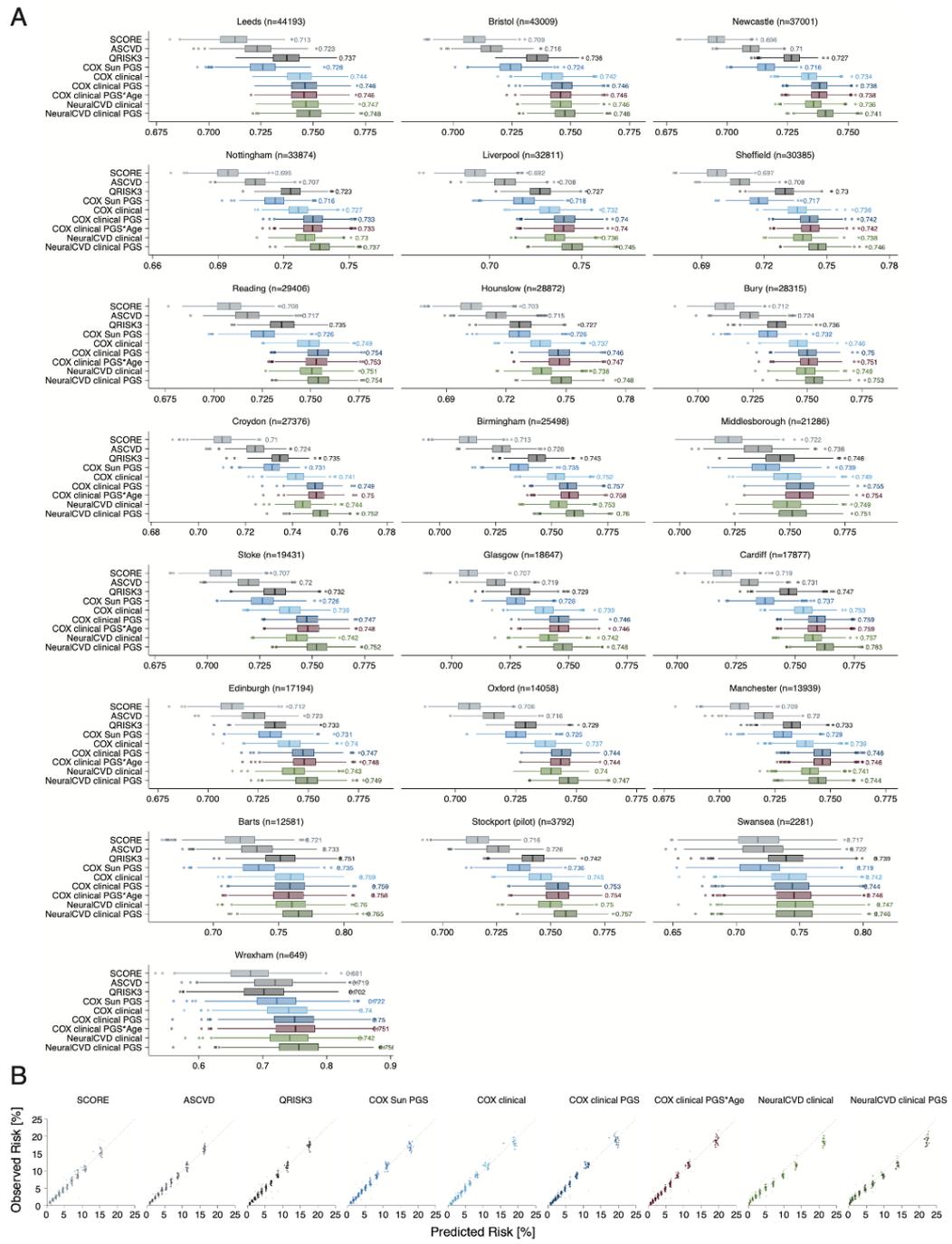


Figure 1: Spatial validation of the NeuralCVD model. a, Displayed is the discriminatory performance of the Neural CVD model. The NeuralCVD score outperforms existing approaches in discrimination at ten years measured by Harrell's C-index (via bootstrap). Displayed is the discriminatory performance of the Neural CVD model in terms of C-index over the 22 assessment centers of the UK Biobank cohort.

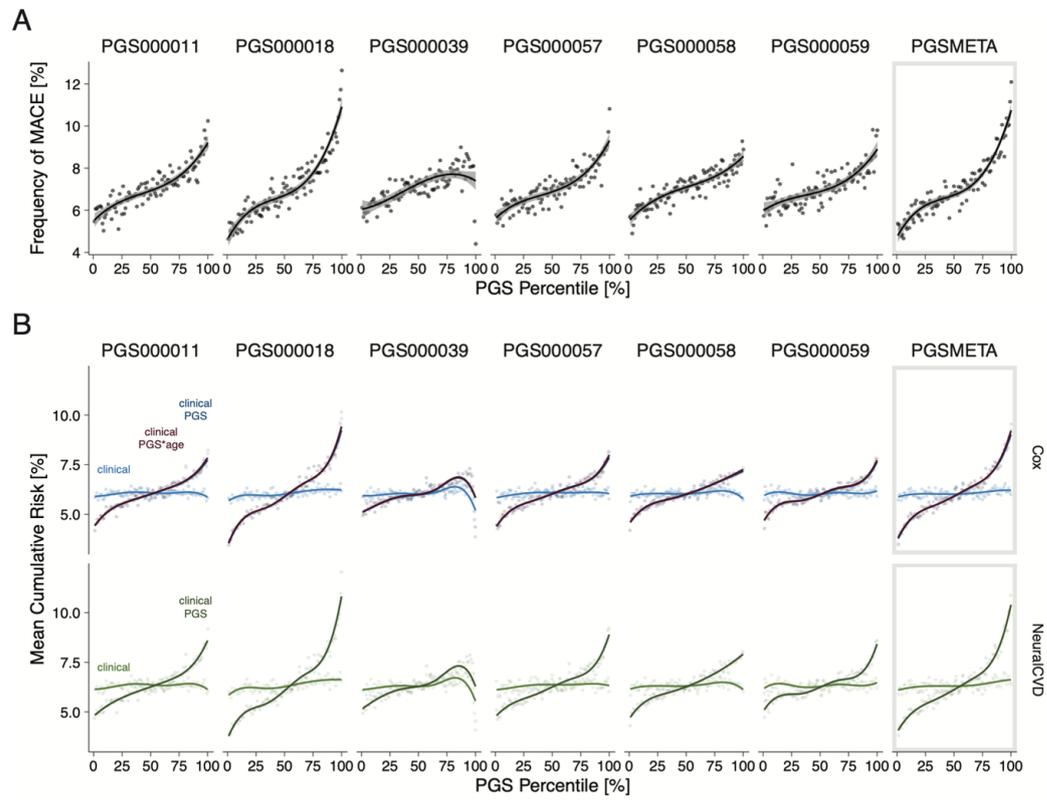


Figure 2: Polygenic risk scores and survival models. **a**, Frequency of Major Adverse Cardiac Events (MACE) over the Polygenic Score (PGS) percentiles for all applied PGS scores, labelled by their PGS catalogue ID. The calculated PGSMETA score summarizes the information from all six scores. **b**, Mean cumulative risk predicted by models trained on solely the clinical predictors (light colour) and models trained on the clinical and PGS information (dark colour) for the Cox and NeuralCVD models, respectively, over the PGS percentiles. All models exploit the information provided by the six PGS variables.

Supplementary Figures

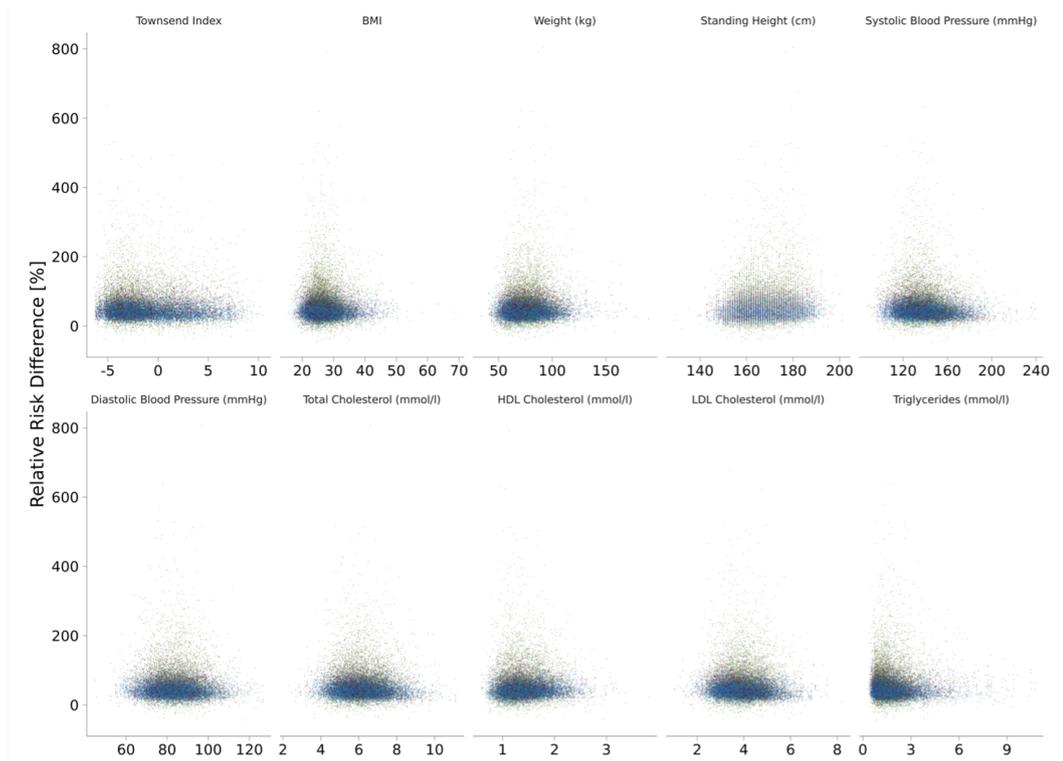


Figure 3: Relative risk differences show no relevant univariate correlation with individual risk factors. Relative risk differences between individual predictions of models trained on clinical variables alone and clinical variables including PGS for the Cox model (linear blue, with interactions red) and the NeuralCVD score (green). Individual clinical factors do not correlate notably with the observed relative risk differences.

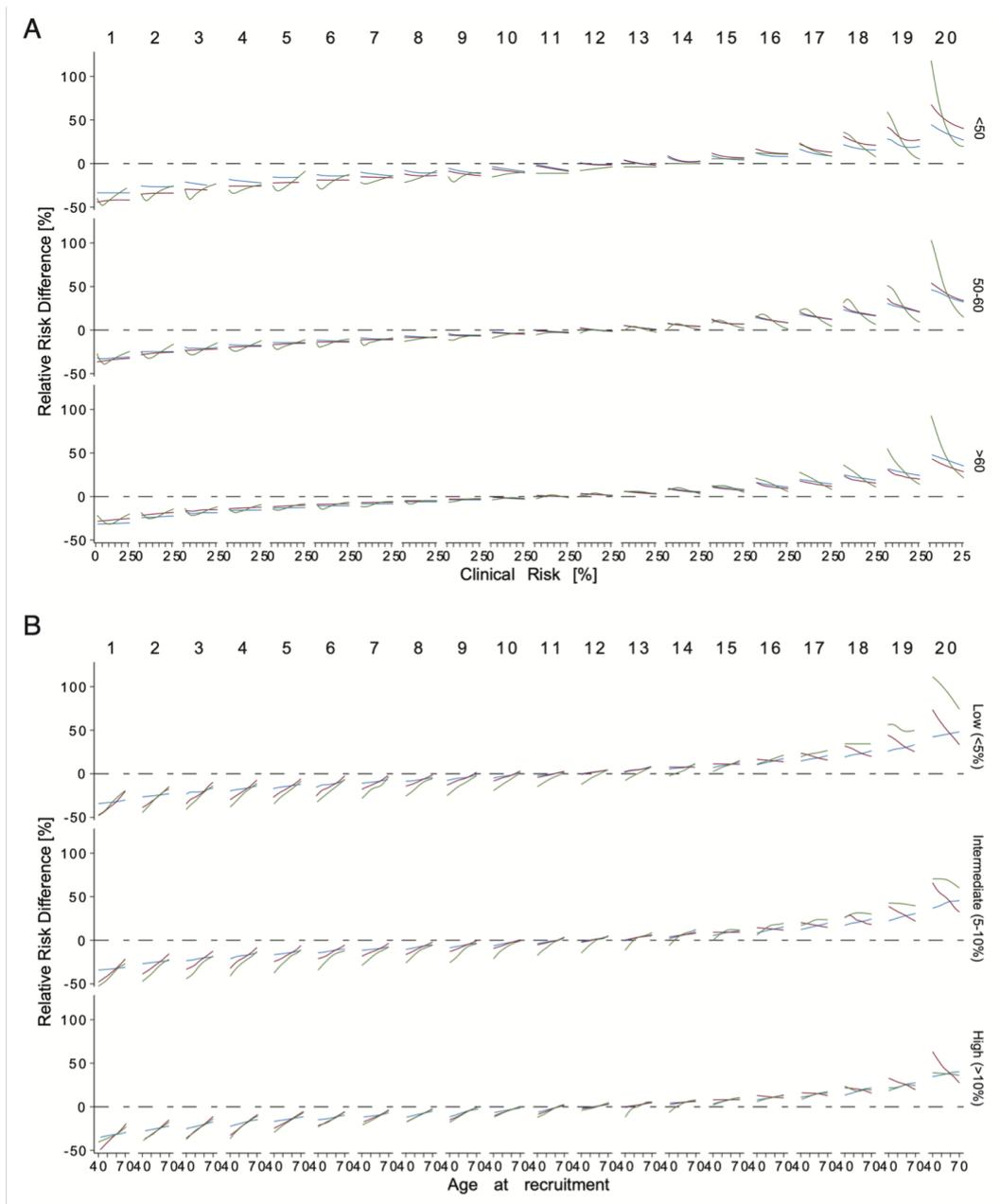


Figure 4: Relative risk differences are influenced by clinical phenotype and age over the entire spectrum of polygenic risk. a, Relative risk differences within the Cox model (linear blue, with interactions red) and the NeuralCVD (green) score upon PGS addition vs clinical risk stratified by age. **b,** Relative risk differences within the Cox model (linear blue, with interactions red) and the NeuralCVD (green) score upon PGS addition vs. age stratified by clinical risk.

Supplementary Figures

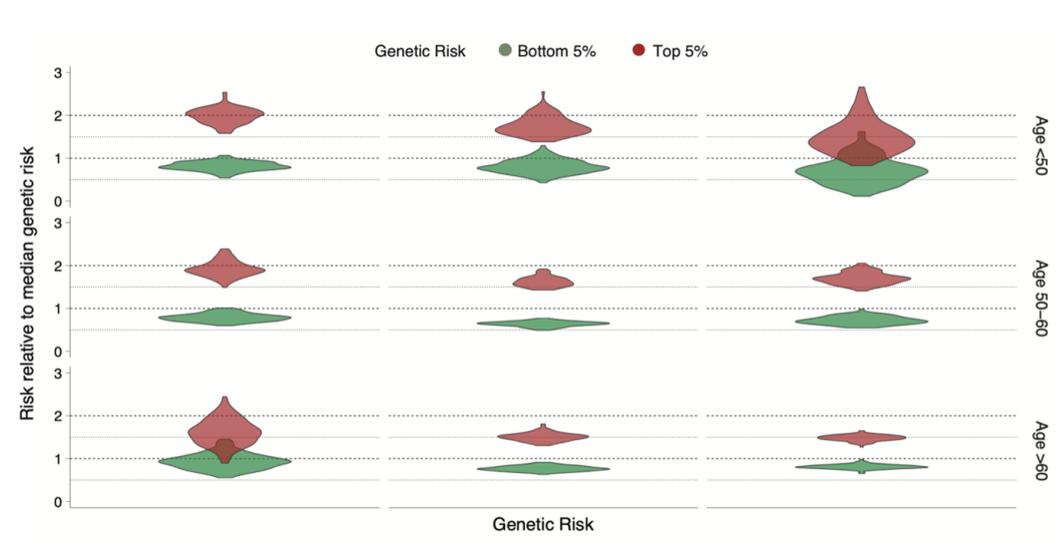


Figure 5: Stratified relative risks of genetic predisposition. Risk (based on event rates) of high and low genetic predisposition relative to median genetic risk stratified by predicted clinical risk (recalibrated QRISK3) and age. High genetic risk individuals face an increase in relative risk at age<50 and otherwise low risk. At age>60 and already high risk, there is a decrease in relative risk.

2 Metabolomic profiles predict individual multi-disease outcomes

Supplementary Figures

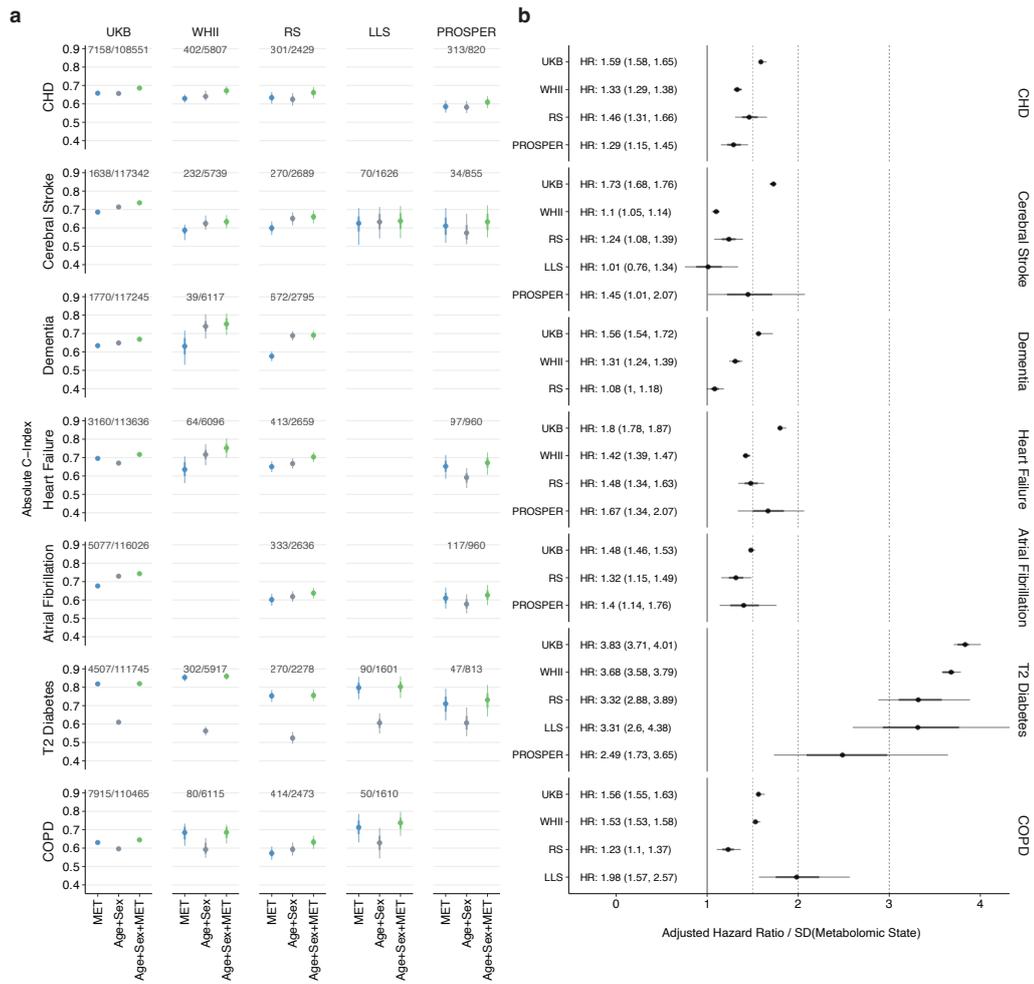


Figure 6: External validation in four independent cohorts. **a**, Displayed are discriminative performances quantified by the C-index for the UK Biobank and the four external validation cohorts, Whitehall II (WHII), Rotterdam Study (RS), Leiden Longevity Study (LLS) and the PROSPER trial (PROSPER). CPH models were trained on the metabolomic state model (MET) as fitted on UK Biobank and applied to each cohort, as well as on Age+Sex and Age+Sex+MET. The metabolomic state is predictive in the replication cohorts for all assessed endpoints. Dots indicate the median performance, while whiskers indicate the 95 % Confidence Interval (CI) determined by bootstrapping over 1000 iterations. **b**, Age+Sex adjusted Hazard Ratio (HR) for the metabolomic state in all five cohorts. A unit standard deviation increase in the metabolomic state corresponds to a HR increase in predicted risk. 95 % confidence interval is indicated and estimated by bootstrapping over 1000 iterations.

2 Metabolomic profiles predict individual multi-disease outcomes

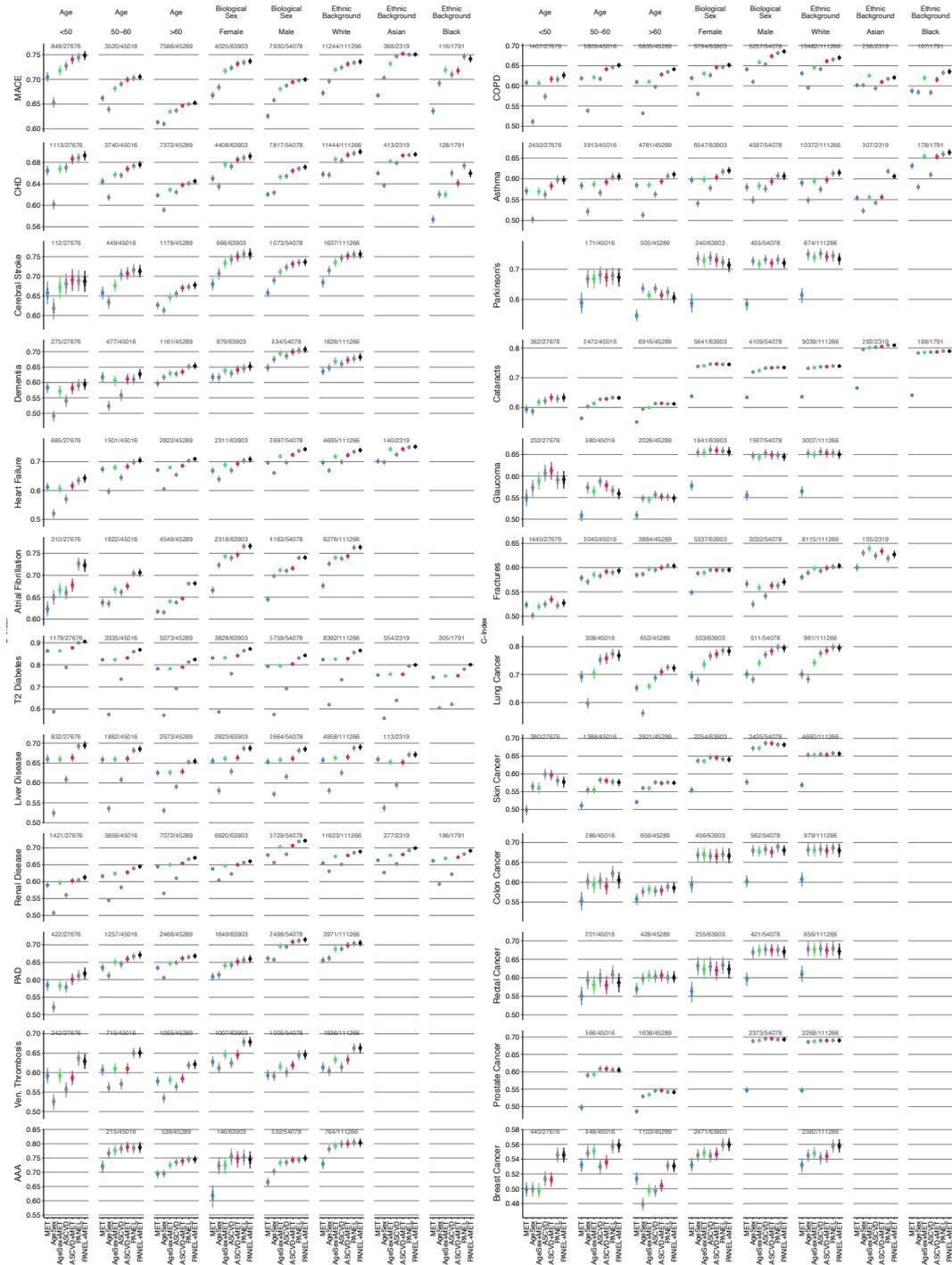


Figure 7: The discriminative performance is largely comparable over multiple subgroups. Displayed are the discriminative performances of the metabolomic state, the three clinical predictor sets and their combinations stratified by endpoint, age at recruitment, biological sex, and self-reported ethnic background. The number of events and eligible individuals is indicated at the top of each panel. As the concordance index is only reliable if a sufficient number of events are recorded, subgroups with less than 100 events were excluded. All performances are reported as medians (dots) and whiskers extend to the 95 % confidence intervals as estimated by bootstrapping over 1000 iterations.

Supplementary Figures

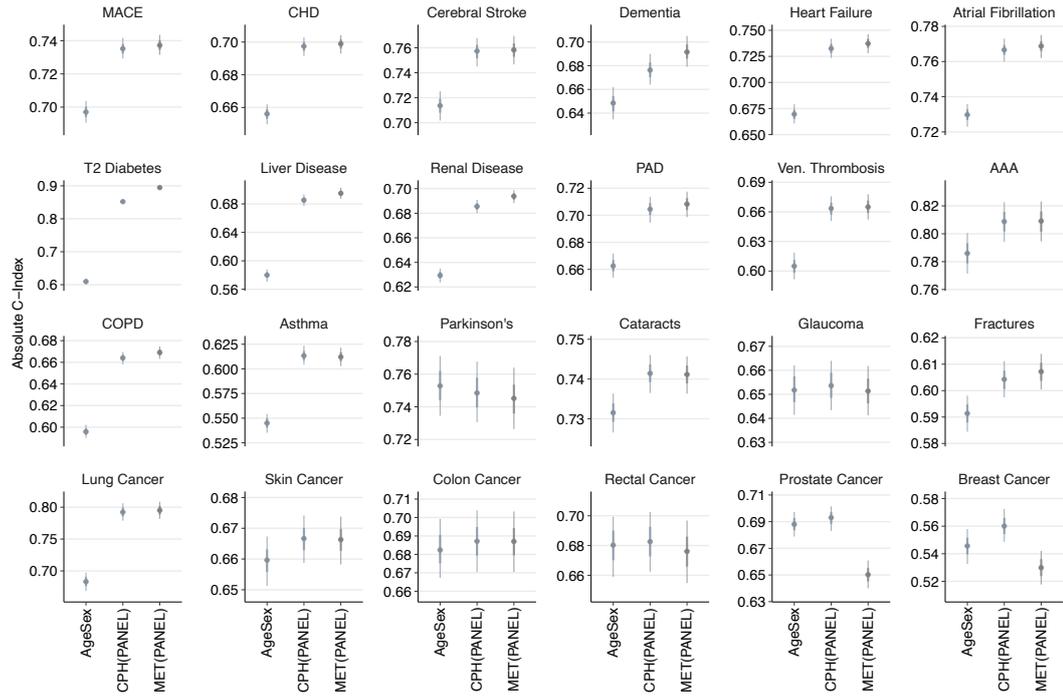


Figure 8: Comparison of predictive performance of the PANEL predictors in a Cox Proportional Hazards (CPH) model and the neural network. Comparison of discriminative performances of the Cox Proportional Hazards (CPH) models and Metabolomic State Model (MSM) trained on the PANEL covariates. The discriminative performance of the PANEL predictors is either similar or can be further improved by modeling with the same architecture as the metabolomic state model for most (non-cancer) endpoints. All performances are reported as medians (dots) and whiskers extend to the 95 % confidence intervals as estimated by bootstrapping over 1000 iterations.

2 Metabolomic profiles predict individual multi-disease outcomes

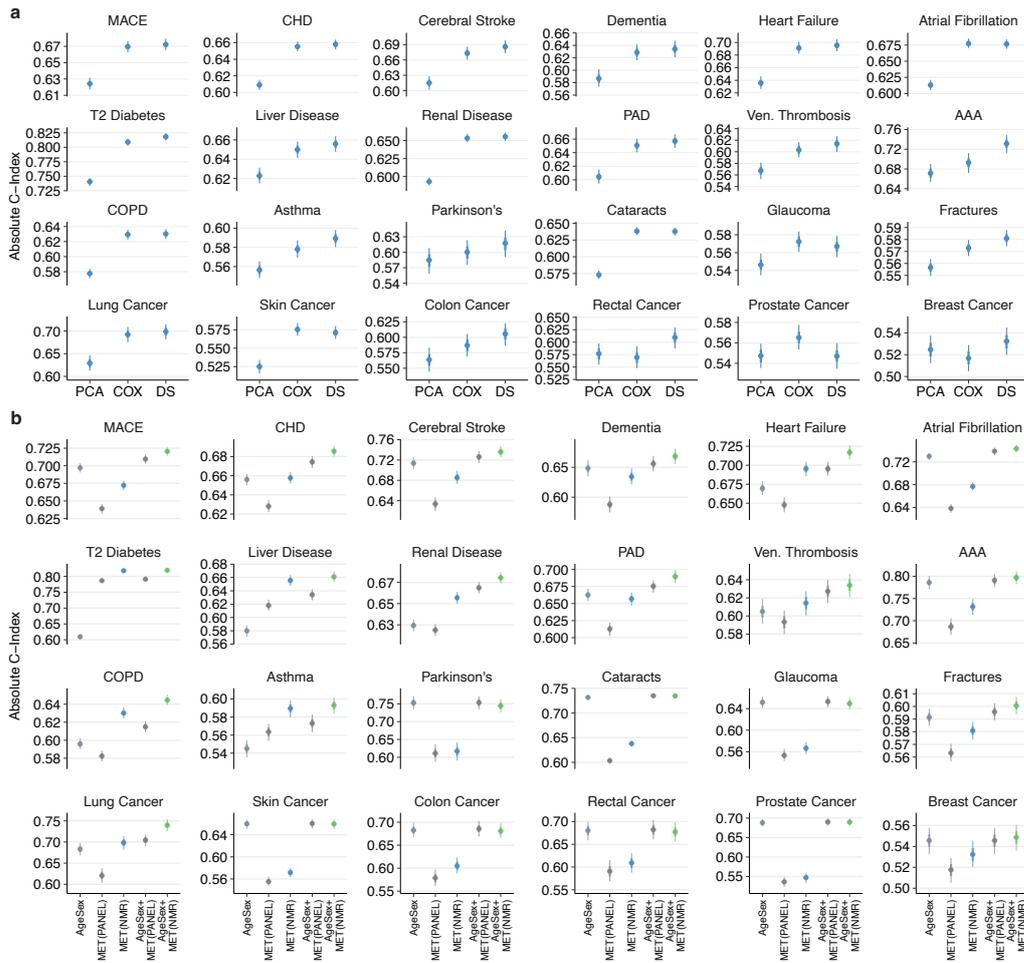
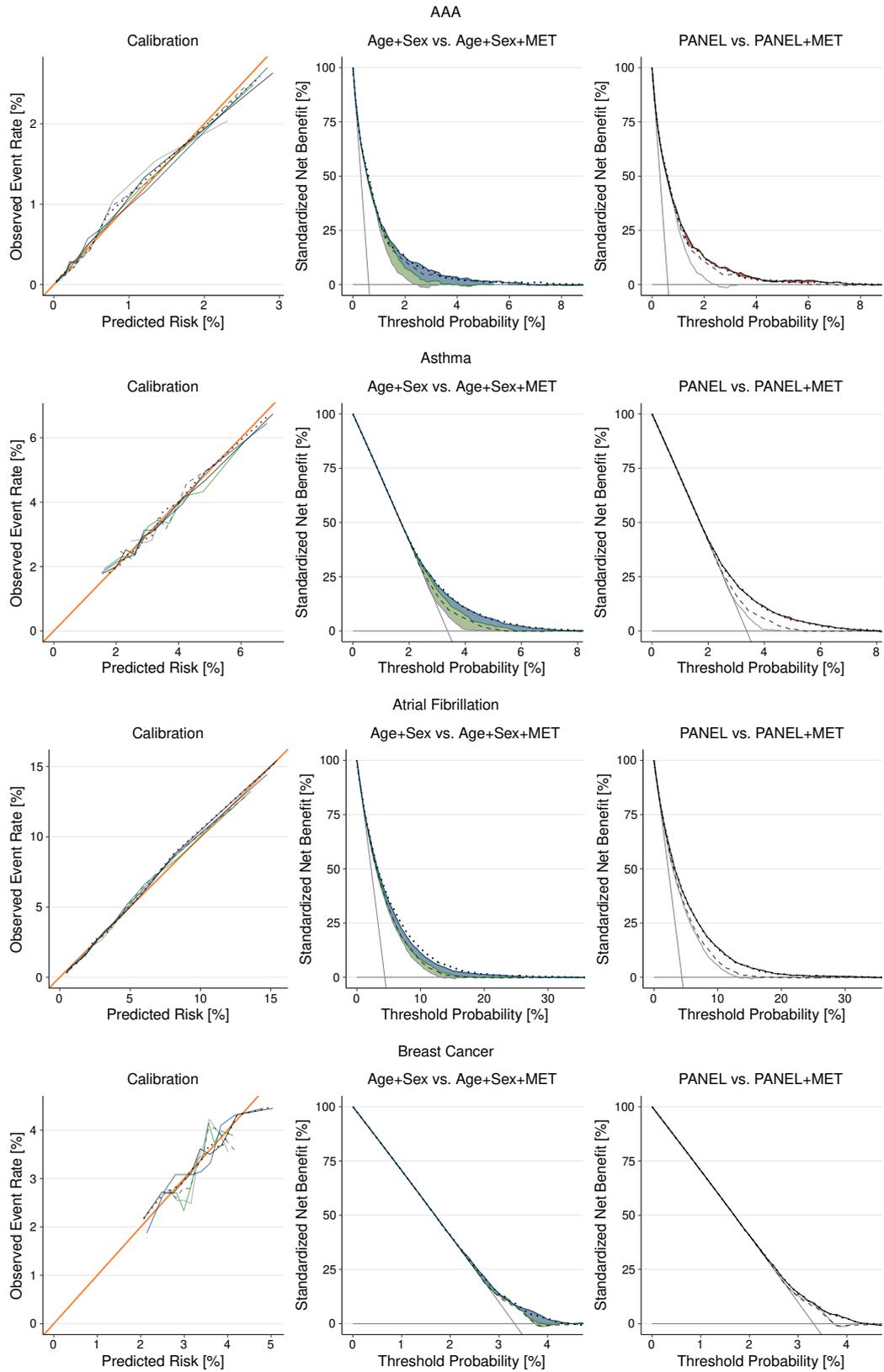
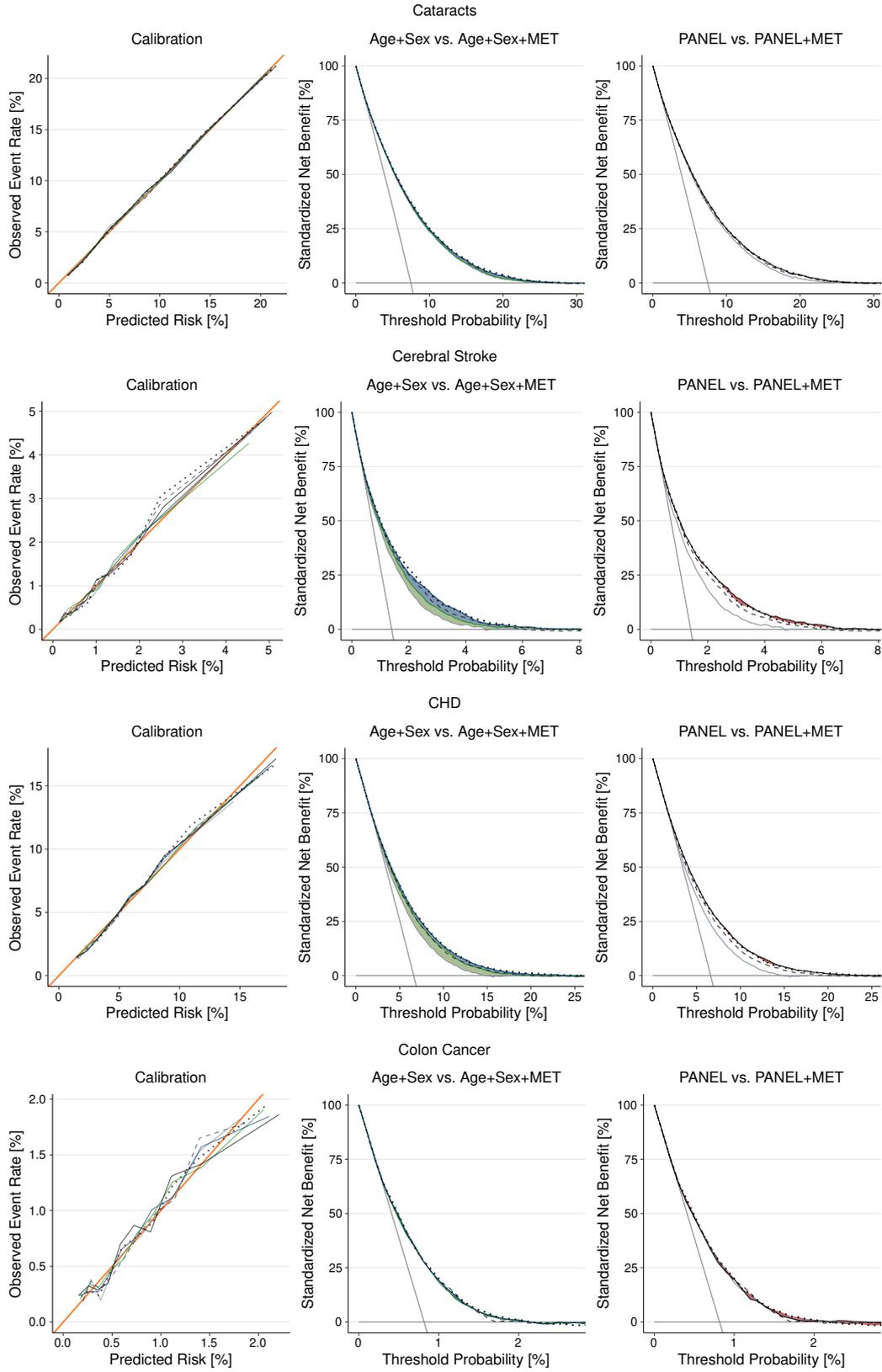


Figure 9: The metabolomic state model outperforms linear baselines on NMR-derived metabolite profiles and NMR-derived metabolite profiles are more predictive than PANEL metabolites. **a**, Displayed are C-indices for the Cox Proportional Hazards (CPH) models trained on the metabolomic state (MET), the 168 metabolites (CPH) as well as on the first 10 components of a PCA-reduction of the 168 metabolites (PCA) for each of the 24 investigated endpoints. The metabolomic state performs comparably or better than both the CPH and PCA models for all endpoints, except prostate cancer. **b**, Displayed are C-indices for Cox Proportional Hazards (CPH) models trained on Age+Sex (Age+Sex), the metabolomic states derived from NMR metabolomics (MET(NMR)), the metabolomic states derived from the PANEL metabolites (MET(PANEL)) and combinations of Age+Sex and the metabolomic states respectively. NMR metabolomic profiles provide predictive information comparable or superior to the PANEL metabolites for all investigated endpoints, also reflected in the predictive performance over the Age+Sex covariates. The MET(PANEL) set included albumin, cholesterol, hdl-cholesterol, ldl direct, triglycerides, glucose and creatinine. All performances are reported as medians (dots) and whiskers extend to the 95 % confidence intervals as estimated by bootstrapping over 1000 iterations. Major Adverse Cardiac Events (MACE), Coronary Heart Disease (CHD), Peripheral Artery Disease (PAD), Abdominal Aortic Aneurysm (AAA), Chronic Obstructive Pulmonary Disease (COPD).

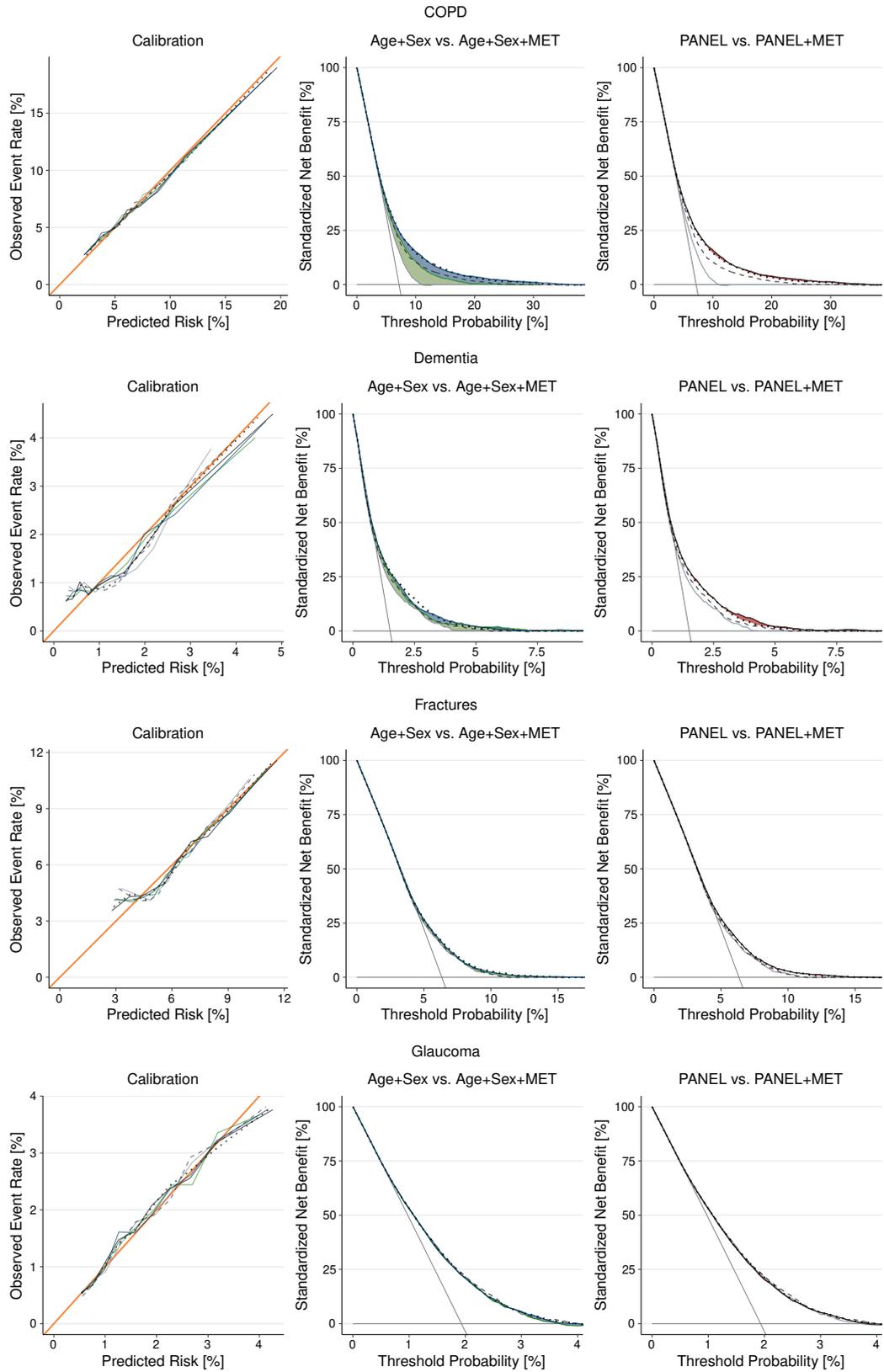
Supplementary Figures



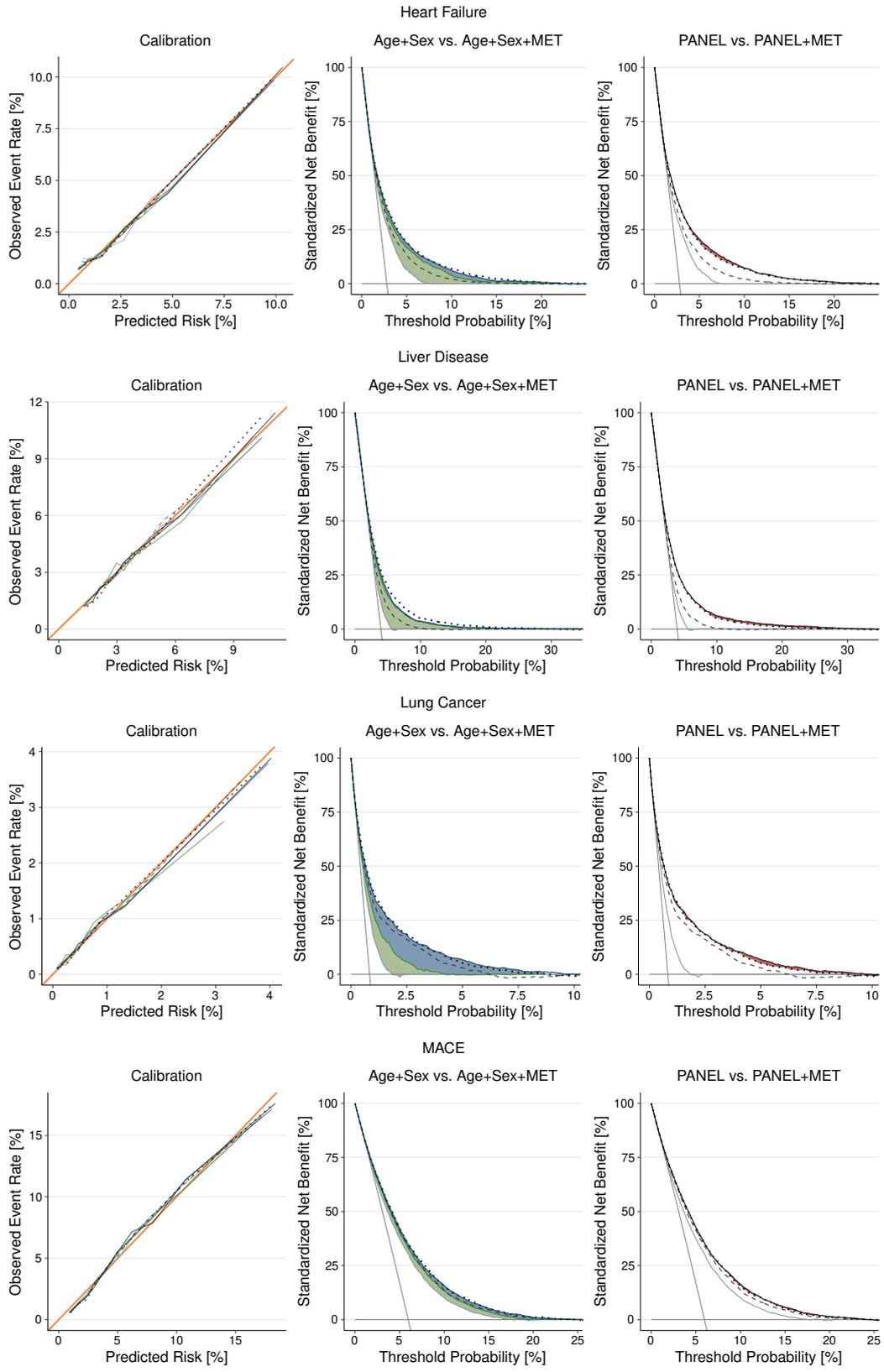
2 Metabolomic profiles predict individual multi-disease outcomes



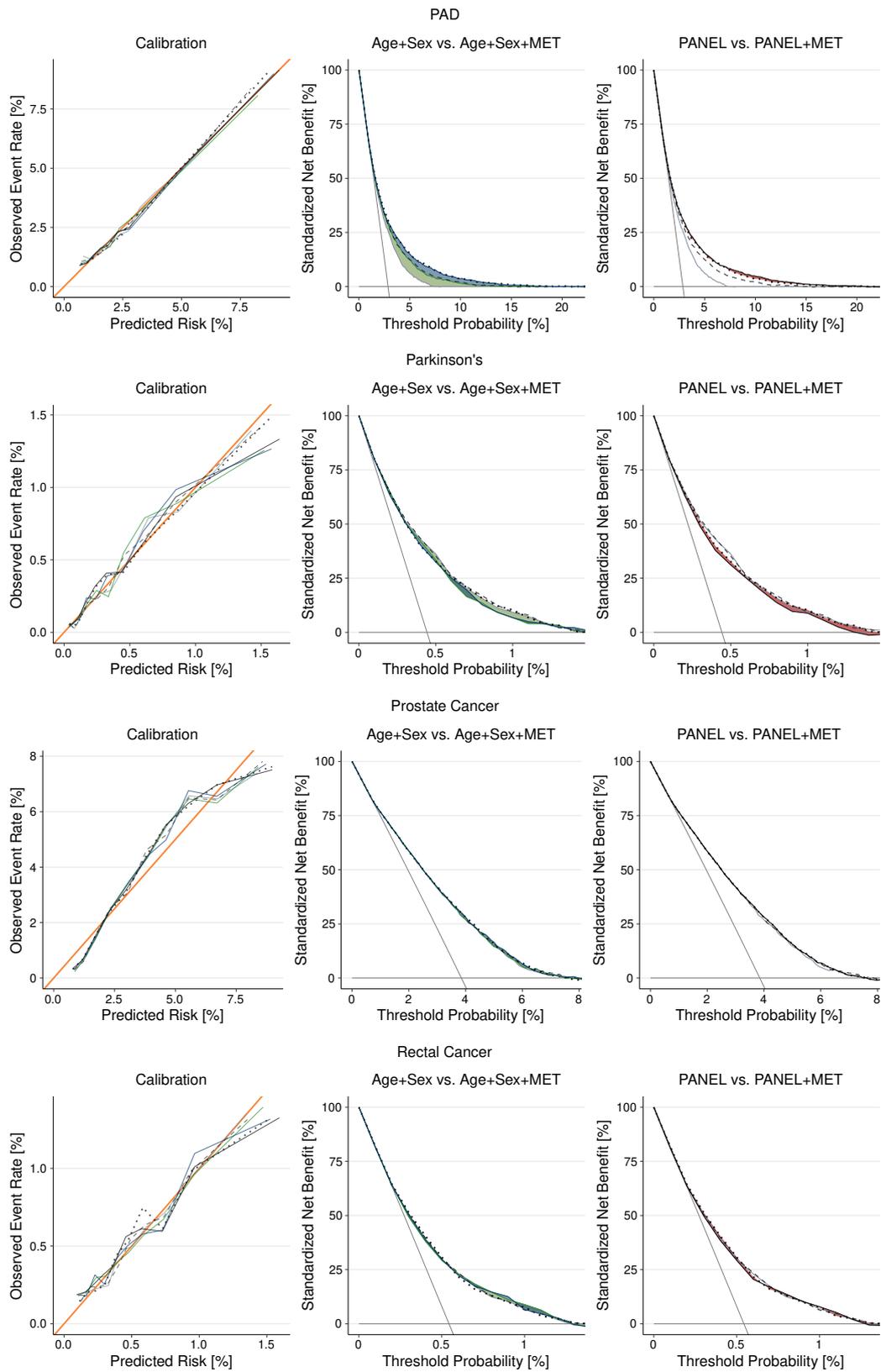
Supplementary Figures



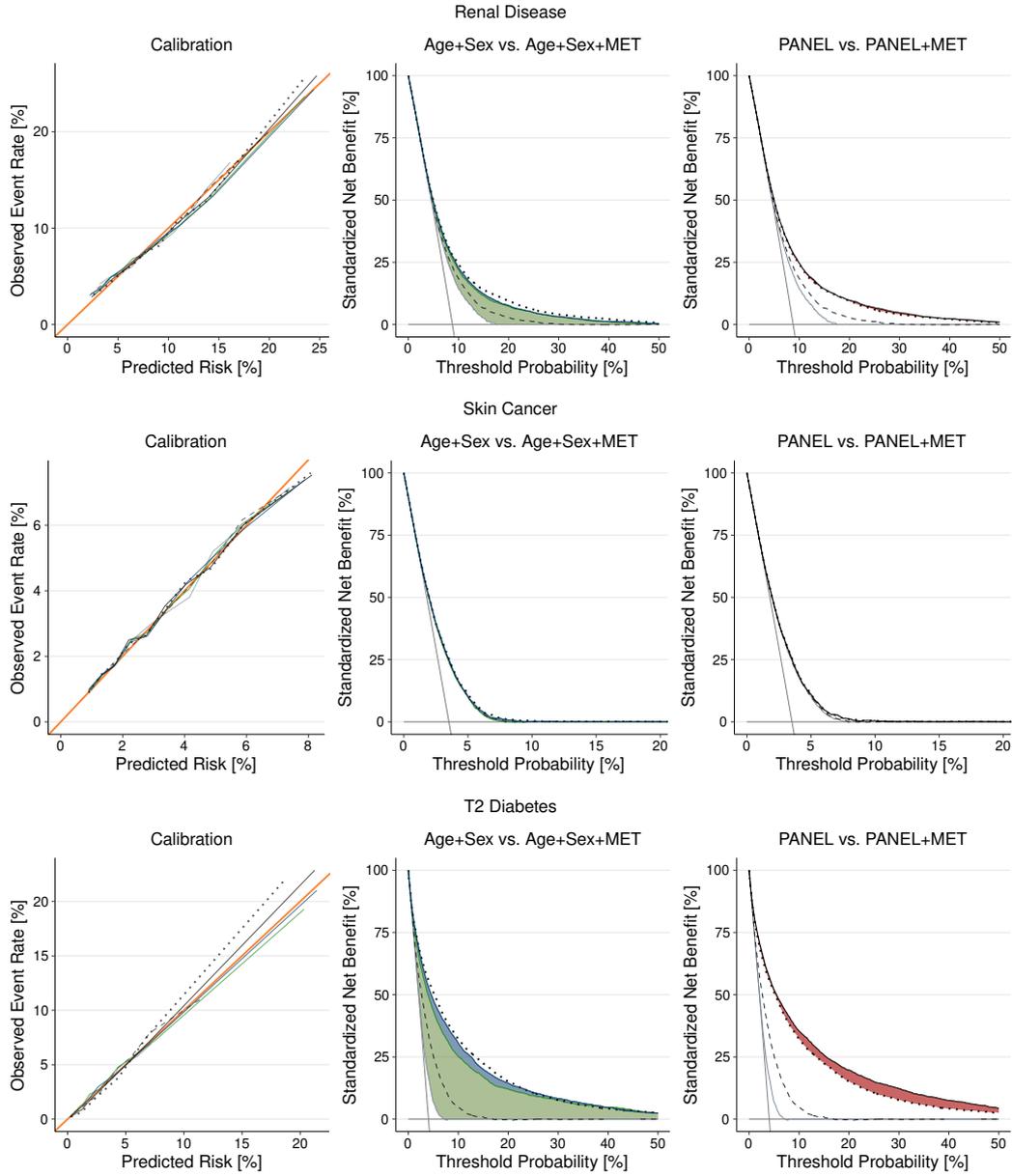
2 Metabolomic profiles predict individual multi-disease outcomes



Supplementary Figures



2 Metabolomic profiles predict individual multi-disease outcomes



Supplementary Figures

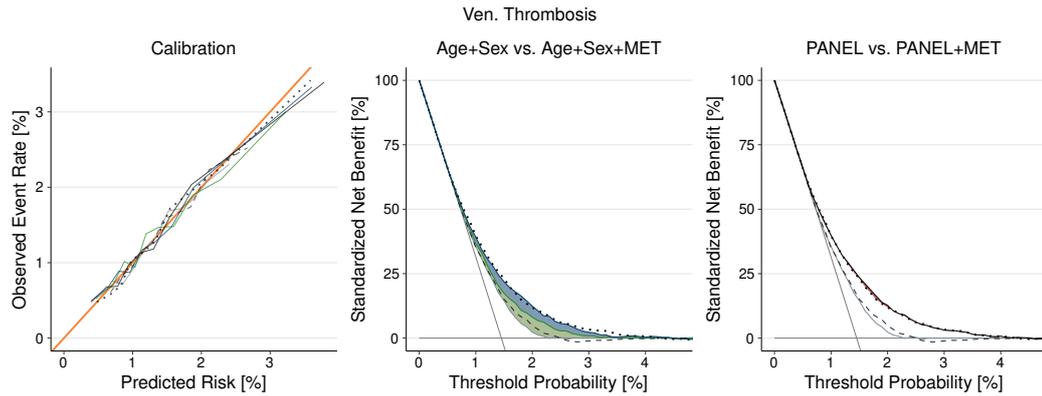


Figure 10: Model calibration and clinical utility for all endpoints. Displayed are calibration curves and decision curves analyzing clinical utility for all endpoints. In each panel, the left figure displays the calibration curve calibration curve for Cox Proportional Hazards (CPH) models including the baseline parameter sets AgeSex, ASCVD and PANEL as well as their extensions with the metabolomic state (i.e. AgeSex+MET). All models are well calibrated. The center figure displays endpoint specific net benefit curves standardized by the endpoint prevalence, where the horizontal solid gray line indicates “treat none” and the vertical solid gray line indicates “treat all”. The standardized net benefit of AgeSex (gray line), ASCVD (dashed dark gray line) and the PANEL set (dotted black line), are compared to AgeSex+MET (green line) and additional non-laboratory predictors of the PANEL (PANELnoLaboratory, blue line). The green alpha indicates the added benefit of the metabolomic state addition over AgeSex, while the blue alpha indicates added performance over PANELnoLaboratory, respectively. The right figure displays the standardized net benefit curves comparing the performance of PANEL+MET (solid black line) against all baselines: AgeSex (gray line), ASCVD (dashed dark gray line) and the PANEL set (dotted black line). Major Adverse Cardiac Events (MACE), Coronary Heart Disease (CHD), Peripheral Artery Disease (PAD), Abdominal Aortic Aneurysm (AAA), Chronic Obstructive Pulmonary Disease (COPD).

2 Metabolomic profiles predict individual multi-disease outcomes

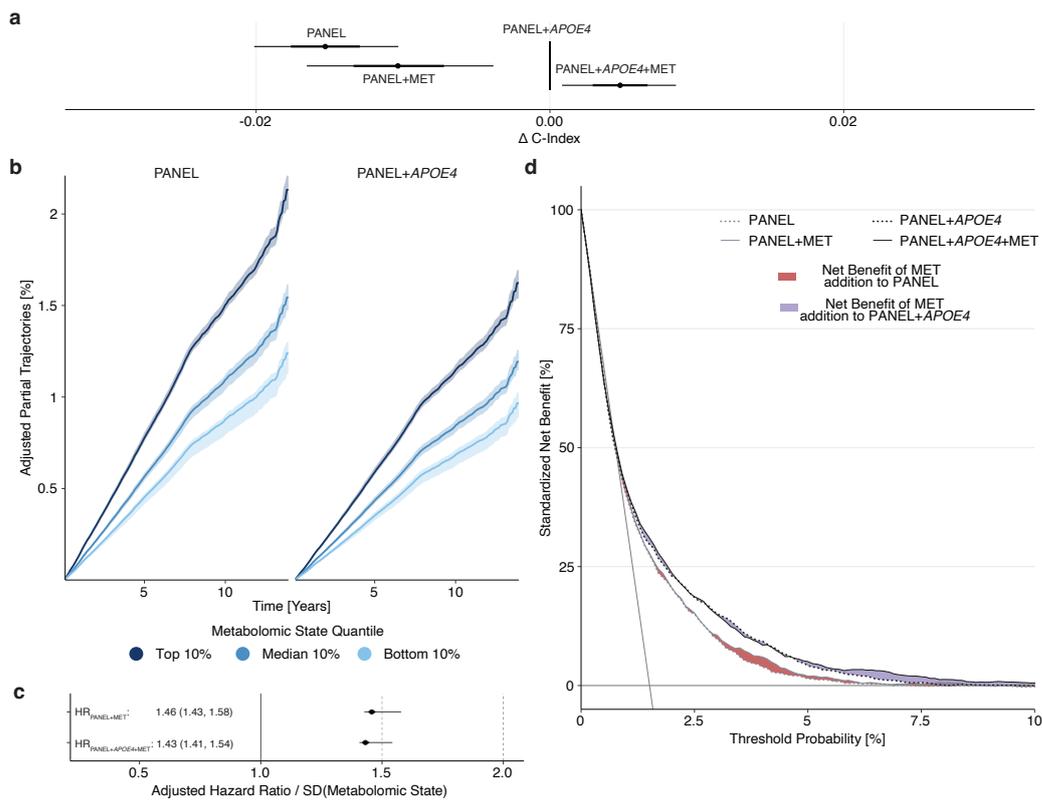


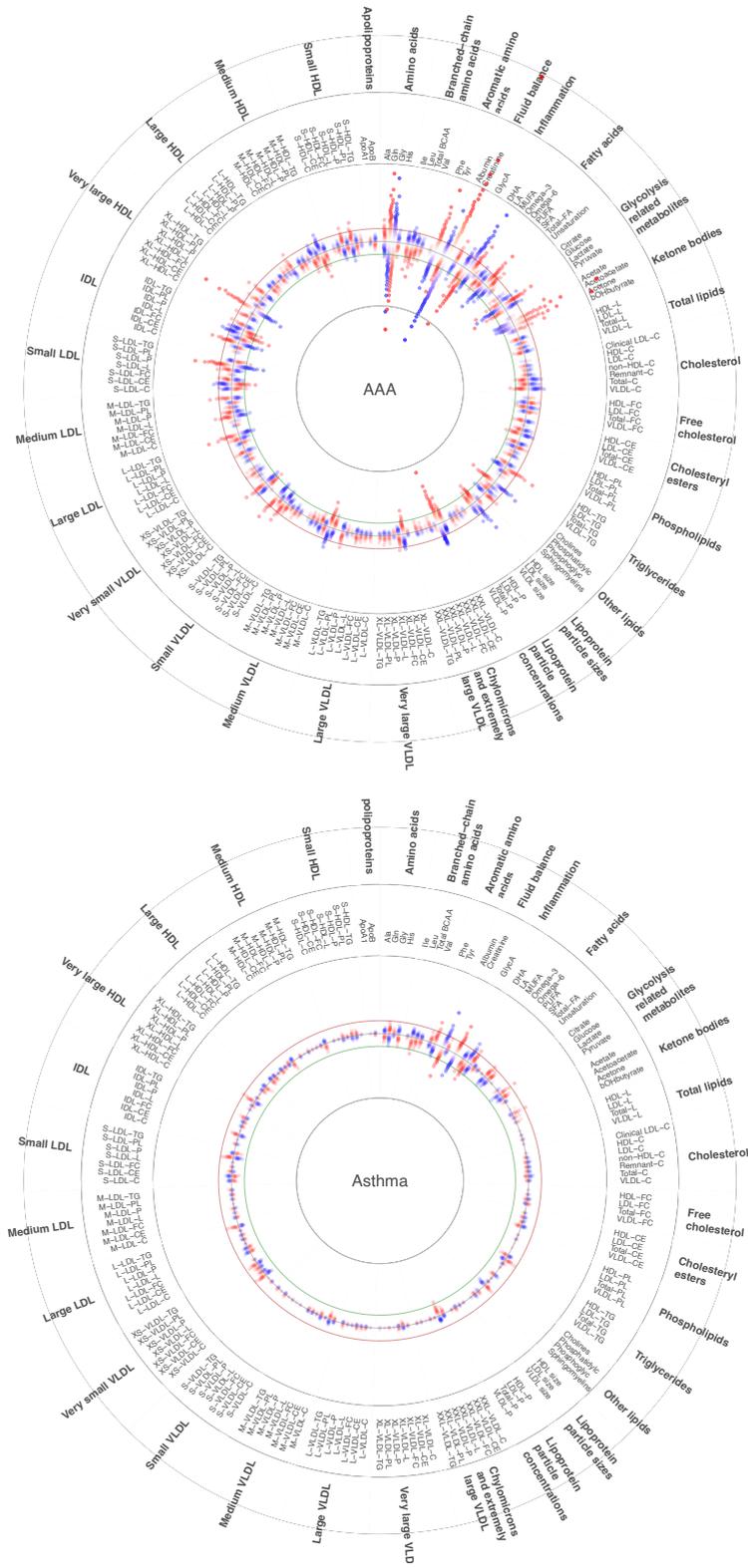
Figure 11: The metabolomic state contains independent predictive information over the *APOE4* carrier status for all-cause dementia. **a**, Displayed are C-index deltas between a CPH model trained on the PANEL+*APOE4* predictor set, its metabolomic state addition (PANEL+*APOE4*+MET) as well as between a CPH model trained on the PANEL set and its respective metabolomic state addition (PANEL+MET). The metabolomic state adds predictive information over the PANEL+*APOE4*. **b**, Displayed are partial trajectories for MET deciles (Top, Median, Bottom 10 %) adjusted for the PANEL covariates and the PANEL+*APOE4* covariates, respectively. **c**, Hazard Ratio (HR) for the metabolomic state adjusted for the predictors of the PANEL and PANEL+*APOE4*. **d**, Decision curve analysis for PANEL/PANEL+MET and PANEL+*APOE4*/PANEL+*APOE4*+MET. The areas in between the solid and dotted lines indicate added net benefits resulting from metabolomic state addition to PANEL (gray lines, red area) and PANEL+*APOE4* (black lines, violett area), respectively. The addition of the metabolomic state to the PANEL predictors improves population net benefit between the 2-8 % decision threshold. In case of PANEL+*APOE4*, the metabolomic state addition improves utility at thresholds between 5-10 %. C-index deltas and HRs are reported as medians (dots) and whiskers extend to the 95 % confidence intervals as estimated by bootstrapping over 1000 iterations.

Supplementary Figures

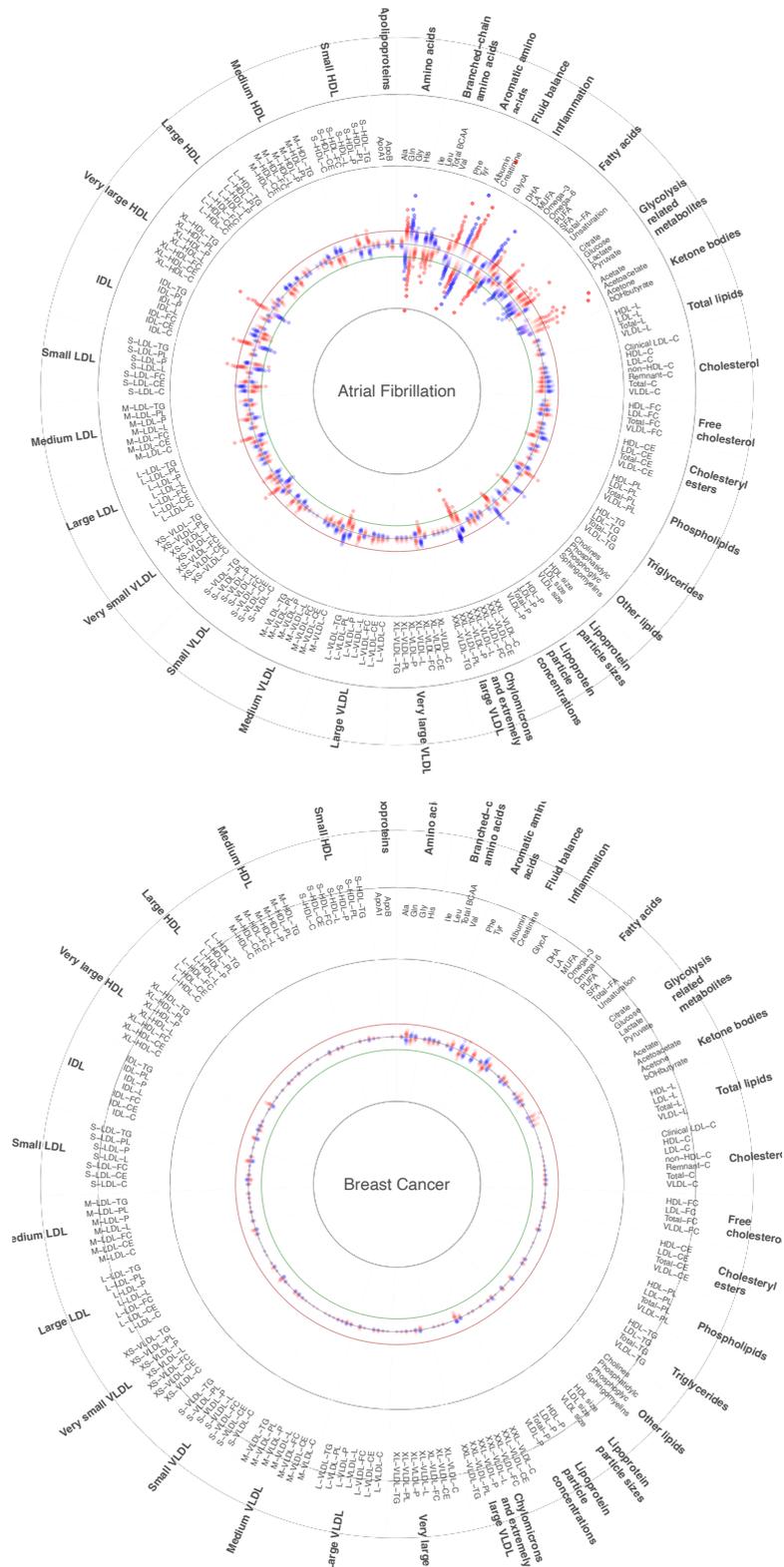


Figure 12: Global metabolite importances for each metabolite and endpoint. Displayed is a heatmap of the metabolite importances, represented by absolute global SHAP value estimates per endpoint for all measured 168 circulating metabolites. The endpoints are sorted by the discriminative performance of the metabolomic state (left to right, see figure 8 on page 210). Major Adverse Cardiac Events (MACE), Coronary Heart Disease (CHD), Peripheral Artery Disease (PAD), Abdominal Aortic Aneurysm (AAA), Chronic Obstructive Pulmonary Disease (COPD).

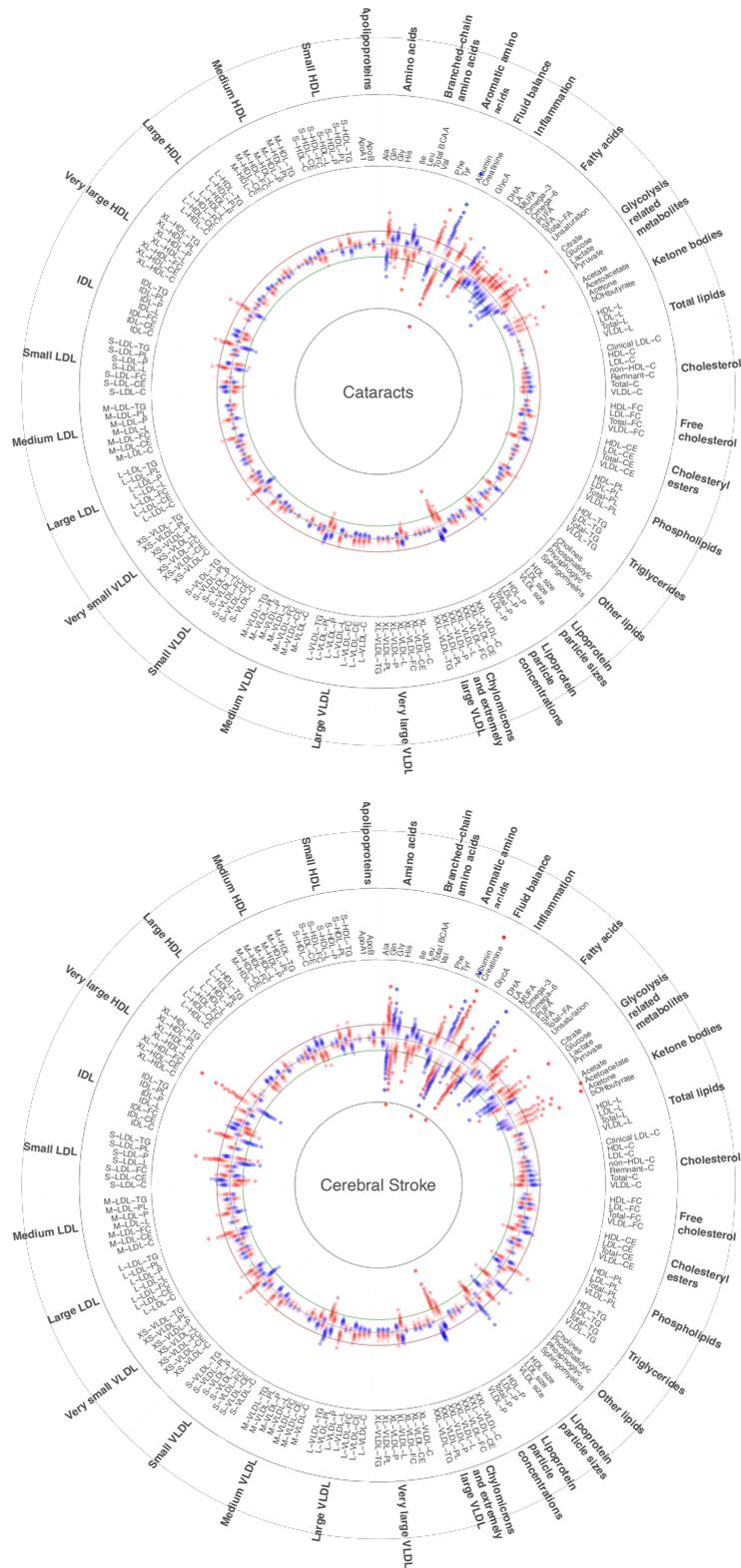
2 Metabolomic profiles predict individual multi-disease outcomes



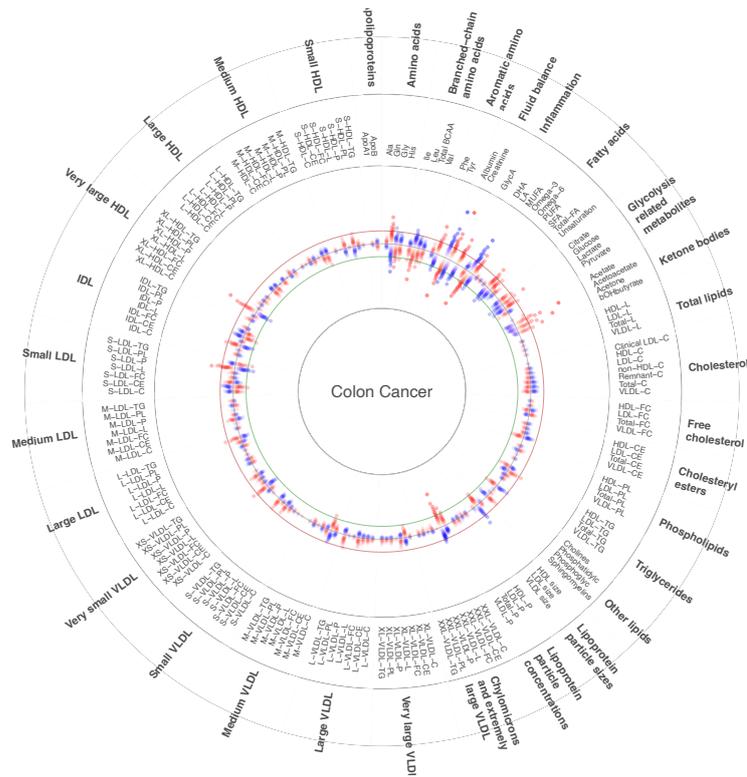
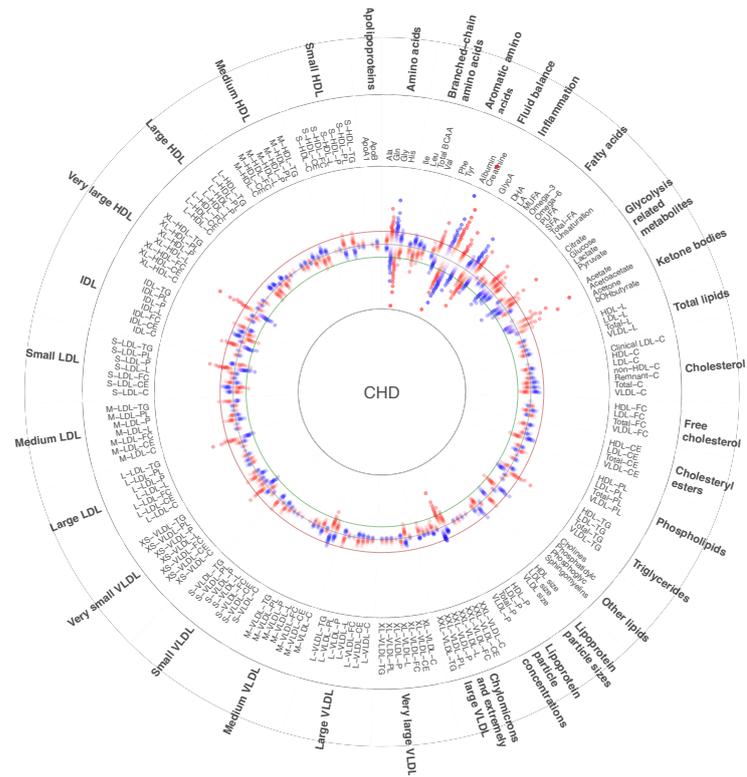
Supplementary Figures



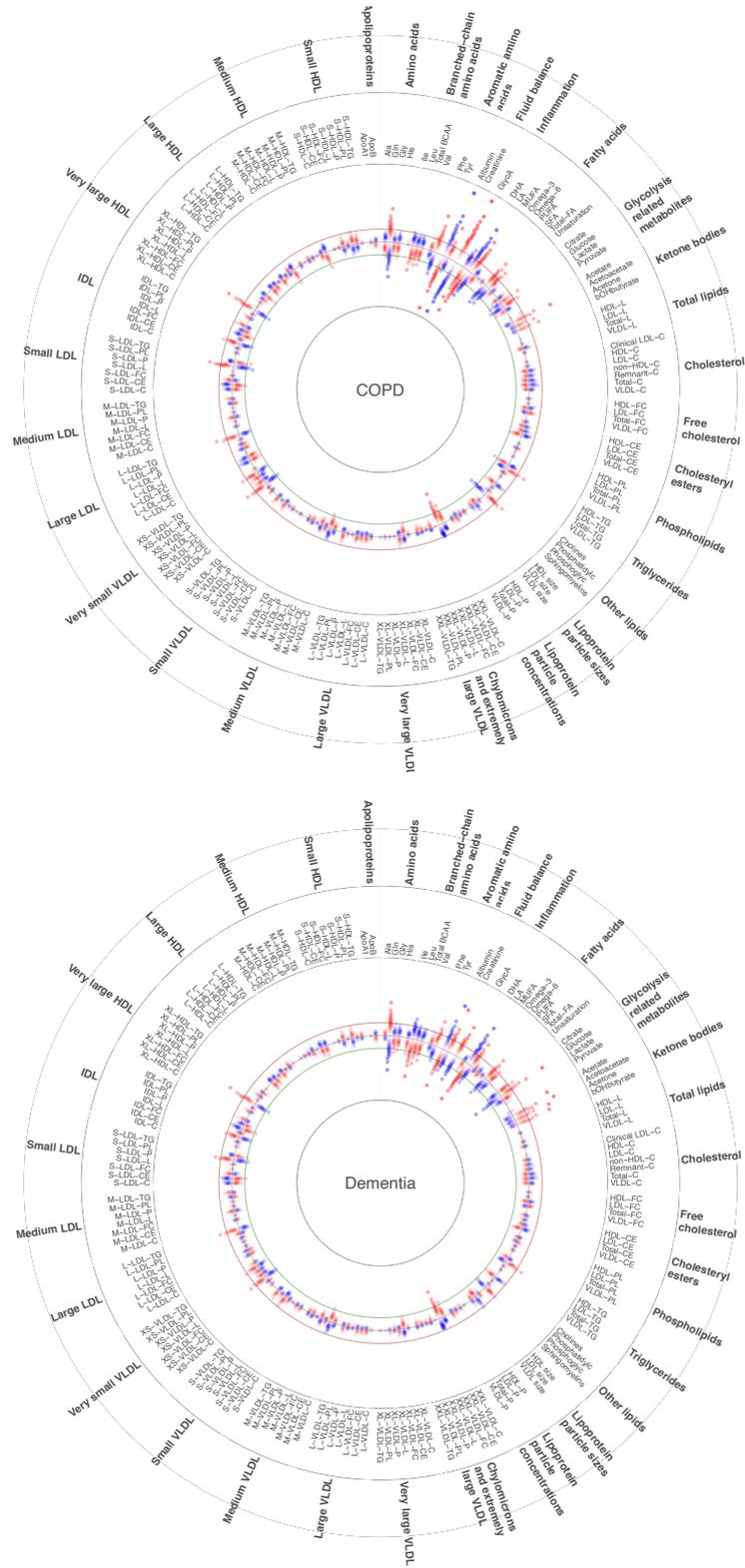
2 Metabolomic profiles predict individual multi-disease outcomes



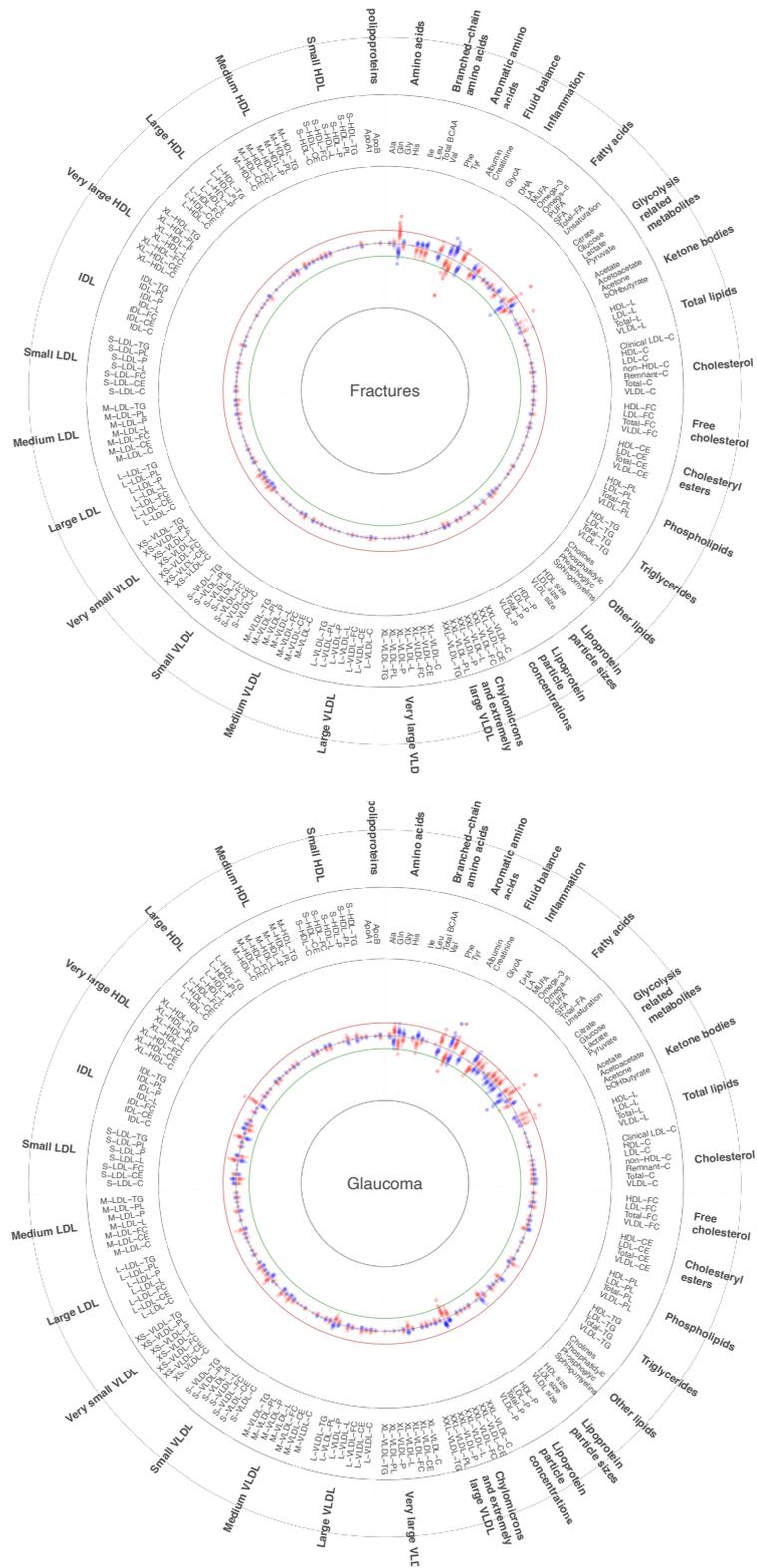
Supplementary Figures



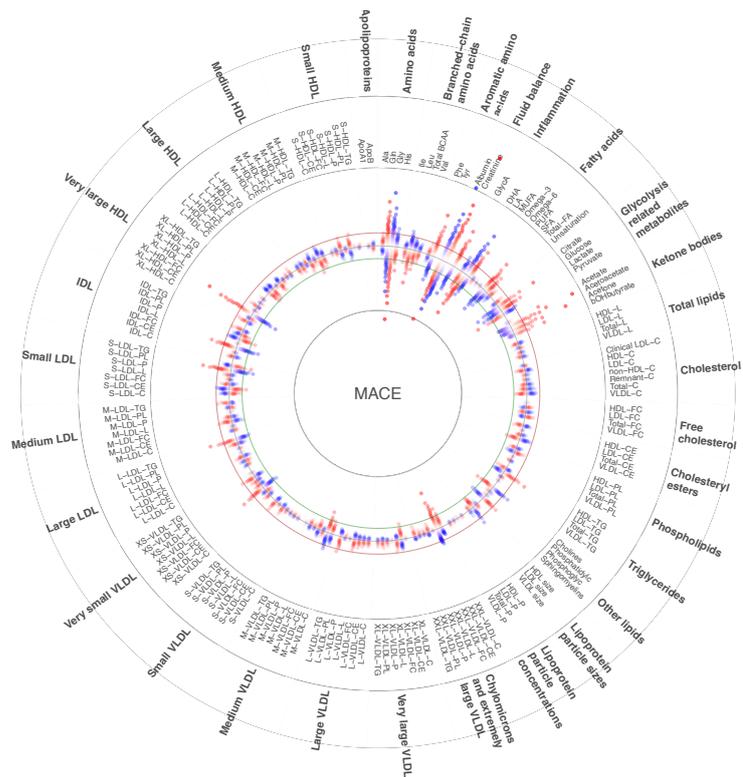
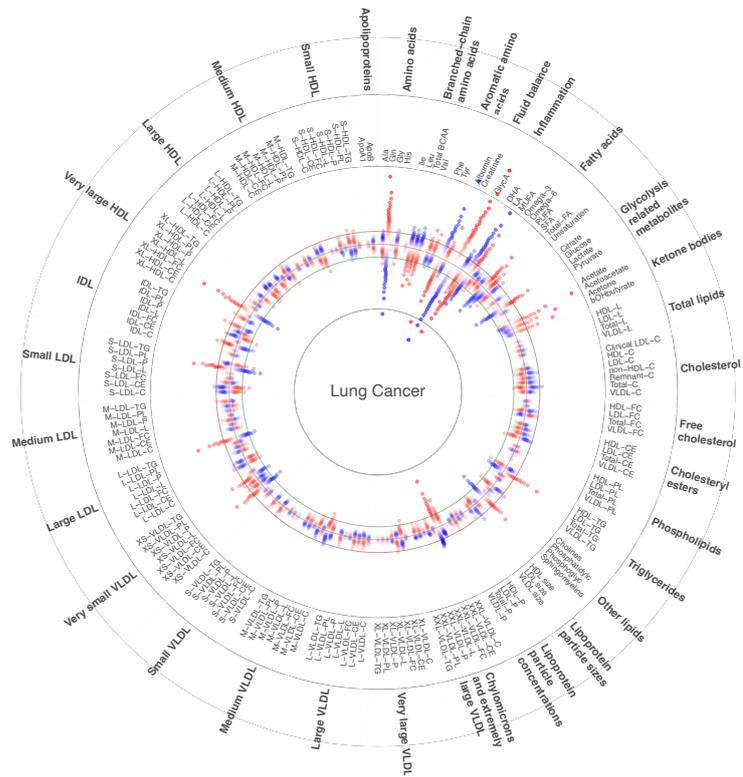
2 Metabolomic profiles predict individual multi-disease outcomes



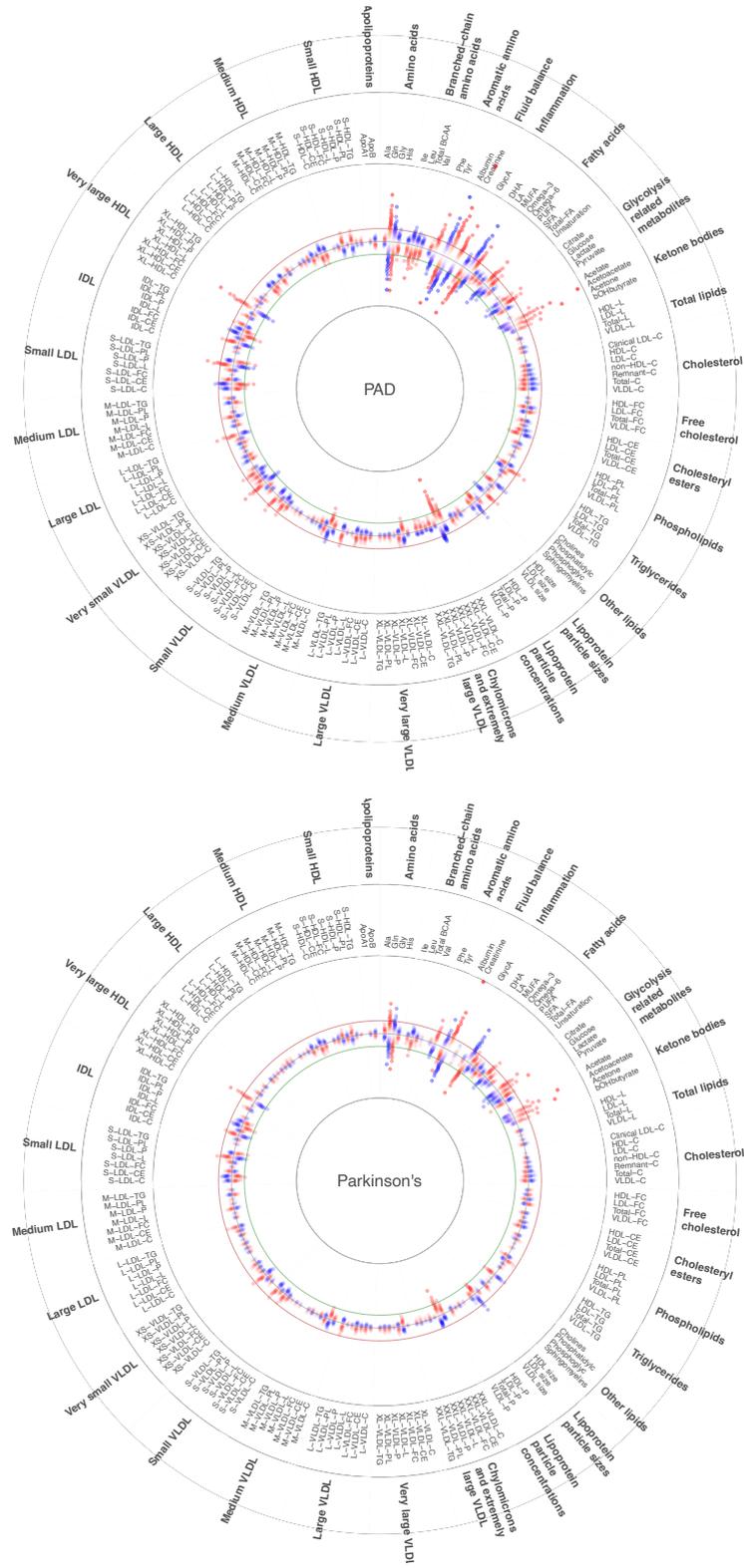
Supplementary Figures



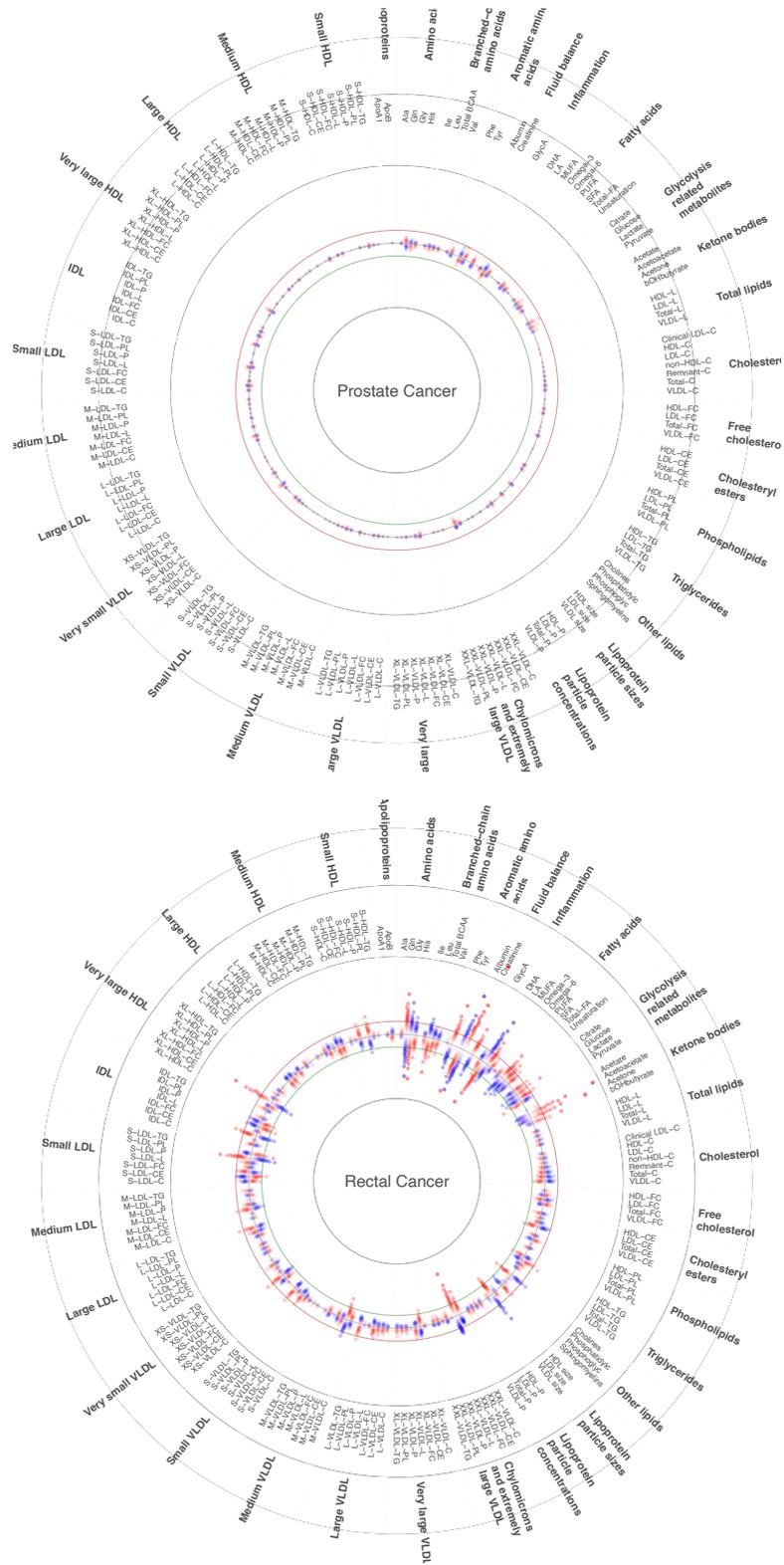
Supplementary Figures



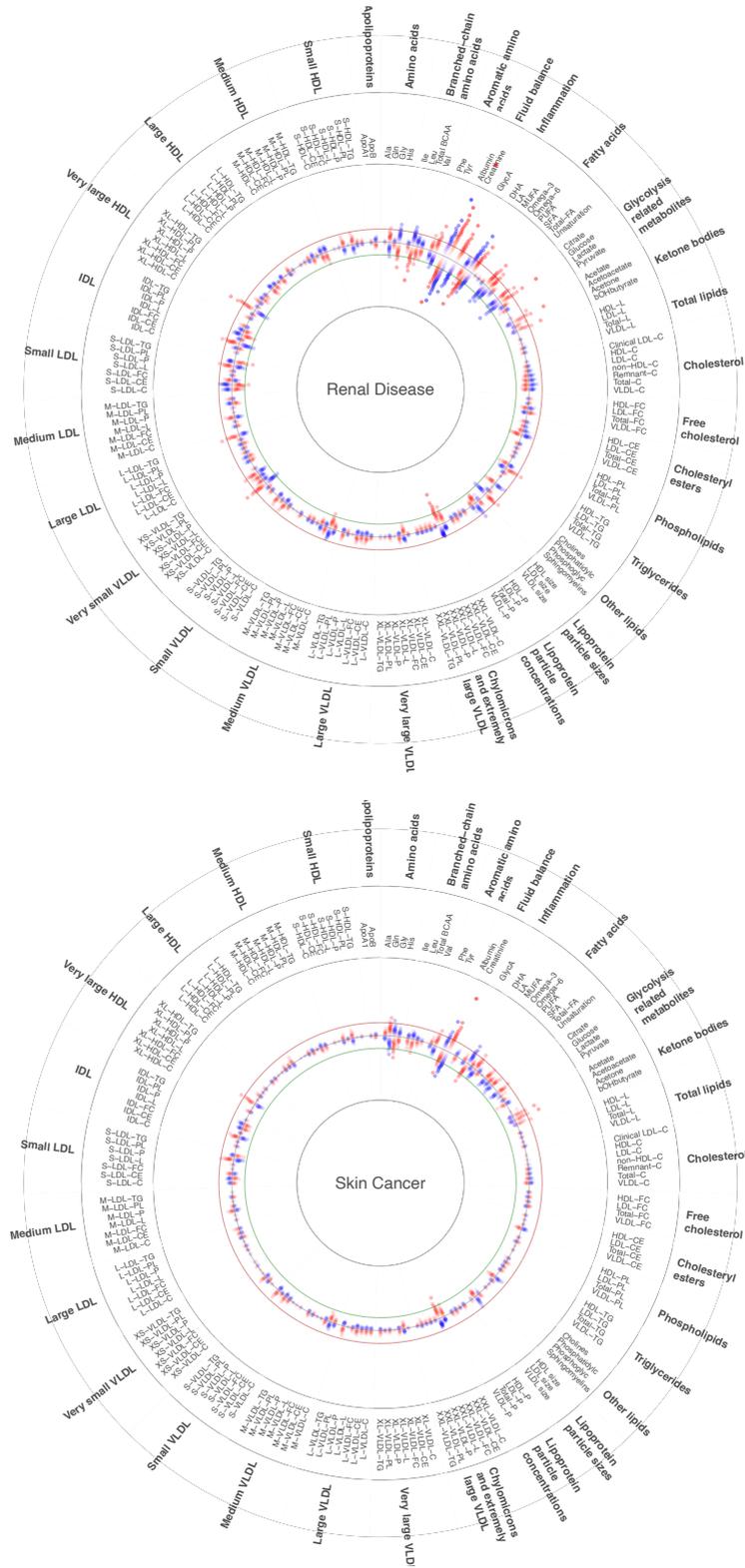
2 Metabolomic profiles predict individual multi-disease outcomes



Supplementary Figures



2 Metabolomic profiles predict individual multi-disease outcomes



2 Metabolomic profiles predict individual multi-disease outcomes

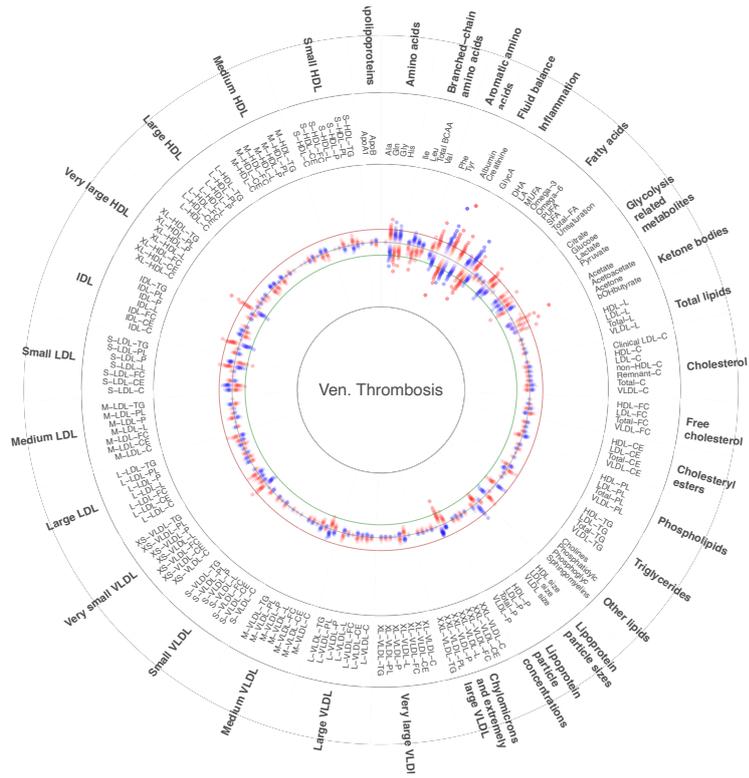


Figure 13: Global attributions for all endpoints. Displayed are global metabolite attribution profiles for all endpoints. Individual attributions are aggregated by percentiles with each dot indicating one percentile. The more distant a dot from the circular baseline, the stronger the absolute attribution for the percentile. Deviations toward the center represent negative, and deviations toward the outside represent positive contributions to the metabolomic state. The color indicates the metabolite's mean normalized plasma value. Major Adverse Cardiac Events (MACE), Coronary Heart Disease (CHD), Peripheral Artery Disease (PAD), Abdominal Aortic Aneurysm (AAA), Chronic Obstructive Pulmonary Disease (COPD).

Supplementary Figures



Figure 14: Metabolites differ throughout the attribution space. Displayed are distributions for all 168 measured metabolites stratified by the region (A, B and C) in the UMAP of the attribution space for type 2 diabetes (see figure 5.7 on page 122). Regions were defined by including all samples with an euclidean distance < 1 to the centroid A, B and C respectively; an euclidean distance of 1 is indicated by the scale bar (see figure 5.7 on page 122). The distributions differ significantly for metabolites, including glucose, fatty acids (i.e., LA and Omega-6) and multiple lipoprotein components (i.e., VLDL cholesterol and very large HDL triglycerides).

Supplementary Tables

1 Data and Implementation

Supplementary Tables

Table 1: Comparison of covariate and metabolite distributions for all cohorts.

	UKB (n=117981)			WHII (n=6117)					Rotterdam Study (n=2949)			Leiden Longevity Study (n=1655)			PROSPER (n=960)		
Basic Charac- teristics	Median (IQR)			44- 49	50- 54	55- 59	60- 64	65- 69	Median (IQR)			Median (IQR)			Median (IQR)		
Age (Years)	58	50	63	1268 (20.7%)	1775 (29.0%)	1293 (21.1%)	1296 (21.2%)	485 (7.9%)	74	70	79	59	54	63	75	73	78
Biological Sex	Male: 54078 (45.8%) Female: 63903 (54.2%)			Male: 4347 (71.1%) Female: 1770 (28.9%)					Male: 1243 (42.1%) Female: 1706 (57.9%)			Male: 734 (44.5%) Female: 921 (55.5%)			Male: 493 (51.4%) Female: 467 (48.6%)		
NMR Metabo- lites	Median (IQR)			Median (IQR)					Median (IQR)			Median (IQR)			Median (IQR)		
Acetate (mmol/l)	0.01	0.01	0.02		0.05	0.03	0.07		0.04	0.03	0.06	0.04	0.03	0.05	0.02	0.01	0.03
Acetoacetate (mmol/l)	0.01	0.01	0.02		0.02	0.01	0.03		0.02	0.02	0.04	0.02	0.01	0.03	0.00	0.00	0.01
Acetone (mmol/l)	0.01	0.01	0.02		0.02	0.02	0.03		0.02	0.01	0.02	0.02	0.01	0.02	0.02	0.02	0.03
Ala (mmol/l)	0.29	0.24	0.34		0.41	0.37	0.46		0.34	0.30	0.39	0.35	0.31	0.41	0.36	0.31	0.42
Albumin (g/l)	38.89	36.88	40.91		44.12	42.14	46.01		38.21	36.55	39.96	38.58	36.79	40.60	36.40	34.46	38.53
ApoA1 (g/l)	1.42	1.27	1.58		1.51	1.39	1.66		1.46	1.31	1.63	1.53	1.36	1.71	1.34	1.20	1.49
ApoB (g/l)	0.82	0.70	0.96		0.98	0.85	1.13		1.00	0.88	1.14	1.01	0.88	1.14	0.89	0.74	1.04
bOH- butyrate (mmol/l)	0.04	0.03	0.07		0.05	0.02	0.12		0.05	0.03	0.08	0.04	0.03	0.07	0.08	0.06	0.13
Cholines (mmol/l)	2.51	2.25	2.78		2.78	2.56	3.02		2.73	2.45	2.97	2.81	2.53	3.06	2.39	2.11	2.64
Citrate (mmol/l)	0.06	0.05	0.07		0.06	0.06	0.07		0.08	0.07	0.09	0.07	0.06	0.08	0.08	0.07	0.09
Clinical LDL-C (mmol/l)	2.47	2.00	2.97		3.18	2.72	3.67		3.33	2.84	3.83	3.37	2.89	3.89	2.80	2.26	3.38
Creatinine (mmol/l)	0.06	0.06	0.07		0.08	0.07	0.09		0.08	0.07	0.09	0.08	0.07	0.09	0.08	0.07	0.09
DHA (mmol/l)	0.22	0.18	0.28		0.24	0.20	0.29		0.20	0.16	0.25	0.22	0.17	0.27	0.19	0.15	0.23
Gln (mmol/l)	0.53	0.48	0.58		0.64	0.59	0.68		0.65	0.61	0.70	0.66	0.61	0.71	0.62	0.57	0.67
Glucose (mmol/l)	3.44	3.01	3.89		5.06	4.75	5.43		5.86	5.49	6.43	5.83	5.36	6.47	4.01	2.93	4.83
Gly (mmol/l)	0.15	0.12	0.19		0.26	0.23	0.29		0.16	0.13	0.20	0.17	0.14	0.22	0.17	0.13	0.21
GlycA (mmol/l)	0.78	0.71	0.86		0.83	0.77	0.90		0.89	0.81	0.97	0.89	0.81	0.97	0.89	0.82	0.97
HDL size (nm)	9.61	9.49	9.77		9.60	9.51	9.73		9.53	9.43	9.67	9.54	9.43	9.67	9.54	9.44	9.66
HDL-C (mmol/l)	1.26	1.06	1.49		1.38	1.23	1.57		1.32	1.13	1.54	1.38	1.18	1.64	1.19	1.02	1.38
HDL- CE (mmol/l)	0.98	0.83	1.17		1.08	0.96	1.23		1.03	0.87	1.21	1.08	0.91	1.28	0.93	0.80	1.08
HDL- FC (mmol/l)	0.28	0.24	0.33		0.30	0.27	0.35		0.29	0.25	0.34	0.31	0.26	0.36	0.26	0.23	0.30
HDL-L (mmol/l)	2.90	2.52	3.36		3.07	2.75	3.47		2.93	2.56	3.38	3.08	2.67	3.55	2.68	2.32	3.04
HDL-P (mmol/l)	0.01	0.01	0.02		0.02	0.02	0.02		0.02	0.01	0.02	0.02	0.02	0.02	0.01	0.01	0.02

1 Data and Implementation

Table 1 continued from previous page

	UKB (n=117981)				WHII (n=6117)				Rotterdam Study (n=2949)			Leiden Longevity Study (n=1655)			PROSPER (n=960)		
HDL-PL (mmol/l)	1.50	1.31	1.73		1.55	1.38	1.75		1.48	1.31	1.71	1.56	1.35	1.79	1.35	1.18	1.53
HDL-TG (mmol/l)	0.14	0.11	0.17		0.13	0.11	0.17		0.12	0.10	0.15	0.13	0.10	0.16	0.13	0.10	0.16
His (mmol/l)	0.06	0.06	0.07		0.08	0.07	0.08		0.07	0.07	0.08	0.07	0.06	0.08	0.07	0.06	0.07
IDL-C (mmol/l)	0.83	0.69	0.97		0.98	0.86	1.11		0.99	0.85	1.14	0.99	0.85	1.15	0.85	0.72	1.01
IDL-CE (mmol/l)	0.61	0.50	0.71		0.73	0.64	0.82		0.73	0.62	0.84	0.74	0.63	0.85	0.62	0.52	0.74
IDL-FC (mmol/l)	0.22	0.18	0.26		0.25	0.22	0.28		0.26	0.23	0.30	0.25	0.22	0.29	0.23	0.19	0.27
IDL-L (mmol/l)	1.21	1.02	1.40		1.42	1.25	1.60		1.42	1.23	1.62	1.42	1.23	1.62	1.26	1.06	1.48
IDL-P (mmol/l)	0.00	0.00	0.00		0.00	0.00	0.00		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
IDL-PL (mmol/l)	0.29	0.25	0.33		0.33	0.29	0.38		0.33	0.29	0.37	0.33	0.29	0.37	0.30	0.26	0.35
IDL-TG (mmol/l)	0.09	0.08	0.11		0.10	0.09	0.12		0.10	0.08	0.11	0.10	0.08	0.12	0.10	0.08	0.12
Ile (mmol/l)	0.05	0.04	0.06		0.06	0.05	0.07		0.05	0.04	0.06	0.05	0.05	0.07	0.06	0.05	0.07
L-HDL-C (mmol/l)	0.26	0.17	0.39		0.29	0.22	0.39		0.24	0.16	0.36	0.26	0.17	0.37	0.22	0.15	0.31
L-HDL-CE (mmol/l)	0.20	0.13	0.30		0.23	0.17	0.30		0.19	0.13	0.28	0.20	0.13	0.29	0.17	0.12	0.24
L-HDL-FC (mmol/l)	0.06	0.04	0.09		0.06	0.05	0.09		0.05	0.04	0.08	0.06	0.04	0.08	0.05	0.03	0.07
L-HDL-L (mmol/l)	0.57	0.41	0.82		0.61	0.47	0.81		0.53	0.37	0.74	0.56	0.38	0.79	0.49	0.35	0.66
L-HDL-P (mmol/l)	0.00	0.00	0.00		0.00	0.00	0.00		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
L-HDL-PL (mmol/l)	0.29	0.21	0.41		0.30	0.23	0.39		0.26	0.19	0.36	0.28	0.19	0.39	0.24	0.17	0.32
L-HDL-TG (mmol/l)	0.03	0.02	0.04		0.03	0.02	0.04		0.02	0.02	0.03	0.02	0.02	0.03	0.02	0.02	0.03
L-LDL-C (mmol/l)	1.10	0.92	1.29		1.40	1.21	1.59		1.45	1.25	1.67	1.49	1.29	1.71	1.23	1.03	1.47
L-LDL-CE (mmol/l)	0.81	0.68	0.95		1.03	0.90	1.18		1.07	0.92	1.23	1.10	0.96	1.27	0.91	0.76	1.08
L-LDL-FC (mmol/l)	0.29	0.24	0.34		0.36	0.32	0.41		0.38	0.32	0.44	0.38	0.33	0.44	0.33	0.27	0.39
L-LDL-L (mmol/l)	1.54	1.30	1.79		1.93	1.69	2.18		2.00	1.74	2.28	2.04	1.79	2.33	1.72	1.45	2.02

Supplementary Tables

Table 1 continued from previous page

	UKB (n=117981)			WHII (n=6117)			Rotterdam Study (n=2949)			Leiden Longevity Study (n=1655)			PROSPER (n=960)		
L- LDL-P (mmol/l)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
L-LDL- PL (mmol/l)	0.35	0.30	0.40	0.43	0.38	0.48	0.44	0.39	0.50	0.45	0.39	0.51	0.39	0.33	0.45
L-LDL- TG (mmol/l)	0.09	0.08	0.11	0.10	0.09	0.12	0.10	0.09	0.11	0.10	0.09	0.12	0.10	0.09	0.12
L- VLDL- C (mmol/l)	0.09	0.06	0.13	0.10	0.07	0.13	0.11	0.08	0.15	0.11	0.08	0.15	0.08	0.06	0.11
L- VLDL- CE (mmol/l)	0.05	0.03	0.07	0.06	0.04	0.07	0.06	0.04	0.08	0.06	0.04	0.08	0.04	0.03	0.06
L- VLDL- FC (mmol/l)	0.04	0.03	0.06	0.05	0.03	0.06	0.05	0.04	0.07	0.05	0.04	0.08	0.04	0.03	0.05
L- VLDL- L (mmol/l)	0.30	0.20	0.43	0.31	0.21	0.43	0.37	0.26	0.53	0.38	0.26	0.56	0.24	0.17	0.33
L- VLDL- P (mmol/l)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
L- VLDL- PL (mmol/l)	0.06	0.04	0.09	0.06	0.04	0.09	0.07	0.05	0.11	0.07	0.05	0.11	0.05	0.03	0.07
L- VLDL- TG (mmol/l)	0.15	0.10	0.22	0.14	0.09	0.21	0.19	0.13	0.27	0.19	0.13	0.30	0.11	0.07	0.15
LA (mmol/l)	3.37	2.95	3.83	3.83	3.45	4.27	3.90	3.44	4.38	3.97	3.49	4.47	3.27	2.82	3.79
Lactate (mmol/l)	3.68	3.05	4.42	2.46	2.10	2.93	1.20	0.95	1.55	1.40	1.04	1.89	5.22	3.72	7.88
LDL size (nm)	23.93	23.86	23.98	23.97	23.94	23.99	23.93	23.85	23.99	23.91	23.84	23.97	23.97	23.92	24.01
LDL-C (mmol/l)	1.68	1.41	1.98	2.14	1.86	2.44	2.25	1.94	2.55	2.30	1.99	2.62	1.90	1.58	2.23
LDL- CE (mmol/l)	1.23	1.03	1.45	1.56	1.35	1.79	1.64	1.41	1.86	1.69	1.45	1.92	1.37	1.14	1.61
LDL- FC (mmol/l)	0.45	0.38	0.53	0.57	0.50	0.65	0.61	0.52	0.69	0.62	0.53	0.71	0.52	0.43	0.61
LDL-L (mmol/l)	2.41	2.04	2.82	3.02	2.64	3.43	3.16	2.74	3.56	3.23	2.81	3.66	2.70	2.27	3.17
LDL-P (mmol/l)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
LDL- PL (mmol/l)	0.59	0.50	0.69	0.73	0.64	0.82	0.76	0.66	0.85	0.77	0.68	0.87	0.66	0.55	0.77
LDL- TG (mmol/l)	0.14	0.12	0.16	0.15	0.13	0.18	0.15	0.13	0.17	0.15	0.13	0.18	0.15	0.13	0.18
Leu (mmol/l)	0.10	0.08	0.11	0.13	0.11	0.14	0.10	0.09	0.12	0.11	0.09	0.13	0.11	0.10	0.13

1 Data and Implementation

Table 1 continued from previous page

	UKB (n=117981)				WHII (n=6117)				Rotterdam Study (n=2949)			Leiden Longevity Study (n=1655)			PROSPER (n=960)		
M-HDL-C (mmol/l)	0.48	0.40	0.56		0.52	0.46	0.59		0.49	0.42	0.59	0.53	0.44	0.63	0.44	0.37	0.52
M-HDL-CE (mmol/l)	0.39	0.33	0.46		0.43	0.38	0.49		0.41	0.35	0.48	0.43	0.36	0.52	0.37	0.31	0.43
M-HDL-FC (mmol/l)	0.08	0.07	0.10		0.09	0.08	0.11		0.09	0.07	0.10	0.09	0.08	0.11	0.08	0.06	0.09
M-HDL-L (mmol/l)	1.01	0.87	1.16		1.06	0.94	1.20		1.01	0.88	1.17	1.07	0.92	1.24	0.91	0.78	1.05
M-HDL-P (mmol/l)	0.00	0.00	0.00		0.00	0.00	0.00		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
M-HDL-PL (mmol/l)	0.47	0.41	0.54		0.49	0.44	0.55		0.47	0.41	0.54	0.50	0.42	0.57	0.42	0.36	0.49
M-HDL-TG (mmol/l)	0.05	0.04	0.06		0.05	0.04	0.06		0.04	0.04	0.06	0.05	0.04	0.06	0.04	0.03	0.06
M-LDL-C (mmol/l)	0.41	0.34	0.49		0.53	0.45	0.61		0.56	0.48	0.64	0.57	0.49	0.66	0.46	0.38	0.54
M-LDL-CE (mmol/l)	0.30	0.24	0.36		0.38	0.32	0.44		0.39	0.34	0.46	0.41	0.35	0.47	0.32	0.26	0.38
M-LDL-FC (mmol/l)	0.12	0.10	0.14		0.15	0.13	0.17		0.16	0.14	0.18	0.17	0.14	0.19	0.14	0.11	0.16
M-LDL-L (mmol/l)	0.60	0.50	0.71		0.76	0.65	0.87		0.80	0.70	0.92	0.83	0.71	0.95	0.68	0.56	0.79
M-LDL-P (mmol/l)	0.00	0.00	0.00		0.00	0.00	0.00		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
M-LDL-PL (mmol/l)	0.16	0.13	0.18		0.20	0.17	0.23		0.21	0.18	0.24	0.22	0.19	0.25	0.18	0.15	0.21
M-LDL-TG (mmol/l)	0.03	0.03	0.04		0.03	0.03	0.04		0.03	0.03	0.04	0.04	0.03	0.04	0.03	0.03	0.04
M-VLDL-C (mmol/l)	0.16	0.12	0.21		0.21	0.17	0.26		0.22	0.17	0.26	0.21	0.17	0.26	0.18	0.13	0.23
M-VLDL-CE (mmol/l)	0.09	0.06	0.12		0.12	0.09	0.14		0.12	0.09	0.14	0.11	0.09	0.14	0.10	0.07	0.13
M-VLDL-FC (mmol/l)	0.08	0.06	0.10		0.09	0.07	0.12		0.10	0.08	0.12	0.10	0.08	0.12	0.08	0.06	0.10

Supplementary Tables

Table 1 continued from previous page

	UKB (n=117981)			WHII (n=6117)			Rotterdam Study (n=2949)			Leiden Longevity Study (n=1655)			PROSPER (n=960)		
M-VLDL-L (mmol/l)	0.55	0.42	0.69	0.64	0.51	0.78	0.69	0.55	0.86	0.71	0.56	0.90	0.52	0.40	0.66
M-VLDL-P (mmol/l)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
M-VLDL-PL (mmol/l)	0.12	0.09	0.16	0.15	0.12	0.19	0.16	0.13	0.19	0.16	0.13	0.20	0.13	0.09	0.16
M-VLDL-TG (mmol/l)	0.25	0.19	0.34	0.27	0.20	0.35	0.31	0.24	0.42	0.33	0.24	0.46	0.22	0.17	0.28
MUFA (mmol/l)	2.68	2.25	3.24	2.91	2.52	3.42	2.83	2.46	3.29	2.91	2.51	3.42	2.49	2.12	2.89
non-HDL-C (mmol/l)	3.21	2.67	3.78	3.91	3.36	4.46	4.06	3.52	4.63	4.09	3.58	4.69	3.49	2.84	4.09
Omega-3 (mmol/l)	0.49	0.37	0.64	0.55	0.43	0.70	0.35	0.25	0.49	0.40	0.28	0.52	0.34	0.25	0.45
Omega-6 (mmol/l)	4.41	3.99	4.87	4.78	4.40	5.21	4.86	4.43	5.33	4.91	4.48	5.41	4.20	3.76	4.69
Phe (mmol/l)	0.04	0.04	0.05	0.08	0.07	0.08	0.05	0.05	0.06	0.05	0.05	0.06	0.06	0.05	0.07
Phosphatidylcholine (mmol/l)	2.06	1.82	2.31	2.27	2.07	2.50	2.27	2.02	2.50	2.34	2.09	2.58	1.90	1.67	2.15
Phosphoglyceride (mmol/l)	2.23	1.98	2.50	2.46	2.24	2.70	2.41	2.15	2.64	2.48	2.22	2.73	2.07	1.81	2.31
PUFA (mmol/l)	4.92	4.42	5.47	5.35	4.89	5.88	5.25	4.75	5.76	5.33	4.84	5.89	4.55	4.07	5.12
Pyruvate (mmol/l)	0.08	0.06	0.09	0.05	0.04	0.06	0.06	0.04	0.07	0.07	0.05	0.09	0.07	0.05	0.12
Remnant-C (mmol/l)	1.52	1.25	1.81	1.76	1.51	2.04	1.81	1.56	2.08	1.80	1.55	2.07	1.56	1.26	1.88
S-HDL-C (mmol/l)	0.44	0.40	0.48	0.49	0.46	0.52	0.50	0.46	0.54	0.52	0.47	0.57	0.45	0.41	0.49
S-HDL-CE (mmol/l)	0.32	0.29	0.35	0.36	0.34	0.39	0.37	0.34	0.40	0.39	0.35	0.43	0.34	0.30	0.37
S-HDL-FC (mmol/l)	0.11	0.10	0.12	0.13	0.12	0.13	0.13	0.12	0.14	0.13	0.12	0.15	0.11	0.10	0.13
S-HDL-L (mmol/l)	1.14	1.04	1.24	1.23	1.15	1.32	1.24	1.14	1.34	1.29	1.19	1.42	1.12	1.02	1.23
S-HDL-P (mmol/l)	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
S-HDL-PL (mmol/l)	0.65	0.59	0.71	0.68	0.63	0.74	0.69	0.63	0.75	0.72	0.66	0.79	0.62	0.56	0.68
S-HDL-TG (mmol/l)	0.05	0.04	0.06	0.05	0.04	0.06	0.05	0.04	0.06	0.05	0.04	0.06	0.05	0.04	0.06
S-LDL-C (mmol/l)	0.18	0.15	0.21	0.22	0.19	0.24	0.23	0.20	0.26	0.23	0.20	0.26	0.19	0.16	0.23

1 Data and Implementation

Table 1 continued from previous page

	UKB (n=117981)				WHII (n=6117)				Rotterdam Study (n=2949)			Leiden Longevity Study (n=1655)			PROSPER (n=960)		
S-LDL-CE (mmol/l)	0.13	0.11	0.15		0.15	0.13	0.18		0.16	0.14	0.18	0.16	0.14	0.19	0.14	0.11	0.16
S-LDL-FC (mmol/l)	0.05	0.04	0.06		0.06	0.05	0.07		0.07	0.06	0.07	0.07	0.06	0.08	0.06	0.05	0.06
S-LDL-L (mmol/l)	0.28	0.24	0.32		0.33	0.29	0.37		0.35	0.30	0.39	0.36	0.31	0.40	0.30	0.25	0.35
S-LDL-P (mmol/l)	0.00	0.00	0.00		0.00	0.00	0.00		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S-LDL-PL (mmol/l)	0.09	0.07	0.10		0.10	0.09	0.11		0.10	0.09	0.12	0.11	0.09	0.12	0.09	0.08	0.11
S-LDL-TG (mmol/l)	0.01	0.01	0.02		0.01	0.01	0.02		0.01	0.01	0.02	0.02	0.01	0.02	0.01	0.01	0.02
S-VLDL-C (mmol/l)	0.15	0.12	0.19		0.18	0.15	0.22		0.18	0.15	0.21	0.18	0.15	0.22	0.17	0.13	0.21
S-VLDL-CE (mmol/l)	0.09	0.07	0.12		0.11	0.09	0.14		0.11	0.09	0.13	0.11	0.09	0.13	0.10	0.08	0.13
S-VLDL-FC (mmol/l)	0.06	0.04	0.07		0.07	0.06	0.08		0.07	0.06	0.08	0.07	0.06	0.09	0.06	0.05	0.08
S-VLDL-L (mmol/l)	0.39	0.31	0.49		0.45	0.37	0.54		0.46	0.37	0.55	0.47	0.38	0.57	0.40	0.33	0.49
S-VLDL-P (mmol/l)	0.00	0.00	0.00		0.00	0.00	0.00		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S-VLDL-PL (mmol/l)	0.09	0.07	0.11		0.11	0.09	0.13		0.11	0.10	0.14	0.12	0.10	0.14	0.10	0.08	0.12
S-VLDL-TG (mmol/l)	0.15	0.11	0.19		0.15	0.12	0.20		0.16	0.12	0.21	0.17	0.13	0.22	0.13	0.11	0.17
SFA (mmol/l)	3.91	3.39	4.55		4.22	3.75	4.82		4.26	3.79	4.83	4.39	3.88	5.05	3.71	3.25	4.29
Sphingomyelins (mmol/l)	0.44	0.39	0.49		0.50	0.46	0.54		0.50	0.45	0.54	0.51	0.47	0.56	0.46	0.41	0.51
Total BCAA (mmol/l)	0.34	0.30	0.40		0.42	0.38	0.47		0.38	0.34	0.43	0.39	0.34	0.45	0.40	0.36	0.45
Total-C (mmol/l)	4.53	3.93	5.15		5.33	4.79	5.91		5.42	4.83	6.06	5.53	4.94	6.19	4.66	4.04	5.39
Total-CE (mmol/l)	3.29	2.86	3.74		3.87	3.49	4.28		3.96	3.51	4.42	4.04	3.60	4.51	3.38	2.95	3.91
Total-FA (mmol/l)	11.57	10.18	13.19		12.52	11.25	14.04		12.39	11.11	13.85	12.70	11.31	14.35	10.73	9.55	12.22

Supplementary Tables

Table 1 continued from previous page

	UKB (n=117981)				WHII (n=6117)				Rotterdam Study (n=2949)			Leiden Longevity Study (n=1655)			PROSPER (n=960)		
Total-FC (mmol/l)	1.24	1.06	1.42		1.45	1.29	1.63		1.47	1.31	1.65	1.49	1.33	1.68	1.27	1.08	1.48
Total-L (mmol/l)	8.65	7.62	9.75		9.67	8.73	10.69		9.94	8.97	11.02	10.20	9.14	11.43	8.50	7.41	9.66
Total-P (mmol/l)	0.02	0.01	0.02		0.02	0.02	0.02		0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.01	0.02
Total-PL (mmol/l)	2.87	2.58	3.19		3.13	2.87	3.42		3.15	2.84	3.44	3.24	2.93	3.53	2.77	2.47	3.08
Total-TG (mmol/l)	1.18	0.87	1.61		1.14	0.87	1.52		1.27	0.98	1.71	1.32	1.00	1.85	1.02	0.78	1.30
Tyr (mmol/l)	0.06	0.05	0.07		0.06	0.05	0.07		0.07	0.06	0.08	0.07	0.06	0.08	0.07	0.06	0.08
Unsaturation (degree)	1.36	1.31	1.41		1.37	1.33	1.41		1.32	1.28	1.36	1.31	1.27	1.36	1.30	1.26	1.35
Val (mmol/l)	0.20	0.17	0.23		0.24	0.22	0.26		0.22	0.20	0.25	0.23	0.20	0.26	0.23	0.20	0.26
VLDL size (nm)	38.60	37.70	39.52		38.49	37.83	39.10		38.87	38.17	39.74	38.92	38.19	39.86	38.04	37.46	38.66
VLDL-C (mmol/l)	0.69	0.54	0.87		0.79	0.63	0.95		0.81	0.66	0.98	0.80	0.65	0.98	0.70	0.55	0.89
VLDL-CE (mmol/l)	0.41	0.32	0.52		0.48	0.39	0.58		0.48	0.39	0.58	0.48	0.39	0.57	0.43	0.34	0.55
VLDL-FC (mmol/l)	0.28	0.21	0.35		0.31	0.24	0.38		0.32	0.26	0.41	0.32	0.26	0.41	0.27	0.21	0.34
VLDL-L (mmol/l)	1.97	1.46	2.60		2.06	1.55	2.61		2.22	1.74	2.91	2.26	1.72	3.02	1.76	1.36	2.26
VLDL-P (mmol/l)	0.00	0.00	0.00		0.00	0.00	0.00		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
VLDL-PL (mmol/l)	0.45	0.34	0.58		0.49	0.37	0.60		0.51	0.40	0.64	0.51	0.40	0.65	0.42	0.32	0.54
VLDL-TG (mmol/l)	0.81	0.54	1.17		0.76	0.53	1.07		0.90	0.64	1.29	0.94	0.66	1.41	0.64	0.45	0.86
XL-HDL-C (mmol/l)	0.07	0.06	0.10		0.08	0.06	0.10		0.07	0.06	0.09	0.07	0.06	0.09	0.07	0.06	0.09
XL-HDL-CE (mmol/l)	0.05	0.04	0.07		0.06	0.04	0.07		0.05	0.04	0.07	0.05	0.04	0.07	0.05	0.04	0.06
XL-HDL-FC (mmol/l)	0.02	0.02	0.03		0.02	0.02	0.02		0.02	0.02	0.03	0.02	0.02	0.02	0.02	0.02	0.03
XL-HDL-L (mmol/l)	0.15	0.11	0.21		0.15	0.12	0.19		0.13	0.10	0.18	0.13	0.10	0.18	0.13	0.10	0.17
XL-HDL-P (mmol/l)	0.00	0.00	0.00		0.00	0.00	0.00		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

1 Data and Implementation

Table 1 continued from previous page

	UKB (n=117981)				WHII (n=6117)				Rotterdam Study (n=2949)			Leiden Longevity Study (n=1655)			PROSPER (n=960)		
XL- HDL- PL (mmol/l)	0.07	0.05	0.10		0.06	0.05	0.09		0.06	0.04	0.08	0.05	0.03	0.08	0.06	0.04	0.08
XL- HDL- TG (mmol/l)	0.01	0.01	0.01		0.01	0.01	0.01		0.01	0.00	0.01	0.01	0.00	0.01	0.01	0.01	0.01
XL- VLDL- C (mmol/l)	0.05	0.03	0.07		0.05	0.04	0.07		0.06	0.04	0.08	0.06	0.04	0.08	0.05	0.03	0.06
XL- VLDL- CE (mmol/l)	0.03	0.02	0.04		0.03	0.02	0.04		0.03	0.03	0.04	0.04	0.03	0.05	0.03	0.02	0.04
XL- VLDL- FC (mmol/l)	0.02	0.01	0.03		0.02	0.01	0.03		0.02	0.02	0.03	0.02	0.01	0.04	0.02	0.01	0.03
XL- VLDL- L (mmol/l)	0.18	0.10	0.28		0.17	0.10	0.25		0.20	0.13	0.31	0.20	0.13	0.33	0.14	0.09	0.20
XL- VLDL- P (mmol/l)	0.00	0.00	0.00		0.00	0.00	0.00		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
XL- VLDL- PL (mmol/l)	0.03	0.02	0.05		0.03	0.02	0.05		0.04	0.02	0.06	0.04	0.02	0.06	0.03	0.02	0.04
XL- VLDL- TG (mmol/l)	0.09	0.05	0.15		0.08	0.04	0.13		0.11	0.06	0.17	0.11	0.06	0.19	0.06	0.04	0.10
XS- VLDL- C (mmol/l)	0.18	0.15	0.22		0.20	0.17	0.23		0.20	0.17	0.23	0.19	0.16	0.22	0.18	0.15	0.22
XS- VLDL- CE (mmol/l)	0.12	0.10	0.15		0.14	0.12	0.16		0.14	0.11	0.16	0.13	0.11	0.15	0.13	0.10	0.15
XS- VLDL- FC (mmol/l)	0.06	0.05	0.07		0.06	0.05	0.07		0.06	0.05	0.07	0.06	0.05	0.07	0.06	0.05	0.07
XS- VLDL- L (mmol/l)	0.35	0.30	0.41		0.38	0.32	0.44		0.36	0.31	0.42	0.36	0.31	0.41	0.35	0.29	0.42
XS- VLDL- P (mmol/l)	0.00	0.00	0.00		0.00	0.00	0.00		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
XS- VLDL- PL (mmol/l)	0.10	0.09	0.12		0.11	0.09	0.13		0.10	0.09	0.12	0.10	0.08	0.12	0.10	0.08	0.12
XS- VLDL- TG (mmol/l)	0.07	0.05	0.08		0.07	0.06	0.08		0.06	0.05	0.08	0.07	0.05	0.08	0.06	0.05	0.08

Supplementary Tables

Table 1 continued from previous page

	UKB (n=117981)				WHII (n=6117)				Rotterdam Study (n=2949)			Leiden Longevity Study (n=1655)			PROSPER (n=960)		
XXL- VLDL- C (mmol/l)	0.05	0.02	0.08		0.03	0.01	0.05		0.04	0.02	0.06	0.04	0.02	0.07	0.04	0.02	0.06
XXL- VLDL- CE (mmol/l)	0.03	0.01	0.04		0.02	0.01	0.03		0.02	0.01	0.04	0.02	0.01	0.04	0.03	0.02	0.04
XXL- VLDL- FC (mmol/l)	0.02	0.01	0.04		0.01	0.00	0.02		0.01	0.01	0.03	0.01	0.00	0.03	0.01	0.01	0.02
XXL- VLDL- L (mmol/l)	0.18	0.07	0.34		0.07	0.02	0.18		0.12	0.05	0.24	0.11	0.03	0.25	0.10	0.03	0.19
XXL- VLDL- P (mmol/l)	0.00	0.00	0.00		0.00	0.00	0.00		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
XXL- VLDL- PL (mmol/l)	0.03	0.01	0.05		0.01	0.00	0.03		0.02	0.01	0.04	0.02	0.00	0.04	0.02	0.01	0.03
XXL- VLDL- TG (mmol/l)	0.10	0.04	0.20		0.03	0.00	0.10		0.06	0.02	0.14	0.06	0.01	0.15	0.04	0.00	0.09

Table 2: Metabolites measured in the 1H-NMR metabolomics assay. All metabolomic predictors were log_{1p} transformed and then standard scaled. All listed predictors are available in the UK Biobank cohort under Field ID 220 and were measured at recruitment.

Abbreviation	Description	Subgroup	Group
Ala	Alanine	Amino acids	Amino acids
Gln	Glutamine	Amino acids	Amino acids
Gly	Glycine	Amino acids	Amino acids
His	Histidine	Amino acids	Amino acids
Ile	Isoleucine	Branched-chain amino acids	Amino acids
Leu	Leucine	Branched-chain amino acids	Amino acids
Total BCAA	Total concentration of branched-chain amino acids (leucine + isoleucine + valine)	Branched-chain amino acids	Amino acids
Val	Valine	Branched-chain amino acids	Amino acids
Phe	Phenylalanine	Aromatic amino acids	Amino acids
Tyr	Tyrosine	Aromatic amino acids	Amino acids
ApoA1	Apolipoprotein A1	Apolipoproteins	Apolipoproteins
ApoB	Apolipoprotein B	Apolipoproteins	Apolipoproteins
Clinical LDL-C	Clinical LDL cholesterol	Cholesterol	Cholesterol
HDL-C	HDL cholesterol	Cholesterol	Cholesterol
LDL-C	LDL cholesterol	Cholesterol	Cholesterol
non-HDL-C	Total cholesterol minus HDL-C	Cholesterol	Cholesterol
Remnant-C	Remnant cholesterol (non-HDL, non-LDL-cholesterol)	Cholesterol	Cholesterol
Total-C	Total cholesterol	Cholesterol	Cholesterol
VLDL-C	VLDL cholesterol	Cholesterol	Cholesterol
HDL-CE	Cholesteryl esters in HDL	Cholesteryl esters	Cholesteryl esters
LDL-CE	Cholesteryl esters in LDL	Cholesteryl esters	Cholesteryl esters
Total-CE	Total esterified cholesterol	Cholesteryl esters	Cholesteryl esters
VLDL-CE	Cholesteryl esters in VLDL	Cholesteryl esters	Cholesteryl esters
DHA	Docosahexaenoic acid	Fatty acids	Fatty acids
LA	Linoleic acid	Fatty acids	Fatty acids
MUFA	Monounsaturated fatty acids	Fatty acids	Fatty acids
Omega-3	Omega-3 fatty acids	Fatty acids	Fatty acids
Omega-6	Omega-6 fatty acids	Fatty acids	Fatty acids
PUFA	Polyunsaturated fatty acids	Fatty acids	Fatty acids
SFA	Saturated fatty acids	Fatty acids	Fatty acids
Total-FA	Total fatty acids	Fatty acids	Fatty acids
Unsaturation	Degree of unsaturation	Fatty acids	Fatty acids
Albumin	Albumin	Fluid balance	Fluid balance
Creatinine	Creatinine	Fluid balance	Fluid balance
HDL-FC	Free cholesterol in HDL	Free cholesterol	Free cholesterol
LDL-FC	Free cholesterol in LDL	Free cholesterol	Free cholesterol
Total-FC	Total free cholesterol	Free cholesterol	Free cholesterol
VLDL-FC	Free cholesterol in VLDL	Free cholesterol	Free cholesterol
Citrate	Citrate	Glycolysis related metabolites	Glycolysis related metabolites
Glucose	Glucose	Glycolysis related metabolites	Glycolysis related metabolites
Lactate	Lactate	Glycolysis related metabolites	Glycolysis related metabolites
Pyruvate	Pyruvate	Glycolysis related metabolites	Glycolysis related metabolites
GlycA	Glycoprotein acetyls	Inflammation	Inflammation
Acetate	Acetate	Ketone bodies	Ketone bodies
Acetoacetate	Acetoacetate	Ketone bodies	Ketone bodies
Acetone	Acetone	Ketone bodies	Ketone bodies
bOHbutyrate	3-Hydroxybutyrate	Ketone bodies	Ketone bodies
HDL-P	Concentration of HDL particles	Lipoprotein particle concentrations	Lipoprotein particle concentrations
LDL-P	Concentration of LDL particles	Lipoprotein particle concentrations	Lipoprotein particle concentrations
Total-P	Total concentration of lipoprotein particles	Lipoprotein particle concentrations	Lipoprotein particle concentrations

Supplementary Tables

Table 2 continued from previous page

Abbreviation	Description	Subgroup	Group
VLDL-P	Concentration of VLDL particles	Lipoprotein particle concentrations	Lipoprotein particle concentrations
HDL size	Average diameter for HDL particles	Lipoprotein particle sizes	Lipoprotein particle sizes
LDL size	Average diameter for LDL particles	Lipoprotein particle sizes	Lipoprotein particle sizes
VLDL size	Average diameter for VLDL particles	Lipoprotein particle sizes	Lipoprotein particle sizes
XXL-VLDL-C	Cholesterol in chylomicrons and extremely large VLDL	Chylomicrons and extremely large VLDL	Lipoprotein subclasses
XXL-VLDL-CE	Cholesteryl esters in chylomicrons and extremely large VLDL	Chylomicrons and extremely large VLDL	Lipoprotein subclasses
XXL-VLDL-FC	Free cholesterol in chylomicrons and extremely large VLDL	Chylomicrons and extremely large VLDL	Lipoprotein subclasses
XXL-VLDL-L	Total lipids in chylomicrons and extremely large VLDL	Chylomicrons and extremely large VLDL	Lipoprotein subclasses
XXL-VLDL-P	Concentration of chylomicrons and extremely large VLDL particles	Chylomicrons and extremely large VLDL	Lipoprotein subclasses
XXL-VLDL-PL	Phospholipids in chylomicrons and extremely large VLDL	Chylomicrons and extremely large VLDL	Lipoprotein subclasses
XXL-VLDL-TG	Triglycerides in chylomicrons and extremely large VLDL	Chylomicrons and extremely large VLDL	Lipoprotein subclasses
XL-VLDL-C	Cholesterol in very large VLDL	Very large VLDL	Lipoprotein subclasses
XL-VLDL-CE	Cholesteryl esters in very large VLDL	Very large VLDL	Lipoprotein subclasses
XL-VLDL-FC	Free cholesterol in very large VLDL	Very large VLDL	Lipoprotein subclasses
XL-VLDL-L	Total lipids in very large VLDL	Very large VLDL	Lipoprotein subclasses
XL-VLDL-P	Concentration of very large VLDL particles	Very large VLDL	Lipoprotein subclasses
XL-VLDL-PL	Phospholipids in very large VLDL	Very large VLDL	Lipoprotein subclasses
XL-VLDL-TG	Triglycerides in very large VLDL	Very large VLDL	Lipoprotein subclasses
L-VLDL-C	Cholesterol in large VLDL	Large VLDL	Lipoprotein subclasses
L-VLDL-CE	Cholesteryl esters in large VLDL	Large VLDL	Lipoprotein subclasses
L-VLDL-FC	Free cholesterol in large VLDL	Large VLDL	Lipoprotein subclasses
L-VLDL-L	Total lipids in large VLDL	Large VLDL	Lipoprotein subclasses
L-VLDL-P	Concentration of large VLDL particles	Large VLDL	Lipoprotein subclasses
L-VLDL-PL	Phospholipids in large VLDL	Large VLDL	Lipoprotein subclasses
L-VLDL-TG	Triglycerides in large VLDL	Large VLDL	Lipoprotein subclasses
M-VLDL-C	Cholesterol in medium VLDL	Medium VLDL	Lipoprotein subclasses
M-VLDL-CE	Cholesteryl esters in medium VLDL	Medium VLDL	Lipoprotein subclasses
M-VLDL-FC	Free cholesterol in medium VLDL	Medium VLDL	Lipoprotein subclasses
M-VLDL-L	Total lipids in medium VLDL	Medium VLDL	Lipoprotein subclasses
M-VLDL-P	Concentration of medium VLDL particles	Medium VLDL	Lipoprotein subclasses
M-VLDL-PL	Phospholipids in medium VLDL	Medium VLDL	Lipoprotein subclasses
M-VLDL-TG	Triglycerides in medium VLDL	Medium VLDL	Lipoprotein subclasses
S-VLDL-C	Cholesterol in small VLDL	Small VLDL	Lipoprotein subclasses
S-VLDL-CE	Cholesteryl esters in small VLDL	Small VLDL	Lipoprotein subclasses
S-VLDL-FC	Free cholesterol in small VLDL	Small VLDL	Lipoprotein subclasses
S-VLDL-L	Total lipids in small VLDL	Small VLDL	Lipoprotein subclasses
S-VLDL-P	Concentration of small VLDL particles	Small VLDL	Lipoprotein subclasses
S-VLDL-PL	Phospholipids in small VLDL	Small VLDL	Lipoprotein subclasses
S-VLDL-TG	Triglycerides in small VLDL	Small VLDL	Lipoprotein subclasses
XS-VLDL-C	Cholesterol in very small VLDL	Very small VLDL	Lipoprotein subclasses
XS-VLDL-CE	Cholesteryl esters in very small VLDL	Very small VLDL	Lipoprotein subclasses
XS-VLDL-FC	Free cholesterol in very small VLDL	Very small VLDL	Lipoprotein subclasses
XS-VLDL-L	Total lipids in very small VLDL	Very small VLDL	Lipoprotein subclasses
XS-VLDL-P	Concentration of very small VLDL particles	Very small VLDL	Lipoprotein subclasses
XS-VLDL-PL	Phospholipids in very small VLDL	Very small VLDL	Lipoprotein subclasses
XS-VLDL-TG	Triglycerides in very small VLDL	Very small VLDL	Lipoprotein subclasses
IDL-C	Cholesterol in IDL	IDL	Lipoprotein subclasses
IDL-CE	Cholesteryl esters in IDL	IDL	Lipoprotein subclasses
IDL-FC	Free cholesterol in IDL	IDL	Lipoprotein subclasses
IDL-L	Total lipids in IDL	IDL	Lipoprotein subclasses
IDL-P	Concentration of IDL particles	IDL	Lipoprotein subclasses
IDL-PL	Phospholipids in IDL	IDL	Lipoprotein subclasses
IDL-TG	Triglycerides in IDL	IDL	Lipoprotein subclasses
L-LDL-C	Cholesterol in large LDL	Large LDL	Lipoprotein subclasses

Table 2 continued from previous page

Abbreviation	Description	Subgroup	Group
L-LDL-CE	Cholesteryl esters in large LDL	Large LDL	Lipoprotein subclasses
L-LDL-FC	Free cholesterol in large LDL	Large LDL	Lipoprotein subclasses
L-LDL-L	Total lipids in large LDL	Large LDL	Lipoprotein subclasses
L-LDL-P	Concentration of large LDL particles	Large LDL	Lipoprotein subclasses
L-LDL-PL	Phospholipids in large LDL	Large LDL	Lipoprotein subclasses
L-LDL-TG	Triglycerides in large LDL	Large LDL	Lipoprotein subclasses
M-LDL-C	Cholesterol in medium LDL	Medium LDL	Lipoprotein subclasses
M-LDL-CE	Cholesteryl esters in medium LDL	Medium LDL	Lipoprotein subclasses
M-LDL-FC	Free cholesterol in medium LDL	Medium LDL	Lipoprotein subclasses
M-LDL-L	Total lipids in medium LDL	Medium LDL	Lipoprotein subclasses
M-LDL-P	Concentration of medium LDL particles	Medium LDL	Lipoprotein subclasses
M-LDL-PL	Phospholipids in medium LDL	Medium LDL	Lipoprotein subclasses
M-LDL-TG	Triglycerides in medium LDL	Medium LDL	Lipoprotein subclasses
S-LDL-C	Cholesterol in small LDL	Small LDL	Lipoprotein subclasses
S-LDL-CE	Cholesteryl esters in small LDL	Small LDL	Lipoprotein subclasses
S-LDL-FC	Free cholesterol in small LDL	Small LDL	Lipoprotein subclasses
S-LDL-L	Total lipids in small LDL	Small LDL	Lipoprotein subclasses
S-LDL-P	Concentration of small LDL particles	Small LDL	Lipoprotein subclasses
S-LDL-PL	Phospholipids in small LDL	Small LDL	Lipoprotein subclasses
S-LDL-TG	Triglycerides in small LDL	Small LDL	Lipoprotein subclasses
XL-HDL-C	Cholesterol in very large HDL	Very large HDL	Lipoprotein subclasses
XL-HDL-CE	Cholesteryl esters in very large HDL	Very large HDL	Lipoprotein subclasses
XL-HDL-FC	Free cholesterol in very large HDL	Very large HDL	Lipoprotein subclasses
XL-HDL-L	Total lipids in very large HDL	Very large HDL	Lipoprotein subclasses
XL-HDL-P	Concentration of very large HDL particles	Very large HDL	Lipoprotein subclasses
XL-HDL-PL	Phospholipids in very large HDL	Very large HDL	Lipoprotein subclasses
XL-HDL-TG	Triglycerides in very large HDL	Very large HDL	Lipoprotein subclasses
L-HDL-C	Cholesterol in large HDL	Large HDL	Lipoprotein subclasses
L-HDL-CE	Cholesteryl esters in large HDL	Large HDL	Lipoprotein subclasses
L-HDL-FC	Free cholesterol in large HDL	Large HDL	Lipoprotein subclasses
L-HDL-L	Total lipids in large HDL	Large HDL	Lipoprotein subclasses
L-HDL-P	Concentration of large HDL particles	Large HDL	Lipoprotein subclasses
L-HDL-PL	Phospholipids in large HDL	Large HDL	Lipoprotein subclasses
L-HDL-TG	Triglycerides in large HDL	Large HDL	Lipoprotein subclasses
M-HDL-C	Cholesterol in medium HDL	Medium HDL	Lipoprotein subclasses
M-HDL-CE	Cholesteryl esters in medium HDL	Medium HDL	Lipoprotein subclasses
M-HDL-FC	Free cholesterol in medium HDL	Medium HDL	Lipoprotein subclasses
M-HDL-L	Total lipids in medium HDL	Medium HDL	Lipoprotein subclasses
M-HDL-P	Concentration of medium HDL particles	Medium HDL	Lipoprotein subclasses
M-HDL-PL	Phospholipids in medium HDL	Medium HDL	Lipoprotein subclasses
M-HDL-TG	Triglycerides in medium HDL	Medium HDL	Lipoprotein subclasses
S-HDL-C	Cholesterol in small HDL	Small HDL	Lipoprotein subclasses
S-HDL-CE	Cholesteryl esters in small HDL	Small HDL	Lipoprotein subclasses
S-HDL-FC	Free cholesterol in small HDL	Small HDL	Lipoprotein subclasses
S-HDL-L	Total lipids in small HDL	Small HDL	Lipoprotein subclasses
S-HDL-P	Concentration of small HDL particles	Small HDL	Lipoprotein subclasses
S-HDL-PL	Phospholipids in small HDL	Small HDL	Lipoprotein subclasses
S-HDL-TG	Triglycerides in small HDL	Small HDL	Lipoprotein subclasses
Cholines	Total cholines	Other lipids	Other lipids
Phosphatidylc	Phosphatidylcholines	Other lipids	Other lipids
Phosphoglyc	Phosphoglycerides	Other lipids	Other lipids
Sphingomyelins	Sphingomyelins	Other lipids	Other lipids
HDL-PL	Phospholipids in HDL	Phospholipids	Phospholipids
LDL-PL	Phospholipids in LDL	Phospholipids	Phospholipids
Total-PL	Total phospholipids in lipoprotein particles	Phospholipids	Phospholipids
VLDL-PL	Phospholipids in VLDL	Phospholipids	Phospholipids
HDL-L	Total lipids in HDL	Total lipids	Total lipids
LDL-L	Total lipids in LDL	Total lipids	Total lipids
Total-L	Total lipids in lipoprotein particles	Total lipids	Total lipids
VLDL-L	Total lipids in VLDL	Total lipids	Total lipids
HDL-TG	Triglycerides in HDL	Triglycerides	Triglycerides
LDL-TG	Triglycerides in LDL	Triglycerides	Triglycerides
Total-TG	Total triglycerides	Triglycerides	Triglycerides
VLDL-TG	Triglycerides in VLDL	Triglycerides	Triglycerides

Table 3: Information on EHR-derived predictors used for the medical history model.

Domain ID	Concept Class	Concept ID	Concept Name	Total Number of Observations	Observation Frequency
Condition	Clinical Finding	132344	Gingival and periodontal disease	223	0.04%
Condition	Clinical Finding	132391	Furuncle of face	433	0.09%
Condition	Clinical Finding	132408	Back problem	142	0.03%
Condition	Clinical Finding	132412	Post-laminectomy syndrome	56	0.01%
Condition	Clinical Finding	132446	Congenital anomaly of skin	60	0.01%
Condition	Clinical Finding	132466	Lumbar sprain	3190	0.63%
Condition	Clinical Finding	132579	Subacute thyroiditis	192	0.04%
Condition	Clinical Finding	132583	Postablative hypothyroidism	113	0.02%
Condition	Clinical Finding	132702	Erythema multiforme	571	0.11%
Condition	Clinical Finding	132703	Lichen planus	3470	0.69%
Condition	Clinical Finding	132706	Disorder of nail	2763	0.55%
Condition	Clinical Finding	132797	Sepsis	1243	0.25%
Condition	Clinical Finding	132835	Pediculosis capitis	582	0.12%
Condition	Clinical Finding	132932	Chronic ethmoidal sinusitis	163	0.03%
Condition	Clinical Finding	132943	Acute periodontitis	73	0.01%
Condition	Clinical Finding	132982	Keloid scar	1131	0.23%
Condition	Clinical Finding	132983	Dermatographic urticaria	250	0.05%
Condition	Clinical Finding	133002	Acute osteomyelitis	244	0.05%
Condition	Clinical Finding	133141	Tinea pedis	10223	2.03%
Condition	Clinical Finding	133147	Primary malignant neoplasm of skin of trunk	863	0.17%
Condition	Clinical Finding	133228	Dental caries	2198	0.44%
Condition	Clinical Finding	133280	Alopecia	3007	0.60%
Condition	Clinical Finding	133283	Pyoderma gangrenosum	58	0.01%
Condition	Clinical Finding	133284	Impetigo herpetiformis	85	0.02%
Condition	Clinical Finding	133285	Dermatitis factitia	105	0.02%
Condition	Clinical Finding	133294	Prepatellar bursitis	2976	0.59%
Condition	Clinical Finding	133299	Swelling of limb	504	0.10%
Condition	Clinical Finding	133327	Viremia	369	0.07%
Condition	Clinical Finding	133384	Burn any degree involving less than 10 percent of body surface	174	0.03%
Condition	Clinical Finding	133424	Primary malignant neoplasm of thyroid gland	235	0.05%
Condition	Clinical Finding	133444	Thyroiditis	418	0.08%
Condition	Clinical Finding	133468	Disorder of extremity	244	0.05%
Condition	Clinical Finding	133547	Pyoderma	2271	0.45%
Condition	Clinical Finding	133569	Fasciitis	116	0.02%
Condition	Clinical Finding	133637	Partial thickness burn of lower limb	63	0.01%
Condition	Clinical Finding	133646	Injury of digital nerve	76	0.02%
Condition	Clinical Finding	133685	Salmonella infection	721	0.14%
Condition	Clinical Finding	133691	Meningococemia	54	0.01%
Condition	Clinical Finding	133713	Malignant melanoma of skin of face	72	0.01%
Condition	Clinical Finding	133726	Carcinoma in situ of skin of face	62	0.01%
Condition	Clinical Finding	133728	Congenital hypothyroidism	72	0.01%
Condition	Clinical Finding	133729	Hyperparathyroidism	410	0.08%
Condition	Clinical Finding	133737	Acute thyroiditis	58	0.01%
Condition	Clinical Finding	133794	Chronic gingivitis	221	0.04%
Condition	Clinical Finding	133803	Periapical abscess with sinus tract	242	0.05%
Condition	Clinical Finding	133834	Atopic dermatitis	19472	3.88%
Condition	Clinical Finding	133835	Eczema	27942	5.56%
Condition	Clinical Finding	133848	Slipped upper femoral epiphysis	77	0.02%
Condition	Clinical Finding	133853	Osteomyelitis of ankle AND/OR foot	87	0.02%
Condition	Clinical Finding	133857	Acquired deformity of limb	152	0.03%
Condition	Clinical Finding	133859	Preauricular cyst	57	0.01%
Condition	Clinical Finding	133861	Cleft lip	52	0.01%
Condition	Clinical Finding	134057	Disorder of cardiovascular system	8757	1.74%
Condition	Clinical Finding	134122	Contact dermatitis due to oil and grease	72	0.01%
Condition	Clinical Finding	134144	Cleft palate with cleft lip	54	0.01%
Condition	Clinical Finding	134159	Precordial pain	1723	0.34%
Condition	Clinical Finding	134222	Injury of forearm	267	0.05%
Condition	Clinical Finding	134304	Benign neoplasm of parathyroid gland	178	0.04%
Condition	Clinical Finding	134308	Carcinoma in situ of skin of lower limb	97	0.02%
Condition	Clinical Finding	134312	Iatrogenic hypothyroidism	56	0.01%
Condition	Clinical Finding	134398	Periodontal disease	121	0.02%

1 Data and Implementation

Table 3 continued from previous page

Domain ID	Concept Class	Concept ID	Concept Name	Total Number of Observations	Observation Frequency
Condition	Clinical Finding	134404	Anal spasm	350	0.07%
Condition	Clinical Finding	134438	Contact dermatitis	7012	1.40%
Condition	Clinical Finding	134441	Chronic ulcer of skin	248	0.05%
Condition	Clinical Finding	134442	Systemic sclerosis	212	0.04%
Condition	Clinical Finding	134452	Olecranon bursitis	5313	1.06%
Condition	Clinical Finding	134453	Bursitis	1234	0.25%
Condition	Clinical Finding	134460	Plantar fascial fibromatosis	6708	1.34%
Condition	Clinical Finding	134461	Tietze's disease	955	0.19%
Condition	Clinical Finding	134603	Chronic myeloid leukemia	125	0.02%
Condition	Clinical Finding	134668	Chronic maxillary sinusitis	990	0.20%
Condition	Clinical Finding	134681	Diffuse spasm of esophagus	590	0.12%
Condition	Clinical Finding	134718	Hirsutism	120	0.02%
Condition	Clinical Finding	134735	Chronic osteomyelitis	126	0.03%
Condition	Clinical Finding	134736	Backache	44489	8.85%
Condition	Clinical Finding	134743	Congenital ichthyosis of skin	92	0.02%
Condition	Clinical Finding	134870	Pityriasis versicolor	2556	0.51%
Condition	Clinical Finding	134898	Non-toxic uninodular goiter	323	0.06%
Condition	Clinical Finding	135030	Discoloration of skin	55	0.01%
Condition	Clinical Finding	135061	Congenital anomaly of musculoskeletal system	90	0.02%
Condition	Clinical Finding	135208	Benign neoplasm of lip	69	0.01%
Condition	Clinical Finding	135214	Polycythemia vera (clinical)	414	0.08%
Condition	Clinical Finding	135215	Hashimoto thyroiditis	429	0.09%
Condition	Clinical Finding	135287	Non-neoplastic nevus	436	0.09%
Condition	Clinical Finding	135303	Cyst of oral soft tissue	340	0.07%
Condition	Clinical Finding	135333	Pressure ulcer	207	0.04%
Condition	Clinical Finding	135360	Syncope	3359	0.67%
Condition	Clinical Finding	135473	Dermatophytosis	8237	1.64%
Condition	Clinical Finding	135526	Spinal cord disease	220	0.04%
Condition	Clinical Finding	135615	Pityriasis	287	0.06%
Condition	Clinical Finding	135618	Pruritic rash	496	0.10%
Condition	Clinical Finding	135772	Goiter	996	0.20%
Condition	Clinical Finding	135777	Neoplasm of uncertain behavior of skin	183	0.04%
Condition	Clinical Finding	135778	Toxic multinodular goiter	148	0.03%
Condition	Clinical Finding	135858	Disorder of lip	511	0.10%
Condition	Clinical Finding	135923	Cleft palate	58	0.01%
Condition	Clinical Finding	135930	Musculoskeletal finding	95	0.02%
Condition	Clinical Finding	136030	Sweating fever	86	0.02%
Condition	Clinical Finding	136057	Benign neoplasm of skin of trunk	656	0.13%
Condition	Clinical Finding	136184	Pruritus of skin	6554	1.30%
Condition	Clinical Finding	136185	Abnormal granulation tissue	294	0.06%
Condition	Clinical Finding	136198	Cervical spondylosis with myelopathy	799	0.16%
Condition	Clinical Finding	136368	Non-toxic multinodular goiter	1198	0.24%
Condition	Clinical Finding	136496	Cellulitis and abscess of face	593	0.12%
Condition	Clinical Finding	136497	Scar conditions and fibrosis of skin	1623	0.32%
Condition	Clinical Finding	136516	Congenital pes planus	131	0.03%
Condition	Clinical Finding	136575	Nonvenomous insect bite with infection	183	0.04%
Condition	Clinical Finding	136580	Dehiscence of surgical wound	860	0.17%
Condition	Clinical Finding	136661	Non-toxic nodular goiter	352	0.07%
Condition	Clinical Finding	136710	Perichondritis of pinna	129	0.03%
Condition	Clinical Finding	136773	Rosacea	12130	2.41%
Condition	Clinical Finding	136775	Contact dermatitis due to solar radiation	197	0.04%
Condition	Clinical Finding	136779	Disorder of forearm	666	0.13%
Condition	Clinical Finding	136788	Spinal stenosis of lumbar region	1205	0.24%
Condition	Clinical Finding	136865	Injury of radial nerve	72	0.01%
Condition	Clinical Finding	136915	Primary malignant neoplasm of skin of lip	160	0.03%
Condition	Clinical Finding	136917	Primary malignant neoplasm of skin of face	2592	0.52%
Condition	Clinical Finding	136934	Primary hyperparathyroidism	347	0.07%
Condition	Clinical Finding	137016	Retained dental root	906	0.18%
Condition	Clinical Finding	137053	Seborrheic dermatitis	11511	2.29%
Condition	Clinical Finding	137057	Paronychia	1995	0.40%
Condition	Clinical Finding	137193	Erythema	462	0.09%
Condition	Clinical Finding	137213	Dermal mycosis	6797	1.35%
Condition	Clinical Finding	137275	Disorder of muscle	403	0.08%

Supplementary Tables

Table 3 continued from previous page

Domain ID	Concept Class	Concept ID	Concept Name	Total Number of Observations	Observation Frequency
Condition	Clinical Finding	137351	Epidermoid cyst of skin	15099	3.01%
Condition	Clinical Finding	137379	Musculoskeletal symptom	3142	0.63%
Condition	Clinical Finding	137548	Cervical radiculopathy	154	0.03%
Condition	Clinical Finding	137613	Eclampsia in pregnancy	143	0.03%
Condition	Clinical Finding	137626	Seborrheic dermatitis of scalp	3181	0.63%
Condition	Clinical Finding	137627	Prickly heat	699	0.14%
Condition	Clinical Finding	137682	Skin sensation disturbance	1231	0.24%
Condition	Clinical Finding	137785	Verruca plantaris	9697	1.93%
Condition	Clinical Finding	137813	Benign neoplasm of mouth region	136	0.03%
Condition	Clinical Finding	137820	Postoperative hypothyroidism	898	0.18%
Condition	Clinical Finding	137829	Aplastic anemia	130	0.03%
Condition	Clinical Finding	137856	Atypical facial pain	902	0.18%
Condition	Clinical Finding	137944	Bullous dermatosis	63	0.01%
Condition	Clinical Finding	137951	Acquired keratoderma	93	0.02%
Condition	Clinical Finding	137967	Muscle, ligament and fascia disorders	513	0.10%
Condition	Clinical Finding	137977	Jaundice	1283	0.26%
Condition	Clinical Finding	137989	Abnormal results of cardiovascular function studies	294	0.06%
Condition	Clinical Finding	138102	Benign neoplasm of skin	10318	2.05%
Condition	Clinical Finding	138113	Cyst of thyroid	665	0.13%
Condition	Clinical Finding	138148	Plantar nerve lesion	712	0.14%
Condition	Clinical Finding	138176	Acute gingivitis	300	0.06%
Condition	Clinical Finding	138181	Radicular cyst	472	0.09%
Condition	Clinical Finding	138225	Disorder of sebaceous gland	250	0.05%
Condition	Clinical Finding	138239	Finding of limb structure	212	0.04%
Condition	Clinical Finding	138278	Sprains and strains of joints and adjacent muscles	3220	0.64%
Condition	Clinical Finding	138346	Erysipelas	728	0.14%
Condition	Clinical Finding	138379	Chronic lymphoid leukemia, disease	279	0.06%
Condition	Clinical Finding	138384	Acquired hypothyroidism	9965	1.98%
Condition	Clinical Finding	138387	Thyrotoxicosis	4300	0.86%
Condition	Clinical Finding	138388	Secondary hyperparathyroidism	55	0.01%
Condition	Clinical Finding	138455	Stomatitis	532	0.11%
Condition	Clinical Finding	138463	Aphthous ulcer of mouth	2123	0.42%
Condition	Clinical Finding	138466	Leukoplakia of oral mucosa	103	0.02%
Condition	Clinical Finding	138501	Cholinergic urticaria	126	0.03%
Condition	Clinical Finding	138502	Vitiligo	1238	0.25%
Condition	Clinical Finding	138525	Pain in limb	7381	1.47%
Condition	Clinical Finding	138565	Hyperhidrosis	745	0.15%
Condition	Clinical Finding	138612	Injury of cervical nerve roots	63	0.01%
Condition	Clinical Finding	138682	Herpes zoster auricularis	116	0.02%
Condition	Clinical Finding	138717	Toxic diffuse goiter	381	0.08%
Condition	Clinical Finding	138790	Chronic periodontitis	411	0.08%
Condition	Clinical Finding	138825	Actinic keratosis	11028	2.19%
Condition	Clinical Finding	138841	Disorder of sacrum	269	0.05%
Condition	Clinical Finding	138845	Rheumatism	1522	0.30%
Condition	Clinical Finding	138994	Myelodysplastic syndrome (clinical)	117	0.02%
Condition	Clinical Finding	139057	Disorder of oral soft tissues	230	0.05%
Condition	Clinical Finding	139099	Ingrowing nail	4215	0.84%
Condition	Clinical Finding	139803	Acute transverse myelitis	60	0.01%
Condition	Clinical Finding	139841	Chronic frontal sinusitis	216	0.04%
Condition	Clinical Finding	139850	Acute frontal sinusitis	2152	0.43%
Condition	Clinical Finding	139899	Pemphigoid	51	0.01%
Condition	Clinical Finding	139900	Urticaria	11408	2.27%
Condition	Clinical Finding	139902	Allergic urticaria	2270	0.45%
Condition	Clinical Finding	139906	Intervertebral disc disorder of cervical region with myelopathy	234	0.05%
Condition	Clinical Finding	139980	Epidermal burn of forearm	61	0.01%
Condition	Clinical Finding	140020	Viral exanthem	461	0.09%
Condition	Clinical Finding	140027	Herpetic gingivostomatitis	680	0.14%
Condition	Clinical Finding	140062	Iodine hypothyroidism	100	0.02%
Condition	Clinical Finding	140090	Lesion of median nerve	75	0.01%
Condition	Clinical Finding	140168	Psoriasis	12098	2.41%
Condition	Clinical Finding	140173	Telogen effluvium	93	0.02%
Condition	Clinical Finding	140176	Erythema nodosum	1115	0.22%

Table 3 continued from previous page

Domain ID	Concept Class	Concept ID	Concept Name	Total Number of Observations	Observation Frequency
Condition	Clinical Finding	140190	Disorder of back	897	0.18%
Condition	Clinical Finding	140214	Eruption	21097	4.20%
Condition	Clinical Finding	140266	Contusion of back	123	0.02%
Condition	Clinical Finding	140273	Contusion of forearm	64	0.01%
Condition	Clinical Finding	140352	Acute myeloid leukemia, disease	133	0.03%
Condition	Clinical Finding	140362	Hypoparathyroidism	134	0.03%
Condition	Clinical Finding	140480	Impetigo	6052	1.20%
Condition	Clinical Finding	140487	Dermatitis herpetiformis	200	0.04%
Condition	Clinical Finding	140526	Congenital anomaly of endocrine gland	62	0.01%
Condition	Clinical Finding	140641	Verruca vulgaris	11879	2.36%
Condition	Clinical Finding	140648	Onychomycosis due to dermatophyte	7661	1.52%
Condition	Clinical Finding	140673	Hypothyroidism	17769	3.54%
Condition	Clinical Finding	140708	Myelopathy due to another disorder	450	0.09%
Condition	Clinical Finding	140803	Idiopathic urticaria	508	0.10%
Condition	Clinical Finding	140821	Spasm	1736	0.35%
Condition	Clinical Finding	140842	Changes in skin texture	175	0.03%
Condition	Clinical Finding	140949	Infestation by <i>Sarcoptes scabiei</i> var <i>hominis</i>	3566	0.71%
Condition	Clinical Finding	140961	Infestation by <i>Phthirus pubis</i>	131	0.03%
Condition	Clinical Finding	140966	Benign neoplasm of skin of face	682	0.14%
Condition	Clinical Finding	141004	Lesion of radial nerve	51	0.01%
Condition	Clinical Finding	141056	Acute ethmoidal sinusitis	165	0.03%
Condition	Clinical Finding	141090	Pyogenic granuloma of skin	755	0.15%
Condition	Clinical Finding	141094	Lichen	377	0.08%
Condition	Clinical Finding	141095	Acne	11076	2.20%
Condition	Clinical Finding	141213	Scarlet fever	1272	0.25%
Condition	Clinical Finding	141216	Molluscum contagiosum infection	972	0.19%
Condition	Clinical Finding	141232	Malignant melanoma of skin	1372	0.27%
Condition	Clinical Finding	141248	Benign neoplasm of skin of lip	150	0.03%
Condition	Clinical Finding	141249	Benign neoplasm of thyroid gland	297	0.06%
Condition	Clinical Finding	141253	Disorder of thyroid gland	1211	0.24%
Condition	Clinical Finding	141323	Acute maxillary sinusitis	6710	1.34%
Condition	Clinical Finding	141371	Pityriasis rosea	2556	0.51%
Condition	Clinical Finding	141375	Hypertrophic condition of skin	1121	0.22%
Condition	Clinical Finding	141456	Chilblains	791	0.16%
Condition	Clinical Finding	141650	Disseminated superficial actinic porokeratosis	77	0.02%
Condition	Clinical Finding	141651	Stevens-Johnson syndrome	80	0.02%
Condition	Clinical Finding	141652	Localized hyperhidrosis	84	0.02%
Condition	Clinical Finding	141663	Osteomyelitis	312	0.06%
Condition	Clinical Finding	141667	Laxity of ligament	79	0.02%
Condition	Clinical Finding	141825	Simple goiter	1366	0.27%
Condition	Clinical Finding	141917	Balanitis xerotica obliterans	388	0.08%
Condition	Clinical Finding	141932	Senile hyperkeratosis	33370	6.64%
Condition	Clinical Finding	141933	Alopecia areata	1316	0.26%
Condition	Clinical Finding	141941	Cervical spondylosis without myelopathy	5276	1.05%
Condition	Clinical Finding	141960	Skin finding	107	0.02%
Condition	Clinical Finding	192242	Acute hepatitis C	117	0.02%
Condition	Clinical Finding	192259	Ascariasis	63	0.01%
Condition	Clinical Finding	192273	Benign neoplasm of adrenal gland	95	0.02%
Condition	Clinical Finding	192275	Alpha-1-antitrypsin deficiency	83	0.02%
Condition	Clinical Finding	192279	Disorder of kidney due to diabetes mellitus	169	0.03%
Condition	Clinical Finding	192353	Disorder of gallbladder	763	0.15%
Condition	Clinical Finding	192357	Paralytic ileus	96	0.02%
Condition	Clinical Finding	192359	Renal failure syndrome	891	0.18%
Condition	Clinical Finding	192367	Dysplasia of cervix	1321	0.26%
Condition	Clinical Finding	192370	Secondary female infertility	1587	0.32%
Condition	Clinical Finding	192433	Congenital hypertrophic pyloric stenosis	69	0.01%
Condition	Clinical Finding	192438	Abdominal mass	746	0.15%
Condition	Clinical Finding	192450	Retention of urine	3576	0.71%
Condition	Clinical Finding	192568	Secondary malignant neoplasm of intra-abdominal lymph nodes	372	0.07%
Condition	Clinical Finding	192606	Paraplegia	149	0.03%
Condition	Clinical Finding	192667	Atrophic gastritis	323	0.06%
Condition	Clinical Finding	192671	Gastrointestinal hemorrhage	1239	0.25%

Supplementary Tables

Table 3 continued from previous page

Domain ID	Concept Class	Concept ID	Concept Name	Total Number of Observations	Observation Frequency
Condition	Clinical Finding	192675	Biliary cirrhosis	86	0.02%
Condition	Clinical Finding	192676	Cervical intraepithelial neoplasia grade 1	6737	1.34%
Condition	Clinical Finding	192680	Portal hypertension	203	0.04%
Condition	Clinical Finding	192683	Uterovaginal prolapse	2357	0.47%
Condition	Clinical Finding	192731	Digestive symptom	117	0.02%
Condition	Clinical Finding	192819	Typhoid fever	142	0.03%
Condition	Clinical Finding	192854	Intramural leiomyoma of uterus	964	0.19%
Condition	Clinical Finding	192855	Cancer in situ of urinary bladder	173	0.03%
Condition	Clinical Finding	192885	Tarsal tunnel syndrome	98	0.02%
Condition	Clinical Finding	192953	Intestinal adhesions with obstruction	488	0.10%
Condition	Clinical Finding	192956	Cholecystitis	1027	0.20%
Condition	Clinical Finding	192960	Noninfectious gastroenteritis	74	0.01%
Condition	Clinical Finding	192963	Disorder of pancreas	180	0.04%
Condition	Clinical Finding	192964	Infectious disorder of kidney	288	0.06%
Condition	Clinical Finding	192974	Secondary uterine inertia	53	0.01%
Condition	Clinical Finding	192979	Obstetric high vaginal laceration	93	0.02%
Condition	Clinical Finding	193138	Primary malignant neoplasm of lower third of esophagus	108	0.02%
Condition	Clinical Finding	193165	Disorder of adrenal gland	85	0.02%
Condition	Clinical Finding	193170	Renal glycosuria	121	0.02%
Condition	Clinical Finding	193238	Acute appendicitis with peritoneal abscess	189	0.04%
Condition	Clinical Finding	193242	Perforation of intestine	149	0.03%
Condition	Clinical Finding	193249	Acute hemorrhagic gastritis	335	0.07%
Condition	Clinical Finding	193252	Diverticulosis of small intestine	161	0.03%
Condition	Clinical Finding	193253	Nephritis	126	0.03%
Condition	Clinical Finding	193254	Disorder of vagina	124	0.02%
Condition	Clinical Finding	193256	Alcoholic fatty liver	203	0.04%
Condition	Clinical Finding	193261	Vaginospasm	216	0.04%
Condition	Clinical Finding	193262	Inflammatory disorder of penis	109	0.02%
Condition	Clinical Finding	193277	Deliveries by cesarean	3214	0.64%
Condition	Clinical Finding	193326	Urge incontinence of urine	3234	0.64%
Condition	Clinical Finding	193402	Bacterial enteritis	1365	0.27%
Condition	Clinical Finding	193439	Benign neoplasm of body of uterus	139	0.03%
Condition	Clinical Finding	193460	Disorder of lower extremity	420	0.08%
Condition	Clinical Finding	193518	Intestinal obstruction	1442	0.29%
Condition	Clinical Finding	193519	Impaired renal function disorder	126	0.03%
Condition	Clinical Finding	193520	Urinary bladder stone	481	0.10%
Condition	Clinical Finding	193522	Acute prostatitis	643	0.13%
Condition	Clinical Finding	193528	Third degree uterine prolapse	320	0.06%
Condition	Clinical Finding	193530	Follicular cyst of ovary	681	0.14%
Condition	Clinical Finding	193666	Injury of hip region	207	0.04%
Condition	Clinical Finding	193688	Clostridioides difficile infection	450	0.09%
Condition	Clinical Finding	193739	Ovarian failure	255	0.05%
Condition	Clinical Finding	193782	End-stage renal disease	347	0.07%
Condition	Clinical Finding	193807	Anorectal fistula	717	0.14%
Condition	Clinical Finding	193818	Calculus of prostate	55	0.01%
Condition	Clinical Finding	193830	Third degree perineal laceration	249	0.05%
Condition	Clinical Finding	193871	Genuine stress incontinence	10056	2.00%
Condition	Clinical Finding	193874	Nocturnal enuresis	1218	0.24%
Condition	Clinical Finding	194081	Acute cystitis	5455	1.09%
Condition	Clinical Finding	194091	Calculus in urethra	60	0.01%
Condition	Clinical Finding	194092	Uterine prolapse without vaginal wall prolapse	172	0.03%
Condition	Clinical Finding	194093	Postcoital bleeding	3532	0.70%
Condition	Clinical Finding	194109	Delayed AND/OR secondary postpartum hemorrhage	373	0.07%
Condition	Clinical Finding	194133	Low back pain	61590	12.26%
Condition	Clinical Finding	194175	Left upper quadrant pain	95	0.02%
Condition	Clinical Finding	194229	Crushing injury of lower limb	66	0.01%
Condition	Clinical Finding	194382	External hemorrhoids	533	0.11%
Condition	Clinical Finding	194406	Urinary tract obstruction	321	0.06%
Condition	Clinical Finding	194408	Hydroureter	62	0.01%
Condition	Clinical Finding	194418	Urethral caruncle	352	0.07%
Condition	Clinical Finding	194420	Endometriosis of fallopian tube	137	0.03%
Condition	Clinical Finding	194439	Obstetric perineal wound disruption	60	0.01%
Condition	Clinical Finding	194475	Hiccoughs	164	0.03%

Table 3 continued from previous page

Domain ID	Concept Class	Concept ID	Concept Name	Total Number of Observations	Observation Frequency
Condition	Clinical Finding	194491	Tenderness of epigastrium	184	0.04%
Condition	Clinical Finding	194526	Injury of trunk	235	0.05%
Condition	Clinical Finding	194611	Carcinoma in situ of uterine cervix	2631	0.52%
Condition	Clinical Finding	194684	Crohn's disease of large bowel	403	0.08%
Condition	Clinical Finding	194686	Acquired renal cystic disease	350	0.07%
Condition	Clinical Finding	194693	Postprocedural pelvic peritoneal adhesions	147	0.03%
Condition	Clinical Finding	194696	Dysmenorrhea	7793	1.55%
Condition	Clinical Finding	194702	Premature rupture of membranes	679	0.14%
Condition	Clinical Finding	194710	Placenta previa without hemorrhage	125	0.02%
Condition	Clinical Finding	194871	Trichomonal vulvovaginitis	146	0.03%
Condition	Clinical Finding	194913	Idiopathic peripheral autonomic neuropathy	77	0.02%
Condition	Clinical Finding	194984	Disease of liver	713	0.14%
Condition	Clinical Finding	194985	Postgastric surgery syndrome	69	0.01%
Condition	Clinical Finding	194990	Inflammatory disease of liver	763	0.15%
Condition	Clinical Finding	194991	Calculus of gallbladder with acute cholecystitis	950	0.19%
Condition	Clinical Finding	194992	Celiac disease	2112	0.42%
Condition	Clinical Finding	194993	Intestinal malabsorption	306	0.06%
Condition	Clinical Finding	194995	Disorder of ureter	277	0.06%
Condition	Clinical Finding	194997	Prostatitis	1976	0.39%
Condition	Clinical Finding	195002	Ulceration of intestine	269	0.05%
Condition	Clinical Finding	195007	Female stress incontinence	4630	0.92%
Condition	Clinical Finding	195012	Intermenstrual bleeding - irregular	4474	0.89%
Condition	Clinical Finding	195072	Medullary sponge kidney	61	0.01%
Condition	Clinical Finding	195074	Longitudinal deficiency of lower limb	54	0.01%
Condition	Clinical Finding	195197	Primary malignant neoplasm of vulva	50	0.01%
Condition	Clinical Finding	195212	Hypercortisolism	99	0.02%
Condition	Clinical Finding	195259	Meralgia paresthetica	1094	0.22%
Condition	Clinical Finding	195294	Thrombosed hemorrhoids	727	0.14%
Condition	Clinical Finding	195300	Alcoholic gastritis	195	0.04%
Condition	Clinical Finding	195306	Gastroduodenitis	3042	0.61%
Condition	Clinical Finding	195307	Hydrops of gallbladder	82	0.02%
Condition	Clinical Finding	195312	Functional disorder of bladder	99	0.02%
Condition	Clinical Finding	195314	Nephrotic syndrome	250	0.05%
Condition	Clinical Finding	195321	Postmenopausal bleeding	15110	3.01%
Condition	Clinical Finding	195363	Congenital anomaly of the kidney	50	0.01%
Condition	Clinical Finding	195401	Contusion of hip	60	0.01%
Condition	Clinical Finding	195464	Epidemic pleurodynia	196	0.04%
Condition	Clinical Finding	195500	Benign neoplasm of uterus	97	0.02%
Condition	Clinical Finding	195562	Hemorrhoids	32784	6.52%
Condition	Clinical Finding	195581	Peritoneal adhesion	1545	0.31%
Condition	Clinical Finding	195585	Crohn's disease of small intestine	279	0.06%
Condition	Clinical Finding	195588	Cystitis	16723	3.33%
Condition	Clinical Finding	195590	Urethral stricture	2488	0.50%
Condition	Clinical Finding	195596	Chronic pancreatitis	403	0.08%
Condition	Clinical Finding	195598	Chronic uterine inflammatory disease	90	0.02%
Condition	Clinical Finding	195599	Prolapse of vaginal vault after hysterectomy	302	0.06%
Condition	Clinical Finding	195603	Vulval and/or perineal noninflammatory disorders	855	0.17%
Condition	Clinical Finding	195632	Rupture of quadriceps tendon	119	0.02%
Condition	Clinical Finding	195769	Submucous leiomyoma of uterus	1185	0.24%
Condition	Clinical Finding	195770	Subserous leiomyoma of uterus	570	0.11%
Condition	Clinical Finding	195834	Atherosclerosis of renal artery	78	0.02%
Condition	Clinical Finding	195856	Cholangitis	259	0.05%
Condition	Clinical Finding	195861	Small kidney	59	0.01%
Condition	Clinical Finding	195862	Urethritis	482	0.10%
Condition	Clinical Finding	195864	Diverticulum of bladder	425	0.08%
Condition	Clinical Finding	195867	Noninflammatory disorder of the vagina	2375	0.47%
Condition	Clinical Finding	195873	Leukorrhea	294	0.06%
Condition	Clinical Finding	195906	Disorder of lumbar spine	167	0.03%
Condition	Clinical Finding	195926	Slowing of urinary stream	240	0.05%
Condition	Clinical Finding	195977	Contusion of trunk	164	0.03%
Condition	Clinical Finding	196044	Primary malignant neoplasm of stomach	76	0.02%
Condition	Clinical Finding	196061	Benign neoplasm of stomach	1014	0.20%
Condition	Clinical Finding	196149	Acute appendicitis with generalized peritonitis	578	0.12%

Supplementary Tables

Table 3 continued from previous page

Domain ID	Concept Class	Concept ID	Concept Name	Total Number of Observations	Observation Frequency
Condition	Clinical Finding	196151	Functional disorder of intestine	347	0.07%
Condition	Clinical Finding	196152	Peritonitis	548	0.11%
Condition	Clinical Finding	196157	Induratio penis plastica	2041	0.41%
Condition	Clinical Finding	196158	Disorder of penis	1267	0.25%
Condition	Clinical Finding	196160	Fistula of intestine	146	0.03%
Condition	Clinical Finding	196162	Inflammatory disease of the uterus	148	0.03%
Condition	Clinical Finding	196163	Cervicitis and endocervicitis	280	0.06%
Condition	Clinical Finding	196165	Cervical intraepithelial neoplasia grade 2	3301	0.66%
Condition	Clinical Finding	196168	Irregular periods	9432	1.88%
Condition	Clinical Finding	196328	Salmonella gastroenteritis	785	0.16%
Condition	Clinical Finding	196359	Primary malignant neoplasm of uterine cervix	198	0.04%
Condition	Clinical Finding	196360	Primary malignant neoplasm of bladder	997	0.20%
Condition	Clinical Finding	196364	Benign neoplasm of uterine cervix	162	0.03%
Condition	Clinical Finding	196436	Internal hemorrhoids	142	0.03%
Condition	Clinical Finding	196454	Colostomy and enterostomy malfunction	262	0.05%
Condition	Clinical Finding	196456	Gallstone	5832	1.16%
Condition	Clinical Finding	196463	Alcoholic cirrhosis	208	0.04%
Condition	Clinical Finding	196468	Acute salpingo-oophoritis	103	0.02%
Condition	Clinical Finding	196473	Hypertrophy of uterus	975	0.19%
Condition	Clinical Finding	196523	Diarrhea	17234	3.43%
Condition	Clinical Finding	196528	Hypospadias	154	0.03%
Condition	Clinical Finding	196569	Contusion of abdominal wall	112	0.02%
Condition	Clinical Finding	196620	Viral enteritis	127	0.03%
Condition	Clinical Finding	196653	Malignant tumor of kidney	106	0.02%
Condition	Clinical Finding	196733	Urethral syndrome	243	0.05%
Condition	Clinical Finding	196734	Disorder of prostate	739	0.15%
Condition	Clinical Finding	196738	Disorder of male genital organ	1777	0.35%
Condition	Clinical Finding	196751	Placenta previa with hemorrhage	144	0.03%
Condition	Clinical Finding	196758	Tumor of body of uterus affecting pregnancy	149	0.03%
Condition	Clinical Finding	196813	Simple renal cyst	84	0.02%
Condition	Clinical Finding	196821	Urethral discharge	202	0.04%
Condition	Clinical Finding	196925	Secondary malignant neoplasm of retroperitoneum and peritoneum	164	0.03%
Condition	Clinical Finding	197006	Residual hemorrhoidal skin tags	2419	0.48%
Condition	Clinical Finding	197023	Bilateral inguinal hernia	1039	0.21%
Condition	Clinical Finding	197032	Hyperplasia of prostate	5427	1.08%
Condition	Clinical Finding	197033	Endometriosis of pelvic peritoneum	829	0.16%
Condition	Clinical Finding	197034	Intussusception of intestine	226	0.04%
Condition	Clinical Finding	197036	Vesicoureteric reflux	99	0.02%
Condition	Clinical Finding	197039	Male genital organ vascular diseases	647	0.13%
Condition	Clinical Finding	197050	Abnormality of organs AND/OR soft tissues of pelvis affecting pregnancy	246	0.05%
Condition	Clinical Finding	197150	Late effect of fracture of lower extremities	214	0.04%
Condition	Clinical Finding	197163	Burn of trunk	103	0.02%
Condition	Clinical Finding	197223	Enterobiasis	1303	0.26%
Condition	Clinical Finding	197230	Malignant neoplasm of uterus	124	0.02%
Condition	Clinical Finding	197236	Uterine leiomyoma	17740	3.53%
Condition	Clinical Finding	197237	Benign neoplasm of prostate	58	0.01%
Condition	Clinical Finding	197304	Ulcer of lower extremity	1174	0.23%
Condition	Clinical Finding	197318	Cholesterosis of gallbladder	275	0.05%
Condition	Clinical Finding	197320	Acute renal failure syndrome	836	0.17%
Condition	Clinical Finding	197324	Disorder of peritoneum	79	0.02%
Condition	Clinical Finding	197331	Disorder of urinary tract	96	0.02%
Condition	Clinical Finding	197332	Incomplete uterovaginal prolapse	2860	0.57%
Condition	Clinical Finding	197343	First degree perineal laceration	2833	0.56%
Condition	Clinical Finding	197349	Perineal hematoma	250	0.05%
Condition	Clinical Finding	197381	Epigastric pain	15159	3.02%
Condition	Clinical Finding	197484	Diarrhea of presumed infectious origin	1624	0.32%
Condition	Clinical Finding	197494	Viral hepatitis C	273	0.05%
Condition	Clinical Finding	197500	Primary malignant neoplasm of colon	488	0.10%
Condition	Clinical Finding	197508	Malignant tumor of urinary bladder	656	0.13%
Condition	Clinical Finding	197593	Impaction of intestine	109	0.02%
Condition	Clinical Finding	197601	Spermatocele	215	0.04%

Table 3 continued from previous page

Domain ID	Concept Class	Concept ID	Concept Name	Total Number of Observations	Observation Frequency
Condition	Clinical Finding	197602	Tube-ovarian inflammatory disease	59	0.01%
Condition	Clinical Finding	197603	Intestinal volvulus	251	0.05%
Condition	Clinical Finding	197605	Inflammatory disorder of male genital organ	611	0.12%
Condition	Clinical Finding	197606	Female infertility of tubal origin	166	0.03%
Condition	Clinical Finding	197607	Excessive and frequent menstruation	9006	1.79%
Condition	Clinical Finding	197609	Cervical, vaginal and vulval inflammatory diseases	214	0.04%
Condition	Clinical Finding	197610	Cyst of ovary	7622	1.52%
Condition	Clinical Finding	197625	Third stage hemorrhage	269	0.05%
Condition	Clinical Finding	197654	Congenital cystic disease of liver	72	0.01%
Condition	Clinical Finding	197672	Urinary incontinence	6564	1.31%
Condition	Clinical Finding	197675	Incontinence of feces	1633	0.33%
Condition	Clinical Finding	197676	Large liver	234	0.05%
Condition	Clinical Finding	197684	Dysuria	12970	2.58%
Condition	Clinical Finding	197751	Burn of lower limb	315	0.06%
Condition	Clinical Finding	197795	Acute type B viral hepatitis	74	0.01%
Condition	Clinical Finding	197917	Disorder of biliary tract	467	0.09%
Condition	Clinical Finding	197925	Hemorrhage of rectum and anus	12576	2.50%
Condition	Clinical Finding	197927	Intrauterine synechiae	67	0.01%
Condition	Clinical Finding	197938	Uterine inertia	274	0.05%
Condition	Clinical Finding	197950	Cellulitis and abscess of trunk	169	0.03%
Condition	Clinical Finding	197981	Abdominal tenderness	223	0.04%
Condition	Clinical Finding	197988	Generalized abdominal pain	685	0.14%
Condition	Clinical Finding	198075	Condyloma acuminatum of the anogenital region	412	0.08%
Condition	Clinical Finding	198101	Lipoma of skin and subcutaneous tissue (excluding face)	418	0.08%
Condition	Clinical Finding	198124	Kidney disease	207	0.04%
Condition	Clinical Finding	198177	Abdominal aortic aneurysm	335	0.07%
Condition	Clinical Finding	198185	Chronic renal failure	473	0.09%
Condition	Clinical Finding	198195	Ovulation bleeding	1341	0.27%
Condition	Clinical Finding	198197	Male infertility	346	0.07%
Condition	Clinical Finding	198198	Polyp of vagina	264	0.05%
Condition	Clinical Finding	198199	Pyelonephritis	637	0.13%
Condition	Clinical Finding	198201	Chronic salpingo-oophoritis	50	0.01%
Condition	Clinical Finding	198202	Cystocele	6345	1.26%
Condition	Clinical Finding	198263	Right upper quadrant pain	810	0.16%
Condition	Clinical Finding	198334	Enteric campylobacteriosis	3253	0.65%
Condition	Clinical Finding	198337	Infectious diarrheal disease	1948	0.39%
Condition	Clinical Finding	198363	Candidiasis of vagina	7307	1.45%
Condition	Clinical Finding	198400	Hemangioma of skin and subcutaneous tissue	113	0.02%
Condition	Clinical Finding	198401	Neoplasm of uncertain behavior of bladder	336	0.07%
Condition	Clinical Finding	198402	Neoplasm of uncertain behavior of kidney	77	0.02%
Condition	Clinical Finding	198464	Incisional hernia	1006	0.20%
Condition	Clinical Finding	198465	Megacolon, not Hirschsprung's	86	0.02%
Condition	Clinical Finding	198483	Stricture or atresia of the vagina	117	0.02%
Condition	Clinical Finding	198488	Placental abruption	96	0.02%
Condition	Clinical Finding	198492	Second degree perineal laceration	4129	0.82%
Condition	Clinical Finding	198520	Lumbar spondylosis with myelopathy	93	0.02%
Condition	Clinical Finding	198572	Cervical intraepithelial neoplasia	611	0.12%
Condition	Clinical Finding	198678	Intestinal infectious disease	1113	0.22%
Condition	Clinical Finding	198700	Secondary malignant neoplasm of liver	79	0.02%
Condition	Clinical Finding	198715	Premature menopause	968	0.19%
Condition	Clinical Finding	198803	Benign prostatic hyperplasia	4565	0.91%
Condition	Clinical Finding	198809	Acute cholecystitis	2111	0.42%
Condition	Clinical Finding	198822	Furuncle of buttock	341	0.07%
Condition	Clinical Finding	198964	Chronic hepatitis C	269	0.05%
Condition	Clinical Finding	198984	Malignant tumor of cervix	431	0.09%
Condition	Clinical Finding	198985	Primary malignant neoplasm of kidney	418	0.08%
Condition	Clinical Finding	199067	Female pelvic inflammatory disease	2436	0.48%
Condition	Clinical Finding	199074	Acute pancreatitis	1704	0.34%
Condition	Clinical Finding	199075	Neurogenic bladder	1118	0.22%
Condition	Clinical Finding	199076	Tubal pregnancy	355	0.07%
Condition	Clinical Finding	199078	Vaginal wall prolapse	492	0.10%
Condition	Clinical Finding	199754	Primary malignant neoplasm of pancreas	53	0.01%
Condition	Clinical Finding	199764	Benign neoplasm of ovary	1097	0.22%

Supplementary Tables

Table 3 continued from previous page

Domain ID	Concept Class	Concept ID	Concept Name	Total Number of Observations	Observation Frequency
Condition	Clinical Finding	199837	Portal vein thrombosis	54	0.01%
Condition	Clinical Finding	199839	External hemorrhoids without complication	3782	0.75%
Condition	Clinical Finding	199856	Varicose veins of lower extremity with inflammation	185	0.04%
Condition	Clinical Finding	199860	Hernia of abdominal cavity	544	0.11%
Condition	Clinical Finding	199861	Perianal abscess	1954	0.39%
Condition	Clinical Finding	199866	Acute gastritis	2952	0.59%
Condition	Clinical Finding	199875	Prolapse of vaginal walls without uterine prolapse	1118	0.22%
Condition	Clinical Finding	199876	Prolapse of female genital organs	1782	0.35%
Condition	Clinical Finding	199877	Mucous polyp of cervix	1422	0.28%
Condition	Clinical Finding	199878	Light and infrequent menstruation	114	0.02%
Condition	Clinical Finding	199881	Endometriosis of ovary	1091	0.22%
Condition	Clinical Finding	199891	Rhesus isoimmunization affecting pregnancy	369	0.07%
Condition	Clinical Finding	199913	Acquired deformity of ankle AND/OR foot	141	0.03%
Condition	Clinical Finding	199978	Contusion of lower limb	469	0.09%
Condition	Clinical Finding	200051	Primary malignant neoplasm of ovary	470	0.09%
Condition	Clinical Finding	200169	Pruritus ani	9139	1.82%
Condition	Clinical Finding	200174	Disorder of skin and/or subcutaneous tissue	12023	2.39%
Condition	Clinical Finding	200219	Abdominal pain	49033	9.76%
Condition	Clinical Finding	200445	Chronic prostatitis	968	0.19%
Condition	Clinical Finding	200447	Gastrointestinal complication	349	0.07%
Condition	Clinical Finding	200450	Disorder of urethra	285	0.06%
Condition	Clinical Finding	200452	Disorder of female genital organs	931	0.19%
Condition	Clinical Finding	200458	Acquired atrophy of ovary and fallopian tube	87	0.02%
Condition	Clinical Finding	200459	Endometrial cystic hyperplasia	196	0.04%
Condition	Clinical Finding	200461	Endometriosis of uterus	1884	0.37%
Condition	Clinical Finding	200527	Splenomegaly	263	0.05%
Condition	Clinical Finding	200528	Ascites	466	0.09%
Condition	Clinical Finding	200675	Neoplasm of uncertain behavior of ovary	136	0.03%
Condition	Clinical Finding	200680	Neoplasm of bladder	266	0.05%
Condition	Clinical Finding	200687	Renal disorder due to type 1 diabetes mellitus	66	0.01%
Condition	Clinical Finding	200762	Autoimmune hepatitis	110	0.02%
Condition	Clinical Finding	200763	Chronic hepatitis	200	0.04%
Condition	Clinical Finding	200765	Chronic cholecystitis	1324	0.26%
Condition	Clinical Finding	200773	Stenosis of rectum and anus	191	0.04%
Condition	Clinical Finding	200775	Endometrial hyperplasia	960	0.19%
Condition	Clinical Finding	200779	Polyp of corpus uteri	7646	1.52%
Condition	Clinical Finding	200780	Disorder of uterus	66	0.01%
Condition	Clinical Finding	200789	Retained placenta, without hemorrhage	231	0.05%
Condition	Clinical Finding	200843	Finding of frequency of urination	9860	1.96%
Condition	Clinical Finding	200845	Urgent desire to urinate	3360	0.67%
Condition	Clinical Finding	200848	Dribbling of urine	226	0.04%
Condition	Clinical Finding	200959	Secondary malignant neoplasm of intrapelvic lymph nodes	69	0.01%
Condition	Clinical Finding	200962	Primary malignant neoplasm of prostate	2377	0.47%
Condition	Clinical Finding	200970	Carcinoma in situ of prostate	346	0.07%
Condition	Clinical Finding	201045	Thrombosed external hemorrhoids	3460	0.69%
Condition	Clinical Finding	201061	Diaphragmatic hernia	17363	3.46%
Condition	Clinical Finding	201066	Obstruction of bile duct	323	0.06%
Condition	Clinical Finding	201078	Atrophic vaginitis	6560	1.31%
Condition	Clinical Finding	201093	Infection of skin and/or subcutaneous tissue	10323	2.05%
Condition	Clinical Finding	201131	Pelvic mass	392	0.08%
Condition	Clinical Finding	201211	Herpetic vulvovaginitis	285	0.06%
Condition	Clinical Finding	201254	Type 1 diabetes mellitus	1686	0.34%
Condition	Clinical Finding	201257	Disorder of endocrine ovary	75	0.01%
Condition	Clinical Finding	201313	Hypertensive renal disease	187	0.04%
Condition	Clinical Finding	201322	Internal hemorrhoids without complication	1767	0.35%
Condition	Clinical Finding	201332	Alveolitis of jaw	55	0.01%
Condition	Clinical Finding	201337	Disorder of bladder	4178	0.83%
Condition	Clinical Finding	201338	Urethral fistula	69	0.01%
Condition	Clinical Finding	201340	Gastritis	16505	3.28%
Condition	Clinical Finding	201341	Incisional hernia with obstruction	112	0.02%
Condition	Clinical Finding	201344	Intestinovesical fistula	69	0.01%
Condition	Clinical Finding	201347	Pelvic congestion syndrome	58	0.01%
Condition	Clinical Finding	201418	Flatulence, eructation and gas pain	268	0.05%

Table 3 continued from previous page

Domain ID	Concept Class	Concept ID	Concept Name	Total Number of Observations	Observation Frequency
Condition	Clinical Finding	201461	Disorder affecting transplanted structure	111	0.02%
Condition	Clinical Finding	201518	Primary malignant neoplasm of lower limb	739	0.15%
Condition	Clinical Finding	201595	Thrombosed internal hemorrhoids	247	0.05%
Condition	Clinical Finding	201603	Disorder of teeth AND/OR supporting structures	876	0.17%
Condition	Clinical Finding	201606	Crohn's disease	1828	0.36%
Condition	Clinical Finding	201612	Alcoholic liver damage	495	0.10%
Condition	Clinical Finding	201614	Calculus of bile duct with cholecystitis	151	0.03%
Condition	Clinical Finding	201618	Disorder of intestine	12961	2.58%
Condition	Clinical Finding	201619	Gastric diverticulum	61	0.01%
Condition	Clinical Finding	201620	Kidney stone	4343	0.86%
Condition	Clinical Finding	201621	Chronic cystitis	476	0.09%
Condition	Clinical Finding	201622	Prostatic congestion or hemorrhage	114	0.02%
Condition	Clinical Finding	201625	Malposition of uterus	624	0.12%
Condition	Clinical Finding	201626	Mittelschmerz	464	0.09%
Condition	Clinical Finding	201627	Abnormal vaginal bleeding	685	0.14%
Condition	Clinical Finding	201642	Retained portions of placenta AND/OR membranes without hemorrhage	61	0.01%
Condition	Clinical Finding	201688	Delay when starting to pass urine	512	0.10%
Condition	Clinical Finding	201690	Renal colic	5705	1.14%
Condition	Clinical Finding	201728	Foreign body in digestive tract	55	0.01%
Condition	Clinical Finding	201780	Shigellosis	182	0.04%
Condition	Clinical Finding	201792	Nongonococcal urethritis	1100	0.22%
Condition	Clinical Finding	201820	Diabetes mellitus	7861	1.56%
Condition	Clinical Finding	201824	Benign neoplasm of bladder	54	0.01%
Condition	Clinical Finding	201826	Type 2 diabetes mellitus	17447	3.47%
Condition	Clinical Finding	201891	Disorder of stomach	280	0.06%
Condition	Clinical Finding	201894	Acute vascular insufficiency of intestine	106	0.02%
Condition	Clinical Finding	201901	Abscess of liver	106	0.02%
Condition	Clinical Finding	201909	Female infertility	2630	0.52%
Condition	Clinical Finding	201916	Ureteric stone	2603	0.52%
Condition	Clinical Finding	201956	Congenital anomaly of lower limb	54	0.01%
Condition	Clinical Finding	22340	Esophageal varices without bleeding	283	0.06%
Condition	Clinical Finding	22350	Edema of larynx	127	0.03%
Condition	Clinical Finding	22492	Foreign body in pharynx	413	0.08%
Condition	Clinical Finding	23220	Chronic tonsillitis	734	0.15%
Condition	Clinical Finding	23325	Heartburn	8184	1.63%
Condition	Clinical Finding	23653	Foreign body in esophagus	236	0.05%
Condition	Clinical Finding	23731	Benign neoplasm of larynx	64	0.01%
Condition	Clinical Finding	23798	Acute laryngopharyngitis	104	0.02%
Condition	Clinical Finding	23986	Disorder of pituitary gland	96	0.02%
Condition	Clinical Finding	24134	Neck pain	36693	7.30%
Condition	Clinical Finding	24148	Congenital diverticulum of pharynx	156	0.03%
Condition	Clinical Finding	24230	Foreign body in larynx	69	0.01%
Condition	Clinical Finding	24602	Benign neoplasm of esophagus	84	0.02%
Condition	Clinical Finding	24609	Hypoglycemia	1057	0.21%
Condition	Clinical Finding	24660	Acute tonsillitis	15379	3.06%
Condition	Clinical Finding	24818	Injury of neck	1052	0.21%
Condition	Clinical Finding	24909	Hereditary spherocytosis	105	0.02%
Condition	Clinical Finding	24966	Esophageal varices	70	0.01%
Condition	Clinical Finding	24970	Chronic laryngitis	591	0.12%
Condition	Clinical Finding	24977	Mucocele of salivary gland	158	0.03%
Condition	Clinical Finding	25056	Congenital anomaly of neck	60	0.01%
Condition	Clinical Finding	252758	Late effect of fracture of skull AND/OR face bones	174	0.03%
Condition	Clinical Finding	25297	Acute pharyngitis	25350	5.05%
Condition	Clinical Finding	253009	Thoracogenic scoliosis	76	0.02%
Condition	Clinical Finding	253321	Stridor	141	0.03%
Condition	Clinical Finding	253506	Pneumonitis	56	0.01%
Condition	Clinical Finding	253549	Disorder of connective tissue	334	0.07%
Condition	Clinical Finding	253796	Pneumothorax	1338	0.27%
Condition	Clinical Finding	253815	Ulceration of vulva	307	0.06%
Condition	Clinical Finding	253896	Traumatic pneumothorax without open wound into thorax	125	0.02%
Condition	Clinical Finding	253954	Pulmonary tuberculosis	1182	0.24%
Condition	Clinical Finding	254061	Pleural effusion	2083	0.41%

Supplementary Tables

Table 3 continued from previous page

Domain ID	Concept Class	Concept ID	Concept Name	Total Number of Observations	Observation Frequency
Condition	Clinical Finding	254068	Disorder of upper respiratory system	435	0.09%
Condition	Clinical Finding	254443	Sjögren's syndrome	744	0.15%
Condition	Clinical Finding	254591	Secondary malignant neoplasm of lung	179	0.04%
Condition	Clinical Finding	254662	Pulmonary infarction	190	0.04%
Condition	Clinical Finding	254669	Disorder of vocal cord	440	0.09%
Condition	Clinical Finding	254761	Cough	48541	9.66%
Condition	Clinical Finding	25518	Sickle cell trait	499	0.10%
Condition	Clinical Finding	255302	Spontaneous pneumothorax	569	0.11%
Condition	Clinical Finding	255348	Polymyalgia rheumatica	1500	0.30%
Condition	Clinical Finding	255573	Chronic obstructive lung disease	5007	1.00%
Condition	Clinical Finding	25572	Disorder of salivary gland	430	0.09%
Condition	Clinical Finding	255841	Chronic bronchitis	293	0.06%
Condition	Clinical Finding	255848	Pneumonia	1835	0.37%
Condition	Clinical Finding	255891	Lupus erythematosus	122	0.02%
Condition	Clinical Finding	255919	Finding of head and neck region	65	0.01%
Condition	Clinical Finding	256439	Allergic rhinitis due to pollen	21162	4.21%
Condition	Clinical Finding	256448	Chronic asthmatic bronchitis	135	0.03%
Condition	Clinical Finding	256449	Bronchiectasis	2069	0.41%
Condition	Clinical Finding	256450	Asbestosis	126	0.03%
Condition	Clinical Finding	256451	Bronchitis	32776	6.52%
Condition	Clinical Finding	256531	Head and neck swelling	522	0.10%
Condition	Clinical Finding	256633	Primary malignant neoplasm of base of tongue	50	0.01%
Condition	Clinical Finding	256717	Bronchospasm	559	0.11%
Condition	Clinical Finding	256722	Bronchopneumonia	901	0.18%
Condition	Clinical Finding	256723	Pneumonia and influenza	1075	0.21%
Condition	Clinical Finding	257004	Acute exacerbation of chronic obstructive airways disease	1149	0.23%
Condition	Clinical Finding	257007	Allergic rhinitis	18287	3.64%
Condition	Clinical Finding	257011	Acute upper respiratory infection	4919	0.98%
Condition	Clinical Finding	257012	Chronic sinusitis	7130	1.42%
Condition	Clinical Finding	257315	Bacterial pneumonia	350	0.07%
Condition	Clinical Finding	257581	Exacerbation of asthma	1928	0.38%
Condition	Clinical Finding	257628	Systemic lupus erythematosus	573	0.11%
Condition	Clinical Finding	257683	Posterior rhinorrhea	913	0.18%
Condition	Clinical Finding	257778	Pertussis	1090	0.22%
Condition	Clinical Finding	257907	Disorder of lung	343	0.07%
Condition	Clinical Finding	258369	Primary malignant neoplasm of lung	72	0.01%
Condition	Clinical Finding	25844	Ulcer of esophagus	3680	0.73%
Condition	Clinical Finding	258482	Abscess of Bartholin's gland	1385	0.28%
Condition	Clinical Finding	258540	Marfan's syndrome	119	0.02%
Condition	Clinical Finding	258780	Emphysematous bronchitis	1623	0.32%
Condition	Clinical Finding	258785	Pneumococcal pneumonia	166	0.03%
Condition	Clinical Finding	259153	Pain in throat	15581	3.10%
Condition	Clinical Finding	259848	Chronic rhinitis	10489	2.09%
Condition	Clinical Finding	259852	Staphylococcal pneumonia	51	0.01%
Condition	Clinical Finding	259865	Cyst of Bartholin's gland duct	1775	0.35%
Condition	Clinical Finding	259995	Foreign body in orifice	52	0.01%
Condition	Clinical Finding	260123	Acute sinusitis	30004	5.97%
Condition	Clinical Finding	260125	Acute bronchiolitis	336	0.07%
Condition	Clinical Finding	260131	Disorder of bronchus	140	0.03%
Condition	Clinical Finding	260139	Acute bronchitis	23221	4.62%
Condition	Clinical Finding	260336	Primary malignant neoplasm of glottis	78	0.02%
Condition	Clinical Finding	260427	Common cold	9052	1.80%
Condition	Clinical Finding	26052	Primary malignant neoplasm of larynx	57	0.01%
Condition	Clinical Finding	260759	Malocclusion of teeth	60	0.01%
Condition	Clinical Finding	261116	Acquired deformity of nose	257	0.05%
Condition	Clinical Finding	261236	Primary malignant neoplasm of upper lobe, bronchus or lung	127	0.03%
Condition	Clinical Finding	261324	Pneumonia due to Streptococcus	74	0.01%
Condition	Clinical Finding	261325	Pulmonary emphysema	834	0.17%
Condition	Clinical Finding	261326	Viral pneumonia	409	0.08%
Condition	Clinical Finding	261687	Hemoptysis	4024	0.80%
Condition	Clinical Finding	261880	Atelectasis	829	0.16%
Condition	Clinical Finding	26638	Primary malignant neoplasm of esophagus	155	0.03%

Table 3 continued from previous page

Domain ID	Concept Class	Concept ID	Concept Name	Total Number of Observations	Observation Frequency
Condition	Clinical Finding	26662	Testicular hypofunction	98	0.02%
Condition	Clinical Finding	26711	Chronic pharyngitis	668	0.13%
Condition	Clinical Finding	26727	Hematemesis	2233	0.44%
Condition	Clinical Finding	26935	Disorder of endocrine testis	109	0.02%
Condition	Clinical Finding	27674	Nausea and vomiting	6637	1.32%
Condition	Clinical Finding	28060	Streptococcal sore throat	623	0.12%
Condition	Clinical Finding	28179	Polyp of vocal cord or larynx	265	0.05%
Condition	Clinical Finding	28403	Gigantism and acromegaly	96	0.02%
Condition	Clinical Finding	28457	Hypertrophy of tonsils	223	0.04%
Condition	Clinical Finding	28461	Hypertrophy of salivary gland	80	0.02%
Condition	Clinical Finding	28779	Bleeding esophageal varices	54	0.01%
Condition	Clinical Finding	28974	Candidiasis of the esophagus	85	0.02%
Condition	Clinical Finding	29056	Sialoadenitis	702	0.14%
Condition	Clinical Finding	29735	Candidiasis of mouth	3122	0.62%
Condition	Clinical Finding	30133	Acute laryngitis	9756	1.94%
Condition	Clinical Finding	30163	Cellulitis and abscess of neck	100	0.02%
Condition	Clinical Finding	30234	Neck sprain	9014	1.79%
Condition	Clinical Finding	30284	Motion sickness	586	0.12%
Condition	Clinical Finding	30365	Panhypopituitarism	58	0.01%
Condition	Clinical Finding	30437	Gastro-esophageal reflux disease with esophagitis	15191	3.02%
Condition	Clinical Finding	30441	Sialolithiasis	412	0.08%
Condition	Clinical Finding	30679	Benign neoplasm of major salivary gland	68	0.01%
Condition	Clinical Finding	30683	Hemoglobin SS disease without crisis	69	0.01%
Condition	Clinical Finding	30753	Esophagitis	13812	2.75%
Condition	Clinical Finding	30968	Diabetes insipidus	121	0.02%
Condition	Clinical Finding	30978	Thalassemia	129	0.03%
Condition	Clinical Finding	31057	Disorder of pharynx	408	0.08%
Condition	Clinical Finding	312327	Acute myocardial infarction	6053	1.20%
Condition	Clinical Finding	312337	Arterial embolus and thrombosis	216	0.04%
Condition	Clinical Finding	312349	Venous varices	560	0.11%
Condition	Clinical Finding	312358	Chronic glomerulonephritis	106	0.02%
Condition	Clinical Finding	312437	Dyspnea	17774	3.54%
Condition	Clinical Finding	312622	Venous retinal branch occlusion	461	0.09%
Condition	Clinical Finding	312648	Benign essential hypertension	807	0.16%
Condition	Clinical Finding	312653	Acute myocarditis	65	0.01%
Condition	Clinical Finding	312723	Congenital heart disease	115	0.02%
Condition	Clinical Finding	312902	Benign intracranial hypertension	120	0.02%
Condition	Clinical Finding	312934	Atherosclerosis of aorta	128	0.03%
Condition	Clinical Finding	312950	IgE-mediated allergic asthma	83	0.02%
Condition	Clinical Finding	312998	Pain in thoracic spine	2063	0.41%
Condition	Clinical Finding	31317	Dysphagia	6207	1.24%
Condition	Clinical Finding	313217	Atrial fibrillation	3669	0.73%
Condition	Clinical Finding	313219	Phlebitis and thrombophlebitis	1382	0.28%
Condition	Clinical Finding	313223	Granulomatosis with polyangiitis	117	0.02%
Condition	Clinical Finding	313459	Sleep apnea	3825	0.76%
Condition	Clinical Finding	313504	Osler hemorrhagic telangiectasia syndrome	101	0.02%
Condition	Clinical Finding	313761	Central retinal vein occlusion	314	0.06%
Condition	Clinical Finding	313791	Bundle branch block	111	0.02%
Condition	Clinical Finding	313792	Paroxysmal tachycardia	392	0.08%
Condition	Clinical Finding	313878	Respiratory symptom	3591	0.71%
Condition	Clinical Finding	314054	Aortic valve disorder	346	0.07%
Condition	Clinical Finding	314059	Right bundle branch block	835	0.17%
Condition	Clinical Finding	314090	Mild pre-eclampsia	113	0.02%
Condition	Clinical Finding	314131	Degeneration of cervical intervertebral disc	1744	0.35%
Condition	Clinical Finding	314171	Low blood pressure reading	414	0.08%
Condition	Clinical Finding	314379	First degree atrioventricular block	359	0.07%
Condition	Clinical Finding	314383	Myocarditis	115	0.02%
Condition	Clinical Finding	314457	Congenital stenosis of aortic valve	64	0.01%
Condition	Clinical Finding	314617	Basophilia	93	0.02%
Condition	Clinical Finding	314658	Cardiomegaly	1190	0.24%
Condition	Clinical Finding	314659	Arteritis	160	0.03%
Condition	Clinical Finding	314665	Atrial flutter	531	0.11%
Condition	Clinical Finding	314666	Old myocardial infarction	2742	0.55%

Supplementary Tables

Table 3 continued from previous page

Domain ID	Concept Class	Concept ID	Concept Name	Total Number of Observations	Observation Frequency
Condition	Clinical Finding	314754	Wheezing	5261	1.05%
Condition	Clinical Finding	314962	Raynaud's disease	3212	0.64%
Condition	Clinical Finding	314965	Embolism and thrombosis of an arm or leg artery	337	0.07%
Condition	Clinical Finding	314971	Chronic respiratory failure	55	0.01%
Condition	Clinical Finding	315078	Palpitations	21656	4.31%
Condition	Clinical Finding	315085	Lymphadenopathy	4419	0.88%
Condition	Clinical Finding	31509	Primary malignant neoplasm of tonsil	102	0.02%
Condition	Clinical Finding	315273	Mitral valve stenosis	176	0.04%
Condition	Clinical Finding	315276	Vitreous hemorrhage	585	0.12%
Condition	Clinical Finding	315286	Chronic ischemic heart disease	7944	1.58%
Condition	Clinical Finding	315296	Preinfarction syndrome	4443	0.88%
Condition	Clinical Finding	315361	Orthopnea	151	0.03%
Condition	Clinical Finding	315558	Atherosclerosis of arteries of the extremities	373	0.07%
Condition	Clinical Finding	315564	Aortic valve regurgitation	694	0.14%
Condition	Clinical Finding	315832	Angina decubitus	55	0.01%
Condition	Clinical Finding	315922	Patent ductus arteriosus	83	0.02%
Condition	Clinical Finding	315947	Thoracic back sprain	1143	0.23%
Condition	Clinical Finding	31597	Acute laryngotracheitis	1019	0.20%
Condition	Clinical Finding	31602	Disorder of the larynx	377	0.08%
Condition	Clinical Finding	316084	Lymphadenitis	318	0.06%
Condition	Clinical Finding	31610	Disorder of esophagus	1139	0.23%
Condition	Clinical Finding	316127	Vitreous opacities	170	0.03%
Condition	Clinical Finding	316139	Heart failure	840	0.17%
Condition	Clinical Finding	316211	Acquired spondylolisthesis	588	0.12%
Condition	Clinical Finding	316428	Hypertrophic obstructive cardiomyopathy	115	0.02%
Condition	Clinical Finding	316429	Premature beats	575	0.11%
Condition	Clinical Finding	316457	Mallory-Weiss syndrome	313	0.06%
Condition	Clinical Finding	316501	Vascular disease of the skin	54	0.01%
Condition	Clinical Finding	316814	Hyperventilation	1047	0.21%
Condition	Clinical Finding	316822	Heart murmur	1601	0.32%
Condition	Clinical Finding	316866	Hypertensive disorder	26570	5.29%
Condition	Clinical Finding	316998	Left bundle branch block	747	0.15%
Condition	Clinical Finding	316999	Conduction disorder of the heart	147	0.03%
Condition	Clinical Finding	317002	Low blood pressure	1503	0.30%
Condition	Clinical Finding	317003	Compression of vein	61	0.01%
Condition	Clinical Finding	317009	Asthma	43294	8.62%
Condition	Clinical Finding	317109	Respiratory arrest	67	0.01%
Condition	Clinical Finding	317248	Disorder of hematopoietic structure	381	0.08%
Condition	Clinical Finding	317299	Disorder of vitreous body	517	0.10%
Condition	Clinical Finding	317305	Stricture of artery	412	0.08%
Condition	Clinical Finding	317309	Peripheral arterial occlusive disease	153	0.03%
Condition	Clinical Finding	317576	Coronary arteriosclerosis	14286	2.84%
Condition	Clinical Finding	317585	Aortic aneurysm	278	0.06%
Condition	Clinical Finding	317893	Paroxysmal supraventricular tachycardia	579	0.12%
Condition	Clinical Finding	318096	Secondary malignant neoplasm of lymph node	442	0.09%
Condition	Clinical Finding	318169	Varicose veins of lower extremity	18247	3.63%
Condition	Clinical Finding	318174	Phlebitis	716	0.14%
Condition	Clinical Finding	318186	Achalasia of esophagus	185	0.04%
Condition	Clinical Finding	31821	Disorder of endocrine system	339	0.07%
Condition	Clinical Finding	318222	Acute lymphadenitis	870	0.17%
Condition	Clinical Finding	318443	Arteriosclerotic vascular disease	191	0.04%
Condition	Clinical Finding	318448	Second degree atrioventricular block	171	0.03%
Condition	Clinical Finding	318566	Flushing	768	0.15%
Condition	Clinical Finding	318712	Peripheral circulatory disorder due to type 1 diabetes mellitus	76	0.02%
Condition	Clinical Finding	318736	Migraine	26542	5.28%
Condition	Clinical Finding	318772	Disorder of pericardium	290	0.06%
Condition	Clinical Finding	318776	Rheumatic mitral regurgitation	64	0.01%
Condition	Clinical Finding	318800	Gastroesophageal reflux disease	13584	2.70%
Condition	Clinical Finding	31884	Acquired diverticulum of esophagus	84	0.02%
Condition	Clinical Finding	319016	Retinal lattice degeneration	109	0.02%
Condition	Clinical Finding	319034	Hypertensive heart disease without congestive heart failure	82	0.02%
Condition	Clinical Finding	319041	Orthostatic hypotension	1459	0.29%

Table 3 continued from previous page

Domain ID	Concept Class	Concept ID	Concept Name	Total Number of Observations	Observation Frequency
Condition	Clinical Finding	319049	Acute respiratory failure	89	0.02%
Condition	Clinical Finding	31967	Nausea	7251	1.44%
Condition	Clinical Finding	319825	Rheumatic heart disease	72	0.01%
Condition	Clinical Finding	319826	Secondary hypertension	143	0.03%
Condition	Clinical Finding	319835	Congestive heart failure	799	0.16%
Condition	Clinical Finding	319843	Mitral valve disorder	236	0.05%
Condition	Clinical Finding	319844	Acute ischemic heart disease	344	0.07%
Condition	Clinical Finding	320073	Neutropenia	1236	0.25%
Condition	Clinical Finding	320116	Acute pericarditis	711	0.14%
Condition	Clinical Finding	320128	Essential hypertension	83173	16.55%
Condition	Clinical Finding	320136	Disorder of respiratory system	3575	0.71%
Condition	Clinical Finding	320416	Diseases of mitral and aortic valves	212	0.04%
Condition	Clinical Finding	320420	Vitreous degeneration	95	0.02%
Condition	Clinical Finding	320425	Heart block	260	0.05%
Condition	Clinical Finding	320536	Electrocardiogram abnormal	6011	1.20%
Condition	Clinical Finding	320739	Dissection of aorta	62	0.01%
Condition	Clinical Finding	320741	Thrombophlebitis	1110	0.22%
Condition	Clinical Finding	320744	Complete atrioventricular block	252	0.05%
Condition	Clinical Finding	320749	Polyarteritis nodosa	58	0.01%
Condition	Clinical Finding	321042	Cardiac arrest	178	0.04%
Condition	Clinical Finding	321052	Peripheral vascular disease	2118	0.42%
Condition	Clinical Finding	321074	Pre-existing hypertension complicating pregnancy, childbirth and puerperium	75	0.01%
Condition	Clinical Finding	321080	Hypertension complicating pregnancy, childbirth and the puerperium	188	0.04%
Condition	Clinical Finding	321107	Congenital insufficiency of aortic valve	64	0.01%
Condition	Clinical Finding	321119	Coarctation of aorta	67	0.01%
Condition	Clinical Finding	321318	Angina pectoris	13688	2.72%
Condition	Clinical Finding	321319	Cardiomyopathy	571	0.11%
Condition	Clinical Finding	321533	Immature white blood cells	643	0.13%
Condition	Clinical Finding	321588	Heart disease	903	0.18%
Condition	Clinical Finding	321596	Peripheral venous insufficiency	300	0.06%
Condition	Clinical Finding	321689	Apnea	82	0.02%
Condition	Clinical Finding	321876	Disorder of capillaries	124	0.02%
Condition	Clinical Finding	321887	Disorder of artery	272	0.05%
Condition	Clinical Finding	35609842	Reactive depression, prolonged single episode	1727	0.34%
Condition	Clinical Finding	35622958	Disorder in remission	155	0.03%
Condition	Clinical Finding	35623051	Autosomal dominant polycystic kidney disease	155	0.03%
Condition	Clinical Finding	35623148	Vomiting co-occurrent and due to infectious disease	78	0.02%
Condition	Clinical Finding	35624213	Secondary cataract	770	0.15%
Condition	Clinical Finding	35624440	Synovial cyst of ankle and foot	59	0.01%
Condition	Clinical Finding	35624442	Synovial cyst of hand	89	0.02%
Condition	Clinical Finding	35624449	Synovial cyst of wrist	101	0.02%
Condition	Clinical Finding	35624549	Mucous retention cyst of cervix uteri	489	0.10%
Condition	Clinical Finding	35624868	Lesion of face	235	0.05%
Condition	Clinical Finding	35626087	Disorder of right macula due to diabetes mellitus	195	0.04%
Condition	Clinical Finding	35626088	Disorder of left macula due to diabetes mellitus	212	0.04%
Condition	Clinical Finding	3654438	Dislocation of wrist	81	0.02%
Condition	Clinical Finding	3654543	Dislocation of joint of foot	86	0.02%
Condition	Clinical Finding	3654603	Cellulitis and abscess of hand	300	0.06%
Condition	Clinical Finding	3655355	Erectile dysfunction	18722	3.73%
Condition	Clinical Finding	3655359	Erectile dysfunction due to psychophysiologic disorder	155	0.03%
Condition	Clinical Finding	3656146	Strain of rectus abdominus muscle	58	0.01%
Condition	Clinical Finding	3657563	First trimester bleeding	289	0.06%
Condition	Clinical Finding	3657564	Burn of wrist and hand	426	0.08%
Condition	Clinical Finding	3661861	Allergic reaction to insect bite	674	0.13%
Condition	Clinical Finding	3663197	Glaucoma suspect	854	0.17%
Condition	Clinical Finding	36674599	Asthma never causes night symptoms	163	0.03%
Condition	Clinical Finding	36675035	Prematurity of infant	215	0.04%
Condition	Clinical Finding	36676308	Laryngeal stridor	53	0.01%
Condition	Clinical Finding	36676420	Paronychia due to ingrown nail	570	0.11%
Condition	Clinical Finding	36683296	Polycystic ovary	1546	0.31%
Condition	Clinical Finding	36683372	Delayed healing of surgical wound	68	0.01%

Supplementary Tables

Table 3 continued from previous page

Domain ID	Concept Class	Concept ID	Concept Name	Total Number of Observations	Observation Frequency
Condition	Clinical Finding	36712695	Suspected fetal abnormality affecting management of mother	153	0.03%
Condition	Clinical Finding	36712756	Accidental needle stick injury	81	0.02%
Condition	Clinical Finding	36712821	Postprocedural infection	3019	0.60%
Condition	Clinical Finding	36712845	Isolated proteinuria	216	0.04%
Condition	Clinical Finding	36712846	Persistent proteinuria	61	0.01%
Condition	Clinical Finding	36713107	Lumbar spondylolisthesis	287	0.06%
Condition	Clinical Finding	36713108	Lumbosacral spondylolisthesis	129	0.03%
Condition	Clinical Finding	36713630	Cystoscopy normal	2052	0.41%
Condition	Clinical Finding	36714126	Difficulty walking	81	0.02%
Condition	Clinical Finding	36714559	Disorder caused by alcohol	80	0.02%
Condition	Clinical Finding	36715479	Hydronephrosis due to ureteral obstruction	59	0.01%
Condition	Clinical Finding	36715575	Hypopituitarism following procedure	54	0.01%
Condition	Clinical Finding	36715577	Disorder of patella	423	0.08%
Condition	Clinical Finding	36715792	Acquired absence of breast	1192	0.24%
Condition	Clinical Finding	36715916	Internal hemorrhoids grade I	2185	0.43%
Condition	Clinical Finding	36716270	Cyst of kidney	548	0.11%
Condition	Clinical Finding	36716565	Multiple open wounds of head	65	0.01%
Condition	Clinical Finding	36716659	Recurrent aphthous stomatitis	60	0.01%
Condition	Clinical Finding	36716672	Strain of tendon of wrist	97	0.02%
Condition	Clinical Finding	36716700	Perforation and abscess of large intestine co-occurrent and due to diverticulitis	219	0.04%
Condition	Clinical Finding	36716712	Calculus of gallbladder without cholecystitis or cholangitis	5774	1.15%
Condition	Clinical Finding	36716939	Recurrent ulcer of mouth	389	0.08%
Condition	Clinical Finding	37016130	Lesion of oral mucosa	1005	0.20%
Condition	Clinical Finding	37016195	Leiomyoma	775	0.15%
Condition	Clinical Finding	37016310	Complicated internal hemorrhoid	1319	0.26%
Condition	Clinical Finding	37016775	Traumatic and/or non-traumatic injury of back	1732	0.34%
Condition	Clinical Finding	37017182	Joint pain of pelvic region	718	0.14%
Condition	Clinical Finding	37017184	On examination retinal arteriovenous nicking	101	0.02%
Condition	Clinical Finding	37017314	Lack of interest	269	0.05%
Condition	Clinical Finding	37018096	Infection of tooth socket	101	0.02%
Condition	Clinical Finding	37018196	Prediabetes	411	0.08%
Condition	Clinical Finding	37018293	Inflammatory disorder of jaw region	58	0.01%
Condition	Clinical Finding	37018674	Increased aspartate transaminase level	136	0.03%
Condition	Clinical Finding	37109843	Absence of lower limb	174	0.03%
Condition	Clinical Finding	37110172	Acquired mallet toe	56	0.01%
Condition	Clinical Finding	37110280	Incomplete legal abortion without complication	103	0.02%
Condition	Clinical Finding	37110367	Acquired varus deformity of limb	53	0.01%
Condition	Clinical Finding	37110444	Harmful pattern of use of nicotine	443	0.09%
Condition	Clinical Finding	37110545	Disorder of nervous system following procedure	55	0.01%
Condition	Clinical Finding	37110634	Strain of tendon of biceps brachii	68	0.01%
Condition	Clinical Finding	37116361	Accidental wound during procedure	778	0.15%
Condition	Clinical Finding	37116845	Acute severe refractory exacerbation of asthma	266	0.05%
Condition	Clinical Finding	37117206	Foreign body granuloma of soft tissue	118	0.02%
Condition	Clinical Finding	37118667	Disorder of anterior uveal tract	91	0.02%
Condition	Clinical Finding	37119138	Iron deficiency anemia due to blood loss	386	0.08%
Condition	Clinical Finding	37119234	Disorder of musculoskeletal system following procedure	83	0.02%
Condition	Clinical Finding	37203984	Anal fistula	1025	0.20%
Condition	Clinical Finding	37206139	Acute bronchitis co-occurrent with wheeze	2131	0.42%
Condition	Clinical Finding	37206607	Overactive bladder	155	0.03%
Condition	Clinical Finding	37206611	Idiopathic detrusor overactivity	3641	0.72%
Condition	Clinical Finding	37206940	Human papillomavirus deoxyribonucleic acid test positive	184	0.04%
Condition	Clinical Finding	37206941	Human papillomavirus deoxyribonucleic acid test negative	208	0.04%
Condition	Clinical Finding	372247	Candidal otitis externa	118	0.02%
Condition	Clinical Finding	372324	Eustachian tube disorder	372	0.07%
Condition	Clinical Finding	372328	Otitis media	10365	2.06%
Condition	Clinical Finding	372409	Sciatica	25451	5.07%
Condition	Clinical Finding	372448	Loss of consciousness	5219	1.04%
Condition	Clinical Finding	372479	Laceration of eye	90	0.02%
Condition	Clinical Finding	372547	Viral encephalitis	169	0.03%
Condition	Clinical Finding	372604	Movement disorder	78	0.02%
Condition	Clinical Finding	372610	Postconcussion syndrome	230	0.05%

Table 3 continued from previous page

Domain ID	Concept Class	Concept ID	Concept Name	Total Number of Observations	Observation Frequency
Condition	Clinical Finding	372615	Post-infectious encephalitis	53	0.01%
Condition	Clinical Finding	372629	Nonexudative age-related macular degeneration	216	0.04%
Condition	Clinical Finding	372635	Corneal degeneration	81	0.02%
Condition	Clinical Finding	372828	Herpes zoster ophthalmicus	269	0.05%
Condition	Clinical Finding	372887	Disorder of brain	197	0.04%
Condition	Clinical Finding	372888	Sciatic nerve lesion	94	0.02%
Condition	Clinical Finding	372894	Central serous chorioretinopathy	452	0.09%
Condition	Clinical Finding	372896	Stimulus deprivation amblyopia	200	0.04%
Condition	Clinical Finding	372897	Homonymous hemianopia	168	0.03%
Condition	Clinical Finding	372906	Blindness - both eyes	113	0.02%
Condition	Clinical Finding	372907	Scarred macula	125	0.02%
Condition	Clinical Finding	372914	Optic atrophy	195	0.04%
Condition	Clinical Finding	372924	Cerebral artery occlusion	123	0.02%
Condition	Clinical Finding	373035	Closed fracture of malar AND/OR maxillary bones	364	0.07%
Condition	Clinical Finding	373101	Injury of ulnar nerve	213	0.04%
Condition	Clinical Finding	37311123	Dislocation of digit of hand	729	0.15%
Condition	Clinical Finding	37311250	Bilateral lower limb edema	512	0.10%
Condition	Clinical Finding	37311319	Hypovolemia	724	0.14%
Condition	Clinical Finding	37312105	Strain of rotator cuff of shoulder	1130	0.22%
Condition	Clinical Finding	37312249	Pimple of skin	3355	0.67%
Condition	Clinical Finding	37312366	Repetitive motion disorder	843	0.17%
Condition	Clinical Finding	373127	Herpes simplex disciform keratitis	81	0.02%
Condition	Clinical Finding	373202	Corneal ulcer	1199	0.24%
Condition	Clinical Finding	373212	Esophoria	60	0.01%
Condition	Clinical Finding	373216	Disorder of mastoid	110	0.02%
Condition	Clinical Finding	373404	Ophthalmic herpes simplex	178	0.04%
Condition	Clinical Finding	373432	Benign neoplasm of nervous system	542	0.11%
Condition	Clinical Finding	373474	Diplopia	2112	0.42%
Condition	Clinical Finding	373478	Presbyopia	150	0.03%
Condition	Clinical Finding	373487	Ischemic optic neuropathy	105	0.02%
Condition	Clinical Finding	373488	Partial oculomotor nerve palsy	59	0.01%
Condition	Clinical Finding	373489	Congenital nystagmus	75	0.01%
Condition	Clinical Finding	373499	Disorder of eye region	1799	0.36%
Condition	Clinical Finding	373503	Transient cerebral ischemia	3492	0.69%
Condition	Clinical Finding	373638	Superficial injury of cornea	192	0.04%
Condition	Clinical Finding	373747	Extrapyramidal disease	223	0.04%
Condition	Clinical Finding	373748	Grand mal status	58	0.01%
Condition	Clinical Finding	373769	Nuclear cataract	100	0.02%
Condition	Clinical Finding	373785	Disorder of lens	72	0.01%
Condition	Clinical Finding	373786	Abnormal vision	235	0.05%
Condition	Clinical Finding	373852	Neuralgia	3397	0.68%
Condition	Clinical Finding	37397422	Asymptomatic periapical periodontitis	479	0.10%
Condition	Clinical Finding	37397686	Increased urine output	280	0.06%
Condition	Clinical Finding	37399074	Sample microscopy: crystals	215	0.04%
Condition	Clinical Finding	37399393	Symptomatic periapical periodontitis	175	0.03%
Condition	Clinical Finding	373995	Delirium	199	0.04%
Condition	Clinical Finding	374022	Hemiplegia	969	0.19%
Condition	Clinical Finding	374025	Retinal hemorrhage	764	0.15%
Condition	Clinical Finding	374027	Lesion of ulnar nerve	1088	0.22%
Condition	Clinical Finding	374028	Age related macular degeneration	598	0.12%
Condition	Clinical Finding	374034	Visual disturbance	3411	0.68%
Condition	Clinical Finding	374035	Keratitis	1585	0.32%
Condition	Clinical Finding	374036	Chronic conjunctivitis	645	0.13%
Condition	Clinical Finding	374043	Hereditary corneal dystrophy	89	0.02%
Condition	Clinical Finding	374044	Ptosis of eyelid	1346	0.27%
Condition	Clinical Finding	374053	Sudden hearing loss	160	0.03%
Condition	Clinical Finding	374055	Basilar artery syndrome	104	0.02%
Condition	Clinical Finding	374202	Penetrating wound of eye	86	0.02%
Condition	Clinical Finding	374221	Injury of brachial plexus	80	0.02%
Condition	Clinical Finding	374313	Neoplasm of uncertain behavior of soft tissues	163	0.03%
Condition	Clinical Finding	374317	Alcohol-induced psychosis	71	0.01%
Condition	Clinical Finding	374347	Blepharoconjunctivitis	1312	0.26%
Condition	Clinical Finding	374360	Disorder of optic nerve	78	0.02%

Supplementary Tables

Table 3 continued from previous page

Domain ID	Concept Class	Concept ID	Concept Name	Total Number of Observations	Observation Frequency
Condition	Clinical Finding	374362	Keratoconus	478	0.10%
Condition	Clinical Finding	374364	Disorder of middle ear	87	0.02%
Condition	Clinical Finding	374366	Sensorineural hearing loss	3979	0.79%
Condition	Clinical Finding	374375	Impacted cerumen	12450	2.48%
Condition	Clinical Finding	374631	Motor neuron disease	58	0.01%

1 Data and Implementation

Table 4: PheCode endpoints for the retinal risk and medical history models. The number of eligible participants, event counts and frequencies were calculated on the respective study populations.

Endpoint	Category	Phecode string	Number eligible (EHR)	Number of observations (EHR)	Endpoint frequency (EHR)	Number eligible (Retina)	Number of observations (Retina)	Endpoint frequency (Retina)
OMOP_4306655	Death	All-Cause Death	502453	37697	7.50%	61256	3490	5.70%
phecode_001	ID	Salmonella	500882	377	0.08%			
phecode_002	ID	Staphylococcus	499671	6587	1.32%	60945	658	1.08%
phecode_002-1	ID	Staphylococcus aureus	500332	4909	0.98%	61010	486	0.80%
phecode_003	ID	Escherichia coli	499141	9687	1.94%	60757	959	1.58%
phecode_004	ID	Streptococcus	497273	4847	0.97%	60584	494	0.82%
phecode_004-1	ID	Streptococcus pneumoniae	501390	863	0.17%			
phecode_004-2	ID	Group A Streptococcus	502357	332	0.07%			
phecode_004-3	ID	Group B Streptococcus	502252	394	0.08%			
phecode_005	ID	Mycobacteria	499403	787	0.16%	60899	113	0.19%
phecode_005-1	ID	Mycobacterium tuberculosis	499443	593	0.12%			
phecode_005-2	ID	Nontuberculous mycobacteria	502404	214	0.04%			
phecode_007	ID	Hemophilus infection	502164	927	0.18%	61220	107	0.17%
phecode_007-1	ID	Hemophilus influenzae	502196	907	0.18%	61223	103	0.17%
phecode_008	ID	Helicobacter [H. pylori]	498210	3253	0.65%	60757	364	0.60%
phecode_009	ID	Pseudomonas	502111	2075	0.41%	61210	231	0.38%
phecode_010	ID	Corynebacterium	500862	568	0.11%			
phecode_011	ID	Klebsiella	502336	1329	0.26%	61234	157	0.26%
phecode_012	ID	Proteus	502375	785	0.16%			
phecode_015	ID	Clostridium	499922	1839	0.37%	60986	198	0.32%
phecode_015-2	ID	Clostridium difficile	500932	1662	0.33%	61131	166	0.27%
phecode_019	ID	Treponema	502208	170	0.03%			
phecode_020	ID	Borrelia	502182	714	0.14%			
phecode_020-1	ID	Lyme disease	502184	711	0.14%			
phecode_024	ID	Pertussis	501358	329	0.07%			
phecode_025	ID	Enterococcus	502417	748	0.15%			
phecode_030	ID	Campylobacter	499207	3857	0.77%	60829	407	0.67%
phecode_050	ID	Enterovirus	501677	272	0.05%			
phecode_050-4	ID	Hand, foot, and mouth disease	502269	181	0.04%			
phecode_052	ID	Herpesvirus	463710	31789	6.86%	56105	3872	6.90%
phecode_052-1	ID	Herpes simplex	492140	6836	1.39%	59849	797	1.33%
phecode_052-3	ID	Varicella zoster virus	477534	27230	5.70%	57980	3342	5.76%
phecode_052-31	ID	Varicella [chickenpox]	495469	636	0.13%			
phecode_052-32	ID	Herpes zoster	484218	27057	5.59%	58885	3301	5.61%
phecode_052-4	ID	Infectious mononucleosis	497301	218	0.04%			
phecode_052-5	ID	Cytomegalovirus [CMV]	502349	261	0.05%			
phecode_054	ID	Hepatovirus	498714	912	0.18%	60742	111	0.18%
phecode_054-2	ID	Hepatitis B	501770	317	0.06%			
phecode_054-3	ID	Hepatitis C	501981	280	0.06%			
phecode_054-31	ID	Chronic hepatitis C	502019	283	0.06%			
phecode_054-5	ID	Hepatitis E	502451	104	0.02%			
phecode_055	ID	Poxvirus	501380	428	0.09%			
phecode_055-1	ID	Molluscum contagiosum	501472	406	0.08%			
phecode_056	ID	Human papillomavirus	472643	18560	3.93%	57047	2397	4.20%
phecode_056-1	ID	Plantar wart	492759	6265	1.27%	59921	840	1.40%
phecode_057	ID	Retrovirus	502025	197	0.04%			
phecode_057-1	ID	Human immunodeficiency virus	502026	190	0.04%			
phecode_058	ID	Pneumoviridae	502441	283	0.06%			
phecode_058-1	ID	Respiratory syncytial virus	502441	258	0.05%			
phecode_059	ID	Coronavirus	502459	8127	1.62%	61256	1055	1.72%
phecode_059-1	ID	COVID-19*	502460	7911	1.57%	61256	1017	1.66%
phecode_060	ID	Adenovirus	502334	131	0.03%			
phecode_061	ID	Influenza virus	491767	3866	0.79%	59986	362	0.60%
phecode_066	ID	Orthorubulavirus [Mumps]	499354	152	0.03%			
phecode_069	ID	Other specified viral infections	502191	192	0.04%			
phecode_070	ID	Candidiasis	465697	27144	5.83%	56152	3219	5.73%
phecode_074	ID	Aspergillosis	502280	411	0.08%			
phecode_076	ID	Pneumocystosis	502429	185	0.04%			
phecode_084	ID	Parasites	496120	2216	0.45%	60406	267	0.44%

Table 4 continued from previous page

Endpoint	Category	Phecode string	Number eligible (EHR)	Number of observations (EHR)	Endpoint frequency (EHR)	Number eligible (Retina)	Number of observations (Retina)	Endpoint frequency (Retina)
phecode_084-2	ID	Malaria [Plasmodium]	501375	306	0.06%			
phecode_084-4	ID	Trichomoniasis	500442	379	0.08%			
phecode_084-6	ID	Enterobiasis	501157	542	0.11%			
phecode_084-7	ID	Giardiasis	501773	409	0.08%			
phecode_086	ID	Pediculosis, acariasis and other infestations	498101	1921	0.39%	60691	247	0.41%
phecode_088	ID	Sexually transmitted disease	493998	2200	0.45%	60219	248	0.41%
phecode_089	ID	Infections	324948	99346	30.57%	37131	11862	31.95%
phecode_089-1	ID	Bacterial infections	464449	37907	8.16%	56465	4047	7.17%
phecode_089-2	ID	Viral infections	396152	69917	17.65%	46932	8466	18.04%
phecode_089-3	ID	Fungal infections	423350	59579	14.07%	50074	7223	14.42%
phecode_091	ID	Gangrene	502341	298	0.06%			
phecode_092	ID	Bacteremia, Sepsis, and SIRS	500511	16551	3.31%	60969	1652	2.71%
phecode_092-1	ID	Systemic inflammatory response syndrome	502460	771	0.15%			
phecode_092-2	ID	Sepsis	500592	16290	3.25%	60983	1635	2.68%
phecode_092-8	ID	Bacteremia	502443	132	0.03%			
phecode_095	ID	Sequela of infection	498668	2255	0.45%	60746	241	0.40%
phecode_096	ID	Contact or exposure to infectious agent	499867	3171	0.63%	60909	467	0.77%
phecode_097	ID	Drug resistant microorganisms	501975	1227	0.24%	61178	137	0.22%
phecode_097-1	ID	Methicillin resistant Staphylococcus aureus	501985	1094	0.22%	61179	117	0.19%
phecode_098	ID	Carrier or suspected carrier of infectious diseases	501765	3116	0.62%	61130	289	0.47%
phecode_098-2	ID	Carrier or suspected carrier of Staphylococcus aureus	502246	154	0.03%			
phecode_099	ID	Lab findings related to infections	496792	7397	1.49%	60344	776	1.29%
phecode_100	Neoplasms	Malignant neoplasm of the head and neck	501324	2282	0.46%	61126	250	0.41%
phecode_100-1	Neoplasms	Malignant neoplasm of the oral cavity	502131	746	0.15%			
phecode_100-12	Neoplasms	Malignant neoplasm of the tongue	502255	481	0.10%			
phecode_100-2	Neoplasms	Malignant neoplasm of the oropharynx	502307	434	0.09%			
phecode_100-5	Neoplasms	Malignant neoplasm of nasal cavities, middle ear, and accessory sinuses	502369	105	0.02%			
phecode_100-6	Neoplasms	Malignant neoplasm of the larynx	502230	361	0.07%			
phecode_100-7	Neoplasms	Malignant neoplasm of the pharynx	502435	128	0.03%			
phecode_100-8	Neoplasms	Malignant neoplasm of the lip	502415	101	0.02%			
phecode_100-9	Neoplasms	Malignant neoplasm of the salivary glands	502332	203	0.04%			
phecode_101	Neoplasms	Malignant neoplasm of the digestive organs	498324	15535	3.12%	60713	1652	2.72%
phecode_101-1	Neoplasms	Malignant neoplasm of the esophagus	502222	1654	0.33%	61222	176	0.29%
phecode_101-2	Neoplasms	Malignant neoplasm of stomach	502242	1407	0.28%	61222	138	0.23%
phecode_101-21	Neoplasms	Malignant neoplasm of cardia	502381	647	0.13%			
phecode_101-3	Neoplasms	Malignant neoplasm of the small intestine	502363	491	0.10%			
phecode_101-4	Neoplasms	Malignant neoplasm of the colon and rectum	499666	8051	1.61%	60906	882	1.45%
phecode_101-41	Neoplasms	Malignant neoplasm of the colon	500526	6050	1.21%	61007	670	1.10%
phecode_101-42	Neoplasms	Malignant neoplasm of the rectum	501377	2656	0.53%	61120	277	0.45%
phecode_101-5	Neoplasms	Malignant neoplasm of the anus and anal canal	502321	382	0.08%			
phecode_101-6	Neoplasms	Malignant neoplasm of the liver and intrahepatic bile ducts	502374	1099	0.22%	61238	118	0.19%
phecode_101-61	Neoplasms	Malignant neoplasm of the liver	502432	500	0.10%			
phecode_101-62	Neoplasms	Malignant neoplasm of the intrahepatic bile ducts	502435	566	0.11%			
phecode_101-7	Neoplasms	Malignant neoplasm of the gallbladder and extrahepatic bile ducts	502417	510	0.10%			
phecode_101-71	Neoplasms	Malignant neoplasm of the gallbladder	502451	181	0.04%			
phecode_101-8	Neoplasms	Malignant neoplasm of the pancreas	502371	1908	0.38%	61244	181	0.30%
phecode_102	Neoplasms	Malignant neoplasm of the thoracic and respiratory organs	501748	6810	1.36%	61165	686	1.12%
phecode_102-1	Neoplasms	Malignant neoplasm of the of bronchus and lung	502066	5832	1.16%	61199	594	0.97%
phecode_102-3	Neoplasms	Malignant neoplasm of the trachea	502401	313	0.06%			
phecode_102-5	Neoplasms	Malignant neoplasm of the heart, mediastinum, thymus, and pleura	502424	188	0.04%			
phecode_103	Neoplasms	Malignant neoplasm of the skin	488663	31041	6.35%	59332	3620	6.10%
phecode_103-1	Neoplasms	Melanomas of skin	499701	5002	1.00%	60872	574	0.94%

1 Data and Implementation

Table 4 continued from previous page

Endpoint	Category	Phecode string	Number eligible (EHR)	Number of observations (EHR)	Endpoint frequency (EHR)	Number eligible (Retina)	Number of observations (Retina)	Endpoint frequency (Retina)
phecode_103-2	Neoplasms	Keratinocyte carcinoma	496218	16640	3.35%	60277	1997	3.31%
phecode_103-21	Neoplasms	Basal cell carcinoma	496402	14745	2.97%	60306	1737	2.88%
phecode_103-22	Neoplasms	Squamous cell carcinoma of the skin	502253	2631	0.52%	61225	368	0.60%
phecode_103-3	Neoplasms	Carcinoma in situ of skin	501228	4852	0.97%	61116	597	0.98%
phecode_104	Neoplasms	Malignant sarcoma-related cancers	501446	2165	0.43%	61123	224	0.37%
phecode_104-1	Neoplasms	Malignant neoplasm of the bone and/or cartilage	502198	431	0.09%			
phecode_104-2	Neoplasms	Malignant neoplasm of retroperitoneum and peritoneum	502402	479	0.10%			
phecode_104-3	Neoplasms	Malignant neoplasm of other connective and soft tissue	501901	986	0.20%			
phecode_104-5	Neoplasms	Gastrointestinal stromal tumor*	502447	159	0.03%			
phecode_105	Neoplasms	Malignant neoplasm of the breast	491370	12724	2.59%	59833	1420	2.37%
phecode_105-1	Neoplasms	Malignant neoplasm of the breast, female	265414	8953	3.37%	32179	1080	3.36%
phecode_106	Neoplasms	Gynecological malignant neoplasms	267075	4215	1.58%	32501	442	1.36%
phecode_106-1	Neoplasms	Malignant neoplasm of external female genital organs and cervix	268852	854	0.32%			
phecode_106-11	Neoplasms	Malignant neoplasm of the vulva	273166	290	0.11%			
phecode_106-13	Neoplasms	Malignant neoplasm of the cervix	269034	531	0.20%			
phecode_106-2	Neoplasms	Malignant neoplasm of the uterus	272382	1966	0.72%	33158	226	0.68%
phecode_106-21	Neoplasms	Malignant neoplasm of endometrium	272466	1877	0.69%	33170	215	0.65%
phecode_106-3	Neoplasms	Malignant neoplasm of the ovary	272514	1646	0.60%	33167	169	0.51%
phecode_106-4	Neoplasms	Malignant neoplasm of the fallopian tube and uterine adnexa	273322	172	0.06%			
phecode_107	Neoplasms	Malignant neoplasm of male genitalia	224915	13780	6.13%	27396	1747	6.38%
phecode_107-1	Neoplasms	Malignant neoplasm of the penis	228915	330	0.14%			
phecode_107-2	Neoplasms	Malignant neoplasm of the prostate	225740	13373	5.92%	27506	1699	6.18%
phecode_107-3	Neoplasms	Malignant neoplasm of the testis	228455	164	0.07%			
phecode_108	Neoplasms	Malignant neoplasm of the urinary tract	500054	6664	1.33%	60947	641	1.05%
phecode_108-4	Neoplasms	Malignant neoplasm of the kidney	501887	2195	0.44%	61172	227	0.37%
phecode_108-41	Neoplasms	Malignant neoplasm of kidney, except pelvis	501960	2004	0.40%	61179	201	0.33%
phecode_108-42	Neoplasms	Malignant neoplasm of renal pelvis	502428	179	0.04%			
phecode_108-5	Neoplasms	Malignant neoplasm of the bladder	500998	3745	0.75%	61076	344	0.56%
phecode_108-7	Neoplasms	Malignant neoplasm of ureter	502424	257	0.05%			
phecode_109	Neoplasms	Malignant neoplasm of the eye, brain and other parts of central nervous system	502056	1653	0.33%	61204	188	0.31%
phecode_109-1	Neoplasms	Malignant neoplasm of eye	502301	194	0.04%			
phecode_109-16	Neoplasms	Malignant neoplasm of choroid	502364	117	0.02%			
phecode_109-3	Neoplasms	Malignant neoplasm of brain	502269	1324	0.26%	61235	151	0.25%
phecode_110	Neoplasms	Malignant neoplasm of the endocrine glands	501903	910	0.18%	61172	102	0.17%
phecode_110-1	Neoplasms	Malignant neoplasm of the thyroid	502050	573	0.11%			
phecode_110-4	Neoplasms	Malignant neoplasm of the pituitary gland and craniopharyngeal duct	502359	140	0.03%			
phecode_112	Neoplasms	Malignant neoplasm of other and ill-defined sites	491673	37928	7.71%	59854	4502	7.52%
phecode_112-1	Neoplasms	Mesothelioma*	502451	531	0.11%			
phecode_114	Neoplasms	Neuroendocrine tumors	502198	1002	0.20%	61226	124	0.20%
phecode_114-4	Neoplasms	Carcinoid tumors	502415	128	0.03%			
phecode_114-6	Neoplasms	Pheochromocytoma (including adrenal gland neoplasms)	502300	728	0.14%			
phecode_116	Neoplasms	Secondary malignant neoplasm	499427	21179	4.24%	60894	2120	3.48%
phecode_116-1	Neoplasms	Secondary malignancy of lymph nodes	500269	10438	2.09%	61015	909	1.49%
phecode_116-2	Neoplasms	Secondary malignancy of respiratory organs	502251	6536	1.30%	61229	621	1.01%
phecode_116-3	Neoplasms	Secondary malignant neoplasm of digestive systems	502233	4324	0.86%	61227	416	0.68%
phecode_116-4	Neoplasms	Secondary malignant neoplasm of liver	502146	7266	1.45%	61226	735	1.20%
phecode_116-5	Neoplasms	Secondary malignancy of brain/spine	502417	2284	0.45%	61250	250	0.41%
phecode_116-6	Neoplasms	Secondary malignancy of bone	502221	6141	1.22%	61214	681	1.11%
phecode_116-7	Neoplasms	Secondary malignant neoplasm of skin	502416	476	0.09%			
phecode_120	Neoplasms	Hemo onc - by cell of origin	499041	8133	1.63%	60774	897	1.48%
phecode_120-1	Neoplasms	Myeloid	501207	3748	0.75%	61099	429	0.70%
phecode_120-11	Neoplasms	Plasma cell	502389	182	0.04%			

Supplementary Tables

Table 4 continued from previous page

Endpoint	Category	Phecode string	Number eligible (EHR)	Number of observations (EHR)	Endpoint frequency (EHR)	Number eligible (Retina)	Number of observations (Retina)	Endpoint frequency (Retina)
phecode_120-12	Neoplasms	Monocyte	502451	151	0.03%			
phecode_120-13	Neoplasms	Erythroid	502142	409	0.08%			
phecode_120-2	Neoplasms	Lymphoid	500803	3995	0.80%	61011	443	0.73%
phecode_120-21	Neoplasms	Mature B-cell	501241	3586	0.72%	61081	381	0.62%
phecode_120-22	Neoplasms	Mature T-Cell	502357	310	0.06%			
phecode_121	Neoplasms	Leukemia	501694	2222	0.44%	61142	263	0.43%
phecode_121-1	Neoplasms	Acute leukemia	502283	766	0.15%			
phecode_121-11	Neoplasms	Acute lymphoid leukemia	502414	104	0.02%			
phecode_121-12	Neoplasms	Acute myeloid leukemia	502326	661	0.13%			
phecode_121-2	Neoplasms	Chronic leukemia	501911	1406	0.28%	61178	165	0.27%
phecode_121-21	Neoplasms	Chronic lymphoid leukemia	502022	1167	0.23%	61191	142	0.23%
phecode_121-22	Neoplasms	Chronic myeloid leukemia	502335	202	0.04%			
phecode_121-23	Neoplasms	Chronic myelomonocytic (monocytic) leukemia	502457	123	0.02%			
phecode_122	Neoplasms	Lymphoma	500922	3174	0.63%	61026	329	0.54%
phecode_122-1	Neoplasms	Hodgkin lymphoma	502038	289	0.06%			
phecode_122-2	Neoplasms	Non-Hodgkin lymphoma	501268	2946	0.59%	61071	296	0.48%
phecode_122-21	Neoplasms	Follicular lymphoma	502170	669	0.13%			
phecode_122-22	Neoplasms	Diffuse large B-cell lymphoma*	502273	1155	0.23%	61234	121	0.20%
phecode_122-24	Neoplasms	T-cell lymphoma	502441	158	0.03%			
phecode_123	Neoplasms	Multiple myeloma and malignant plasma cell neoplasms	502222	1253	0.25%	61225	147	0.24%
phecode_123-1	Neoplasms	Multiple myeloma	502251	1226	0.24%	61228	140	0.23%
phecode_124	Neoplasms	Myeloproliferative disorder	501593	2154	0.43%	61148	246	0.40%
phecode_124-3	Neoplasms	Polycythemia vera	502158	393	0.08%			
phecode_124-5	Neoplasms	Essential thrombocythemia	502156	861	0.17%	61204	103	0.17%
phecode_124-6	Neoplasms	Myelodysplastic syndrome	502373	705	0.14%			
phecode_124-7	Neoplasms	Chronic myeloproliferative disease	502321	434	0.09%			
phecode_125	Neoplasms	Other malignant neoplasms of lymphoid, hematopoietic and related tissue	502220	322	0.06%			
phecode_130	Neoplasms	Cancer (solid tumor, excluding BCC)	467151	69687	14.92%	56709	7801	13.76%
phecode_132	Neoplasms	Sequelae of cancer	501000	5493	1.10%	61101	667	1.09%
phecode_135	Neoplasms	Benign neoplasm of the head and neck	495301	5890	1.19%	60373	677	1.12%
phecode_135-1	Neoplasms	Benign neoplasm of the oral cavity	501304	1414	0.28%	61134	145	0.24%
phecode_135-12	Neoplasms	Benign neoplasm of the tongue	502303	127	0.03%			
phecode_135-16	Neoplasms	Benign neoplasm of the salivary glands	501978	645	0.13%			
phecode_135-5	Neoplasms	Benign neoplasm of the paranasal sinus and nasal cavity	496862	4203	0.85%	60541	496	0.82%
phecode_135-6	Neoplasms	Benign neoplasm of vocal cord or larynx	502019	311	0.06%			
phecode_136	Neoplasms	Benign neoplasm of the digestive organs	487512	52354	10.74%	59179	5933	10.03%
phecode_136-1	Neoplasms	Benign neoplasm of the esophagus	502356	163	0.03%			
phecode_136-2	Neoplasms	Benign neoplasm of stomach	500164	11273	2.25%	60935	1299	2.13%
phecode_136-3	Neoplasms	Benign neoplasm of the small intestine	502300	254	0.05%			
phecode_136-4	Neoplasms	Benign neoplasm of colon, rectum, anus and anal canal	489805	44166	9.02%	59473	5004	8.41%
phecode_136-41	Neoplasms	Benign neoplasm of the colon	493775	37524	7.60%	60008	4231	7.05%
phecode_136-42	Neoplasms	Benign neoplasm of rectum and anus	496403	17530	3.53%	60451	1990	3.29%
phecode_136-6	Neoplasms	Benign neoplasm of the liver and intrahepatic bile ducts	502390	183	0.04%			
phecode_136-8	Neoplasms	Benign neoplasm of the pancreas	502448	151	0.03%			
phecode_137	Neoplasms	Benign neoplasm of the thoracic and respiratory organs	502372	248	0.05%			
phecode_137-5	Neoplasms	Benign neoplasm of the heart, mediastinum, thymus, and pleura	502411	145	0.03%			
phecode_138	Neoplasms	Benign neoplasm of the skin	455462	34419	7.56%	54546	4328	7.93%
phecode_138-1	Neoplasms	Nevus, non-neoplastic	499732	3635	0.73%	60891	464	0.76%
phecode_138-2	Neoplasms	Melanocytic nevi*	473750	27003	5.70%	56817	3466	6.10%
phecode_139	Neoplasms	Benign sarcoma-related cancers	482142	18854	3.91%	58504	2174	3.72%
phecode_139-1	Neoplasms	Benign neoplasms of the bone and/or cartilage	501840	425	0.08%			
phecode_139-3	Neoplasms	Benign neoplasm of other connective and soft tissue	500712	1103	0.22%	61019	103	0.17%

Table 4 continued from previous page

Endpoint	Category	Phecode string	Number eligible (EHR)	Number of observations (EHR)	Endpoint frequency (EHR)	Number eligible (Retina)	Number of observations (Retina)	Endpoint frequency (Retina)
phecode_139-4	Neoplasms	Benign neoplasm of peripheral nerves*	502044	509	0.10%			
phecode_139-5	Neoplasms	Lipoma	487886	13458	2.76%	59248	1585	2.68%
phecode_139-51	Neoplasms	Lipomatosis*	502274	146	0.03%			
phecode_139-52	Neoplasms	Lipoma of intrathoracic organs	502089	1020	0.20%	61215	117	0.19%
phecode_139-53	Neoplasms	Lipoma of other skin subcutaneous tissue	496405	4832	0.97%	60407	520	0.86%
phecode_139-54	Neoplasms	Testicular lipoma	228722	1124	0.49%	27926	129	0.46%
phecode_139-6	Neoplasms	Hemangioma and lymphangioma	499054	4410	0.88%	60832	494	0.81%
phecode_139-61	Neoplasms	Hemangioma	499130	4312	0.86%	60843	488	0.80%
phecode_139-62	Neoplasms	Lymphangioma	502373	109	0.02%			
phecode_140	Neoplasms	Benign neoplasm of the breast	496908	1469	0.30%	60535	170	0.28%
phecode_142	Neoplasms	Lump or mass in breast or nonspecific abnormal breast exam	465844	21214	4.55%	56243	2813	5.00%
phecode_142-1	Neoplasms	Lump or mass in breast	472819	13671	2.89%	57209	1725	3.02%
phecode_142-2	Neoplasms	Abnormal mammogram	502256	186	0.04%			
phecode_142-21	Neoplasms	Mammographic microcalcification	502256	186	0.04%			
phecode_144	Neoplasms	Gynecological benign neoplasms	241780	15460	6.39%	29060	1869	6.43%
phecode_144-1	Neoplasms	Benign neoplasms of external female genital organs and cervix	265110	3365	1.27%	32266	373	1.16%
phecode_144-11	Neoplasms	Benign neoplasms of the vulva	273114	138	0.05%			
phecode_144-12	Neoplasms	Benign neoplasms of the vagina	272987	220	0.08%			
phecode_144-13	Neoplasms	Benign neoplasms of the cervix	265667	3065	1.15%	32334	337	1.04%
phecode_144-2	Neoplasms	Benign neoplasms of the uterus	249187	12639	5.07%	29962	1575	5.26%
phecode_144-21	Neoplasms	Leiomyoma of uterus	254139	8978	3.53%	30607	1165	3.81%
phecode_144-3	Neoplasms	Benign neoplasms of the ovary	272034	1565	0.58%	33112	158	0.48%
phecode_146	Neoplasms	Benign neoplasm of the genitourinary system	500174	6392	1.28%	60887	842	1.38%
phecode_146-2	Neoplasms	Benign neoplasm of the prostate	227092	5961	2.62%	27640	797	2.88%
phecode_146-4	Neoplasms	Benign neoplasm of the kidney	502414	192	0.04%			
phecode_146-5	Neoplasms	Benign neoplasm of the bladder	502352	183	0.04%			
phecode_148	Neoplasms	Benign neoplasm of the eye, brain and other parts of central nervous system	501101	3257	0.65%	61055	394	0.65%
phecode_148-1	Neoplasms	Benign neoplasm of eye	502087	1743	0.35%	61188	216	0.35%
phecode_148-16	Neoplasms	Benign neoplasm of choroid	502214	1599	0.32%	61212	206	0.34%
phecode_148-2	Neoplasms	Benign neoplasm of meninges (Meningioma)	502040	1011	0.20%	61202	121	0.20%
phecode_148-3	Neoplasms	Benign neoplasm of brain	502363	121	0.02%			
phecode_148-5	Neoplasms	Benign neoplasm of cranial nerve	502070	451	0.09%			
phecode_149	Neoplasms	Benign neoplasm of the endocrine glands	501385	1623	0.32%	61113	219	0.36%
phecode_149-1	Neoplasms	Benign neoplasm of the thyroid	502026	326	0.06%			
phecode_149-3	Neoplasms	Benign neoplasm of the parathyroid gland	502274	756	0.15%	61236	109	0.18%
phecode_149-4	Neoplasms	Benign neoplasm of the pituitary gland and craniopharyngeal duct	502015	540	0.11%			
phecode_153	Neoplasms	Benign neoplasm of other or unspecified sites	501562	580	0.12%			
phecode_159	Neoplasms	Genetic susceptibility to malignant neoplasm	502412	354	0.07%			
phecode_159-1	Neoplasms	Genetic susceptibility to malignant neoplasm of breast	502417	271	0.05%			
phecode_160	Blood	Nutritional anemias	485092	25919	5.34%	59133	2876	4.86%
phecode_160-1	Blood	Iron deficiency anemia	486933	23536	4.83%	59385	2600	4.38%
phecode_160-2	Blood	Megaloblastic anemia	500443	3728	0.74%	60985	417	0.68%
phecode_160-4	Blood	Other deficiency anemia	501912	642	0.13%			
phecode_161	Blood	Hemolytic anemias	502110	382	0.08%			
phecode_161-1	Blood	Intrinsic (hereditary) hemolytic anemias	502232	193	0.04%			
phecode_161-2	Blood	Extrinsic (acquired) hemolytic anemias	502308	258	0.05%			
phecode_161-21	Blood	Autoimmune hemolytic anemias [AIHA]	502377	185	0.04%			
phecode_162	Blood	Aplastic anemia	502198	1380	0.27%	61227	151	0.25%
phecode_162-8	Blood	Pancytopenia	502397	336	0.07%			
phecode_164	Blood	Anemia	474002	43808	9.24%	57646	4887	8.48%
phecode_164-1	Blood	Microcytic anemia	486928	23551	4.84%	59384	2603	4.38%
phecode_164-2	Blood	Macrocytic anemia	500091	3938	0.79%	60941	444	0.73%
phecode_164-3	Blood	Acute posthemorrhagic anemia	502185	215	0.04%			
phecode_164-6	Blood	Anemia secondary to chronic diseases and conditions	500455	1692	0.34%	60967	154	0.25%
phecode_164-62	Blood	Anemia in neoplastic disease	502310	777	0.15%			

Supplementary Tables

Table 4 continued from previous page

Endpoint	Category	Phecode string	Number eligible (EHR)	Number of observations (EHR)	Endpoint frequency (EHR)	Number eligible (Retina)	Number of observations (Retina)	Endpoint frequency (Retina)
phecode_165	Blood	Hemoglobinopathies	501169	771	0.15%	61042	136	0.22%
phecode_165-2	Blood	Thalassemia	501810	352	0.07%			
phecode_165-25	Blood	Thalassemia minor	501962	320	0.06%			
phecode_165-3	Blood	Hemoglobin C trait [Sickle-cell trait]	501961	308	0.06%			
phecode_168	Blood	Coagulation defects, purpura and other hemorrhagic conditions	498730	6510	1.31%	60713	812	1.34%
phecode_168-1	Blood	Hypo-coagulability	500190	3501	0.70%	60956	446	0.73%
phecode_168-11	Blood	Hereditary hypo-coagulability	502202	168	0.03%			
phecode_168-12	Blood	Hemorrhagic disorder due to intrinsic circulating anticoagulants	502388	165	0.03%			
phecode_168-15	Blood	Acquired coagulation factor deficiency	502449	103	0.02%			
phecode_168-18	Blood	Other nonthrombocytopenic purpura	502147	716	0.14%			
phecode_168-19	Blood	Spontaneous ecchymoses	500838	2271	0.45%	61045	258	0.42%
phecode_168-2	Blood	Hyper-coagulability	502061	834	0.17%	61180	123	0.20%
phecode_168-21	Blood	Primary hypercoagulable state [Primary thrombophilia]	502074	644	0.13%			
phecode_168-211	Blood	Activated protein C resistance*	502343	130	0.03%			
phecode_168-214	Blood	Antiphospholipid syndrome*	502345	134	0.03%			
phecode_168-3	Blood	Hereditary deficiency of other clotting factors	502154	373	0.07%			
phecode_168-4	Blood	Abnormal coagulation profile	502268	1197	0.24%	61214	147	0.24%
phecode_169	Blood	Platelet defects	500885	4918	0.98%	61004	675	1.11%
phecode_169-1	Blood	Thrombocytopenia	500918	4890	0.98%	61006	672	1.10%
phecode_169-11	Blood	Immune thrombocytopenic purpura [ITP]	501997	683	0.14%	61187	102	0.17%
phecode_169-14	Blood	Secondary thrombocytopenia	502373	282	0.06%			
phecode_170	Blood	Decreased white blood cell count	499695	8125	1.63%	60823	906	1.49%
phecode_170-1	Blood	Neutropenia	499774	7908	1.58%	60832	882	1.45%
phecode_170-13	Blood	Neutropenia due to infection	502444	291	0.06%			
phecode_170-19	Blood	Neutropenia NOS	501109	2424	0.48%	61046	282	0.46%
phecode_170-2	Blood	Lymphocytopenia	502370	270	0.05%			
phecode_171	Blood	Increased white blood cell count	501496	1398	0.28%	61081	179	0.29%
phecode_171-1	Blood	Lymphocytosis (symptomatic)	502043	740	0.15%	61148	102	0.17%
phecode_171-7	Blood	Eosinophilia	502155	460	0.09%			
phecode_171-9	Blood	Elevated white blood cell count [Leukocytosis] NOS	502277	148	0.03%			
phecode_172	Blood	Other disorders of white blood cells	500925	2100	0.42%	60871	267	0.44%
phecode_172-2	Blood	Genetic anomalies of leukocytes	502365	244	0.05%			
phecode_174	Blood	Diseases of spleen	501864	2260	0.45%	61174	254	0.42%
phecode_174-1	Blood	Hyposplenism*	502272	662	0.13%			
phecode_174-2	Blood	Splenomegaly	502136	1237	0.25%	61206	144	0.24%
phecode_174-6	Blood	Cyst of spleen*	502436	102	0.02%			
phecode_174-7	Blood	Infarction of spleen*	502402	236	0.05%			
phecode_175	Blood	Polycythemias	501961	1042	0.21%	61207	139	0.23%
phecode_175-2	Blood	Secondary polycythemia	502189	909	0.18%	61230	125	0.20%
phecode_176	Blood	Other diseases of blood and blood-forming organs	500275	2605	0.52%	60998	446	0.73%
phecode_177	Blood	Abnormality of the lymph nodes	490447	15769	3.22%	59576	1984	3.33%
phecode_177-1	Blood	Lymphadenitis	500388	705	0.14%			
phecode_177-13	Blood	Acute lymphadenitis	501186	315	0.06%			
phecode_177-2	Blood	Enlargement of lymph nodes [Lymphadenopathy]	493937	11966	2.42%	60041	1482	2.47%
phecode_177-3	Blood	Lymphangitis	502177	324	0.06%			
phecode_177-4	Blood	Lymphedema	501322	3172	0.63%	61098	417	0.68%
phecode_179	Blood	Immunodeficiencies	501978	1588	0.32%	61172	187	0.31%
phecode_179-1	Blood	Hypogammaglobulinemia NOS	502419	306	0.06%			
phecode_179-9	Blood	Immunodeficiency NOS	502137	1053	0.21%	61187	115	0.19%
phecode_180	Blood	Other disorders involving the immune mechanism	501834	3425	0.68%	61147	390	0.64%
phecode_180-3	Blood	Paraproteinemias	502346	2068	0.41%	61242	242	0.40%
phecode_180-31	Blood	Monoclonal gammopathy	502382	1926	0.38%	61245	220	0.36%
phecode_180-33	Blood	Macroglobulinemia [Waldenstrom macroglobulinemia]	502424	176	0.04%			

Table 4 continued from previous page

Endpoint	Category	Phecode string	Number eligible (EHR)	Number of observations (EHR)	Endpoint frequency (EHR)	Number eligible (Retina)	Number of observations (Retina)	Endpoint frequency (Retina)
phecode_181	Blood	Autoimmune disease	486677	14610	3.00%	59193	1578	2.67%
phecode_200	Endo	Disorders of thyroid gland	474595	28006	5.90%	57383	3122	5.44%
phecode_200-1	Endo	Hypothyroidism	483612	23512	4.86%	58570	2538	4.33%
phecode_200-12	Endo	Hypothyroidism due to drugs or iatrogenic causes	502385	146	0.03%			
phecode_200-13	Endo	Postprocedural hypothyroidism	501427	2149	0.43%	61124	239	0.39%
phecode_200-14	Endo	Hypothyroidism, not specified as secondary	484555	20668	4.27%	58685	2233	3.81%
phecode_200-2	Endo	Goiter	496277	5721	1.15%	60432	704	1.16%
phecode_200-21	Endo	Diffuse goiter	500008	1405	0.28%	60952	157	0.26%
phecode_200-22	Endo	Uninodular goiter [single thyroid nodule]	500825	2387	0.48%	61008	315	0.52%
phecode_200-23	Endo	Multinodular goiter	501066	1913	0.38%	61084	243	0.40%
phecode_200-3	Endo	Thyrotoxicosis [hyperthyroidism]	497060	4344	0.87%	60536	499	0.82%
phecode_200-31	Endo	Graves' disease [Toxic diffuse goiter]	501154	1071	0.21%	61096	117	0.19%
phecode_200-4	Endo	Thyroiditis	501012	938	0.19%	61062	107	0.18%
phecode_200-41	Endo	Hashimoto thyroiditis [Chronic lymphocytic thyroiditis]	501700	631	0.13%			
phecode_200-7	Endo	Iodine-deficiency related thyroid disorders*	502174	3621	0.72%	61221	367	0.60%
phecode_200-9	Endo	Abnormal thyroid function studies	500237	2200	0.44%	60940	268	0.44%
phecode_202	Endo	Diabetes mellitus	477657	37953	7.95%	57936	4259	7.35%
phecode_202-1	Endo	Type 1 diabetes	499103	3722	0.75%	60870	450	0.74%
phecode_202-2	Endo	Type 2 diabetes	478574	37460	7.83%	58037	4163	7.17%
phecode_202-3	Endo	Secondary diabetes	502420	197	0.04%			
phecode_202-32	Endo	Drug or chemical induced diabetes mellitus*	502427	172	0.03%			
phecode_202-4	Endo	Other specified diabetes*	496341	38221	7.70%	60444	5006	8.28%
phecode_203	Endo	Metabolic syndrome [Dysmetabolic syndrome X]	502326	191	0.04%			
phecode_204	Endo	Elevated blood glucose level	489096	37190	7.60%	59259	4823	8.14%
phecode_204-1	Endo	Impaired fasting glucose	499340	6011	1.20%	60651	720	1.19%
phecode_204-2	Endo	Impaired glucose tolerance (oral)	498520	13587	2.73%	60635	1698	2.80%
phecode_204-4	Endo	Prediabetes*	502049	570	0.11%	61194	108	0.18%
phecode_205	Endo	Hypoglycemia	501289	4032	0.80%	61110	416	0.68%
phecode_205-3	Endo	Drug-induced hypoglycemia*	502430	143	0.03%			
phecode_206	Endo	Disorders of pancreatic internal secretion (excl. DM)	502439	127	0.03%			
phecode_208	Endo	Disorders of parathyroid gland	501591	2953	0.59%	61134	374	0.61%
phecode_208-1	Endo	Hypoparathyroidism	502316	196	0.04%			
phecode_208-2	Endo	Hyperparathyroidism	501739	2844	0.57%	61147	362	0.59%
phecode_208-21	Endo	Primary hyperparathyroidism	502108	1771	0.35%	61195	232	0.38%
phecode_208-22	Endo	Secondary hyperparathyroidism	502367	687	0.14%			
phecode_209	Endo	Disorders of the pituitary gland and its hypothalamic control	501436	1609	0.32%	61120	197	0.32%
phecode_209-1	Endo	Pituitary hyperfunction	501901	882	0.18%	61174	102	0.17%
phecode_209-12	Endo	Syndrome of inappropriate secretion of antidiuretic hormone	502408	674	0.13%			
phecode_209-13	Endo	Hyperprolactinemia*	502087	119	0.02%			
phecode_209-2	Endo	Hypopituitarism	502003	643	0.13%			
phecode_209-21	Endo	Diabetes insipidus	502332	123	0.02%			
phecode_209-22	Endo	Hypopituitarism NOS	502114	535	0.11%			
phecode_209-23	Endo	Iatrogenic hypopituitarism*	502391	114	0.02%			
phecode_210	Endo	Cushing's syndrome	502342	129	0.03%			
phecode_211	Endo	Disorders of adrenal glands	501939	1430	0.28%	61205	175	0.29%
phecode_211-1	Endo	Hyperaldosteronism	502335	101	0.02%			
phecode_211-2	Endo	Adrenocortical insufficiency	502165	774	0.15%			
phecode_211-21	Endo	Primary adrenocortical insufficiency	502274	277	0.06%			
phecode_211-22	Endo	Drug-induced adrenocortical insufficiency*	502452	107	0.02%			
phecode_214	Endo	Ovarian dysfunction	272065	479	0.18%			
phecode_214-1	Endo	Primary ovarian failure	272177	455	0.17%			
phecode_214-11	Endo	Premature menopause	272385	422	0.15%			
phecode_215	Endo	Testicular dysfunction	228738	641	0.28%	27923	103	0.37%
phecode_215-1	Endo	Testicular hypofunction	228810	612	0.27%			
phecode_229	Endo	Other endocrine disorders	502037	336	0.07%			

Supplementary Tables

Table 4 continued from previous page

Endpoint	Category	Phecode string	Number eligible (EHR)	Number of observations (EHR)	Endpoint frequency (EHR)	Number eligible (Retina)	Number of observations (Retina)	Endpoint frequency (Retina)
phecode_230	Metab	Malnutrition and underweight	495351	26056	5.26%	60371	3065	5.08%
phecode_230-1	Metab	Protein-calorie malnutrition	502320	719	0.14%			
phecode_230-2	Metab	Abnormal loss of weight and underweight	496512	21092	4.25%	60500	2452	4.05%
phecode_230-21	Metab	Abnormal weight loss	497079	20015	4.03%	60533	2341	3.87%
phecode_230-22	Metab	Underweight	501869	1477	0.29%	61223	154	0.25%
phecode_230-3	Metab	Anorexia	501209	6442	1.29%	61116	776	1.27%
phecode_230-4	Metab	Cachexia	502446	313	0.06%			
phecode_230-5	Metab	Early satiety	502460	128	0.03%			
phecode_232	Metab	Vitamin deficiencies	497025	35843	7.21%	60415	5117	8.47%
phecode_232-1	Metab	Vitamin A deficiency	501753	113	0.02%			
phecode_232-2	Metab	Vitamin B group deficiency	498666	13489	2.71%	60705	1585	2.61%
phecode_232-27	Metab	Vitamin B12 deficiency	500830	2562	0.51%	61031	272	0.45%
phecode_232-29	Metab	Folate deficiency [Vitamin B9]	502091	1230	0.24%	61208	153	0.25%
phecode_232-4	Metab	Vitamin D deficiency	501521	25142	5.01%	61062	3911	6.40%
phecode_234	Metab	Other nutritional deficiencies	500482	4532	0.91%	61057	1280	2.10%
phecode_236	Metab	Overweight and obesity	471690	52586	11.15%	57314	6340	11.06%
phecode_236-1	Metab	Obesity	471819	52464	11.12%	57338	6324	11.03%
phecode_236-11	Metab	Morbid obesity	501516	2668	0.53%	61109	283	0.46%
phecode_236-2	Metab	Localized adiposity	502292	244	0.05%			
phecode_237	Metab	Abnormal weight gain	498418	3510	0.70%	60755	369	0.61%
phecode_239	Metab	Hyperlipidemia	442105	78585	17.78%	52389	10050	19.18%
phecode_239-1	Metab	Hypercholesterolemia	451646	73101	16.19%	53958	9407	17.43%
phecode_239-11	Metab	Pure hypercholesterolemia	453668	73199	16.13%	54293	9463	17.43%
phecode_239-12	Metab	Familial hypercholesterolemia*	501395	1535	0.31%	61068	186	0.30%
phecode_239-2	Metab	Hyperglyceridemia	499001	2375	0.48%	60716	258	0.42%
phecode_239-21	Metab	Pure hyperglyceridemia	502105	904	0.18%			
phecode_239-3	Metab	Mixed hyperlipidemia	499330	1529	0.31%	60770	195	0.32%
phecode_240	Metab	Disorders of amino-acid transport and metabolism	502298	168	0.03%			
phecode_241	Metab	Disorders of carbohydrate metabolism	501854	493	0.10%			
phecode_242	Metab	Lipid storage disorders	502344	155	0.03%			
phecode_244	Metab	Disorders of lipoprotein metabolism and other lipidemias	501101	2859	0.57%	61043	431	0.71%
phecode_244-4	Metab	Lipodystrophy, not elsewhere classified	502395	242	0.05%			
phecode_247	Metab	Disorders of mineral metabolism and mineral deficiencies	484323	34180	7.06%	59033	3877	6.57%
phecode_247-3	Metab	Disorder of phosphorus metabolism	502407	1805	0.36%	61252	218	0.36%
phecode_247-4	Metab	Disorders of magnesium metabolism	502406	2999	0.60%	61244	358	0.58%
phecode_247-42	Metab	Hypomagnesemia*	502431	576	0.11%			
phecode_247-5	Metab	Disorders of calcium metabolism	501174	6268	1.25%	61060	737	1.21%
phecode_247-51	Metab	Hypocalcemia	501933	1293	0.26%	61187	132	0.22%
phecode_247-52	Metab	Hypercalcemia	502113	1681	0.33%	61189	222	0.36%
phecode_247-7	Metab	Disorders of iron metabolism	485646	26868	5.53%	59229	2985	5.04%
phecode_247-71	Metab	Hemochromatosis	502159	743	0.15%			
phecode_247-711	Metab	Hereditary hemochromatosis	502324	222	0.04%			
phecode_247-72	Metab	Iron deficiency	486123	25754	5.30%	59285	2852	4.81%
phecode_248	Metab	Disorders of plasma-protein metabolism, NEC	502262	1136	0.23%	61231	129	0.21%
phecode_248-1	Metab	Alpha-1-antitrypsin deficiency	502375	132	0.03%			
phecode_249	Metab	Amyloidosis	502390	550	0.11%			
phecode_249-1	Metab	Cerebral amyloid angiopathy*	502455	211	0.04%			
phecode_251	Metab	Disorders of bilirubin excretion	500937	1590	0.32%	61032	182	0.30%
phecode_251-1	Metab	Gilbert syndrome*	500996	1499	0.30%	61042	173	0.28%
phecode_252	Metab	Other and unspecified disorders of metabolism	502263	345	0.07%			
phecode_256	Metab	Disorders of fluid, electrolyte and acid-base balance	495716	36496	7.36%	60493	4817	7.96%
phecode_256-1	Metab	Hyperosmolality and/or hypernatremia	502058	2042	0.41%	61205	319	0.52%
phecode_256-2	Metab	Hyposmolality and/or hyponatremia	499993	14641	2.93%	61002	2315	3.79%
phecode_256-3	Metab	Mixed disorder of acid-base balance	502192	4656	0.93%	61220	492	0.80%
phecode_256-31	Metab	Acidosis	502208	4057	0.81%	61222	437	0.71%
phecode_256-32	Metab	Alkalosis	502445	680	0.14%			
phecode_256-4	Metab	Hyperkalemia [Hyperpotassemia]	501796	5418	1.08%	61159	616	1.01%

Table 4 continued from previous page

Endpoint	Category	Phecode string	Number eligible (EHR)	Number of observations (EHR)	Endpoint frequency (EHR)	Number eligible (Retina)	Number of observations (Retina)	Endpoint frequency (Retina)
phecode_256-5	Metab	Hypokalemia [Hypopotassemia]	501339	8409	1.68%	61144	960	1.57%
phecode_256-6	Metab	Fluid overload	502369	3073	0.61%	61244	343	0.56%
phecode_256-7	Metab	Volume depletion	501618	10584	2.11%	61134	1169	1.91%
phecode_256-71	Metab	Dehydration	502302	845	0.17%			
phecode_257	Metab	Polydipsia	500601	1547	0.31%	61018	190	0.31%
phecode_280	Mental	Substance related disorders	489557	20707	4.23%	59584	2413	4.05%
phecode_280-1	Mental	Alcohol use disorders	490806	19603	3.99%	59746	2266	3.79%
phecode_280-11	Mental	Alcohol abuse	497292	11089	2.23%	60633	1212	2.00%
phecode_280-12	Mental	Alcohol dependence	497114	6285	1.26%	60514	684	1.13%
phecode_280-13	Mental	Alcoholic liver disease	501590	1776	0.35%	61146	186	0.30%
phecode_280-14	Mental	Alcoholic pancreatitis	502334	215	0.04%			
phecode_280-2	Mental	Opioid related disorders	502246	248	0.05%			
phecode_280-22	Mental	Opioid dependence	502277	181	0.04%			
phecode_280-3	Mental	Cannabis related disorders	502262	386	0.08%			
phecode_280-31	Mental	Cannabis abuse	502374	308	0.06%			
phecode_280-4	Mental	Sedative, hypnotic, or anxiolytic related disorders	502041	278	0.06%			
phecode_280-42	Mental	Sedative, hypnotic or anxiolytic-related dependence	502054	264	0.05%			
phecode_280-8	Mental	Other psychoactive substance related disorders	501606	904	0.18%	61147	145	0.24%
phecode_280-81	Mental	Other psychoactive substance abuse	502156	235	0.05%			
phecode_280-82	Mental	Other psychoactive substance dependence	501833	710	0.14%	61165	106	0.17%
phecode_281	Mental	Substance abuse, dependence, and withdrawal	400986	31733	7.91%	47197	3749	7.94%
phecode_281-1	Mental	Substance abuse	496902	11529	2.32%	60581	1274	2.10%
phecode_281-2	Mental	Substance dependence	403510	26850	6.65%	47476	3215	6.77%
phecode_281-21	Mental	Substance withdrawal	502091	440	0.09%			
phecode_282-1	Mental	Current tobacco use and nicotine dependence	428880	21120	4.92%	51019	2415	4.73%
phecode_283	Mental	Other behavioral problems	379487	79939	21.07%	43721	9143	20.91%
phecode_283-3	Mental	High risk sexual behavior	501986	267	0.05%			
phecode_283-4	Mental	Patient's noncompliance with medical treatment and regimen	496888	19329	3.89%	60438	2414	3.99%
phecode_283-8	Mental	Other problems related to lifestyle	388872	71273	18.33%	45188	7971	17.64%
phecode_284	Mental	Suicide ideation and attempt or self harm	498884	5253	1.05%	60757	664	1.09%
phecode_284-1	Mental	Suicidal ideations	499748	3959	0.79%	60875	536	0.88%
phecode_284-2	Mental	Suicide and self-inflicted harm	501423	1931	0.39%	61109	206	0.34%
phecode_284-29	Mental	Intentional self-harm*	501458	1914	0.38%	61112	203	0.33%
phecode_286	Mental	Mood [affective] disorders	428189	38738	9.05%	51563	4730	9.17%
phecode_286-1	Mental	Bipolar disorder	501038	1048	0.21%	61050	110	0.18%
phecode_286-2	Mental	Major depressive disorder	435448	39083	8.98%	52388	4781	9.13%
phecode_286-21	Mental	Major depressive disorder, recurrent	497154	2007	0.40%	60532	268	0.44%
phecode_286-3	Mental	Premenstrual dysphoric disorder	266025	489	0.18%			
phecode_286-4	Mental	Dysthymic disorder	498047	537	0.11%			
phecode_287	Mental	Psychotic disorder	500446	1687	0.34%	60982	169	0.28%
phecode_287-1	Mental	Schizophrenia	501435	545	0.11%			
phecode_287-2	Mental	Schizoaffective disorder	502194	176	0.04%			
phecode_287-4	Mental	Delusional disorders	501832	803	0.16%			
phecode_287-5	Mental	Drug-induced psychotic disorder	502173	245	0.05%			
phecode_288	Mental	Anxiety disorders	450706	38089	8.45%	54383	4845	8.91%
phecode_288-2	Mental	Panic disorder [episodic paroxysmal anxiety]	495730	4220	0.85%	60359	546	0.90%
phecode_288-3	Mental	Generalized anxiety disorder	488225	8475	1.74%	59535	1120	1.88%
phecode_288-4	Mental	Phobic disorders	497607	5257	1.06%	60610	673	1.11%
phecode_288-41	Mental	Agoraphobia	501817	174	0.03%			
phecode_289	Mental	Obsessive-compulsive disorder	501112	536	0.11%			
phecode_290	Mental	Reaction to severe stress, and adjustment disorders	478418	18280	3.82%	58246	2233	3.83%
phecode_290-1	Mental	Posttraumatic stress disorder	501526	961	0.19%	61119	110	0.18%
phecode_291	Mental	Dissociative, conversion and factitious disorders	501847	603	0.12%			
phecode_292	Mental	Somatiform disorders	500197	1009	0.20%	61016	126	0.21%
phecode_293	Mental	Eating disorders	500840	519	0.10%			
phecode_293-1	Mental	Anorexia nervosa	501725	109	0.02%			

Supplementary Tables

Table 4 continued from previous page

Endpoint	Category	Phecode string	Number eligible (EHR)	Number of observations (EHR)	Endpoint frequency (EHR)	Number eligible (Retina)	Number of observations (Retina)	Endpoint frequency (Retina)
phecode_293-4	Mental	Polyphagia	502214	273	0.05%			
phecode_294	Mental	Sexual dysfunction and disorders	479648	26856	5.60%	57966	3405	5.87%
phecode_296	Mental	Specific personality disorders	501083	494	0.10%			
phecode_296-4	Mental	Borderline personality disorder	502206	260	0.05%			
phecode_299	Mental	Mental disorder, not otherwise specified	498234	2701	0.54%	60647	309	0.51%
phecode_308	Mental	Signs and symptoms involving emotional state	467448	34220	7.32%	56265	4269	7.59%
phecode_308-1	Mental	Irritability	501397	1136	0.23%	61108	135	0.22%
phecode_308-3	Mental	Emotional lability	501161	1253	0.25%	61070	163	0.27%
phecode_308-4	Mental	Demoralization and apathy	502059	812	0.16%	61197	176	0.29%
phecode_308-5	Mental	Nervousness	501127	1480	0.30%	61119	182	0.30%
phecode_308-6	Mental	Excessive crying of child, adolescent, or adult	501587	799	0.16%	61117	111	0.18%
phecode_308-7	Mental	Restlessness and agitation*	501589	2023	0.40%	61145	209	0.34%
phecode_320	Neuro	Meningitis	500464	471	0.09%			
phecode_320-1	Neuro	Infective meningitis	500898	322	0.06%			
phecode_320-11	Neuro	Bacterial meningitis	501951	184	0.04%			
phecode_320-12	Neuro	Viral meningitis	501397	140	0.03%			
phecode_320-3	Neuro	Meningitis NOS	501993	159	0.03%			
phecode_321	Neuro	Encephalitis, myelitis and encephalomyelitis	501641	687	0.14%			
phecode_321-1	Neuro	Encephalitis	501901	291	0.06%			
phecode_321-12	Neuro	Viral encephalitis	502156	231	0.05%			
phecode_321-2	Neuro	Myelitis	502227	130	0.03%			
phecode_321-21	Neuro	Acute (transverse) myelitis	502307	106	0.02%			
phecode_322	Neuro	Other CNS infection	502271	437	0.09%			
phecode_322-4	Neuro	Intracranial and intraspinal abscess	502367	200	0.04%			
phecode_323	Neuro	Systemic atrophies primarily affecting the central nervous system	502127	1128	0.22%	61217	130	0.21%
phecode_323-1	Neuro	Hereditary ataxia	502246	379	0.08%			
phecode_323-3	Neuro	Motor neuron disease	502387	684	0.14%			
phecode_323-31	Neuro	Amyotrophic lateral sclerosis [ALS]	502444	190	0.04%			
phecode_324	Neuro	Extrapyramidal and movement disorders	487693	21890	4.49%	59234	2523	4.26%
phecode_324-1	Neuro	Parkinsonism	501677	4363	0.87%	61158	495	0.81%
phecode_324-11	Neuro	Parkinson's disease (Primary)	501702	4196	0.84%	61159	481	0.79%
phecode_324-12	Neuro	Secondary parkinsonism	502426	343	0.07%			
phecode_324-2	Neuro	Degenerative diseases of the basal ganglia (excluding parkinsons)	502440	407	0.08%			
phecode_324-21	Neuro	Progressive supranuclear ophthalmoplegia [Steele-Richardson-Olszewski]*	502456	215	0.04%			
phecode_324-3	Neuro	Dystonia	493180	7610	1.54%	59966	913	1.52%
phecode_324-34	Neuro	Torticollis	494329	4830	0.98%	60102	614	1.02%
phecode_324-36	Neuro	Blepharospasm	502229	377	0.08%			
phecode_324-4	Neuro	Tremor	499622	8421	1.69%	60896	975	1.60%
phecode_324-41	Neuro	Essential tremor*	501800	2791	0.56%	61159	305	0.50%
phecode_324-5	Neuro	Myoclonus	502320	401	0.08%			
phecode_324-8	Neuro	Restless legs syndrome	500491	3891	0.78%	60976	413	0.68%
phecode_325	Neuro	Symptoms and signs related to movement disorders	496250	22452	4.52%	60560	2492	4.11%
phecode_325-1	Neuro	Abnormal involuntary movements	501769	1279	0.25%	61166	147	0.24%
phecode_325-12	Neuro	Fasciculation*	501977	689	0.14%			
phecode_325-2	Neuro	Abnormality of gait and mobility	499946	20734	4.15%	60926	2272	3.73%
phecode_325-21	Neuro	Ataxic gait*	502363	335	0.07%			
phecode_325-23	Neuro	Unsteadiness on feet*	501606	3016	0.60%	61144	300	0.49%
phecode_325-3	Neuro	Lack of coordination	498822	3863	0.77%	60881	428	0.70%
phecode_326	Neuro	Demyelinating diseases of the central nervous system	500417	887	0.18%			
phecode_326-1	Neuro	Multiple sclerosis	500662	761	0.15%			
phecode_327	Neuro	Other degenerative diseases of nervous system	502156	2813	0.56%	61207	303	0.50%
phecode_328	Neuro	Dementias and cerebral degeneration	502284	10001	1.99%	61233	968	1.58%
phecode_328-1	Neuro	Alzheimer's disease	502395	5432	1.08%	61246	535	0.87%
phecode_328-2	Neuro	Frontotemporal dementia	502454	333	0.07%			
phecode_328-4	Neuro	Dementia with Lewy bodies	502457	570	0.11%			
phecode_328-7	Neuro	Vascular dementia	502432	2304	0.46%	61251	198	0.32%

1 Data and Implementation

Table 4 continued from previous page

Endpoint	Category	Phecode string	Number eligible (EHR)	Number of observations (EHR)	Endpoint frequency (EHR)	Number eligible (Retina)	Number of observations (Retina)	Endpoint frequency (Retina)
phecode_328-8	Neuro	Dementia in conditions classified elsewhere	502392	5056	1.01%	61244	520	0.85%
phecode_328-9	Neuro	Dementia NOS	502383	5185	1.03%	61246	491	0.80%
phecode_329	Neuro	Symptoms and signs involving cognitive functions and awareness	494132	35145	7.11%	60136	3963	6.59%
phecode_329-1	Neuro	Memory loss	499014	15719	3.15%	60736	1916	3.15%
phecode_329-4	Neuro	Other specified cognitive deficit	500427	3224	0.64%	60970	365	0.60%
phecode_329-41	Neuro	Attention and concentration deficit	501758	890	0.18%	61163	123	0.20%
phecode_329-42	Neuro	Cognitive communication deficit	501131	2360	0.47%	61067	246	0.40%
phecode_329-5	Neuro	Mild cognitive impairment, so stated	502390	3121	0.62%	61245	349	0.57%
phecode_329-6	Neuro	Transient global amnesia	502027	1412	0.28%	61199	171	0.28%
phecode_329-8	Neuro	Altered mental status, unspecified	501573	563	0.11%			
phecode_329-9	Neuro	Delirium	502344	6836	1.36%	61237	609	0.99%
phecode_330	Neuro	Epilepsy, recurrent seizures, convulsions	494833	6375	1.29%	60275	648	1.08%
phecode_330-1	Neuro	Epilepsy	496820	3843	0.77%	60529	417	0.69%
phecode_330-11	Neuro	Generalized epilepsy	500993	1123	0.22%	61044	137	0.22%
phecode_330-12	Neuro	Partial epilepsy	501520	904	0.18%	61135	112	0.18%
phecode_330-3	Neuro	Convulsions	498957	4677	0.94%	60798	475	0.78%
phecode_331	Neuro	Headache	431867	45870	10.62%	51822	5529	10.67%
phecode_331-1	Neuro	Tension headache	495365	5637	1.14%	60317	720	1.19%
phecode_331-3	Neuro	Headache syndromes, non migraine	502277	892	0.18%	61221	141	0.23%
phecode_331-4	Neuro	Cluster headaches	501383	516	0.10%			
phecode_331-6	Neuro	Migraine	473384	13329	2.82%	57495	1590	2.77%
phecode_331-61	Neuro	Migraine with aura	499712	2920	0.58%	60938	362	0.59%
phecode_331-62	Neuro	Hemiplegic migraine	502285	127	0.03%			
phecode_331-7	Neuro	Drug induced headache	502360	238	0.05%			
phecode_331-8	Neuro	Headache NOS	456388	40688	8.92%	55021	4880	8.87%
phecode_333	Neuro	Sleep disorders	469527	39471	8.41%	56668	4822	8.51%
phecode_333-1	Neuro	Sleep apnea	497726	9970	2.00%	60614	1226	2.02%
phecode_333-11	Neuro	Obstructive sleep apnea	500744	5036	1.01%	61002	663	1.09%
phecode_333-2	Neuro	Insomnia	484832	17645	3.64%	58769	2197	3.74%
phecode_333-3	Neuro	Hypersomnia	501993	417	0.08%			
phecode_333-4	Neuro	Circadian rhythm sleep disorder	501240	1105	0.22%	61095	117	0.19%
phecode_333-43	Neuro	Circadian rhythm sleep disorder, jet lag or shift work type	502356	116	0.02%			
phecode_333-5	Neuro	Parasomnia and sleep arousal Disorders	502111	671	0.13%			
phecode_334	Neuro	Disorders of other cranial nerves	496122	7112	1.43%	60446	794	1.31%
phecode_334-1	Neuro	Trigeminal nerve disorders [CN5]	499552	3323	0.67%	60908	368	0.60%
phecode_334-11	Neuro	Trigeminal neuralgia	500521	2777	0.55%	61033	298	0.49%
phecode_334-12	Neuro	Atypical face pain	501558	498	0.10%			
phecode_334-2	Neuro	Facial nerve disorders and weakness	499422	2850	0.57%	60843	324	0.53%
phecode_334-21	Neuro	Bell's palsy	499833	2193	0.44%	60906	267	0.44%
phecode_334-23	Neuro	Facial weakness	502259	415	0.08%			
phecode_334-24	Neuro	Clonic hemifacial spasm*	502354	194	0.04%			
phecode_334-4	Neuro	Disorders of oculomotor nerves	502104	899	0.18%			
phecode_334-41	Neuro	Third [oculomotor] nerve palsy	502337	261	0.05%			
phecode_334-42	Neuro	Fourth [trochlear] nerve palsy	502390	234	0.05%			
phecode_334-44	Neuro	Sixth [abducent] nerve palsy	502290	448	0.09%			
phecode_335	Neuro	Nerve root and plexus disorders	499030	9541	1.91%	60847	1020	1.68%
phecode_335-1	Neuro	Nerve plexus lesions	502180	218	0.04%			
phecode_335-11	Neuro	Brachial plexus lesions	502188	197	0.04%			
phecode_335-2	Neuro	Nerve root lesions	502256	270	0.05%			
phecode_335-4	Neuro	Phantom limb (syndrome)	502408	128	0.03%			
phecode_336	Neuro	Mononeuropathies	479190	26221	5.47%	58154	3135	5.39%
phecode_336-1	Neuro	Carpal tunnel syndrome	484865	17138	3.53%	58894	2040	3.46%
phecode_336-2	Neuro	Lesion of median, ulnar, radial nerve	500169	3519	0.70%	60995	416	0.68%
phecode_336-4	Neuro	Mononeuritis of upper limb	502305	253	0.05%			
phecode_336-5	Neuro	Mononeuritis of lower limb	498722	7434	1.49%	60725	939	1.55%
phecode_336-51	Neuro	Lesion of sciatic nerve	502344	639	0.13%			
phecode_336-52	Neuro	Meralgia paresthetica	501365	2026	0.40%	61105	250	0.41%
phecode_336-54	Neuro	Tarsal tunnel syndrome	502362	142	0.03%			
phecode_336-55	Neuro	Lesion of plantar nerve	500276	4296	0.86%	60936	573	0.94%

Supplementary Tables

Table 4 continued from previous page

Endpoint	Category	Phecode string	Number eligible (EHR)	Number of observations (EHR)	Endpoint frequency (EHR)	Number eligible (Retina)	Number of observations (Retina)	Endpoint frequency (Retina)
phecode_337	Neuro	Polyneuropathies	499466	9040	1.81%	60859	1046	1.72%
phecode_337-1	Neuro	Hereditary and idiopathic neuropathy	502248	632	0.13%			
phecode_337-11	Neuro	Hereditary motor and sensory neuropathy	502303	457	0.09%			
phecode_337-2	Neuro	Inflammatory polyneuropathy	501902	500	0.10%			
phecode_337-21	Neuro	Guillain-Barre syndrome [Acute infective polyneuritis]	502181	262	0.05%			
phecode_337-3	Neuro	Toxic neuropathy	502391	404	0.08%			
phecode_337-31	Neuro	Drug-induced polyneuropathy	502437	308	0.06%			
phecode_337-8	Neuro	Polyneuropathy in diseases classified elsewhere	502209	1402	0.28%	61219	169	0.28%
phecode_338	Neuro	Myasthenia gravis and other myoneural disorders	502224	405	0.08%			
phecode_338-1	Neuro	Myasthenia gravis	502241	351	0.07%			
phecode_339	Neuro	Primary disorders of muscles	502266	222	0.04%			
phecode_339-1	Neuro	Muscular dystrophy	502363	121	0.02%			
phecode_340	Neuro	Myopathies	502303	466	0.09%			
phecode_341	Neuro	Cerebral palsy and other paralytic syndromes	500585	4483	0.90%	61048	482	0.79%
phecode_341-1	Neuro	Cerebral palsy	502226	110	0.02%			
phecode_341-2	Neuro	Hemiplegia and hemiparesis	500985	3742	0.75%	61089	402	0.66%
phecode_341-6	Neuro	Cauda equina syndrome	502278	613	0.12%			
phecode_342	Neuro	Plegia and unspecified paralysis	501145	2390	0.48%	61076	277	0.45%
phecode_342-1	Neuro	Paraplegia/Diplegia	502245	369	0.07%			
phecode_342-2	Neuro	Quadriplegia	502361	186	0.04%			
phecode_342-4	Neuro	Monoplegia	501751	1828	0.36%	61145	204	0.33%
phecode_342-5	Neuro	Paralysis NOS	502122	154	0.03%			
phecode_343	Neuro	Disorders of autonomic nervous system	501115	1903	0.38%	61138	212	0.35%
phecode_343-1	Neuro	Autonomic neuropathy	502295	388	0.08%			
phecode_343-3	Neuro	Complex regional pain syndrome	501602	941	0.19%			
phecode_343-5	Neuro	Horner's syndrome*	502236	203	0.04%			
phecode_343-6	Neuro	Multi-system degeneration of the autonomic nervous system*	502440	149	0.03%			
phecode_344	Neuro	Disorders of the circulation of the cerebrospinal fluid	501909	1348	0.27%	61196	148	0.24%
phecode_344-1	Neuro	Hydrocephalus	502175	998	0.20%	61221	110	0.18%
phecode_344-12	Neuro	Obstructive hydrocephalus	502423	158	0.03%			
phecode_344-13	Neuro	(Idiopathic) normal pressure hydrocephalus	502448	219	0.04%			
phecode_344-2	Neuro	Benign intracranial hypertension	502306	153	0.03%			
phecode_344-3	Neuro	Cerebrospinal fluid leak	502336	239	0.05%			
phecode_345	Neuro	Encephalopathy	502403	445	0.09%			
phecode_346	Neuro	Brain damage and brain death	502101	1531	0.30%	61206	165	0.27%
phecode_346-1	Neuro	Postconcussion syndrome	502252	334	0.07%			
phecode_346-3	Neuro	Anoxic brain damage	502435	425	0.08%			
phecode_346-5	Neuro	Compression of brain	502419	207	0.04%			
phecode_346-6	Neuro	Cerebral edema	502429	545	0.11%			
phecode_347	Neuro	Other disorders of the brain and CNS	501835	2464	0.49%	61194	276	0.45%
phecode_347-1	Neuro	Cerebral cysts	502260	433	0.09%			
phecode_347-2	Neuro	Disorders of meninges	502420	174	0.03%			
phecode_348	Neuro	Other diseases of spinal cord	499657	4797	0.96%	60903	540	0.89%
phecode_348-2	Neuro	Myelopathies	500193	3732	0.75%	60978	407	0.67%
phecode_348-21	Neuro	Vascular myelopathies	502399	338	0.07%			
phecode_348-4	Neuro	Spinal cord compression*	502273	848	0.17%			
phecode_349	Neuro	Disorder of nervous system	498071	5924	1.19%	60567	635	1.05%
phecode_349-1	Neuro	Abnormal findings on diagnostic test of central nervous system	500640	3607	0.72%	60984	353	0.58%
phecode_349-12	Neuro	Abnormal electroencephalogram [EEG]	501886	154	0.03%			
phecode_349-13	Neuro	Abnormal findings on diagnostic imaging of skull and head	501382	2692	0.54%	61090	280	0.46%
phecode_349-15	Neuro	Intracranial space-occupying lesion found on diagnostic imaging of central nervous system*	502397	313	0.06%			
phecode_349-2	Neuro	Abnormal results of function studies of peripheral nervous system	501296	1693	0.34%	61079	205	0.34%

1 Data and Implementation

Table 4 continued from previous page

Endpoint	Category	Phecode string	Number eligible (EHR)	Number of observations (EHR)	Endpoint frequency (EHR)	Number eligible (Retina)	Number of observations (Retina)	Endpoint frequency (Retina)
phecode_349-3	Neuro	Nonspecific abnormal electromyogram [EMG]	502375	140	0.03%			
phecode_350	Neuro	Other symptoms involving nervous system	488239	31455	6.44%	59034	3682	6.24%
phecode_350-3	Neuro	Abnormal reflex	501712	381	0.08%			
phecode_350-5	Neuro	Repeated falls*	500777	18645	3.72%	61066	1926	3.15%
phecode_351	Neuro	Disturbances of skin sensation	475577	38133	8.02%	57397	4697	8.18%
phecode_351-1	Neuro	Anesthesia of skin*	495744	8410	1.70%	60286	1091	1.81%
phecode_351-2	Neuro	Hypoesthesia of skin*	501154	3383	0.68%	61035	430	0.70%
phecode_351-3	Neuro	Paresthesia of skin*	487507	18559	3.81%	59156	2260	3.82%
phecode_351-4	Neuro	Hyperesthesia*	502190	351	0.07%			
phecode_352	Neuro	Disturbances of sensation of smell and taste	500473	4198	0.84%	60956	521	0.85%
phecode_352-1	Neuro	Anosmia*	501459	1928	0.38%	61102	227	0.37%
phecode_352-2	Neuro	Parosmia*	502354	288	0.06%			
phecode_352-3	Neuro	Parageusia*	501882	1319	0.26%	61169	169	0.28%
phecode_353	Neuro	Symptoms and signs involving general sensations and perceptions	499399	8927	1.79%	60823	1036	1.70%
phecode_353-1	Neuro	Hallucinations	502025	2164	0.43%	61191	230	0.38%
phecode_353-11	Neuro	Auditory hallucinations*	502319	377	0.08%			
phecode_353-12	Neuro	Visual hallucinations	502360	905	0.18%			
phecode_354	Neuro	Dizziness and giddiness	464289	53900	11.61%	55889	6330	11.33%
phecode_355	Neuro	Coma and other alteration of consciousness	499900	5920	1.18%	60985	670	1.10%
phecode_355-1	Neuro	Coma	501539	1656	0.33%	61134	211	0.35%
phecode_355-2	Neuro	Alteration of consciousness	500809	4376	0.87%	61106	470	0.77%
phecode_355-21	Neuro	Transient alteration of awareness	502431	166	0.03%			
phecode_356	Neuro	Speech disturbance	500761	5792	1.16%	61049	568	0.93%
phecode_356-1	Neuro	Dysarthria	502236	998	0.20%	61231	107	0.17%
phecode_356-2	Neuro	Aphasia and dysphasia	501910	2374	0.47%	61186	235	0.38%
phecode_356-4	Neuro	Slurred speech*	502333	550	0.11%			
phecode_360	Eye	Inflammation of eyelids	477620	30243	6.33%	57946	3710	6.40%
phecode_360-1	Eye	Hordeolum	494424	10558	2.14%	60111	1408	2.34%
phecode_360-11	Eye	Hordeolum externum	494424	10558	2.14%	60111	1408	2.34%
phecode_360-12	Eye	Hordeolum internum	500795	2224	0.44%	61041	239	0.39%
phecode_360-13	Eye	Abscess of eyelid	502195	324	0.06%			
phecode_360-2	Eye	Chalazion	494833	7730	1.56%	60260	952	1.58%
phecode_360-4	Eye	Blepharitis	490900	17709	3.61%	59789	1991	3.33%
phecode_360-5	Eye	Noninfectious dermatoses of eyelid	501445	1106	0.22%	61121	149	0.24%
phecode_360-51	Eye	Eczenematous dermatitis of eyelid	501535	1027	0.20%	61137	139	0.23%
phecode_361	Eye	Disorders of eyelid function	499472	6676	1.34%	60863	710	1.17%
phecode_361-1	Eye	Entropion and trichiasis of eyelid	501922	1706	0.34%	61192	158	0.26%
phecode_361-15	Eye	Trichiasis of eyelid without entropion	502374	374	0.07%			
phecode_361-2	Eye	Lagophthalmos	502397	125	0.02%			
phecode_361-3	Eye	Ptosis of eyelid	500852	3212	0.64%	61027	387	0.63%
phecode_361-4	Eye	Blepharochalasis	502045	1089	0.22%	61196	124	0.20%
phecode_361-9	Eye	Ectropion of eyelid	502006	1336	0.27%	61196	125	0.20%
phecode_362	Eye	Other disorders of the eyelids	496266	7322	1.48%	60401	861	1.43%
phecode_362-1	Eye	Xanthelasma of eyelid	501420	680	0.14%			
phecode_362-5	Eye	Cysts of eyelid	500751	1635	0.33%	61004	180	0.30%
phecode_362-6	Eye	Dermatochalasis of eyelid	502451	151	0.03%			
phecode_363	Eye	Disorders of lacrimal system	489661	26912	5.50%	59474	3265	5.49%
phecode_363-2	Eye	Dry eye syndrome [Tear film insufficiency]	493904	21541	4.36%	59986	2704	4.51%
phecode_363-5	Eye	Epiphora	499342	5230	1.05%	60847	551	0.91%
phecode_363-51	Eye	Epiphora due to excess lacrimation	502299	143	0.03%			
phecode_363-6	Eye	Inflammation of lacrimal passages	502004	503	0.10%			
phecode_363-61	Eye	Dacryocystitis	502087	345	0.07%			
phecode_363-7	Eye	Stenosis and insufficiency of lacrimal passages	501179	1939	0.39%	61109	193	0.32%
phecode_365	Eye	Noninflammatory disorders of the orbit	501713	781	0.16%			
phecode_365-2	Eye	Orbital edema or congestion	502114	345	0.07%			
phecode_365-3	Eye	Exophthalmos [Proptosis]	502165	289	0.06%			
phecode_366	Eye	Noninflammatory disorders of conjunctiva	498647	3681	0.74%	60699	447	0.74%
phecode_366-1	Eye	Pterygium of eye	501439	911	0.18%	61112	113	0.18%
phecode_366-2	Eye	Conjunctival degenerations and deposits	501859	645	0.13%			
phecode_366-21	Eye	Pinguecula	501925	554	0.11%			

Supplementary Tables

Table 4 continued from previous page

Endpoint	Category	Phecode string	Number eligible (EHR)	Number of observations (EHR)	Endpoint frequency (EHR)	Number eligible (Retina)	Number of observations (Retina)	Endpoint frequency (Retina)
phecode_366-4	Eye	Vascular abnormalities of conjunctiva	500767	1861	0.37%	60973	219	0.36%
phecode_366-42	Eye	Conjunctival hyperemia	501211	1135	0.23%	61049	132	0.22%
phecode_366-5	Eye	Conjunctival edema	502408	103	0.02%			
phecode_366-6	Eye	Conjunctival cysts	502199	327	0.07%			
phecode_367	Eye	Inflammation of the eye	453151	33578	7.41%	54765	4039	7.38%
phecode_367-1	Eye	Conjunctivitis	461240	28229	6.12%	55811	3398	6.09%
phecode_367-12	Eye	Allergic [atopic] conjunctivitis	497084	3539	0.71%	60523	481	0.79%
phecode_367-13	Eye	Blepharconjunctivitis	501106	1828	0.36%	61116	220	0.36%
phecode_367-2	Eye	Keratitis	499025	3088	0.62%	60861	309	0.51%
phecode_367-21	Eye	Corneal ulcer	501185	1010	0.20%	61112	101	0.17%
phecode_367-22	Eye	Punctate keratitis	502308	152	0.03%			
phecode_367-3	Eye	Keratoconjunctivitis	501862	429	0.09%			
phecode_367-4	Eye	Inflammation of orbit	502058	587	0.12%			
phecode_367-41	Eye	Cellulitis of orbit	502171	414	0.08%			
phecode_367-5	Eye	Uveitis	498848	2883	0.58%	60777	363	0.60%
phecode_367-52	Eye	Iridocyclitis	498874	2878	0.58%	60786	363	0.60%
phecode_367-6	Eye	Episcleritis	500443	1363	0.27%	61005	168	0.28%
phecode_367-7	Eye	Scleritis	501689	365	0.07%			
phecode_367-9	Eye	Chorioretinal inflammation	502168	217	0.04%			
phecode_369	Eye	Noninflammatory disorders of the cornea	499983	5281	1.06%	60972	647	1.06%
phecode_369-1	Eye	Corneal scars and opacities	502043	874	0.17%	61212	103	0.17%
phecode_369-2	Eye	Corneal edema	502353	414	0.08%			
phecode_369-4	Eye	Corneal degenerations	501905	2269	0.45%	61190	290	0.47%
phecode_369-42	Eye	Recurrent erosion of cornea	502104	247	0.05%			
phecode_369-44	Eye	Senile corneal changes including arcus senilis	502352	136	0.03%			
phecode_369-5	Eye	Hereditary corneal dystrophies	501986	1310	0.26%	61197	173	0.28%
phecode_369-51	Eye	Fuchs' dystrophy	502260	577	0.11%			
phecode_369-6	Eye	Corneal deformities	501963	170	0.03%			
phecode_369-62	Eye	Keratoconus	501976	139	0.03%			
phecode_370	Eye	Disorders of iris and ciliary body	502021	1895	0.38%	61209	227	0.37%
phecode_370-1	Eye	Degeneration of iris and ciliary body	502391	315	0.06%			
phecode_370-3	Eye	Vascular disorders of iris and ciliary body	502249	657	0.13%			
phecode_370-4	Eye	Adhesions of iris	502381	269	0.05%			
phecode_371	Eye	Cataract	490051	61484	12.55%	59805	6798	11.37%
phecode_371-3	Eye	Nuclear cataract	501112	19200	3.83%	61146	2023	3.31%
phecode_371-31	Eye	Age-related nuclear cataract	501112	19200	3.83%	61146	2023	3.31%
phecode_372	Eye	Disorders of lens (excluding cataracts)	502200	583	0.12%			
phecode_372-1	Eye	Aphakia	502304	234	0.05%			
phecode_372-2	Eye	Dislocation of lens	502398	110	0.02%			
phecode_373	Eye	Noninflammatory disorders of choroid	501900	1382	0.28%	61167	189	0.31%
phecode_373-1	Eye	Chorioretinal scars	502220	551	0.11%			
phecode_373-2	Eye	Choroidal degenerations	502378	107	0.02%			
phecode_374	Eye	Disorders of the retina	486877	38291	7.86%	59019	4472	7.58%
phecode_374-1	Eye	Retinal detachments and breaks	498824	5436	1.09%	60784	672	1.11%
phecode_374-11	Eye	Serous retinal detachment	500169	2641	0.53%	60959	324	0.53%
phecode_374-12	Eye	Traction detachment of retina	502415	161	0.03%			
phecode_374-13	Eye	Horseshoe tear of retina without detachment	502434	111	0.02%			
phecode_374-14	Eye	Round hole of retina without detachment	502399	121	0.02%			
phecode_374-2	Eye	Retinoschisis and retinal cysts	502160	643	0.13%			
phecode_374-21	Eye	Retinoschisis	502259	457	0.09%			
phecode_374-3	Eye	Retinal vascular changes and occlusions	499516	6192	1.24%	60848	736	1.21%
phecode_374-32	Eye	Retinal microaneurysms	502176	258	0.05%			
phecode_374-33	Eye	Retinal neovascularization	502348	531	0.11%			
phecode_374-37	Eye	Retinal arterial occlusions	502183	764	0.15%			
phecode_374-38	Eye	Retinal vein occlusions	501584	2651	0.53%	61125	314	0.51%
phecode_374-39	Eye	Transient retinal arterial occlusion [Amaurosis fugax]	502031	1082	0.22%	61197	131	0.21%
phecode_374-4	Eye	Retinal disorders in diseases classified elsewhere	496739	16415	3.30%	60356	1839	3.05%
phecode_374-41	Eye	Hypertensive retinopathy	502249	166	0.03%			
phecode_374-42	Eye	Diabetic retinopathy	497268	15571	3.13%	60400	1760	2.91%

Table 4 continued from previous page

Endpoint	Category	Phecode string	Number eligible (EHR)	Number of observations (EHR)	Endpoint frequency (EHR)	Number eligible (Retina)	Number of observations (Retina)	Endpoint frequency (Retina)
phecode_374-5	Eye	Macular degeneration	500060	14908	2.98%	60918	1816	2.98%
phecode_374-51	Eye	Age-related macular degeneration	501304	5717	1.14%	61073	704	1.15%
phecode_374-511	Eye	Nonexudative (dry) age-related macular degeneration	502244	1821	0.36%	61203	246	0.40%
phecode_374-512	Eye	Exudative (wet) age-related macular degeneration	502416	1210	0.24%	61247	150	0.24%
phecode_374-52	Eye	Macular cyst, hole, or pseudohole	502094	1819	0.36%	61216	251	0.41%
phecode_374-54	Eye	Drusen (degenerative) of macula	502251	639	0.13%			
phecode_374-55	Eye	Puckering of macula	502198	3171	0.63%	61217	418	0.68%
phecode_374-6	Eye	Peripheral retinal degeneration	502028	546	0.11%			
phecode_374-61	Eye	Lattice degeneration of retina	502351	162	0.03%			
phecode_374-7	Eye	Hereditary retinal dystrophy	502026	462	0.09%			
phecode_374-8	Eye	Retinal edema	501996	2250	0.45%	61192	267	0.44%
phecode_374-9	Eye	Central serous chorioretinopathy	502007	345	0.07%			
phecode_375	Eye	Abnormal intraocular pressure	494435	19829	4.01%	60078	2465	4.10%
phecode_375-1	Eye	Glaucoma	496839	14012	2.82%	60486	1664	2.75%
phecode_375-11	Eye	Open angle glaucoma	500869	6228	1.24%	61057	752	1.23%
phecode_375-112	Eye	Pigmentary glaucoma	502394	138	0.03%			
phecode_375-113	Eye	Primary open angle glaucoma	501901	2605	0.52%	61174	327	0.53%
phecode_375-12	Eye	Angle-Closure Glaucoma	502054	2111	0.42%	61222	222	0.36%
phecode_375-14	Eye	Low-tension glaucoma (Normal-tension glaucoma)	502227	953	0.19%	61224	120	0.20%
phecode_375-6	Eye	Ocular hypertension	499959	5669	1.13%	60822	748	1.23%
phecode_375-7	Eye	Hypotony of eye	502436	114	0.02%			
phecode_376	Eye	Disorders of vitreous body	493291	20094	4.07%	59957	2501	4.17%
phecode_376-1	Eye	Vitreous degeneration	496807	14419	2.90%	60473	1855	3.07%
phecode_376-2	Eye	Vitreous opacities	493513	19976	4.05%	59995	2488	4.15%
phecode_376-21	Eye	Crystalline deposits in vitreous body	498547	5715	1.15%	60672	646	1.06%
phecode_376-4	Eye	Vitreomacular adhesion	502459	446	0.09%			
phecode_377	Eye	Hemorrhage of the eye	493772	14189	2.87%	60115	1621	2.70%
phecode_377-2	Eye	Conjunctival hemorrhage	495626	11218	2.26%	60353	1321	2.19%
phecode_377-4	Eye	Retinal hemorrhage	501641	1775	0.35%	61158	171	0.28%
phecode_377-5	Eye	Vitreous hemorrhage	501746	1344	0.27%	61157	144	0.24%
phecode_377-8	Eye	Hyphema	502084	170	0.03%			
phecode_379	Eye	Infection of the eye	500423	2473	0.49%	61009	311	0.51%
phecode_379-2	Eye	Eye infection, viral	500633	2321	0.46%	61036	284	0.47%
phecode_379-21	Eye	Infection of the eye, herpes	501288	1257	0.25%	61126	155	0.25%
phecode_380	Eye	Disorders of optic nerve and visual pathways	500347	2881	0.58%	60969	305	0.50%
phecode_380-1	Eye	Optic neuropathy	501519	776	0.15%			
phecode_380-11	Eye	Optic neuritis	501686	262	0.05%			
phecode_380-12	Eye	Ischemic optic neuropathy	502350	288	0.06%			
phecode_380-2	Eye	Disorders of optic disc	501483	1908	0.38%	61114	205	0.34%
phecode_380-21	Eye	Papilledema	502327	133	0.03%			
phecode_380-22	Eye	Optic disc drusen	502360	160	0.03%			
phecode_380-3	Eye	Optic atrophy	502230	357	0.07%			
phecode_381	Eye	Strabismus	498337	2681	0.54%	60761	302	0.50%
phecode_381-1	Eye	Paralytic strabismus [Neurogenic strabismus]	502023	1221	0.24%	61202	132	0.22%
phecode_381-11	Eye	Ophthalmoplegia	502408	282	0.06%			
phecode_381-3	Eye	Esotropia	500885	374	0.07%			
phecode_381-4	Eye	Exotropia	501689	541	0.11%			
phecode_381-6	Eye	Duane's syndrome [Duane anomaly]	501031	522	0.10%			
phecode_381-8	Eye	Heterophoria	502241	408	0.08%			
phecode_381-81	Eye	Esophoria	502400	105	0.02%			
phecode_381-82	Eye	Exophoria	502331	205	0.04%			
phecode_382	Eye	Other disorders of binocular movement	502129	341	0.07%			
phecode_383	Eye	Irregular eye movements	501885	1148	0.23%	61189	108	0.18%
phecode_383-1	Eye	Nystagmus	501980	689	0.14%			
phecode_384	Eye	Anomalies of pupillary function	501760	1035	0.21%	61191	102	0.17%
phecode_384-1	Eye	Anisocoria	502275	208	0.04%			
phecode_384-3	Eye	Mydriasis	502223	232	0.05%			
phecode_384-4	Eye	Tonic pupil	502238	102	0.02%			

Supplementary Tables

Table 4 continued from previous page

Endpoint	Category	Phecode string	Number eligible (EHR)	Number of observations (EHR)	Endpoint frequency (EHR)	Number eligible (Retina)	Number of observations (Retina)	Endpoint frequency (Retina)
phecode_385	Eye	Abnormal results of function studies of eye	502423	1848	0.37%	61248	186	0.30%
phecode_386	Eye	Visual disturbances	490087	19438	3.97%	59699	2201	3.69%
phecode_386-1	Eye	Amblyopia	501355	2765	0.55%	61130	305	0.50%
phecode_386-2	Eye	Diplopia	500338	3379	0.68%	61034	376	0.62%
phecode_386-3	Eye	Visual discomfort	502323	175	0.03%			
phecode_386-4	Eye	Visual field defects	500661	2712	0.54%	61030	307	0.50%
phecode_386-41	Eye	Scotoma	501838	563	0.11%			
phecode_386-42	Eye	Hemianopia	502123	572	0.11%			
phecode_386-8	Eye	Sudden or transient visual loss	502359	211	0.04%			
phecode_386-9	Eye	Visual distortions and subjective visual disturbances	501568	1323	0.26%	61103	156	0.26%
phecode_387	Eye	Disorders of refraction and accommodation	497979	10839	2.18%	60676	1363	2.25%
phecode_387-1	Eye	Hypermetropia	501833	1065	0.21%	61170	128	0.21%
phecode_387-2	Eye	Myopia	499531	4671	0.94%	60911	557	0.91%
phecode_387-21	Eye	Progressive high (degenerative) myopia	502256	428	0.09%			
phecode_387-3	Eye	Astigmatism	501666	4841	0.96%	61118	685	1.12%
phecode_387-4	Eye	Presbyopia	502310	221	0.04%			
phecode_387-5	Eye	Anisometropia	502370	178	0.04%			
phecode_388	Eye	Blindness and low vision	500517	6563	1.31%	61022	884	1.45%
phecode_388-1	Eye	Blindness of both eyes	502347	611	0.12%			
phecode_389	Eye	Other disorders of eye	472649	38333	8.11%	56945	4472	7.85%
phecode_389-1	Eye	Ocular pain	499349	3351	0.67%	60797	379	0.62%
phecode_390	Hearing	Disorders of external ear	430811	72104	16.74%	51278	8946	17.45%
phecode_390-1	Hearing	Otitis externa	465507	33953	7.29%	56257	4157	7.39%
phecode_390-4	Hearing	Impacted cerumen	460447	54778	11.90%	55248	6818	12.34%
phecode_390-5	Hearing	Stenosis of external ear canal	502333	104	0.02%			
phecode_390-6	Hearing	Perichondritis and chondritis of pinna	501702	2076	0.41%	61160	244	0.40%
phecode_391	Hearing	Disorders of the middle ear	462567	33890	7.33%	55937	4151	7.42%
phecode_391-1	Hearing	Otitis media	477272	18214	3.82%	57852	2207	3.81%
phecode_391-11	Hearing	Acute otitis media	496606	1603	0.32%	60555	189	0.31%
phecode_391-12	Hearing	Chronic otitis media	497486	2778	0.56%	60659	321	0.53%
phecode_391-2	Hearing	Eustachian tube disorders	488905	15187	3.11%	59444	1883	3.17%
phecode_391-21	Hearing	Eustachian salpingitis	499679	1446	0.29%	60918	184	0.30%
phecode_391-22	Hearing	Obstruction of Eustachian tube	500941	454	0.09%			
phecode_391-4	Hearing	Tympanosclerosis	502238	219	0.04%			
phecode_391-6	Hearing	Cholesteatoma of middle ear	501694	549	0.11%			
phecode_391-7	Hearing	Perforation of tympanic membrane	498660	2741	0.55%	60809	307	0.50%
phecode_391-8	Hearing	Otosclerosis	501360	353	0.07%			
phecode_391-9	Hearing	Otorrhea	499737	2565	0.51%	60856	308	0.51%
phecode_392	Hearing	Otalgia and effusion of ear	480506	23160	4.82%	58063	2814	4.85%
phecode_393	Hearing	Mastoiditis and related conditions	502061	374	0.07%			
phecode_394	Hearing	Disorders of vestibular function	493053	16057	3.26%	59925	1916	3.20%
phecode_394-1	Hearing	Meniere disease	500368	1668	0.33%	60980	211	0.35%
phecode_394-2	Hearing	Vertigo	496531	13479	2.71%	60371	1582	2.62%
phecode_394-21	Hearing	Paroxysmal vertigo	498022	11441	2.30%	60560	1344	2.22%
phecode_394-22	Hearing	Vestibular neuronitis	500961	2371	0.47%	61061	274	0.45%
phecode_394-4	Hearing	Abnormal vestibular function study	501725	983	0.20%	61142	119	0.19%
phecode_395	Hearing	Other diseases of inner ear	488618	14213	2.91%	59508	1614	2.71%
phecode_395-1	Hearing	Labyrinthitis	490571	13402	2.73%	59703	1511	2.53%
phecode_395-3	Hearing	Noise effects on inner ear	502079	124	0.02%			
phecode_396	Hearing	Hearing impairment	469854	53896	11.47%	56785	6075	10.70%
phecode_396-1	Hearing	Conductive hearing loss	500766	1498	0.30%	61050	153	0.25%
phecode_396-11	Hearing	Conductive hearing loss, bilateral	502351	182	0.04%			
phecode_396-2	Hearing	Sensorineural hearing loss	495834	11314	2.28%	60363	1204	1.99%
phecode_396-21	Hearing	Sensorineural hearing loss, bilateral	501072	3888	0.78%	61046	416	0.68%
phecode_396-22	Hearing	Presbycusis	501376	3135	0.63%	61115	313	0.51%
phecode_396-3	Hearing	Mixed conductive and sensorineural hearing loss	502218	685	0.14%			
phecode_396-5	Hearing	Sudden idiopathic hearing loss	502287	140	0.03%			
phecode_397	Hearing	Other hearing abnormality	487635	18878	3.87%	59152	2376	4.02%
phecode_397-1	Hearing	Tinnitus	487926	18592	3.81%	59215	2333	3.94%

1 Data and Implementation

Table 4 continued from previous page

Endpoint	Category	Phecode string	Number eligible (EHR)	Number of observations (EHR)	Endpoint frequency (EHR)	Number eligible (Retina)	Number of observations (Retina)	Endpoint frequency (Retina)
phecode_397-3	Hearing	Hyperacusis	502366	129	0.03%			
phecode_398	Hearing	Other disorders of ear	496836	7828	1.58%	60497	977	1.61%
phecode_398-1	Hearing	Abnormal auditory function study	501382	1968	0.39%	61094	238	0.39%
phecode_400	Cardio	Rheumatic fever and chronic rheumatic heart diseases	500483	8897	1.78%	61009	1034	1.69%
phecode_400-2	Cardio	Chronic rheumatic heart diseases	501248	8905	1.78%	61099	1030	1.69%
phecode_401	Cardio	Hypertension	407287	99909	24.53%	48591	11114	22.87%
phecode_401-1	Cardio	Essential hypertension	407510	99799	24.49%	48612	11111	22.86%
phecode_401-2	Cardio	Hypertensive heart disease	502224	1044	0.21%			
phecode_401-3	Cardio	Hypertensive chronic kidney disease	501603	1543	0.31%	61133	115	0.19%
phecode_401-6	Cardio	Secondary hypertension	501208	2995	0.60%	61056	450	0.74%
phecode_402	Cardio	Elevated blood pressure reading without diagnosis of hypertension	471799	29538	6.26%	57523	3111	5.41%
phecode_403	Cardio	Angina pectoris	484383	21945	4.53%	59052	2385	4.04%
phecode_403-1	Cardio	Coronary artery spasm [Coronary vasospasm]	502271	312	0.06%			
phecode_404	Cardio	Ischemic heart disease	481544	36540	7.59%	58791	4055	6.90%
phecode_404-1	Cardio	Myocardial infarction [Heart attack]	493387	17790	3.61%	60213	1800	2.99%
phecode_404-11	Cardio	Acute myocardial infarction	494415	12879	2.60%	60303	1265	2.10%
phecode_404-2	Cardio	Coronary atherosclerosis [Atherosclerotic heart disease]	487408	27309	5.60%	59460	3044	5.12%
phecode_406	Cardio	Chronic pulmonary heart disease	502141	2921	0.58%	61221	316	0.52%
phecode_406-1	Cardio	Pulmonary hypertension	502186	2572	0.51%	61226	292	0.48%
phecode_406-11	Cardio	Primary pulmonary hypertension	502250	572	0.11%			
phecode_406-13	Cardio	Secondary pulmonary arterial hypertension*	502426	142	0.03%			
phecode_408	Cardio	Diseases of pulmonary vessels	502415	118	0.02%			
phecode_410	Cardio	Inflammation of the heart [Carditis]	498360	4195	0.84%	60762	498	0.82%
phecode_410-1	Cardio	Pericarditis	501540	788	0.16%			
phecode_410-2	Cardio	Endocarditis	499427	3115	0.62%	60892	376	0.62%
phecode_410-3	Cardio	Myocarditis	502272	395	0.08%			
phecode_411	Cardio	Other diseases of pericardium	501905	2937	0.59%	61175	358	0.59%
phecode_411-1	Cardio	Hemopericardium NEC	502455	206	0.04%			
phecode_411-2	Cardio	Pericardial effusion (noninflammatory)*	502217	2078	0.41%	61215	257	0.42%
phecode_413	Cardio	Heart valve disorders	496622	21548	4.34%	60553	2375	3.92%
phecode_413-1	Cardio	Mitral valve disorders	499104	13381	2.68%	60828	1464	2.41%
phecode_413-11	Cardio	Mitral valve insufficiency	500364	8343	1.67%	60984	899	1.47%
phecode_413-12	Cardio	Mitral valve prolapse*	501445	952	0.19%	61134	108	0.18%
phecode_413-13	Cardio	Mitral valve stenosis	502237	362	0.07%			
phecode_413-2	Cardio	Aortic valve disorders	499841	11941	2.39%	60957	1249	2.05%
phecode_413-21	Cardio	Aortic stenosis	501283	6212	1.24%	61131	625	1.02%
phecode_413-22	Cardio	Aortic insufficiency	501383	4423	0.88%	61127	469	0.77%
phecode_413-3	Cardio	Tricuspid valve disorders	502085	6939	1.38%	61195	803	1.31%
phecode_413-32	Cardio	Tricuspid valve insufficiency*	502355	1872	0.37%	61245	207	0.34%
phecode_413-4	Cardio	Pulmonary valve disorders	502146	964	0.19%	61207	119	0.19%
phecode_413-42	Cardio	Pulmonary valve insufficiency*	502411	923	0.18%	61248	114	0.19%
phecode_413-6	Cardio	Heart valve replaced	501709	2911	0.58%	61191	313	0.51%
phecode_414	Cardio	Cardiomyopathy	501389	3749	0.75%	61116	414	0.68%
phecode_414-1	Cardio	Hypertrophic cardiomyopathy	502245	646	0.13%			
phecode_414-11	Cardio	Obstructive hypertrophic cardiomyopathy	502345	348	0.07%			
phecode_414-2	Cardio	Dilated cardiomyopathy*	502049	1276	0.25%	61197	132	0.22%
phecode_414-5	Cardio	Ischemic cardiomyopathy*	502373	823	0.16%			
phecode_414-9	Cardio	Takotsubo syndrome [Stress cardiomyopathy]	502459	167	0.03%			
phecode_416	Cardio	Cardiac arrhythmia and conduction disorders	484929	52637	10.85%	58954	5938	10.07%
phecode_416-1	Cardio	Paroxysmal tachycardia	498738	6196	1.24%	60761	769	1.27%
phecode_416-11	Cardio	Supraventricular tachycardia	499529	4559	0.91%	60850	574	0.94%
phecode_416-12	Cardio	Ventricular tachycardia	501885	1927	0.38%	61183	230	0.38%
phecode_416-2	Cardio	Atrial fibrillation and flutter	494218	32911	6.66%	60144	3692	6.14%
phecode_416-21	Cardio	Atrial fibrillation	497277	22131	4.45%	60476	2707	4.48%
phecode_416-211	Cardio	Paroxysmal atrial fibrillation*	500568	9583	1.91%	60969	1166	1.91%
phecode_416-212	Cardio	Persistent atrial fibrillation*	502448	1571	0.31%	61253	194	0.32%
phecode_416-213	Cardio	Chronic atrial fibrillation*	502460	616	0.12%			
phecode_416-214	Cardio	Permanent atrial fibrillation*	502454	126	0.03%			

Supplementary Tables

Table 4 continued from previous page

Endpoint	Category	Phecode string	Number eligible (EHR)	Number of observations (EHR)	Endpoint frequency (EHR)	Number eligible (Retina)	Number of observations (Retina)	Endpoint frequency (Retina)
phecode_416-22	Cardio	Atrial flutter	501221	3815	0.76%	61034	460	0.75%
phecode_416-221	Cardio	Typical atrial flutter*	502460	224	0.04%			
phecode_416-222	Cardio	Atypical atrial flutter*	502460	102	0.02%			
phecode_416-3	Cardio	Ventricular fibrillation and flutter	502284	827	0.16%			
phecode_416-31	Cardio	Ventricular fibrillation	502403	166	0.03%			
phecode_416-4	Cardio	Heart block	499315	18187	3.64%	60848	2026	3.33%
phecode_416-41	Cardio	Atrioventricular block	501602	7736	1.54%	61145	879	1.44%
phecode_416-42	Cardio	Left bundle branch block	501418	5964	1.19%	61113	622	1.02%
phecode_416-43	Cardio	Right bundle branch block	501233	6305	1.26%	61107	734	1.20%
phecode_416-5	Cardio	Premature depolarization [Premature beats]	499787	6849	1.37%	60918	857	1.41%
phecode_416-51	Cardio	Atrial premature depolarization [Supraventricular premature beats]	502148	1467	0.29%	61206	181	0.30%
phecode_416-52	Cardio	Ventricular premature depolarization*	501281	3948	0.79%	61113	531	0.87%
phecode_416-6	Cardio	Long QT syndrome	502439	125	0.02%			
phecode_416-7	Cardio	Sinoatrial node dysfunction	502229	1180	0.23%	61225	116	0.19%
phecode_416-71	Cardio	Sick sinus syndrome*	502229	1180	0.23%	61225	116	0.19%
phecode_416-8	Cardio	Pre-excitation syndrome [Anomalous atrioventricular excitation]	502157	158	0.03%			
phecode_417	Cardio	Abnormalities of heart beat	475187	58820	12.38%	57387	6912	12.04%
phecode_417-1	Cardio	Palpitations	480022	27313	5.69%	58118	3307	5.69%
phecode_417-2	Cardio	Tachycardia	500372	8174	1.63%	60974	965	1.58%
phecode_417-3	Cardio	Bradycardia*	499968	11922	2.38%	60894	1453	2.39%
phecode_418	Cardio	Abnormal results of cardiovascular function studies	492031	19270	3.92%	59668	2227	3.73%
phecode_418-1	Cardio	Abnormal electrocardiogram [ECG] [EKG]	497517	13964	2.81%	60472	1649	2.73%
phecode_419	Cardio	Presence of cardiac device	501380	6524	1.30%	61138	688	1.13%
phecode_419-2	Cardio	Presence of cardiac defibrillator	502452	136	0.03%			
phecode_420	Cardio	Cardiac arrest	502105	2930	0.58%	61211	341	0.56%
phecode_423	Cardio	Abnormal cardiac sounds	496771	6216	1.25%	60604	706	1.16%
phecode_423-1	Cardio	Cardiac murmurs	496983	5294	1.07%	60631	602	0.99%
phecode_424	Cardio	Heart failure	499141	19432	3.89%	60846	1996	3.28%
phecode_424-1	Cardio	Left heart failure	500217	9413	1.88%	61001	884	1.45%
phecode_424-2	Cardio	Systolic heart failure	502093	3174	0.63%	61201	320	0.52%
phecode_424-3	Cardio	Diastolic heart failure	502402	1498	0.30%	61242	154	0.25%
phecode_424-5	Cardio	Right heart failure*	502429	106	0.02%			
phecode_424-6	Cardio	Hypertensive heart disease with heart failure	502425	282	0.06%			
phecode_425	Cardio	Cardiomegaly	500037	11640	2.33%	60922	1318	2.16%
phecode_426	Cardio	Other heart disorders in diseases NEC	500855	11384	2.27%	61028	1163	1.91%
phecode_430	Cardio	Nontraumatic Intracranial hemorrhage	500766	4343	0.87%	61055	443	0.73%
phecode_430-1	Cardio	Nontraumatic subarachnoid hemorrhage	501326	1824	0.36%	61118	159	0.26%
phecode_430-2	Cardio	Nontraumatic intracerebral hemorrhage	501944	2167	0.43%	61192	236	0.39%
phecode_430-3	Cardio	Nontraumatic subdural hemorrhage	502312	1019	0.20%	61234	118	0.19%
phecode_431	Cardio	Stroke and transient cerebral ischemic attacks	494276	20901	4.23%	60256	2197	3.65%
phecode_431-1	Cardio	Stroke	497707	13787	2.77%	60692	1388	2.29%
phecode_431-11	Cardio	Cerebral infarction [Ischemic stroke]	499013	11434	2.29%	60849	1157	1.90%
phecode_431-12	Cardio	Hemorrhagic stroke	500936	3334	0.67%	61074	333	0.55%
phecode_431-14	Cardio	Cerebellar stroke syndrome*	502356	391	0.08%			
phecode_431-15	Cardio	Lacunar syndrome*	502411	734	0.15%			
phecode_431-2	Cardio	Transient cerebral ischemic attacks	498793	10481	2.10%	60783	1126	1.85%
phecode_433	Cardio	Other cerebrovascular disease	499648	13315	2.66%	60908	1534	2.52%
phecode_433-1	Cardio	Occlusion and stenosis of cerebral arteries	501889	2671	0.53%	61192	283	0.46%

2 A Neural Survival model integrates polygenic information for cardiovascular risk prediction

2 A Neural Survival model integrates polygenic information for cardiovascular risk prediction

2 *A Neural Survival model integrates polygenic information for cardiovascular risk prediction*

Figure 2: Comparison of the NeuralCVD score with established risk scores. The NeuralCVD score outperforms existing approaches in discrimination at ten years measured by Harrell's C-index (via bootstrapping with 1000 iterations).

Differences in C-Index: NeuralCVD clinical vs. Other Models			
SCORE	Clinical Baselines		Clinical
	ASCVD	QRISK3	Cox clinical
0.037 (CI 0.034-0.039)	0.024 (CI 0.023-0.026)	0.01 (CI 0.009-0.011)	0.003 (CI 0.002-0.004)

Supplementary Tables

Figure 3: Comparison after addition of PGS. PGS addition increases the C-index for the NeuralCVD score. Furthermore, it improves upon the Cox model without PGS and all other tested scores with added PGS information incl. the Cox with interaction terms between PGS and age.

Differences in C-Index: NeuralCVD clinical + PGS vs. Other Models				
Clinical + PGS			Clinical	
Cox Sun PGS	Cox clinical PGS	COX clinical PGS*age	Cox clinical	NeuralCVD clinical
0.022 (CI 0.021 -0.024)	0.002 (CI 0.002 -0.003)	0.002 (CI 0.002 -0.003)	0.009 (CI 0.008 -0.01)	0.006 (CI 0.005 -0.007)

2 *A Neural Survival model integrates polygenic information for cardiovascular risk prediction*

Figure 4: Discriminative Performance of the NeuralCVD score after addition of PGS on spatially distinct assessment centres. The NeuralCVD score improves upon all other scores with and without PGS. The improvements in discrimination (via bootstrapped C-Index, 1000 iterations) are stable over the 22 spatially distinct assessment centre test sets, not visible at training time.

Differences in C-Index by Assessment Centre: NeuralCVD clinical + PGS vs. Other Models								
	Clinical Baselines			Clinical		Clinical + PGS		
	SCORE	ASCVD	QRISK3	Cox clinical	NeuralCVD clinical	Cox Sun PGS	Cox clinical PGS	Cox clinical PGS*age
Leeds	0.041 (CI 0.035-0.048)	0.028 (CI 0.022-0.034)	0.017 (CI 0.013-0.022)	0.01 (CI 0.007-0.014)	0.007 (CI 0.005-0.01)	0.02 (CI 0.016-0.026)	0.002 (CI 0-0.004)	0.002 (CI 0-0.004)
Bristol	0.045 (CI 0.038-0.052)	0.031 (CI 0.025-0.037)	0.014 (CI 0.01-0.019)	0.007 (CI 0.004-0.01)	0.005 (CI 0.002-0.008)	0.025 (CI 0.019-0.03)	0.003 (CI 0-0.005)	0.003 (CI 0-0.005)
Newcastle	0.04 (CI 0.033-0.048)	0.029 (CI 0.022-0.036)	0.018 (CI 0.013-0.023)	0.008 (CI 0.004-0.012)	0.006 (CI 0.002-0.01)	0.02 (CI 0.014-0.026)	0.002 (CI -0.001-0.004)	0.002 (CI -0.001-0.004)
Nottingham	0.044 (CI 0.035-0.052)	0.032 (CI 0.025-0.039)	0.015 (CI 0.01-0.021)	0.009 (CI 0.005-0.013)	0.005 (CI 0.002-0.008)	0.025 (CI 0.019-0.032)	0.003 (CI 0.001-0.006)	0.003 (CI 0.001-0.006)
Liverpool	0.047 (CI 0.04-0.055)	0.032 (CI 0.026-0.04)	0.017 (CI 0.012-0.022)	0.008 (CI 0.005-0.012)	0.007 (CI 0.004-0.01)	0.025 (CI 0.019-0.031)	0.003 (CI 0.001-0.005)	0.002 (CI 0-0.005)
Sheffield	0.035 (CI 0.027-0.044)	0.024 (CI 0.017-0.033)	0.012 (CI 0.006-0.017)	0.006 (CI 0.002-0.009)	0.004 (CI 0-0.007)	0.016 (CI 0.009-0.023)	-0.002 (CI -0.005-0.001)	-0.002 (CI -0.005-0.001)
Reading	0.041 (CI 0.032-0.049)	0.031 (CI 0.023-0.038)	0.018 (CI 0.012-0.024)	0.01 (CI 0.006-0.014)	0.007 (CI 0.003-0.011)	0.022 (CI 0.015-0.028)	0.003 (CI 0.001-0.006)	0.003 (CI 0.001-0.006)
Hounslow	0.041 (CI 0.032-0.05)	0.029 (CI 0.022-0.038)	0.017 (CI 0.011-0.023)	0.007 (CI 0.004-0.011)	0.004 (CI 0-0.007)	0.021 (CI 0.014-0.028)	0.003 (CI 0.001-0.005)	0.002 (CI 0-0.005)
Bury	0.042 (CI 0.033-0.052)	0.03 (CI 0.022-0.038)	0.013 (CI 0.008-0.019)	0.01 (CI 0.006-0.014)	0.007 (CI 0.003-0.01)	0.021 (CI 0.014-0.028)	0.003 (CI 0.001-0.006)	0.003 (CI 0.001-0.006)
Croydon	0.048 (CI 0.039-0.058)	0.037 (CI 0.029-0.046)	0.016 (CI 0.01-0.022)	0.01 (CI 0.005-0.014)	0.007 (CI 0.003-0.011)	0.028 (CI 0.021-0.035)	0.004 (CI 0.001-0.007)	0.004 (CI 0.001-0.006)
Birmingham	0.038 (CI 0.029-0.049)	0.031 (CI 0.023-0.04)	0.012 (CI 0.006-0.018)	0.005 (CI 0.001-0.01)	0.002 (CI -0.002-0.006)	0.023 (CI 0.016-0.031)	0.001 (CI -0.001-0.004)	0.002 (CI -0.001-0.005)
Middlesborough	0.046 (CI 0.035-0.056)	0.033 (CI 0.024-0.041)	0.02 (CI 0.014-0.027)	0.013 (CI 0.008-0.018)	0.01 (CI 0.006-0.014)	0.026 (CI 0.019-0.033)	0.005 (CI 0.002-0.008)	0.004 (CI 0.001-0.007)
Stoke	0.041 (CI 0.029-0.052)	0.031 (CI 0.021-0.041)	0.015 (CI 0.009-0.022)	0.012 (CI 0.007-0.017)	0.007 (CI 0.003-0.012)	0.022 (CI 0.014-0.03)	0.004 (CI 0-0.007)	0.004 (CI 0-0.007)
Glasgow	0.045 (CI 0.035-0.055)	0.033 (CI 0.024-0.042)	0.021 (CI 0.014-0.028)	0.01 (CI 0.005-0.015)	0.01 (CI 0.005-0.014)	0.021 (CI 0.013-0.029)	0.001 (CI -0.001-0.004)	0.001 (CI -0.002-0.004)
Cardiff	0.053 (CI 0.041-0.066)	0.036 (CI 0.027-0.047)	0.017 (CI 0.01-0.025)	0.012 (CI 0.007-0.018)	0.009 (CI 0.005-0.014)	0.027 (CI 0.018-0.035)	0.004 (CI 0.001-0.008)	0.005 (CI 0.002-0.008)
Edinburgh	0.046 (CI 0.034-0.058)	0.037 (CI 0.026-0.048)	0.019 (CI 0.011-0.027)	0.005 (CI 0-0.011)	0.004 (CI -0.001-0.009)	0.029 (CI 0.019-0.038)	0 (CI -0.003-0.004)	0.001 (CI -0.003-0.005)
Oxford	0.037 (CI 0.024-0.051)	0.027 (CI 0.016-0.039)	0.016 (CI 0.008-0.024)	0.009 (CI 0.003-0.015)	0.007 (CI 0.002-0.012)	0.018 (CI 0.009-0.029)	0.002 (CI -0.001-0.006)	0.001 (CI -0.002-0.005)
Manchester	0.029 (CI 0.017-0.042)	0.016 (CI 0.004-0.027)	0.005 (CI -0.003-0.015)	0.002 (CI -0.004-0.009)	0.003 (CI -0.004-0.009)	0.012 (CI 0.003-0.021)	-0.003 (CI -0.007-0)	-0.003 (CI -0.007-0.001)
Barts	0.036 (CI 0.023-0.05)	0.025 (CI 0.013-0.037)	0.011 (CI 0.003-0.019)	0.005 (CI -0.002-0.011)	0.002 (CI -0.004-0.007)	0.022 (CI 0.013-0.034)	0.002 (CI -0.002-0.007)	0.002 (CI -0.002-0.007)
Stockport (pilot)	0.044 (CI 0.022-0.067)	0.031 (CI 0.012-0.053)	0.013 (CI -0.002-0.028)	0.006 (CI -0.003-0.016)	0.005 (CI -0.004-0.014)	0.03 (CI 0.012-0.049)	0.006 (CI 0-0.013)	0.007 (CI 0.001-0.014)
Swansea	0.027 (CI -0.006-0.063)	0.024 (CI -0.004-0.055)	0.007 (CI -0.014-0.026)	0.004 (CI -0.01-0.017)	-0.001 (CI -0.014-0.011)	0.026 (CI 0-0.057)	0.002 (CI -0.007-0.012)	0 (CI -0.009-0.01)
Wrexham	0.075 (CI 0.024-0.132)	0.038 (CI 0-0.083)	0.054 (CI 0.014-0.101)	0.018 (CI -0.009-0.047)	0.015 (CI -0.009-0.04)	0.034 (CI 0.003-0.069)	0.007 (CI -0.006-0.022)	0.005 (CI -0.007-0.018)

Supplementary Tables

Figure 5: Stratified relative risks of genetic predisposition. Event rates of high and low genetic predisposition relative to median genetic risk stratified by predicted clinical risk (recalibrated QRISK3) and age.

Clinical Risk	Age (Years)	Relative Risks								
		Low Genetic Risk			Intermediate Genetic Risk			High Genetic Risk		
		n	n(Events)	Relative Risk	n	n(Events)	Relative Risk	n	n(Events)	Relative Risk
Low (<5%)	<50	4964	81	0.8 (CI 95% 0.62-1)	5284	108	0.99 (CI 95% 0.87-1.09)	5864	238	1.96 (CI 95% 1.59-2.44)
Low (<5%)	50-60	4224	98	0.79 (CI 95% 0.66-0.99)	3808	110	0.98 (CI 95% 0.86-1.1)	3724	208	1.88 (CI 95% 1.59-2.26)
Low (<5%)	>60	990	35	0.91 (CI 95% 0.59-1.33)	766	30	0.99 (CI 95% 0.75-1.24)	556	33	1.52 (CI 95% 0.96-2.35)
Intermediate (5-10%)	<50	331	25	0.8 (CI 95% 0.51-1.08)	509	44	0.91 (CI 95% 0.75-1.09)	704	115	1.72 (CI 95% 1.34-2.19)
Intermediate (5-10%)	50-60	2208	106	0.62 (CI 95% 0.53-0.79)	2363	186	1.04 (CI 95% 0.95-1.15)	2614	320	1.64 (CI 95% 1.4-1.86)
Intermediate (5-10%)	>60	3356	220	0.78 (CI 95% 0.66-0.88)	3087	249	0.94 (CI 95% 0.87-1.02)	2447	312	1.5 (CI 95% 1.31-1.71)
High (>10%)	<50	50	6	0.61 (CI 95% 0.25-1.35)	58	9	0.8 (CI 95% 0.48-1.21)	106	30	1.47 (CI 95% 0.94-2.23)
High (>10%)	50-60	582	61	0.7 (CI 95% 0.51-0.88)	740	115	1.03 (CI 95% 0.89-1.17)	818	208	1.68 (CI 95% 1.4-2.07)
High (>10%)	>60	2479	340	0.8 (CI 95% 0.72-0.89)	2530	423	0.98 (CI 95% 0.91-1.04)	2342	592	1.49 (CI 95% 1.32-1.6)

3 Metabolomic profiles predict individual multi-disease outcomes

Supplementary Tables

Table 5: Event rate ratios of the metabolomic state with the proportion of individuals with incident disease. Proportion of incident disease in the bottom 10 % and top 10 % quantiles of the metabolomic state and the respective event rate ratios for all endpoints. The numbers in the brackets correspond to the 95 % CIs over 1000 bootstrapping iterations.

	Bottom 10%	Top 10%	Observed Event Rate Ratio
MACE	2.01% (1.77%, 2.26%)	18.59% (18.09%, 19.28%)	9.25 (8.12, 10.53)
CHD	2.34% (2.04%, 2.67%)	16.47% (15.95%, 17.23%)	7.07 (6.24, 8.15)
Cerebral Stroke	0.49% (0.39%, 0.62%)	4.77% (4.40%, 5.16%)	9.66 (7.64, 12.14)
Dementia	0.77% (0.62%, 0.93%)	5.00% (4.60%, 5.33%)	6.39 (5.4, 8.09)
Heart Failure	0.96% (0.81%, 1.16%)	10.80% (10.24%, 11.29%)	11.27 (9.43, 13.5)
Atrial Fibrillation	1.69% (1.48%, 1.96%)	13.74% (13.24%, 14.50%)	8.13 (6.95, 9.37)
T2 Diabetes	0.36% (0.25%, 0.46%)	21.87% (21.07%, 22.65%)	61.45 (47, 86.12)
Liver Disease	2.13% (1.89%, 2.36%)	12.30% (11.76%, 13.02%)	5.79 (5.17, 6.68)
Renal Disease	4.31% (3.96%, 4.69%)	27.19% (26.57%, 28.03%)	6.3 (5.85, 6.87)
PAD	1.80% (1.60%, 1.99%)	8.74% (8.26%, 9.17%)	4.89 (4.4, 5.53)
Ven. Thrombosis	0.81% (0.66%, 0.98%)	3.24% (2.97%, 3.60%)	4.03 (3.26, 5.05)
AAA	0.18% (0.11%, 0.24%)	2.46% (2.15%, 2.72%)	14.1 (9.93, 24.45)
COPD	3.44% (3.09%, 3.81%)	17.06% (16.27%, 17.84%)	4.98 (4.37, 5.62)
Asthma	2.48% (2.20%, 2.72%)	5.52% (5.17%, 5.99%)	2.22 (2.01, 2.57)
Parkinson's	0.25% (0.17%, 0.34%)	1.31% (1.11%, 1.49%)	5.18 (3.41, 7.44)
Cataracts	3.32% (2.95%, 3.65%)	17.79% (17.00%, 18.45%)	5.4 (4.83, 6.08)
Glaucoma	1.57% (1.37%, 1.79%)	3.47% (3.17%, 3.73%)	2.19 (1.91, 2.62)
Fractures	4.40% (4.10%, 4.88%)	11.29% (10.67%, 11.90%)	2.55 (2.26, 2.77)

3 Metabolomic profiles predict individual multi-disease outcomes

Table 6: Absolute discriminatory performances for all scores in the UKB.

	Metabolomics	Age+Sex	Age+Sex+Metabolomics	ASCVD	ASCVD+Metabolomics	PANEL	PANEL+Metabolomics
MACE	0.672 (0.665, 0.679)	0.697 (0.691, 0.704)	0.72 (0.714, 0.727)	0.725 (0.719, 0.731)	0.732 (0.726, 0.738)	0.735 (0.729, 0.741)	0.737 (0.731, 0.743)
CHD	0.658 (0.652, 0.664)	0.656 (0.65, 0.662)	0.686 (0.68, 0.692)	0.684 (0.678, 0.69)	0.694 (0.688, 0.7)	0.697 (0.692, 0.704)	0.7 (0.695, 0.706)
Cerebral Stroke	0.686 (0.673, 0.698)	0.714 (0.703, 0.725)	0.736 (0.726, 0.747)	0.747 (0.736, 0.758)	0.754 (0.742, 0.764)	0.757 (0.746, 0.768)	0.758 (0.748, 0.769)
Dementia	0.634 (0.621, 0.648)	0.649 (0.635, 0.662)	0.669 (0.655, 0.681)	0.661 (0.647, 0.673)	0.673 (0.659, 0.685)	0.677 (0.664, 0.689)	0.682 (0.669, 0.694)
Heart Failure	0.695 (0.686, 0.705)	0.67 (0.66, 0.678)	0.717 (0.708, 0.726)	0.699 (0.69, 0.707)	0.721 (0.712, 0.73)	0.733 (0.724, 0.742)	0.738 (0.729, 0.746)
Atrial Fibrillation	0.676 (0.669, 0.684)	0.73 (0.723, 0.736)	0.743 (0.737, 0.749)	0.741 (0.735, 0.747)	0.747 (0.74, 0.752)	0.766 (0.76, 0.772)	0.767 (0.761, 0.773)
T2 Diabetes	0.818 (0.812, 0.824)	0.61 (0.602, 0.618)	0.82 (0.813, 0.826)	0.724 (0.717, 0.731)	0.821 (0.815, 0.827)	0.852 (0.847, 0.858)	0.861 (0.856, 0.867)
Liver Disease	0.656 (0.648, 0.664)	0.58 (0.572, 0.588)	0.661 (0.653, 0.669)	0.625 (0.616, 0.633)	0.664 (0.655, 0.672)	0.685 (0.678, 0.694)	0.687 (0.68, 0.696)
Renal Disease	0.655 (0.65, 0.661)	0.63 (0.624, 0.635)	0.674 (0.668, 0.68)	0.65 (0.645, 0.656)	0.677 (0.672, 0.683)	0.686 (0.68, 0.691)	0.689 (0.684, 0.695)
PAD	0.657 (0.647, 0.667)	0.663 (0.654, 0.672)	0.69 (0.681, 0.7)	0.69 (0.681, 0.699)	0.7 (0.691, 0.71)	0.705 (0.696, 0.714)	0.707 (0.698, 0.717)
Ven. Thrombosis	0.614 (0.6, 0.627)	0.605 (0.592, 0.616)	0.634 (0.62, 0.646)	0.614 (0.6, 0.626)	0.634 (0.62, 0.647)	0.663 (0.651, 0.676)	0.664 (0.651, 0.677)
AAA	0.731 (0.712, 0.749)	0.786 (0.771, 0.8)	0.797 (0.782, 0.811)	0.803 (0.789, 0.817)	0.805 (0.791, 0.819)	0.808 (0.794, 0.822)	0.808 (0.793, 0.822)
COPD	0.63 (0.624, 0.636)	0.596 (0.59, 0.602)	0.645 (0.638, 0.651)	0.639 (0.633, 0.646)	0.659 (0.653, 0.665)	0.664 (0.658, 0.67)	0.668 (0.662, 0.674)
Asthma	0.589 (0.58, 0.598)	0.545 (0.536, 0.555)	0.593 (0.583, 0.601)	0.574 (0.564, 0.583)	0.596 (0.587, 0.605)	0.613 (0.603, 0.622)	0.615 (0.605, 0.624)
Parkinson's	0.617 (0.591, 0.642)	0.753 (0.735, 0.772)	0.745 (0.727, 0.763)	0.757 (0.738, 0.774)	0.747 (0.729, 0.765)	0.749 (0.731, 0.767)	0.739 (0.72, 0.757)
Cataracts	0.638 (0.632, 0.644)	0.731 (0.727, 0.736)	0.735 (0.73, 0.739)	0.738 (0.733, 0.743)	0.739 (0.734, 0.743)	0.741 (0.737, 0.746)	0.741 (0.737, 0.746)
Glaucoma	0.567 (0.555, 0.579)	0.652 (0.642, 0.663)	0.65 (0.639, 0.661)	0.657 (0.646, 0.667)	0.654 (0.643, 0.665)	0.654 (0.644, 0.665)	0.651 (0.641, 0.662)
Fractures	0.581 (0.574, 0.588)	0.591 (0.584, 0.598)	0.601 (0.594, 0.607)	0.595 (0.588, 0.602)	0.601 (0.594, 0.608)	0.604 (0.598, 0.611)	0.606 (0.6, 0.613)
Lung Cancer	0.698 (0.682, 0.715)	0.683 (0.669, 0.698)	0.74 (0.724, 0.754)	0.771 (0.757, 0.786)	0.78 (0.766, 0.795)	0.793 (0.779, 0.806)	0.79 (0.776, 0.804)
Skin Cancer	0.571 (0.563, 0.58)	0.659 (0.651, 0.667)	0.659 (0.651, 0.667)	0.661 (0.653, 0.67)	0.66 (0.652, 0.669)	0.667 (0.659, 0.675)	0.666 (0.658, 0.674)
Colon Cancer	0.605 (0.587, 0.623)	0.682 (0.665, 0.699)	0.681 (0.664, 0.696)	0.684 (0.668, 0.7)	0.679 (0.662, 0.695)	0.687 (0.669, 0.702)	0.68 (0.662, 0.695)
Rectal Cancer	0.609 (0.587, 0.629)	0.68 (0.66, 0.699)	0.678 (0.657, 0.698)	0.682 (0.661, 0.7)	0.676 (0.655, 0.696)	0.683 (0.663, 0.702)	0.676 (0.654, 0.694)
Prostate Cancer	0.547 (0.535, 0.56)	0.689 (0.679, 0.698)	0.69 (0.68, 0.699)	0.692 (0.682, 0.701)	0.692 (0.682, 0.701)	0.693 (0.683, 0.702)	0.693 (0.683, 0.702)
Breast Cancer	0.532 (0.52, 0.545)	0.546 (0.533, 0.557)	0.548 (0.536, 0.559)	0.542 (0.53, 0.553)	0.544 (0.532, 0.556)	0.56 (0.547, 0.572)	0.56 (0.547, 0.572)
Overall	0.639 (0.637, 0.642)	0.658 (0.656, 0.66)	0.688 (0.686, 0.69)	0.681 (0.679, 0.684)	0.694 (0.691, 0.696)	0.703 (0.701, 0.705)	0.704 (0.701, 0.706)

Supplementary Tables

Table 7: Relative discriminatory performances for all comparisons in the UKB. Relative performances are reported as medians and 95 % CIs, estimated by bootstrapping over 1000 iterations. Differences were considered significant if the 95 % CIs did not cross zero (no difference).

Endpoint	Comparisons	Median Delta C-Index	Lower 95% CI Delta C-Index	Upper 95% CI Delta C-Index	Significance
MACE	Age+Sex vs. Metabolomics	0.025	0.016	0.033	*
MACE	Age+Sex+Metabolomics vs. Age+Sex	0.023	0.020	0.027	*
MACE	Age+Sex+Metabolomics vs. Metabolomics	0.048	0.043	0.053	*
MACE	ASCVD vs. Age+Sex	0.028	0.024	0.032	*
MACE	ASCVD vs. Age+Sex+Metabolomics	0.005	0.001	0.008	*
MACE	ASCVD vs. Metabolomics	0.053	0.046	0.060	*
MACE	ASCVD+Metabolomics vs. Age+Sex	0.035	0.031	0.039	*
MACE	ASCVD+Metabolomics vs. Age+Sex+Metabolomics	0.011	0.009	0.014	*
MACE	ASCVD+Metabolomics vs. ASCVD	0.007	0.005	0.009	*
MACE	ASCVD+Metabolomics vs. Metabolomics	0.060	0.054	0.065	*
MACE	PANEL vs. Age+Sex	0.038	0.034	0.042	*
MACE	PANEL vs. Age+Sex+Metabolomics	0.015	0.011	0.018	*
MACE	PANEL vs. ASCVD	0.010	0.008	0.012	*
MACE	PANEL vs. ASCVD+Metabolomics	0.003	0.001	0.005	*
MACE	PANEL vs. Metabolomics	0.063	0.056	0.069	*
MACE	PANEL+Metabolomics vs. Age+Sex	0.040	0.036	0.045	*
MACE	PANEL+Metabolomics vs. Age+Sex+Metabolomics	0.017	0.014	0.019	*
MACE	PANEL+Metabolomics vs. ASCVD	0.012	0.009	0.015	*
MACE	PANEL+Metabolomics vs. ASCVD+Metabolomics	0.005	0.004	0.006	*
MACE	PANEL+Metabolomics vs. Metabolomics	0.065	0.059	0.070	*
MACE	PANEL+Metabolomics vs. PANEL	0.002	0.000	0.004	*
CHD	Age+Sex vs. Metabolomics	-0.002	-0.009	0.005	ns
CHD	Age+Sex+Metabolomics vs. Age+Sex	0.029	0.026	0.033	*
CHD	Age+Sex+Metabolomics vs. Metabolomics	0.028	0.024	0.031	*
CHD	ASCVD vs. Age+Sex	0.028	0.025	0.031	*
CHD	ASCVD vs. Age+Sex+Metabolomics	-0.002	-0.005	0.002	ns
CHD	ASCVD vs. Metabolomics	0.026	0.021	0.032	*
CHD	ASCVD+Metabolomics vs. Age+Sex	0.038	0.034	0.042	*
CHD	ASCVD+Metabolomics vs. Age+Sex+Metabolomics	0.008	0.007	0.010	*
CHD	ASCVD+Metabolomics vs. ASCVD	0.010	0.007	0.012	*
CHD	ASCVD+Metabolomics vs. Metabolomics	0.036	0.032	0.040	*
CHD	PANEL vs. Age+Sex	0.041	0.037	0.045	*
CHD	PANEL vs. Age+Sex+Metabolomics	0.012	0.009	0.015	*
CHD	PANEL vs. ASCVD	0.013	0.011	0.016	*
CHD	PANEL vs. ASCVD+Metabolomics	0.004	0.001	0.006	*
CHD	PANEL vs. Metabolomics	0.040	0.034	0.045	*
CHD	PANEL+Metabolomics vs. Age+Sex	0.044	0.040	0.048	*
CHD	PANEL+Metabolomics vs. Age+Sex+Metabolomics	0.015	0.012	0.017	*
CHD	PANEL+Metabolomics vs. ASCVD	0.016	0.013	0.019	*
CHD	PANEL+Metabolomics vs. ASCVD+Metabolomics	0.006	0.005	0.008	*
CHD	PANEL+Metabolomics vs. Metabolomics	0.042	0.038	0.046	*
CHD	PANEL+Metabolomics vs. PANEL	0.003	0.001	0.004	*
Cerebral Stroke	Age+Sex vs. Metabolomics	0.028	0.015	0.042	*
Cerebral Stroke	Age+Sex+Metabolomics vs. Age+Sex	0.023	0.015	0.029	*
Cerebral Stroke	Age+Sex+Metabolomics vs. Metabolomics	0.051	0.043	0.059	*
Cerebral Stroke	ASCVD vs. Age+Sex	0.033	0.026	0.040	*
Cerebral Stroke	ASCVD vs. Age+Sex+Metabolomics	0.011	0.003	0.017	*
Cerebral Stroke	ASCVD vs. Metabolomics	0.061	0.049	0.074	*

3 Metabolomic profiles predict individual multi-disease outcomes

Table 7 continued from previous page

Endpoint	Comparisons	Median Delta C-Index	Lower 95% CI Delta C-Index	Upper 95% CI Delta C-Index	Significance
Cerebral Stroke	ASCVD+Metabolomics vs. Age+Sex	0.040	0.032	0.048	*
Cerebral Stroke	ASCVD+Metabolomics vs. Age+Sex+Metabolomics	0.017	0.013	0.022	*
Cerebral Stroke	ASCVD+Metabolomics vs. ASCVD	0.007	0.002	0.011	*
Cerebral Stroke	ASCVD+Metabolomics vs. Metabolomics	0.068	0.060	0.076	*
Cerebral Stroke	PANEL vs. Age+Sex	0.043	0.036	0.051	*
Cerebral Stroke	PANEL vs. Age+Sex+Metabolomics	0.021	0.014	0.027	*
Cerebral Stroke	PANEL vs. ASCVD	0.010	0.007	0.014	*
Cerebral Stroke	PANEL vs. ASCVD+Metabolomics	0.004	-0.001	0.008	ns
Cerebral Stroke	PANEL vs. Metabolomics	0.072	0.060	0.083	*
Cerebral Stroke	PANEL+Metabolomics vs. Age+Sex	0.045	0.036	0.053	*
Cerebral Stroke	PANEL+Metabolomics vs. Age+Sex+Metabolomics	0.022	0.017	0.027	*
Cerebral Stroke	PANEL+Metabolomics vs. ASCVD	0.011	0.006	0.017	*
Cerebral Stroke	PANEL+Metabolomics vs. ASCVD+Metabolomics	0.005	0.002	0.007	*
Cerebral Stroke	PANEL+Metabolomics vs. Metabolomics	0.073	0.064	0.081	*
Cerebral Stroke	PANEL+Metabolomics vs. PANEL	0.001	-0.003	0.005	ns
Dementia	Age+Sex vs. Metabolomics	0.014	0.000	0.030	ns
Dementia	Age+Sex+Metabolomics vs. Age+Sex	0.020	0.013	0.028	*
Dementia	Age+Sex+Metabolomics vs. Metabolomics	0.035	0.026	0.044	*
Dementia	ASCVD vs. Age+Sex	0.012	0.007	0.016	*
Dementia	ASCVD vs. Age+Sex+Metabolomics	-0.008	-0.015	-0.001	*
Dementia	ASCVD vs. Metabolomics	0.026	0.012	0.042	*
Dementia	ASCVD+Metabolomics vs. Age+Sex	0.024	0.017	0.031	*
Dementia	ASCVD+Metabolomics vs. Age+Sex+Metabolomics	0.004	0.002	0.006	*
Dementia	ASCVD+Metabolomics vs. ASCVD	0.012	0.006	0.018	*
Dementia	ASCVD+Metabolomics vs. Metabolomics	0.038	0.030	0.048	*
Dementia	PANEL vs. Age+Sex	0.028	0.021	0.035	*
Dementia	PANEL vs. Age+Sex+Metabolomics	0.008	0.001	0.015	*
Dementia	PANEL vs. ASCVD	0.016	0.011	0.022	*
Dementia	PANEL vs. ASCVD+Metabolomics	0.004	-0.002	0.011	ns
Dementia	PANEL vs. Metabolomics	0.043	0.030	0.056	*
Dementia	PANEL+Metabolomics vs. Age+Sex	0.033	0.025	0.042	*
Dementia	PANEL+Metabolomics vs. Age+Sex+Metabolomics	0.013	0.009	0.017	*
Dementia	PANEL+Metabolomics vs. ASCVD	0.021	0.014	0.029	*
Dementia	PANEL+Metabolomics vs. ASCVD+Metabolomics	0.009	0.005	0.013	*
Dementia	PANEL+Metabolomics vs. Metabolomics	0.048	0.038	0.058	*
Dementia	PANEL+Metabolomics vs. PANEL	0.005	0.000	0.009	*
Heart Failure	Age+Sex vs. Metabolomics	-0.026	-0.036	-0.016	*
Heart Failure	Age+Sex+Metabolomics vs. Age+Sex	0.047	0.041	0.054	*
Heart Failure	Age+Sex+Metabolomics vs. Metabolomics	0.022	0.017	0.026	*
Heart Failure	ASCVD vs. Age+Sex	0.029	0.024	0.034	*
Heart Failure	ASCVD vs. Age+Sex+Metabolomics	-0.018	-0.024	-0.013	*
Heart Failure	ASCVD vs. Metabolomics	0.003	-0.006	0.012	ns
Heart Failure	ASCVD+Metabolomics vs. Age+Sex	0.052	0.045	0.058	*
Heart Failure	ASCVD+Metabolomics vs. Age+Sex+Metabolomics	0.004	0.002	0.006	*
Heart Failure	ASCVD+Metabolomics vs. ASCVD	0.022	0.018	0.027	*
Heart Failure	ASCVD+Metabolomics vs. Metabolomics	0.026	0.021	0.030	*
Heart Failure	PANEL vs. Age+Sex	0.063	0.057	0.070	*
Heart Failure	PANEL vs. Age+Sex+Metabolomics	0.016	0.010	0.021	*
Heart Failure	PANEL vs. ASCVD	0.034	0.030	0.039	*
Heart Failure	PANEL vs. ASCVD+Metabolomics	0.012	0.006	0.016	*

Supplementary Tables

Table 7 continued from previous page

Endpoint	Comparisons	Median Delta C-Index	Lower 95% CI Delta C-Index	Upper 95% CI Delta C-Index	Significance
Heart Failure	PANEL vs. Metabolomics	0.038	0.029	0.045	*
Heart Failure	PANEL+Metabolomics vs. Age+Sex	0.068	0.061	0.075	*
Heart Failure	PANEL+Metabolomics vs. Age+Sex+Metabolomics	0.021	0.017	0.025	*
Heart Failure	PANEL+Metabolomics vs. ASCVD	0.039	0.034	0.045	*
Heart Failure	PANEL+Metabolomics vs. ASCVD+Metabolomics	0.016	0.013	0.020	*
Heart Failure	PANEL+Metabolomics vs. Metabolomics	0.042	0.036	0.048	*
Heart Failure	PANEL+Metabolomics vs. PANEL	0.005	0.003	0.007	*
Atrial Fibrillation	Age+Sex vs. Metabolomics	0.053	0.045	0.061	*
Atrial Fibrillation	Age+Sex+Metabolomics vs. Age+Sex	0.014	0.011	0.017	*
Atrial Fibrillation	Age+Sex+Metabolomics vs. Metabolomics	0.067	0.061	0.072	*
Atrial Fibrillation	ASCVD vs. Age+Sex	0.011	0.009	0.013	*
Atrial Fibrillation	ASCVD vs. Age+Sex+Metabolomics	-0.003	-0.005	0.000	ns
Atrial Fibrillation	ASCVD vs. Metabolomics	0.064	0.057	0.071	*
Atrial Fibrillation	ASCVD+Metabolomics vs. Age+Sex	0.017	0.014	0.020	*
Atrial Fibrillation	ASCVD+Metabolomics vs. Age+Sex+Metabolomics	0.003	0.002	0.004	*
Atrial Fibrillation	ASCVD+Metabolomics vs. ASCVD	0.006	0.004	0.008	*
Atrial Fibrillation	ASCVD+Metabolomics vs. Metabolomics	0.070	0.064	0.075	*
Atrial Fibrillation	PANEL vs. Age+Sex	0.037	0.033	0.041	*
Atrial Fibrillation	PANEL vs. Age+Sex+Metabolomics	0.023	0.020	0.027	*
Atrial Fibrillation	PANEL vs. ASCVD	0.026	0.023	0.029	*
Atrial Fibrillation	PANEL vs. ASCVD+Metabolomics	0.020	0.017	0.023	*
Atrial Fibrillation	PANEL vs. Metabolomics	0.090	0.083	0.097	*
Atrial Fibrillation	PANEL+Metabolomics vs. Age+Sex	0.037	0.033	0.041	*
Atrial Fibrillation	PANEL+Metabolomics vs. Age+Sex+Metabolomics	0.024	0.021	0.027	*
Atrial Fibrillation	PANEL+Metabolomics vs. ASCVD	0.026	0.023	0.030	*
Atrial Fibrillation	PANEL+Metabolomics vs. ASCVD+Metabolomics	0.020	0.017	0.023	*
Atrial Fibrillation	PANEL+Metabolomics vs. Metabolomics	0.090	0.084	0.097	*
Atrial Fibrillation	PANEL+Metabolomics vs. PANEL	0.000	-0.001	0.001	ns
T2 Diabetes	Age+Sex vs. Metabolomics	-0.208	-0.218	-0.199	*
T2 Diabetes	Age+Sex+Metabolomics vs. Age+Sex	0.209	0.200	0.219	*
T2 Diabetes	Age+Sex+Metabolomics vs. Metabolomics	0.001	0.001	0.002	*
T2 Diabetes	ASCVD vs. Age+Sex	0.114	0.105	0.123	*
T2 Diabetes	ASCVD vs. Age+Sex+Metabolomics	-0.095	-0.102	-0.088	*
T2 Diabetes	ASCVD vs. Metabolomics	-0.094	-0.100	-0.087	*
T2 Diabetes	ASCVD+Metabolomics vs. Age+Sex	0.211	0.202	0.221	*
T2 Diabetes	ASCVD+Metabolomics vs. Age+Sex+Metabolomics	0.002	0.001	0.002	*
T2 Diabetes	ASCVD+Metabolomics vs. ASCVD	0.097	0.091	0.103	*
T2 Diabetes	ASCVD+Metabolomics vs. Metabolomics	0.003	0.002	0.004	*

3 Metabolomic profiles predict individual multi-disease outcomes

Table 7 continued from previous page

Endpoint	Comparisons	Median Delta C-Index	Lower 95% CI Delta C-Index	Upper 95% CI Delta C-Index	Significance
T2 Diabetes	PANEL vs. Age+Sex	0.242	0.233	0.252	*
T2 Diabetes	PANEL vs. Age+Sex+Metabolomics	0.033	0.028	0.037	*
T2 Diabetes	PANEL vs. ASCVD	0.128	0.122	0.134	*
T2 Diabetes	PANEL vs. ASCVD+Metabolomics	0.031	0.026	0.035	*
T2 Diabetes	PANEL vs. Metabolomics	0.034	0.029	0.039	*
T2 Diabetes	PANEL+Metabolomics vs. Age+Sex	0.251	0.242	0.261	*
T2 Diabetes	PANEL+Metabolomics vs. Age+Sex+Metabolomics	0.042	0.039	0.045	*
T2 Diabetes	PANEL+Metabolomics vs. ASCVD	0.137	0.131	0.144	*
T2 Diabetes	PANEL+Metabolomics vs. ASCVD+Metabolomics	0.040	0.037	0.043	*
T2 Diabetes	PANEL+Metabolomics vs. Metabolomics	0.043	0.041	0.046	*
T2 Diabetes	PANEL+Metabolomics vs. PANEL	0.009	0.007	0.012	*
Liver Disease	Age+Sex vs. Metabolomics	-0.076	-0.085	-0.065	*
Liver Disease	Age+Sex+Metabolomics vs. Age+Sex	0.081	0.072	0.089	*
Liver Disease	Age+Sex+Metabolomics vs. Metabolomics	0.005	0.004	0.008	*
Liver Disease	ASCVD vs. Age+Sex	0.045	0.038	0.051	*
Liver Disease	ASCVD vs. Age+Sex+Metabolomics	-0.037	-0.044	-0.030	*
Liver Disease	ASCVD vs. Metabolomics	-0.031	-0.039	-0.023	*
Liver Disease	ASCVD+Metabolomics vs. Age+Sex	0.084	0.075	0.091	*
Liver Disease	ASCVD+Metabolomics vs. Age+Sex+Metabolomics	0.002	0.001	0.004	*
Liver Disease	ASCVD+Metabolomics vs. ASCVD	0.039	0.032	0.046	*
Liver Disease	ASCVD+Metabolomics vs. Metabolomics	0.008	0.005	0.010	*
Liver Disease	PANEL vs. Age+Sex	0.105	0.097	0.114	*
Liver Disease	PANEL vs. Age+Sex+Metabolomics	0.024	0.019	0.030	*
Liver Disease	PANEL vs. ASCVD	0.061	0.054	0.068	*
Liver Disease	PANEL vs. ASCVD+Metabolomics	0.022	0.017	0.028	*
Liver Disease	PANEL vs. Metabolomics	0.030	0.024	0.036	*
Liver Disease	PANEL+Metabolomics vs. Age+Sex	0.107	0.098	0.116	*
Liver Disease	PANEL+Metabolomics vs. Age+Sex+Metabolomics	0.026	0.023	0.030	*
Liver Disease	PANEL+Metabolomics vs. ASCVD	0.063	0.056	0.070	*
Liver Disease	PANEL+Metabolomics vs. ASCVD+Metabolomics	0.024	0.021	0.028	*
Liver Disease	PANEL+Metabolomics vs. Metabolomics	0.032	0.028	0.036	*
Liver Disease	PANEL+Metabolomics vs. PANEL	0.002	-0.001	0.005	ns
Renal Disease	Age+Sex vs. Metabolomics	-0.026	-0.032	-0.020	*
Renal Disease	Age+Sex+Metabolomics vs. Age+Sex	0.045	0.041	0.049	*
Renal Disease	Age+Sex+Metabolomics vs. Metabolomics	0.019	0.016	0.021	*
Renal Disease	ASCVD vs. Age+Sex	0.021	0.018	0.023	*
Renal Disease	ASCVD vs. Age+Sex+Metabolomics	-0.024	-0.028	-0.020	*
Renal Disease	ASCVD vs. Metabolomics	-0.005	-0.011	0.001	ns
Renal Disease	ASCVD+Metabolomics vs. Age+Sex	0.048	0.044	0.052	*
Renal Disease	ASCVD+Metabolomics vs. Age+Sex+Metabolomics	0.003	0.002	0.004	*
Renal Disease	ASCVD+Metabolomics vs. ASCVD	0.027	0.024	0.030	*
Renal Disease	ASCVD+Metabolomics vs. Metabolomics	0.022	0.019	0.025	*
Renal Disease	PANEL vs. Age+Sex	0.056	0.052	0.060	*
Renal Disease	PANEL vs. Age+Sex+Metabolomics	0.011	0.008	0.015	*
Renal Disease	PANEL vs. ASCVD	0.035	0.032	0.038	*
Renal Disease	PANEL vs. ASCVD+Metabolomics	0.008	0.005	0.011	*
Renal Disease	PANEL vs. Metabolomics	0.030	0.025	0.035	*
Renal Disease	PANEL+Metabolomics vs. Age+Sex	0.060	0.056	0.064	*
Renal Disease	PANEL+Metabolomics vs. Age+Sex+Metabolomics	0.015	0.013	0.017	*
Renal Disease	PANEL+Metabolomics vs. ASCVD	0.039	0.035	0.042	*

Supplementary Tables

Table 7 continued from previous page

Endpoint	Comparisons	Median Delta C-Index	Lower 95% CI Delta C-Index	Upper 95% CI Delta C-Index	Significance
Renal Disease	PANEL+Metabolomics vs. AS-CVD+Metabolomics	0.012	0.010	0.014	*
Renal Disease	PANEL+Metabolomics vs. Metabolomics	0.034	0.030	0.037	*
Renal Disease	PANEL+Metabolomics vs. PANEL	0.004	0.002	0.005	*
PAD	Age+Sex vs. Metabolomics	0.006	-0.004	0.016	ns
PAD	Age+Sex+Metabolomics vs. Age+Sex	0.027	0.021	0.033	*
PAD	Age+Sex+Metabolomics vs. Metabolomics	0.033	0.028	0.038	*
PAD	ASCVD vs. Age+Sex	0.027	0.022	0.032	*
PAD	ASCVD vs. Age+Sex+Metabolomics	0.000	-0.006	0.005	ns
PAD	ASCVD vs. Metabolomics	0.033	0.024	0.041	*
PAD	ASCVD+Metabolomics vs. Age+Sex	0.037	0.031	0.044	*
PAD	ASCVD+Metabolomics vs. Age+Sex+Metabolomics	0.010	0.007	0.013	*
PAD	ASCVD+Metabolomics vs. ASCVD	0.010	0.006	0.014	*
PAD	ASCVD+Metabolomics vs. Metabolomics	0.043	0.037	0.049	*
PAD	PANEL vs. Age+Sex	0.042	0.036	0.049	*
PAD	PANEL vs. Age+Sex+Metabolomics	0.015	0.009	0.020	*
PAD	PANEL vs. ASCVD	0.015	0.012	0.019	*
PAD	PANEL vs. ASCVD+Metabolomics	0.005	0.001	0.009	*
PAD	PANEL vs. Metabolomics	0.048	0.040	0.056	*
PAD	PANEL+Metabolomics vs. Age+Sex	0.044	0.038	0.052	*
PAD	PANEL+Metabolomics vs. Age+Sex+Metabolomics	0.017	0.013	0.021	*
PAD	PANEL+Metabolomics vs. ASCVD	0.018	0.013	0.022	*
PAD	PANEL+Metabolomics vs. AS-CVD+Metabolomics	0.007	0.005	0.010	*
PAD	PANEL+Metabolomics vs. Metabolomics	0.050	0.044	0.057	*
PAD	PANEL+Metabolomics vs. PANEL	0.002	0.000	0.005	ns
Ven. Thrombosis	Age+Sex vs. Metabolomics	-0.009	-0.025	0.007	ns
Ven. Thrombosis	Age+Sex+Metabolomics vs. Age+Sex	0.029	0.017	0.041	*
Ven. Thrombosis	Age+Sex+Metabolomics vs. Metabolomics	0.020	0.014	0.026	*
Ven. Thrombosis	ASCVD vs. Age+Sex	0.009	0.003	0.015	*
Ven. Thrombosis	ASCVD vs. Age+Sex+Metabolomics	-0.020	-0.030	-0.009	*
Ven. Thrombosis	ASCVD vs. Metabolomics	0.000	-0.015	0.015	ns
Ven. Thrombosis	ASCVD+Metabolomics vs. Age+Sex	0.030	0.018	0.041	*
Ven. Thrombosis	ASCVD+Metabolomics vs. Age+Sex+Metabolomics	0.000	-0.002	0.002	ns
Ven. Thrombosis	ASCVD+Metabolomics vs. ASCVD	0.021	0.010	0.030	*
Ven. Thrombosis	ASCVD+Metabolomics vs. Metabolomics	0.020	0.014	0.026	*
Ven. Thrombosis	PANEL vs. Age+Sex	0.058	0.047	0.071	*
Ven. Thrombosis	PANEL vs. Age+Sex+Metabolomics	0.029	0.018	0.040	*
Ven. Thrombosis	PANEL vs. ASCVD	0.049	0.038	0.060	*
Ven. Thrombosis	PANEL vs. ASCVD+Metabolomics	0.029	0.018	0.039	*
Ven. Thrombosis	PANEL vs. Metabolomics	0.049	0.035	0.063	*
Ven. Thrombosis	PANEL+Metabolomics vs. Age+Sex	0.059	0.047	0.072	*

3 Metabolomic profiles predict individual multi-disease outcomes

Table 7 continued from previous page

Endpoint	Comparisons	Median Delta C-Index	Lower 95% CI Delta C-Index	Upper 95% CI Delta C-Index	Significance
Ven. Thrombosis	PANEL+Metabolomics vs. Age+Sex+Metabolomics	0.030	0.022	0.038	*
Ven. Thrombosis	PANEL+Metabolomics vs. ASCVD	0.050	0.038	0.061	*
Ven. Thrombosis	PANEL+Metabolomics vs. AS-CVD+Metabolomics	0.030	0.022	0.037	*
Ven. Thrombosis	PANEL+Metabolomics vs. Metabolomics	0.050	0.041	0.060	*
Ven. Thrombosis	PANEL+Metabolomics vs. PANEL	0.001	-0.005	0.007	ns
AAA	Age+Sex vs. Metabolomics	0.054	0.039	0.072	*
AAA	Age+Sex+Metabolomics vs. Age+Sex	0.011	0.004	0.018	*
AAA	Age+Sex+Metabolomics vs. Metabolomics	0.066	0.056	0.076	*
AAA	ASCVD vs. Age+Sex	0.018	0.012	0.024	*
AAA	ASCVD vs. Age+Sex+Metabolomics	0.006	-0.001	0.014	ns
AAA	ASCVD vs. Metabolomics	0.072	0.057	0.088	*
AAA	ASCVD+Metabolomics vs. Age+Sex	0.019	0.011	0.027	*
AAA	ASCVD+Metabolomics vs. Age+Sex+Metabolomics	0.008	0.004	0.012	*
AAA	ASCVD+Metabolomics vs. ASCVD	0.002	-0.004	0.007	ns
AAA	ASCVD+Metabolomics vs. Metabolomics	0.074	0.063	0.084	*
AAA	PANEL vs. Age+Sex	0.023	0.015	0.030	*
AAA	PANEL vs. Age+Sex+Metabolomics	0.011	0.003	0.020	*
AAA	PANEL vs. ASCVD	0.005	0.001	0.010	*
AAA	PANEL vs. ASCVD+Metabolomics	0.003	-0.003	0.011	ns
AAA	PANEL vs. Metabolomics	0.076	0.062	0.093	*
AAA	PANEL+Metabolomics vs. Age+Sex	0.022	0.013	0.031	*
AAA	PANEL+Metabolomics vs. Age+Sex+Metabolomics	0.011	0.005	0.017	*
AAA	PANEL+Metabolomics vs. ASCVD	0.005	-0.002	0.011	ns
AAA	PANEL+Metabolomics vs. AS-CVD+Metabolomics	0.003	-0.001	0.007	ns
AAA	PANEL+Metabolomics vs. Metabolomics	0.077	0.065	0.089	*
AAA	PANEL+Metabolomics vs. PANEL	0.000	-0.006	0.005	ns
COPD	Age+Sex vs. Metabolomics	-0.034	-0.042	-0.027	*
COPD	Age+Sex+Metabolomics vs. Age+Sex	0.049	0.044	0.054	*
COPD	Age+Sex+Metabolomics vs. Metabolomics	0.014	0.012	0.017	*
COPD	ASCVD vs. Age+Sex	0.044	0.039	0.049	*
COPD	ASCVD vs. Age+Sex+Metabolomics	-0.005	-0.011	0.000	ns
COPD	ASCVD vs. Metabolomics	0.009	0.002	0.016	*
COPD	ASCVD+Metabolomics vs. Age+Sex	0.064	0.058	0.069	*
COPD	ASCVD+Metabolomics vs. Age+Sex+Metabolomics	0.015	0.012	0.017	*
COPD	ASCVD+Metabolomics vs. ASCVD	0.020	0.016	0.023	*
COPD	ASCVD+Metabolomics vs. Metabolomics	0.029	0.025	0.033	*
COPD	PANEL vs. Age+Sex	0.068	0.062	0.074	*
COPD	PANEL vs. Age+Sex+Metabolomics	0.019	0.015	0.024	*
COPD	PANEL vs. ASCVD	0.024	0.021	0.028	*
COPD	PANEL vs. ASCVD+Metabolomics	0.005	0.001	0.008	*
COPD	PANEL vs. Metabolomics	0.034	0.028	0.039	*
COPD	PANEL+Metabolomics vs. Age+Sex	0.073	0.067	0.078	*
COPD	PANEL+Metabolomics vs. Age+Sex+Metabolomics	0.024	0.020	0.027	*
COPD	PANEL+Metabolomics vs. ASCVD	0.029	0.025	0.033	*
COPD	PANEL+Metabolomics vs. AS-CVD+Metabolomics	0.009	0.007	0.011	*
COPD	PANEL+Metabolomics vs. Metabolomics	0.038	0.034	0.042	*
COPD	PANEL+Metabolomics vs. PANEL	0.005	0.003	0.006	*

Supplementary Tables

Table 7 continued from previous page

Endpoint	Comparisons	Median Delta C-Index	Lower 95% CI Delta C-Index	Upper 95% CI Delta C-Index	Significance
Asthma	Age+Sex vs. Metabolomics	-0.044	-0.055	-0.033	*
Asthma	Age+Sex+Metabolomics vs. Age+Sex	0.047	0.038	0.057	*
Asthma	Age+Sex+Metabolomics vs. Metabolomics	0.003	0.001	0.006	*
Asthma	ASCVD vs. Age+Sex	0.029	0.021	0.037	*
Asthma	ASCVD vs. Age+Sex+Metabolomics	-0.019	-0.028	-0.010	*
Asthma	ASCVD vs. Metabolomics	-0.015	-0.026	-0.006	*
Asthma	ASCVD+Metabolomics vs. Age+Sex	0.051	0.042	0.061	*
Asthma	ASCVD+Metabolomics vs. Age+Sex+Metabolomics	0.004	0.002	0.006	*
Asthma	ASCVD+Metabolomics vs. ASCVD	0.023	0.015	0.030	*
Asthma	ASCVD+Metabolomics vs. Metabolomics	0.007	0.004	0.010	*
Asthma	PANEL vs. Age+Sex	0.068	0.058	0.078	*
Asthma	PANEL vs. Age+Sex+Metabolomics	0.021	0.013	0.028	*
Asthma	PANEL vs. ASCVD	0.039	0.032	0.047	*
Asthma	PANEL vs. ASCVD+Metabolomics	0.017	0.010	0.024	*
Asthma	PANEL vs. Metabolomics	0.024	0.016	0.032	*
Asthma	PANEL+Metabolomics vs. Age+Sex	0.069	0.059	0.080	*
Asthma	PANEL+Metabolomics vs. Age+Sex+Metabolomics	0.022	0.017	0.027	*
Asthma	PANEL+Metabolomics vs. ASCVD	0.041	0.033	0.049	*
Asthma	PANEL+Metabolomics vs. ASCVD+Metabolomics	0.018	0.014	0.023	*
Asthma	PANEL+Metabolomics vs. Metabolomics	0.026	0.020	0.031	*
Asthma	PANEL+Metabolomics vs. PANEL	0.001	-0.002	0.005	ns
Parkinson's	Age+Sex vs. Metabolomics	0.137	0.112	0.163	*
Parkinson's	Age+Sex+Metabolomics vs. Age+Sex	-0.009	-0.016	-0.002	*
Parkinson's	Age+Sex+Metabolomics vs. Metabolomics	0.128	0.109	0.148	*
Parkinson's	ASCVD vs. Age+Sex	0.003	-0.002	0.008	ns
Parkinson's	ASCVD vs. Age+Sex+Metabolomics	0.012	0.005	0.019	*
Parkinson's	ASCVD vs. Metabolomics	0.140	0.117	0.164	*
Parkinson's	ASCVD+Metabolomics vs. Age+Sex	-0.007	-0.015	0.001	ns
Parkinson's	ASCVD+Metabolomics vs. Age+Sex+Metabolomics	0.002	-0.001	0.005	ns
Parkinson's	ASCVD+Metabolomics vs. ASCVD	-0.010	-0.015	-0.004	*
Parkinson's	ASCVD+Metabolomics vs. Metabolomics	0.131	0.111	0.150	*
Parkinson's	PANEL vs. Age+Sex	-0.004	-0.012	0.004	ns
Parkinson's	PANEL vs. Age+Sex+Metabolomics	0.004	-0.005	0.015	ns
Parkinson's	PANEL vs. ASCVD	-0.007	-0.015	-0.001	*
Parkinson's	PANEL vs. ASCVD+Metabolomics	0.002	-0.007	0.012	ns
Parkinson's	PANEL vs. Metabolomics	0.133	0.108	0.158	*
Parkinson's	PANEL+Metabolomics vs. Age+Sex	-0.014	-0.024	-0.004	*
Parkinson's	PANEL+Metabolomics vs. Age+Sex+Metabolomics	-0.006	-0.013	0.002	ns
Parkinson's	PANEL+Metabolomics vs. ASCVD	-0.017	-0.027	-0.008	*
Parkinson's	PANEL+Metabolomics vs. ASCVD+Metabolomics	-0.008	-0.015	0.000	*
Parkinson's	PANEL+Metabolomics vs. Metabolomics	0.123	0.102	0.143	*
Parkinson's	PANEL+Metabolomics vs. PANEL	-0.010	-0.016	-0.004	*
Cataracts	Age+Sex vs. Metabolomics	0.094	0.087	0.100	*
Cataracts	Age+Sex+Metabolomics vs. Age+Sex	0.003	0.002	0.004	*
Cataracts	Age+Sex+Metabolomics vs. Metabolomics	0.097	0.091	0.102	*
Cataracts	ASCVD vs. Age+Sex	0.006	0.005	0.007	*
Cataracts	ASCVD vs. Age+Sex+Metabolomics	0.003	0.002	0.004	*
Cataracts	ASCVD vs. Metabolomics	0.100	0.094	0.106	*
Cataracts	ASCVD+Metabolomics vs. Age+Sex	0.007	0.006	0.008	*
Cataracts	ASCVD+Metabolomics vs. Age+Sex+Metabolomics	0.004	0.003	0.005	*

3 Metabolomic profiles predict individual multi-disease outcomes

Table 7 continued from previous page

Endpoint	Comparisons	Median Delta C-Index	Lower 95% CI Delta C-Index	Upper 95% CI Delta C-Index	Significance
Cataracts	ASCVD+Metabolomics vs. ASCVD	0.001	0.000	0.001	*
Cataracts	ASCVD+Metabolomics vs. Metabolomics	0.101	0.095	0.106	*
Cataracts	PANEL vs. Age+Sex	0.010	0.008	0.012	*
Cataracts	PANEL vs. Age+Sex+Metabolomics	0.007	0.006	0.008	*
Cataracts	PANEL vs. ASCVD	0.004	0.003	0.005	*
Cataracts	PANEL vs. ASCVD+Metabolomics	0.003	0.002	0.004	*
Cataracts	PANEL vs. Metabolomics	0.104	0.098	0.109	*
Cataracts	PANEL+Metabolomics vs. Age+Sex	0.010	0.008	0.011	*
Cataracts	PANEL+Metabolomics vs. Age+Sex+Metabolomics	0.007	0.006	0.008	*
Cataracts	PANEL+Metabolomics vs. ASCVD	0.004	0.003	0.005	*
Cataracts	PANEL+Metabolomics vs. ASCVD+Metabolomics	0.003	0.002	0.004	*
Cataracts	PANEL+Metabolomics vs. Metabolomics	0.104	0.098	0.109	*
Cataracts	PANEL+Metabolomics vs. PANEL	0.000	0.000	0.000	ns
Glaucoma	Age+Sex vs. Metabolomics	0.085	0.074	0.098	*
Glaucoma	Age+Sex+Metabolomics vs. Age+Sex	-0.002	-0.005	0.000	ns
Glaucoma	Age+Sex+Metabolomics vs. Metabolomics	0.083	0.074	0.094	*
Glaucoma	ASCVD vs. Age+Sex	0.004	0.001	0.007	*
Glaucoma	ASCVD vs. Age+Sex+Metabolomics	0.006	0.004	0.009	*
Glaucoma	ASCVD vs. Metabolomics	0.090	0.079	0.102	*
Glaucoma	ASCVD+Metabolomics vs. Age+Sex	0.001	-0.002	0.005	ns
Glaucoma	ASCVD+Metabolomics vs. Age+Sex+Metabolomics	0.004	0.002	0.006	*
Glaucoma	ASCVD+Metabolomics vs. ASCVD	-0.003	-0.005	-0.001	*
Glaucoma	ASCVD+Metabolomics vs. Metabolomics	0.087	0.077	0.097	*
Glaucoma	PANEL vs. Age+Sex	0.002	-0.002	0.006	ns
Glaucoma	PANEL vs. Age+Sex+Metabolomics	0.004	0.000	0.008	*
Glaucoma	PANEL vs. ASCVD	-0.002	-0.005	0.000	ns
Glaucoma	PANEL vs. ASCVD+Metabolomics	0.000	-0.002	0.003	ns
Glaucoma	PANEL vs. Metabolomics	0.087	0.076	0.099	*
Glaucoma	PANEL+Metabolomics vs. Age+Sex	-0.001	-0.004	0.003	ns
Glaucoma	PANEL+Metabolomics vs. Age+Sex+Metabolomics	0.001	-0.002	0.005	ns
Glaucoma	PANEL+Metabolomics vs. ASCVD	-0.005	-0.008	-0.002	*
Glaucoma	PANEL+Metabolomics vs. ASCVD+Metabolomics	-0.002	-0.005	0.000	*
Glaucoma	PANEL+Metabolomics vs. Metabolomics	0.084	0.074	0.095	*
Glaucoma	PANEL+Metabolomics vs. PANEL	-0.003	-0.004	-0.001	*
Fractures	Age+Sex vs. Metabolomics	0.010	0.003	0.017	*
Fractures	Age+Sex+Metabolomics vs. Age+Sex	0.009	0.006	0.014	*
Fractures	Age+Sex+Metabolomics vs. Metabolomics	0.020	0.016	0.023	*
Fractures	ASCVD vs. Age+Sex	0.004	0.002	0.006	*
Fractures	ASCVD vs. Age+Sex+Metabolomics	-0.006	-0.009	-0.002	*
Fractures	ASCVD vs. Metabolomics	0.014	0.007	0.020	*
Fractures	ASCVD+Metabolomics vs. Age+Sex	0.010	0.006	0.014	*
Fractures	ASCVD+Metabolomics vs. Age+Sex+Metabolomics	0.000	0.000	0.001	ns
Fractures	ASCVD+Metabolomics vs. ASCVD	0.006	0.003	0.009	*
Fractures	ASCVD+Metabolomics vs. Metabolomics	0.020	0.016	0.024	*
Fractures	PANEL vs. Age+Sex	0.013	0.009	0.017	*
Fractures	PANEL vs. Age+Sex+Metabolomics	0.003	0.000	0.007	ns
Fractures	PANEL vs. ASCVD	0.009	0.006	0.012	*
Fractures	PANEL vs. ASCVD+Metabolomics	0.003	-0.001	0.007	ns
Fractures	PANEL vs. Metabolomics	0.023	0.017	0.029	*
Fractures	PANEL+Metabolomics vs. Age+Sex	0.015	0.011	0.020	*

Supplementary Tables

Table 7 continued from previous page

Endpoint	Comparisons	Median Delta C-Index	Lower 95% CI Delta C-Index	Upper 95% CI Delta C-Index	Significance
Fractures	PANEL+Metabolomics vs. Age+Sex+Metabolomics	0.006	0.003	0.008	*
Fractures	PANEL+Metabolomics vs. ASCVD	0.011	0.007	0.015	*
Fractures	PANEL+Metabolomics vs. AS-CVD+Metabolomics	0.005	0.003	0.008	*
Fractures	PANEL+Metabolomics vs. Metabolomics	0.025	0.021	0.030	*
Fractures	PANEL+Metabolomics vs. PANEL	0.002	0.000	0.004	*
Lung Cancer	Age+Sex vs. Metabolomics	-0.015	-0.036	0.006	ns
Lung Cancer	Age+Sex+Metabolomics vs. Age+Sex	0.056	0.042	0.071	*
Lung Cancer	Age+Sex+Metabolomics vs. Metabolomics	0.041	0.034	0.050	*
Lung Cancer	ASCVD vs. Age+Sex	0.087	0.074	0.101	*
Lung Cancer	ASCVD vs. Age+Sex+Metabolomics	0.031	0.018	0.044	*
Lung Cancer	ASCVD vs. Metabolomics	0.072	0.054	0.090	*
Lung Cancer	ASCVD+Metabolomics vs. Age+Sex	0.097	0.081	0.111	*
Lung Cancer	ASCVD+Metabolomics vs. Age+Sex+Metabolomics	0.040	0.032	0.049	*
Lung Cancer	ASCVD+Metabolomics vs. ASCVD	0.009	0.002	0.017	*
Lung Cancer	ASCVD+Metabolomics vs. Metabolomics	0.082	0.070	0.094	*
Lung Cancer	PANEL vs. Age+Sex	0.109	0.095	0.123	*
Lung Cancer	PANEL vs. Age+Sex+Metabolomics	0.053	0.041	0.065	*
Lung Cancer	PANEL vs. ASCVD	0.022	0.016	0.029	*
Lung Cancer	PANEL vs. ASCVD+Metabolomics	0.012	0.005	0.020	*
Lung Cancer	PANEL vs. Metabolomics	0.095	0.079	0.110	*
Lung Cancer	PANEL+Metabolomics vs. Age+Sex	0.106	0.091	0.121	*
Lung Cancer	PANEL+Metabolomics vs. Age+Sex+Metabolomics	0.050	0.041	0.060	*
Lung Cancer	PANEL+Metabolomics vs. ASCVD	0.019	0.012	0.027	*
Lung Cancer	PANEL+Metabolomics vs. AS-CVD+Metabolomics	0.010	0.006	0.014	*
Lung Cancer	PANEL+Metabolomics vs. Metabolomics	0.092	0.079	0.105	*
Lung Cancer	PANEL+Metabolomics vs. PANEL	-0.003	-0.008	0.002	ns
Skin Cancer	Age+Sex vs. Metabolomics	0.088	0.079	0.097	*
Skin Cancer	Age+Sex+Metabolomics vs. Age+Sex	0.000	-0.001	0.001	ns
Skin Cancer	Age+Sex+Metabolomics vs. Metabolomics	0.088	0.079	0.096	*
Skin Cancer	ASCVD vs. Age+Sex	0.002	0.000	0.004	*
Skin Cancer	ASCVD vs. Age+Sex+Metabolomics	0.002	0.000	0.004	*
Skin Cancer	ASCVD vs. Metabolomics	0.090	0.080	0.099	*
Skin Cancer	ASCVD+Metabolomics vs. Age+Sex	0.001	-0.001	0.003	ns
Skin Cancer	ASCVD+Metabolomics vs. Age+Sex+Metabolomics	0.001	0.000	0.003	ns
Skin Cancer	ASCVD+Metabolomics vs. ASCVD	-0.001	-0.002	0.000	*
Skin Cancer	ASCVD+Metabolomics vs. Metabolomics	0.089	0.080	0.097	*
Skin Cancer	PANEL vs. Age+Sex	0.007	0.004	0.010	*
Skin Cancer	PANEL vs. Age+Sex+Metabolomics	0.007	0.004	0.010	*
Skin Cancer	PANEL vs. ASCVD	0.005	0.003	0.008	*
Skin Cancer	PANEL vs. ASCVD+Metabolomics	0.006	0.004	0.009	*
Skin Cancer	PANEL vs. Metabolomics	0.095	0.086	0.105	*
Skin Cancer	PANEL+Metabolomics vs. Age+Sex	0.006	0.003	0.009	*
Skin Cancer	PANEL+Metabolomics vs. Age+Sex+Metabolomics	0.007	0.004	0.009	*
Skin Cancer	PANEL+Metabolomics vs. ASCVD	0.004	0.002	0.007	*
Skin Cancer	PANEL+Metabolomics vs. AS-CVD+Metabolomics	0.005	0.003	0.008	*
Skin Cancer	PANEL+Metabolomics vs. Metabolomics	0.094	0.086	0.103	*
Skin Cancer	PANEL+Metabolomics vs. PANEL	-0.001	-0.002	0.000	ns
Colon Cancer	Age+Sex vs. Metabolomics	0.076	0.058	0.097	*
Colon Cancer	Age+Sex+Metabolomics vs. Age+Sex	-0.001	-0.009	0.005	ns

3 Metabolomic profiles predict individual multi-disease outcomes

Table 7 continued from previous page

Endpoint	Comparisons	Median Delta C-Index	Lower 95% CI Delta C-Index	Upper 95% CI Delta C-Index	Significance
Colon Cancer	Age+Sex+Metabolomics vs. Metabolomics	0.075	0.063	0.089	*
Colon Cancer	ASCVD vs. Age+Sex	0.001	-0.002	0.005	ns
Colon Cancer	ASCVD vs. Age+Sex+Metabolomics	0.003	-0.003	0.010	ns
Colon Cancer	ASCVD vs. Metabolomics	0.078	0.060	0.097	*
Colon Cancer	ASCVD+Metabolomics vs. Age+Sex	-0.003	-0.011	0.004	ns
Colon Cancer	ASCVD+Metabolomics vs. Age+Sex+Metabolomics	-0.001	-0.002	-0.001	*
Colon Cancer	ASCVD+Metabolomics vs. ASCVD	-0.004	-0.011	0.002	ns
Colon Cancer	ASCVD+Metabolomics vs. Metabolomics	0.074	0.061	0.088	*
Colon Cancer	PANEL vs. Age+Sex	0.005	-0.002	0.012	ns
Colon Cancer	PANEL vs. Age+Sex+Metabolomics	0.006	-0.003	0.014	ns
Colon Cancer	PANEL vs. ASCVD	0.003	-0.004	0.010	ns
Colon Cancer	PANEL vs. ASCVD+Metabolomics	0.008	-0.002	0.016	ns
Colon Cancer	PANEL vs. Metabolomics	0.081	0.063	0.099	*
Colon Cancer	PANEL+Metabolomics vs. Age+Sex	-0.002	-0.011	0.007	ns
Colon Cancer	PANEL+Metabolomics vs. Age+Sex+Metabolomics	-0.001	-0.007	0.006	ns
Colon Cancer	PANEL+Metabolomics vs. ASCVD	-0.003	-0.013	0.005	ns
Colon Cancer	PANEL+Metabolomics vs. ASCVD+Metabolomics	0.001	-0.005	0.007	ns
Colon Cancer	PANEL+Metabolomics vs. Metabolomics	0.075	0.060	0.089	*
Colon Cancer	PANEL+Metabolomics vs. PANEL	-0.007	-0.013	0.000	*
Rectal Cancer	Age+Sex vs. Metabolomics	0.070	0.050	0.090	*
Rectal Cancer	Age+Sex+Metabolomics vs. Age+Sex	-0.002	-0.011	0.006	ns
Rectal Cancer	Age+Sex+Metabolomics vs. Metabolomics	0.068	0.054	0.081	*
Rectal Cancer	ASCVD vs. Age+Sex	0.001	-0.002	0.005	ns
Rectal Cancer	ASCVD vs. Age+Sex+Metabolomics	0.004	-0.004	0.012	ns
Rectal Cancer	ASCVD vs. Metabolomics	0.072	0.052	0.091	*
Rectal Cancer	ASCVD+Metabolomics vs. Age+Sex	-0.004	-0.013	0.005	ns
Rectal Cancer	ASCVD+Metabolomics vs. Age+Sex+Metabolomics	-0.001	-0.003	0.000	ns
Rectal Cancer	ASCVD+Metabolomics vs. ASCVD	-0.006	-0.014	0.002	ns
Rectal Cancer	ASCVD+Metabolomics vs. Metabolomics	0.067	0.053	0.080	*
Rectal Cancer	PANEL vs. Age+Sex	0.003	-0.006	0.011	ns
Rectal Cancer	PANEL vs. Age+Sex+Metabolomics	0.005	-0.006	0.016	ns
Rectal Cancer	PANEL vs. ASCVD	0.001	-0.006	0.009	ns
Rectal Cancer	PANEL vs. ASCVD+Metabolomics	0.007	-0.004	0.017	ns
Rectal Cancer	PANEL vs. Metabolomics	0.073	0.052	0.093	*
Rectal Cancer	PANEL+Metabolomics vs. Age+Sex	-0.005	-0.016	0.006	ns
Rectal Cancer	PANEL+Metabolomics vs. Age+Sex+Metabolomics	-0.002	-0.010	0.005	ns
Rectal Cancer	PANEL+Metabolomics vs. ASCVD	-0.007	-0.017	0.004	ns
Rectal Cancer	PANEL+Metabolomics vs. ASCVD+Metabolomics	-0.001	-0.008	0.006	ns
Rectal Cancer	PANEL+Metabolomics vs. Metabolomics	0.066	0.051	0.080	*
Rectal Cancer	PANEL+Metabolomics vs. PANEL	-0.008	-0.015	0.000	*
Prostate Cancer	Age+Sex vs. Metabolomics	0.142	0.127	0.154	*
Prostate Cancer	Age+Sex+Metabolomics vs. Age+Sex	0.002	0.001	0.002	*
Prostate Cancer	Age+Sex+Metabolomics vs. Metabolomics	0.143	0.128	0.156	*
Prostate Cancer	ASCVD vs. Age+Sex	0.004	0.001	0.006	*
Prostate Cancer	ASCVD vs. Age+Sex+Metabolomics	0.002	0.000	0.004	ns
Prostate Cancer	ASCVD vs. Metabolomics	0.145	0.131	0.159	*
Prostate Cancer	ASCVD+Metabolomics vs. Age+Sex	0.004	0.002	0.006	*
Prostate Cancer	ASCVD+Metabolomics vs. Age+Sex+Metabolomics	0.002	0.000	0.004	*
Prostate Cancer	ASCVD+Metabolomics vs. ASCVD	0.000	0.000	0.001	ns

Supplementary Tables

Table 7 continued from previous page

Endpoint	Comparisons	Median Delta C-Index	Lower 95% CI Delta C-Index	Upper 95% CI Delta C-Index	Significance
Prostate Cancer	ASCVD+Metabolomics vs. Metabolomics	0.145	0.131	0.159	*
Prostate Cancer	PANEL vs. Age+Sex	0.004	0.001	0.008	*
Prostate Cancer	PANEL vs. Age+Sex+Metabolomics	0.003	-0.001	0.007	ns
Prostate Cancer	PANEL vs. ASCVD	0.001	-0.002	0.004	ns
Prostate Cancer	PANEL vs. ASCVD+Metabolomics	0.001	-0.003	0.004	ns
Prostate Cancer	PANEL vs. Metabolomics	0.146	0.131	0.160	*
Prostate Cancer	PANEL+Metabolomics vs. Age+Sex	0.004	0.001	0.008	*
Prostate Cancer	PANEL+Metabolomics vs. Age+Sex+Metabolomics	0.003	-0.001	0.007	ns
Prostate Cancer	PANEL+Metabolomics vs. ASCVD	0.001	-0.003	0.004	ns
Prostate Cancer	PANEL+Metabolomics vs. ASCVD+Metabolomics	0.000	-0.003	0.003	ns
Prostate Cancer	PANEL+Metabolomics vs. Metabolomics	0.146	0.130	0.160	*
Prostate Cancer	PANEL+Metabolomics vs. PANEL	0.000	0.000	0.000	*
Breast Cancer	Age+Sex vs. Metabolomics	0.013	0.000	0.026	ns
Breast Cancer	Age+Sex+Metabolomics vs. Age+Sex	0.003	-0.002	0.008	ns
Breast Cancer	Age+Sex+Metabolomics vs. Metabolomics	0.016	0.007	0.026	*
Breast Cancer	ASCVD vs. Age+Sex	-0.004	-0.009	0.002	ns
Breast Cancer	ASCVD vs. Age+Sex+Metabolomics	-0.007	-0.012	-0.001	*
Breast Cancer	ASCVD vs. Metabolomics	0.010	-0.003	0.022	ns
Breast Cancer	ASCVD+Metabolomics vs. Age+Sex	-0.001	-0.007	0.005	ns
Breast Cancer	ASCVD+Metabolomics vs. Age+Sex+Metabolomics	-0.004	-0.008	0.000	*
Breast Cancer	ASCVD+Metabolomics vs. ASCVD	0.002	-0.001	0.006	ns
Breast Cancer	ASCVD+Metabolomics vs. Metabolomics	0.012	0.002	0.022	*
Breast Cancer	PANEL vs. Age+Sex	0.014	0.003	0.027	*
Breast Cancer	PANEL vs. Age+Sex+Metabolomics	0.011	0.000	0.023	*
Breast Cancer	PANEL vs. ASCVD	0.018	0.007	0.029	*
Breast Cancer	PANEL vs. ASCVD+Metabolomics	0.016	0.005	0.027	*
Breast Cancer	PANEL vs. Metabolomics	0.028	0.013	0.042	*
Breast Cancer	PANEL+Metabolomics vs. Age+Sex	0.014	0.003	0.027	*
Breast Cancer	PANEL+Metabolomics vs. Age+Sex+Metabolomics	0.011	0.000	0.023	*
Breast Cancer	PANEL+Metabolomics vs. ASCVD	0.018	0.007	0.030	*
Breast Cancer	PANEL+Metabolomics vs. ASCVD+Metabolomics	0.016	0.005	0.027	*
Breast Cancer	PANEL+Metabolomics vs. Metabolomics	0.028	0.013	0.042	*
Breast Cancer	PANEL+Metabolomics vs. PANEL	0.000	-0.001	0.001	ns

3 Metabolomic profiles predict individual multi-disease outcomes

Table 8: Relative discriminatory performances in the four independent cohorts. Relative performances are reported as medians and 95 % CIs, estimated by bootstrapping over 1000 iterations. Differences were considered significant if the 95 % CIs did not cross zero (no difference).

Endpoint	Comparisons	Cohort	Median Delta C-Index	Lower 95% CI Delta C-Index	Upper 95% CI Delta C-Index	Significance
CHD	MET vs. Age+Sex	UKB	0.00	-0.005	0.009	ns
CHD	MET vs. Age+Sex	WHII	-0.01	-0.046	0.017	ns
CHD	MET vs. Age+Sex	RS	0.01	-0.030	0.050	ns
CHD	MET vs. Age+Sex	PROSPER	0.00	-0.038	0.045	ns
CHD	Age+Sex+MET vs. Age+Sex	UKB	0.03	0.026	0.033	*
CHD	Age+Sex+MET vs. Age+Sex	WHII	0.03	0.019	0.042	*
CHD	Age+Sex+MET vs. Age+Sex	RS	0.04	0.016	0.061	*
CHD	Age+Sex+MET vs. Age+Sex	PROSPER	0.03	0.008	0.052	*
Cerebral Stroke	MET vs. Age+Sex	UKB	-0.03	-0.042	-0.015	*
Cerebral Stroke	MET vs. Age+Sex	WHII	-0.04	-0.096	-0.004	*
Cerebral Stroke	MET vs. Age+Sex	RS	-0.05	-0.094	-0.010	*
Cerebral Stroke	MET vs. Age+Sex	LLS	-0.01	-0.117	0.087	ns
Cerebral Stroke	MET vs. Age+Sex	PROSPER	0.03	-0.095	0.147	ns
Cerebral Stroke	Age+Sex+MET vs. Age+Sex	UKB	0.02	0.015	0.029	*
Cerebral Stroke	Age+Sex+MET vs. Age+Sex	WHII	0.01	-0.003	0.015	ns
Cerebral Stroke	Age+Sex+MET vs. Age+Sex	RS	0.01	0.000	0.027	ns
Cerebral Stroke	Age+Sex+MET vs. Age+Sex	LLS	0.00	-0.020	0.040	ns
Cerebral Stroke	Age+Sex+MET vs. Age+Sex	PROSPER	0.05	-0.001	0.153	ns
Dementia	MET vs. Age+Sex	UKB	-0.01	-0.030	0.000	ns
Dementia	MET vs. Age+Sex	WHII	-0.11	-0.198	0.001	ns
Dementia	MET vs. Age+Sex	RS	-0.11	-0.144	-0.082	*
Dementia	Age+Sex+MET vs. Age+Sex	UKB	0.02	0.013	0.028	*
Dementia	Age+Sex+MET vs. Age+Sex	WHII	0.01	-0.009	0.040	ns
Dementia	Age+Sex+MET vs. Age+Sex	RS	0.00	0.000	0.007	ns
Heart Failure	MET vs. Age+Sex	UKB	0.03	0.016	0.036	*
Heart Failure	MET vs. Age+Sex	WHII	-0.08	-0.170	0.012	ns
Heart Failure	MET vs. Age+Sex	RS	-0.02	-0.055	0.020	ns
Heart Failure	MET vs. Age+Sex	PROSPER	0.06	-0.012	0.147	ns
Heart Failure	Age+Sex+MET vs. Age+Sex	UKB	0.05	0.041	0.054	*
Heart Failure	Age+Sex+MET vs. Age+Sex	WHII	0.04	0.009	0.069	*
Heart Failure	Age+Sex+MET vs. Age+Sex	RS	0.04	0.019	0.055	*
Heart Failure	Age+Sex+MET vs. Age+Sex	PROSPER	0.08	0.033	0.148	*
Atrial Fibrillation	MET vs. Age+Sex	UKB	-0.05	-0.061	-0.045	*
Atrial Fibrillation	MET vs. Age+Sex	RS	-0.02	-0.056	0.018	ns
Atrial Fibrillation	MET vs. Age+Sex	PROSPER	0.03	-0.037	0.101	ns
Atrial Fibrillation	Age+Sex+MET vs. Age+Sex	UKB	0.01	0.011	0.017	*
Atrial Fibrillation	Age+Sex+MET vs. Age+Sex	RS	0.02	0.004	0.038	*
Atrial Fibrillation	Age+Sex+MET vs. Age+Sex	PROSPER	0.05	0.012	0.100	*
T2 Diabetes	MET vs. Age+Sex	UKB	0.21	0.199	0.218	*
T2 Diabetes	MET vs. Age+Sex	WHII	0.29	0.258	0.329	*
T2 Diabetes	MET vs. Age+Sex	RS	0.23	0.183	0.274	*
T2 Diabetes	MET vs. Age+Sex	LLS	0.19	0.124	0.261	*
T2 Diabetes	MET vs. Age+Sex	PROSPER	0.10	-0.002	0.203	ns
T2 Diabetes	Age+Sex+MET vs. Age+Sex	UKB	0.21	0.200	0.219	*
T2 Diabetes	Age+Sex+MET vs. Age+Sex	WHII	0.30	0.267	0.333	*
T2 Diabetes	Age+Sex+MET vs. Age+Sex	RS	0.23	0.186	0.275	*
T2 Diabetes	Age+Sex+MET vs. Age+Sex	LLS	0.20	0.134	0.268	*
T2 Diabetes	Age+Sex+MET vs. Age+Sex	PROSPER	0.12	0.044	0.210	*
COPD	MET vs. Age+Sex	UKB	0.03	0.027	0.042	*
COPD	MET vs. Age+Sex	WHII	0.08	0.007	0.155	*
COPD	MET vs. Age+Sex	RS	-0.02	-0.078	0.029	ns
COPD	MET vs. Age+Sex	LLS	0.08	-0.037	0.204	ns
COPD	Age+Sex+MET vs. Age+Sex	UKB	0.05	0.044	0.054	*
COPD	Age+Sex+MET vs. Age+Sex	WHII	0.09	0.043	0.130	*
COPD	Age+Sex+MET vs. Age+Sex	RS	0.04	0.012	0.065	*

Supplementary Tables

Table 8 continued from previous page

Endpoint	Comparisons	Cohort	Median Delta C-Index	Lower 95% CI Delta C-Index	Upper 95% CI Delta C-Index	Significance
COPD	Age+Sex+MET vs. Age+Sex	LLS	0.10	0.034	0.205	*

3 Metabolomic profiles predict individual multi-disease outcomes

Table 9: Correlation matrix between the NMR metabolites and the PANEL predictors. Part 2.

	Systolic Blood Pressure (mmHg)	Total Cholesterol (mmol/L)	LDL Cholesterol (mmol/L)	HDL Cholesterol (mmol/L)	Triglycerides (mmol/L)	Glucose (mmol/L)	Glycated Hemoglobin (%)	Creatinine (umol/L)	Cystatin C (ng/L)	Urea (mmol/L)	Urate (umol/L)	Aspartate Aminotransferase (U/L)	Alanine Aminotransferase (U/L)	Alkaline Phosphatase (U/L)
Ala	0.0	0.0	0.0	-0.1	0.2	0.2	0.2	0.1	0.1	0.0	0.1	0.1	0.1	0.0
Gln	-0.1	0.0	0.0	0.0	-0.1	-0.1	-0.1	0.1	0.0	0.0	-0.1	0.0	-0.1	-0.1
Gly	-0.2	0.0	0.0	0.2	-0.2	-0.1	-0.1	-0.1	-0.1	-0.1	-0.3	-0.1	-0.2	0.0
His	0.0	0.0	0.0	0.0	0.0	0.0	-0.1	0.1	0.0	0.1	0.0	0.0	0.0	-0.1
Ile	0.0	0.0	0.0	-0.2	0.3	0.1	0.1	0.2	0.1	0.3	0.3	0.1	0.1	0.0
Leu	0.0	0.0	0.0	-0.3	0.3	0.1	0.1	0.2	0.1	0.3	0.3	0.1	0.2	0.0
Total_BCAA	0.1	0.0	0.0	-0.3	0.4	0.2	0.1	0.2	0.1	0.3	0.3	0.1	0.2	0.0
Val	0.1	0.0	0.0	-0.3	0.4	0.2	0.2	0.2	0.1	0.3	0.3	0.1	0.2	0.0
Phe	0.0	-0.1	-0.1	-0.1	0.2	0.1	0.1	0.1	0.1	0.2	0.1	0.0	0.1	0.0
Tyr	0.1	0.0	0.0	-0.1	0.2	0.1	0.1	0.0	0.1	0.2	0.1	0.1	0.1	0.1
Albumin	0.0	0.1	0.1	0.1	0.0	0.0	-0.1	0.0	-0.1	0.0	0.0	0.0	0.0	-0.1
Creatinine	0.1	0.0	0.0	-0.3	0.3	0.1	0.1	0.8	0.6	0.4	0.5	0.1	0.1	0.0
GlycA	0.2	0.2	0.2	-0.3	0.5	0.1	0.2	0.0	0.3	0.1	0.2	0.1	0.2	0.3
DHA	0.0	0.3	0.3	0.4	-0.1	0.0	-0.1	-0.1	-0.1	0.1	-0.1	0.0	0.0	0.0
LA	0.1	0.7	0.7	0.3	0.3	-0.1	-0.1	-0.1	0.0	0.0	0.0	0.0	0.0	0.0
MUFA	0.2	0.5	0.5	-0.2	0.8	0.1	0.1	0.0	0.2	0.1	0.2	0.1	0.2	0.1
Omega_3	0.1	0.4	0.3	0.2	0.3	0.0	0.0	-0.1	-0.1	0.1	0.0	0.0	0.1	0.0
Omega_6	0.1	0.8	0.7	0.3	0.3	0.0	-0.1	-0.1	0.0	0.0	0.0	0.0	0.0	0.1
PUFA	0.1	0.8	0.7	0.3	0.3	0.0	0.0	-0.1	-0.1	0.0	0.0	0.0	0.0	0.0
SFA	0.2	0.6	0.6	0.1	0.7	0.1	0.1	0.0	0.1	0.1	0.2	0.1	0.2	0.1
Total_FA	0.2	0.7	0.6	0.1	0.6	0.0	0.0	0.0	0.1	0.1	0.1	0.1	0.1	0.1
Unsaturation	-0.1	0.1	0.0	0.4	-0.6	-0.1	-0.1	-0.1	-0.2	0.0	-0.3	-0.1	-0.2	-0.1
Citrate	0.0	0.0	0.0	0.1	0.0	0.1	0.1	0.1	0.1	0.0	0.0	0.1	0.0	0.1
Glucose	0.1	0.0	0.0	0.0	0.2	0.7	0.5	0.0	0.1	0.1	0.1	0.1	0.1	0.1
Lactate	0.0	0.0	0.0	-0.1	0.1	0.1	0.1	0.0	0.0	-0.1	0.1	0.0	0.1	0.0
Pyruvate	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Acetate	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0
Acetoacetate	0.1	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.1	0.0	0.0
Acetone	0.0	0.0	-0.1	0.1	-0.1	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.0	0.0
bOHbutyrate	0.1	0.1	0.0	0.1	-0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0
HDL_L	0.0	0.3	0.1	0.8	-0.2	0.0	-0.1	-0.3	-0.2	0.0	-0.3	0.0	-0.2	0.0
LDL_L	0.1	0.9	0.9	0.2	0.2	-0.1	-0.1	-0.1	-0.1	0.0	0.0	0.0	0.0	0.0
Total_L	0.1	0.8	0.8	0.3	0.4	0.0	-0.1	-0.1	0.0	0.0	0.0	0.0	0.1	0.0
VLDL_L	0.2	0.5	0.6	-0.4	0.8	0.0	0.1	0.1	0.2	0.1	0.3	0.1	0.2	0.1
Clinical_LDL_C	0.1	0.8	0.9	0.2	0.0	-0.1	-0.1	-0.1	-0.1	0.0	0.0	0.0	0.0	0.0
HDL_C	0.0	0.3	0.1	0.9	-0.4	-0.1	-0.2	-0.3	-0.3	-0.1	-0.3	-0.1	-0.2	-0.1
LDL_C	0.1	0.9	0.9	0.2	0.1	-0.1	-0.1	-0.1	-0.1	0.0	0.0	0.0	0.0	0.0
non_HDL_C	0.1	0.9	0.9	0.2	0.2	-0.1	-0.1	-0.1	0.0	0.0	0.0	0.0	0.0	0.0
Remnant_C	0.1	0.9	0.9	0.1	0.2	-0.1	-0.1	-0.1	0.0	0.0	0.0	0.0	0.0	0.1
Total_C	0.1	0.9	0.8	0.5	0.0	-0.1	-0.2	-0.2	-0.1	0.0	-0.1	0.0	0.0	0.0
VLDL_C	0.1	0.7	0.8	-0.2	0.5	0.0	0.0	0.0	0.1	0.0	0.2	0.0	0.1	0.1
HDL_FC	0.0	0.4	0.1	0.9	-0.2	-0.1	-0.1	-0.3	-0.2	-0.1	-0.3	0.0	-0.2	0.0
LDL_FC	0.0	0.8	0.8	0.3	-0.1	-0.1	-0.1	-0.1	-0.1	0.0	-0.1	0.0	0.0	0.0
Total_FC	0.1	0.9	0.9	0.3	0.1	-0.1	-0.1	-0.1	-0.1	0.0	-0.1	0.0	0.0	0.0
VLDL_FC	0.2	0.7	0.7	-0.3	0.6	0.0	0.0	0.0	0.1	0.1	0.2	0.1	0.2	0.1
HDL_CE	0.0	0.2	0.0	0.9	-0.4	-0.1	-0.2	-0.3	-0.3	-0.1	-0.3	-0.1	-0.2	-0.1
LDL_CE	0.1	0.9	0.9	0.2	0.2	-0.1	-0.1	-0.1	0.0	0.0	0.0	0.0	0.0	0.0
Total_CE	0.1	0.9	0.8	0.5	0.0	-0.1	-0.2	-0.2	-0.1	0.0	-0.1	0.0	-0.1	0.0
VLDL_CE	0.1	0.8	0.9	-0.1	0.4	-0.1	0.0	0.0	0.1	0.0	0.1	0.0	0.1	0.1
HDL_PL	0.0	0.3	0.0	0.8	-0.1	0.0	-0.1	-0.2	-0.2	0.0	-0.3	0.0	-0.1	0.0
LDL_PL	0.1	0.8	0.9	0.1	0.1	-0.1	-0.1	-0.1	0.0	0.0	0.0	0.0	0.0	0.0
Total_PL	0.1	0.8	0.7	0.5	0.2	-0.1	-0.1	-0.2	-0.1	0.0	-0.1	0.0	0.0	0.0
VLDL_PL	0.2	0.6	0.7	-0.3	0.7	0.0	0.1	0.1	0.1	0.1	0.2	0.1	0.2	0.1
HDL_TG	0.2	0.3	0.2	-0.1	0.8	0.1	0.1	0.0	0.1	0.1	0.2	0.1	0.2	0.1

Table 9 continued from previous page

	Systolic Blood Pressure (mmHg)	Total Cholesterol (mmol/L)	LDL Cholesterol (mmol/L)	HDL Cholesterol (mmol/L)	Triglycerides (mmol/L)	Glucose (mmol/L)	Glycated Hemoglobin (%)	Creatinine (umol/L)	Cystatin C (mg/L)	Urea (mmol/L)	Urate (umol/L)	Aspartate Aminotransferase (U/L)	Alanine Aminotransferase (U/L)	Alkaline Phosphatase (U/L)
LDL_TG	0.2	0.5	0.5	-0.2	0.8	0.1	0.1	0.0	0.2	0.1	0.2	0.1	0.2	0.1
Total_TG	0.2	0.3	0.4	-0.4	0.8	0.1	0.1	0.1	0.2	0.1	0.3	0.1	0.3	0.1
VLDL_TG	0.2	0.3	0.4	-0.4	0.8	0.1	0.1	0.1	0.2	0.1	0.3	0.1	0.3	0.1
Cholines	0.1	0.7	0.6	0.6	0.2	0.0	-0.1	-0.2	-0.1	0.0	-0.1	0.0	0.0	0.0
Phosphatidyle	0.1	0.7	0.5	0.6	0.2	0.0	-0.1	-0.2	-0.1	0.0	-0.1	0.0	0.0	0.0
Phosphoglyc	0.1	0.7	0.5	0.6	0.2	0.0	-0.1	-0.2	-0.1	0.0	-0.1	0.0	0.0	0.0
Sphingomyelins	0.1	0.8	0.7	0.6	-0.1	-0.1	-0.1	-0.2	-0.1	0.0	-0.2	0.0	-0.1	0.0
HDL_size	-0.1	0.1	-0.1	0.5	-0.2	-0.1	-0.1	-0.1	-0.1	0.0	-0.3	0.0	-0.1	0.0
LDL_size	-0.1	0.2	0.2	0.3	-0.4	-0.1	-0.1	-0.1	-0.1	0.0	-0.2	-0.1	-0.2	-0.1
VLDL_size	0.1	0.1	0.1	-0.4	0.6	0.1	0.1	0.1	0.1	0.1	0.3	0.1	0.2	0.1
HDL_P	0.1	0.4	0.2	0.7	-0.1	0.0	-0.1	-0.2	-0.2	0.0	-0.2	0.0	-0.1	0.0
LDL_P	0.1	0.8	0.9	0.0	0.2	-0.1	-0.1	0.0	0.0	0.0	0.1	0.0	0.1	0.0
Total_P	0.1	0.5	0.3	0.7	-0.1	-0.1	-0.1	-0.2	-0.2	0.0	-0.2	0.0	-0.1	0.0
VLDL_P	0.1	0.7	0.7	-0.2	0.6	0.0	0.0	0.0	0.1	0.1	0.2	0.1	0.2	0.1
XXL_VLDL_C	0.2	0.3	0.3	-0.4	0.8	0.1	0.1	0.1	0.2	0.1	0.3	0.1	0.3	0.1
XXL_VLDL_CE	0.2	0.3	0.4	-0.4	0.8	0.1	0.1	0.1	0.2	0.1	0.3	0.1	0.2	0.1
XXL_VLDL_FC	0.2	0.2	0.3	-0.4	0.9	0.1	0.1	0.1	0.2	0.1	0.3	0.1	0.3	0.1
XXL_VLDL_L	0.2	0.2	0.2	-0.4	0.8	0.1	0.1	0.1	0.2	0.1	0.3	0.1	0.3	0.1
XXL_VLDL_P	0.2	0.2	0.2	-0.4	0.9	0.1	0.1	0.1	0.2	0.1	0.3	0.1	0.3	0.1
XXL_VLDL_PL	0.2	0.2	0.2	-0.4	0.9	0.1	0.1	0.1	0.2	0.1	0.3	0.1	0.3	0.1
XXL_VLDL_TG	0.2	0.2	0.2	-0.4	0.8	0.1	0.1	0.1	0.2	0.1	0.3	0.1	0.3	0.1
XL_VLDL_C	0.2	0.5	0.6	-0.4	0.7	0.0	0.1	0.1	0.2	0.1	0.3	0.1	0.2	0.1
XL_VLDL_CE	0.2	0.6	0.7	-0.4	0.7	0.0	0.1	0.1	0.2	0.1	0.3	0.1	0.2	0.1
XL_VLDL_FC	0.2	0.4	0.5	-0.4	0.8	0.1	0.1	0.1	0.2	0.1	0.3	0.1	0.2	0.1
XL_VLDL_L	0.2	0.3	0.4	-0.5	0.8	0.1	0.1	0.1	0.2	0.1	0.3	0.1	0.3	0.1
XL_VLDL_P	0.2	0.3	0.4	-0.5	0.8	0.1	0.1	0.1	0.2	0.1	0.3	0.1	0.3	0.1
XL_VLDL_PL	0.2	0.4	0.4	-0.5	0.8	0.1	0.1	0.1	0.2	0.1	0.3	0.1	0.3	0.1
XL_VLDL_TG	0.2	0.2	0.3	-0.5	0.9	0.1	0.1	0.1	0.2	0.1	0.3	0.1	0.3	0.1
L_VLDL_C	0.2	0.5	0.6	-0.4	0.7	0.0	0.1	0.1	0.2	0.1	0.3	0.1	0.2	0.1
L_VLDL_CE	0.1	0.6	0.7	-0.4	0.6	0.0	0.0	0.1	0.1	0.1	0.2	0.1	0.2	0.1
L_VLDL_FC	0.2	0.4	0.5	-0.5	0.8	0.1	0.1	0.1	0.2	0.1	0.3	0.1	0.2	0.1
L_VLDL_L	0.2	0.4	0.5	-0.4	0.8	0.1	0.1	0.1	0.2	0.1	0.3	0.1	0.2	0.1
L_VLDL_P	0.2	0.4	0.4	-0.4	0.8	0.1	0.1	0.1	0.2	0.1	0.3	0.1	0.2	0.1
L_VLDL_PL	0.2	0.4	0.4	-0.5	0.8	0.1	0.1	0.1	0.2	0.1	0.3	0.1	0.2	0.1
L_VLDL_TG	0.2	0.3	0.4	-0.4	0.8	0.1	0.1	0.1	0.2	0.1	0.3	0.1	0.2	0.1
M_VLDL_C	0.1	0.8	0.9	0.1	0.1	-0.1	-0.1	-0.1	0.0	0.0	0.0	0.0	0.0	0.0
M_VLDL_CE	0.0	0.8	0.9	0.2	0.0	-0.1	-0.1	-0.1	-0.1	0.0	-0.1	0.0	0.0	0.0
M_VLDL_FC	0.1	0.8	0.9	-0.1	0.4	-0.1	0.0	0.0	0.0	0.0	0.1	0.0	0.1	0.1
M_VLDL_L	0.1	0.7	0.7	-0.2	0.6	0.0	0.0	0.0	0.1	0.1	0.2	0.0	0.2	0.1
M_VLDL_P	0.1	0.7	0.8	-0.2	0.5	0.0	0.0	0.0	0.1	0.0	0.2	0.0	0.1	0.1
M_VLDL_PL	0.1	0.8	0.8	-0.1	0.5	0.0	0.0	0.0	0.1	0.0	0.1	0.0	0.1	0.1
M_VLDL_TG	0.2	0.4	0.5	-0.4	0.8	0.1	0.1	0.1	0.2	0.1	0.3	0.1	0.2	0.1
S_VLDL_C	0.1	0.7	0.8	-0.2	0.4	0.0	0.0	0.0	0.1	0.0	0.1	0.0	0.1	0.1
S_VLDL_CE	0.1	0.7	0.8	-0.2	0.4	0.0	0.0	0.0	0.1	0.1	0.2	0.1	0.1	0.1
S_VLDL_FC	0.1	0.8	0.9	-0.1	0.3	-0.1	0.0	0.0	0.0	0.0	0.1	0.0	0.1	0.1
S_VLDL_L	0.2	0.6	0.7	-0.3	0.6	0.0	0.1	0.1	0.1	0.1	0.2	0.1	0.2	0.1
S_VLDL_P	0.2	0.6	0.7	-0.3	0.7	0.0	0.1	0.1	0.2	0.1	0.2	0.1	0.2	0.1
S_VLDL_PL	0.1	0.7	0.8	-0.2	0.5	0.0	0.0	0.0	0.1	0.1	0.2	0.0	0.1	0.1
S_VLDL_TG	0.2	0.3	0.4	-0.4	0.8	0.1	0.1	0.1	0.2	0.1	0.3	0.1	0.2	0.1
XS_VLDL_C	0.1	0.8	0.8	0.2	0.1	-0.1	-0.1	-0.1	0.0	0.0	-0.1	0.0	0.0	0.1
XS_VLDL_CE	0.0	0.8	0.8	0.3	0.0	-0.1	-0.1	-0.1	-0.1	0.0	-0.1	0.0	0.0	0.0
XS_VLDL_FC	0.1	0.8	0.8	0.1	0.3	-0.1	0.0	-0.1	0.0	0.0	0.0	0.0	0.1	0.1
XS_VLDL_L	0.1	0.8	0.8	0.1	0.4	0.0	0.0	-0.1	0.0	0.0	0.0	0.1	0.1	0.1
XS_VLDL_P	0.1	0.8	0.8	0.0	0.4	0.0	0.0	-0.1	0.0	0.0	0.0	0.1	0.1	0.1
XS_VLDL_PL	0.1	0.7	0.8	0.0	0.4	0.0	0.0	-0.1	0.1	0.1	0.0	0.1	0.1	0.1

3 Metabolomic profiles predict individual multi-disease outcomes

Table 9 continued from previous page

	Systolic Blood Pressure (mmHg)	Total Cholesterol (mmol/L)	LDL Cholesterol (mmol/L)	HDL Cholesterol (mmol/L)	Triglycerides (mmol/L)	Glucose (mmol/L)	Glycated Hemoglobin (%)	Creatinine (umol/L)	Cystatin C (mg/L)	Urea (mmol/L)	Urate (umol/L)	Aspartate Aminotransferase (U/L)	Alanine Aminotransferase (U/L)	Alkaline Phosphatase (U/L)
XS_VLDL_TG	0.2	0.4	0.4	-0.3	0.8	0.1	0.1	0.1	0.2	0.1	0.2	0.1	0.2	0.1
L_LDL_C	0.1	0.9	0.9	0.3	0.0	-0.1	-0.1	-0.1	-0.1	0.0	-0.1	0.0	0.0	0.0
L_LDL_CE	0.1	0.9	0.9	0.3	0.1	-0.1	-0.1	-0.1	-0.1	0.0	0.0	0.0	0.0	0.0
L_LDL_FC	0.0	0.8	0.8	0.4	-0.1	-0.1	-0.2	-0.1	-0.1	0.0	-0.1	0.0	-0.1	0.0
L_LDL_L	0.1	0.9	0.9	0.3	0.1	-0.1	-0.1	-0.1	-0.1	0.0	0.0	0.0	0.0	0.0
L_LDL_P	0.1	0.8	0.9	0.1	0.2	-0.1	-0.1	-0.1	0.0	0.0	0.0	0.0	0.0	0.0
L_LDL_PL	0.1	0.9	0.9	0.2	0.0	-0.1	-0.1	-0.1	-0.1	0.0	0.0	0.0	0.0	0.0
L_LDL_TG	0.2	0.5	0.5	-0.1	0.7	0.1	0.1	0.0	0.2	0.1	0.2	0.1	0.2	0.1
M_LDL_C	0.1	0.8	0.9	0.0	0.2	-0.1	-0.1	0.0	0.0	0.0	0.1	0.0	0.1	0.0
M_LDL_CE	0.1	0.8	0.8	0.0	0.3	-0.1	0.0	0.0	0.0	0.0	0.1	0.0	0.1	0.0
M_LDL_FC	0.1	0.8	0.8	0.2	0.0	-0.1	-0.1	-0.1	-0.1	0.0	0.0	0.0	0.0	0.0
M_LDL_L	0.1	0.8	0.9	0.0	0.3	-0.1	0.0	0.0	0.0	0.0	0.1	0.0	0.1	0.0
M_LDL_P	0.1	0.8	0.9	0.0	0.3	-0.1	0.0	0.0	0.0	0.0	0.1	0.0	0.1	0.0
M_LDL_PL	0.1	0.8	0.9	0.0	0.3	-0.1	0.0	0.0	0.0	0.0	0.1	0.0	0.1	0.0
M_LDL_TG	0.2	0.5	0.5	-0.2	0.8	0.1	0.1	0.1	0.2	0.1	0.3	0.1	0.2	0.1
S_LDL_C	0.1	0.8	0.9	0.0	0.2	-0.1	-0.1	0.0	0.0	0.0	0.1	0.0	0.1	0.0
S_LDL_CE	0.1	0.8	0.9	0.0	0.3	-0.1	0.0	0.0	0.0	0.0	0.1	0.0	0.1	0.0
S_LDL_FC	0.0	0.7	0.8	0.1	-0.1	-0.1	-0.1	-0.1	-0.1	0.0	0.0	0.0	0.0	0.0
S_LDL_L	0.1	0.8	0.9	0.0	0.3	-0.1	0.0	0.0	0.0	0.0	0.1	0.0	0.1	0.0
S_LDL_P	0.1	0.8	0.9	-0.1	0.4	-0.1	0.0	0.0	0.0	0.0	0.1	0.0	0.1	0.1
S_LDL_PL	0.1	0.8	0.9	0.0	0.2	-0.1	-0.1	0.0	0.0	0.0	0.1	0.0	0.1	0.0
S_LDL_TG	0.2	0.4	0.4	-0.3	0.9	0.1	0.1	0.1	0.2	0.1	0.3	0.1	0.3	0.1
IDL_C	0.0	0.8	0.8	0.4	-0.1	-0.1	-0.2	-0.2	-0.1	0.0	-0.2	0.0	-0.1	0.0
IDL_CE	0.0	0.8	0.8	0.4	-0.1	-0.1	-0.2	-0.2	-0.1	0.0	-0.2	0.0	-0.1	0.0
IDL_FC	0.0	0.8	0.8	0.4	-0.1	-0.1	-0.2	-0.2	-0.1	0.0	-0.1	0.0	-0.1	0.0
IDL_L	0.1	0.9	0.9	0.4	0.0	-0.1	-0.2	-0.2	-0.1	0.0	-0.1	0.0	0.0	0.0
IDL_P	0.1	0.8	0.9	0.2	0.1	-0.1	-0.1	-0.1	0.0	0.0	0.0	0.0	0.0	0.0
IDL_PL	0.0	0.9	0.9	0.4	0.0	-0.1	-0.1	-0.2	-0.1	0.0	-0.1	0.0	0.0	0.0
IDL_TG	0.2	0.5	0.5	-0.1	0.7	0.1	0.1	0.0	0.2	0.1	0.2	0.1	0.2	0.1
XL_HDL_C	-0.1	0.3	0.1	0.8	-0.4	-0.1	-0.2	-0.2	-0.2	-0.1	-0.4	-0.1	-0.2	0.0
XL_HDL_CE	-0.1	0.3	0.1	0.8	-0.4	-0.1	-0.2	-0.2	-0.2	-0.1	-0.4	-0.1	-0.2	-0.1
XL_HDL_FC	-0.1	0.3	0.1	0.6	-0.4	-0.1	-0.2	-0.2	-0.2	-0.1	-0.4	0.0	-0.2	0.0
XL_HDL_L	-0.1	0.2	0.1	0.8	-0.4	-0.1	-0.2	-0.2	-0.2	-0.1	-0.4	-0.1	-0.2	0.0
XL_HDL_P	-0.1	0.3	0.1	0.8	-0.3	-0.1	-0.1	-0.2	-0.2	-0.1	-0.4	0.0	-0.2	0.0
XL_HDL_PL	-0.1	0.2	0.0	0.7	-0.4	-0.1	-0.1	-0.2	-0.2	-0.1	-0.4	-0.1	-0.2	0.0
XL_HDL_TG	0.1	0.5	0.4	0.1	0.7	0.1	0.1	0.0	0.1	0.1	0.1	0.1	0.1	0.1
L_HDL_C	-0.1	0.2	0.0	0.9	-0.5	-0.1	-0.2	-0.2	-0.3	-0.1	-0.4	-0.1	-0.3	-0.1
L_HDL_CE	-0.1	0.2	0.0	0.9	-0.5	-0.1	-0.2	-0.2	-0.3	-0.1	-0.4	-0.1	-0.3	-0.1
L_HDL_FC	-0.1	0.2	0.0	0.9	-0.4	-0.1	-0.2	-0.2	-0.2	-0.1	-0.4	-0.1	-0.2	-0.1
L_HDL_L	-0.1	0.2	0.0	0.9	-0.4	-0.1	-0.2	-0.3	-0.2	-0.1	-0.4	-0.1	-0.2	-0.1
L_HDL_P	-0.1	0.2	0.0	0.9	-0.4	-0.1	-0.2	-0.2	-0.2	-0.1	-0.4	-0.1	-0.2	-0.1
L_HDL_PL	-0.1	0.2	-0.1	0.9	-0.4	-0.1	-0.1	-0.2	-0.2	-0.1	-0.4	-0.1	-0.2	-0.1
L_HDL_TG	0.1	0.3	0.2	0.3	0.5	0.0	0.0	-0.1	0.0	0.1	-0.1	0.1	0.0	0.1
M_HDL_C	0.0	0.2	0.0	0.8	-0.3	0.0	-0.1	-0.2	-0.2	0.0	-0.3	0.0	-0.2	-0.1
M_HDL_CE	0.0	0.2	0.0	0.8	-0.3	0.0	-0.1	-0.2	-0.2	-0.1	-0.3	0.0	-0.2	-0.1
M_HDL_FC	0.0	0.3	0.1	0.8	-0.2	0.0	-0.1	-0.2	-0.2	0.0	-0.3	0.0	-0.1	0.0
M_HDL_L	0.1	0.2	0.0	0.7	-0.1	0.0	-0.1	-0.2	-0.2	0.0	-0.2	0.0	-0.1	0.0
M_HDL_P	0.0	0.2	0.0	0.8	-0.1	0.0	-0.1	-0.2	-0.2	0.0	-0.2	0.0	-0.1	0.0
M_HDL_PL	0.1	0.2	0.0	0.7	0.0	0.0	-0.1	-0.2	-0.2	0.0	-0.2	0.0	-0.1	0.0
M_HDL_TG	0.2	0.3	0.2	-0.1	0.7	0.1	0.1	0.0	0.1	0.1	0.2	0.1	0.2	0.1
S_HDL_C	0.1	0.4	0.3	0.3	0.1	0.0	0.0	-0.1	-0.1	0.0	0.1	0.0	0.0	0.0
S_HDL_CE	0.1	0.3	0.3	0.3	0.0	0.0	0.0	-0.1	-0.1	0.0	0.1	0.0	0.0	0.0
S_HDL_FC	0.2	0.6	0.4	0.4	0.2	0.0	0.0	-0.1	-0.1	0.0	0.0	0.0	0.0	0.0
S_HDL_L	0.2	0.4	0.3	0.2	0.3	0.0	0.0	-0.1	0.0	0.0	0.1	0.0	0.1	0.0
S_HDL_P	0.2	0.4	0.4	0.3	0.1	0.0	0.0	-0.1	-0.1	0.0	0.1	0.0	0.1	0.0

Supplementary Tables

Table 9 continued from previous page

	Systolic Blood Pressure (mmHg)	Total Cholesterol (mmol/L)	LDL Cholesterol (mmol/L)	HDL Cholesterol (mmol/L)	Triglycerides (mmol/L)	Glucose (mmol/L)	Glycated Hemoglobin (%)	Creatinine (umol/L)	Cystatin C (mg/L)	Urea (mmol/L)	Urate (umol/L)	Aspartate Aminotransferase (U/L)	Alanine Aminotransferase (U/L)	Alkaline Phosphatase (U/L)
S_HDL_PL	0.2	0.3	0.2	0.3	0.3	0.0	0.0	-0.1	0.0	0.0	0.1	0.0	0.1	0.0
S_HDL_TG	0.2	0.2	0.3	-0.4	0.9	0.1	0.2	0.1	0.2	0.1	0.3	0.1	0.3	0.1
ApoA1	0.0	0.3	0.1	0.8	-0.1	0.0	-0.1	-0.2	-0.2	0.0	-0.2	0.0	-0.1	0.0
ApoB	0.1	0.9	0.9	0.0	0.3	-0.1	-0.1	0.0	0.0	0.0	0.0	0.0	0.1	0.0

3 Metabolomic profiles predict individual multi-disease outcomes

Table 10: Correlation matrix between the NMR metabolites and the PANEL predictors. Part 3.

	Albumin (g/L)	C-Reactive Protein (mg/L)	Erythrocytes (10 ¹² cells/L)	Leucocytes (10 ⁹ cells/L)	Platelets (10 ⁹ cells/L)	Haemoglobin (g/dL)	Haematocrit (%)	Mean Corpuscular Volume	Mean Corpuscular Haemoglobin (pg)	Mean Corpuscular Haemoglobin (g/dL)	Antihypertensives
Ala	0.0	0.0	0.1	0.0	0.0	0.1	0.1	0.0	0.0	0.0	0.0
Gln	0.1	-0.2	0.0	-0.1	-0.1	0.0	0.0	0.0	0.0	0.0	0.0
Gly	0.0	-0.1	-0.3	-0.1	0.1	-0.3	-0.3	0.0	0.0	-0.1	0.0
His	0.1	-0.1	0.0	-0.1	-0.1	0.0	0.0	0.0	0.0	0.0	0.0
Ile	0.0	0.0	0.1	0.1	0.0	0.1	0.1	0.0	0.0	0.1	0.0
Leu	0.0	0.0	0.2	0.1	-0.1	0.2	0.2	0.0	0.0	0.1	0.0
Total_BCAA	0.0	0.0	0.2	0.1	0.0	0.2	0.2	0.0	0.0	0.1	0.0
Val	0.0	0.0	0.2	0.1	-0.1	0.2	0.2	0.0	0.0	0.1	0.0
Phe	-0.1	0.1	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Tyr	-0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Albumin	0.5	-0.2	0.0	-0.1	-0.1	0.1	0.1	0.0	0.1	0.1	0.0
Creatinine	0.0	0.0	0.2	0.1	-0.1	0.3	0.2	0.0	0.0	0.0	0.1
GlycA	-0.1	0.4	0.1	0.3	0.2	0.0	0.0	-0.1	-0.1	0.0	0.0
DHA	0.1	-0.1	-0.2	-0.1	0.0	-0.1	-0.1	0.1	0.1	0.0	0.0
LA	0.1	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	-0.1
MUFA	0.1	0.1	0.1	0.2	0.1	0.1	0.1	0.0	0.0	0.1	0.0
Omega_3	0.1	0.0	-0.1	0.0	0.0	-0.1	-0.1	0.0	0.1	0.0	0.0
Omega_6	0.1	0.0	0.0	0.0	0.1	0.0	-0.1	0.0	0.0	0.0	-0.1
PUFA	0.1	0.0	-0.1	0.0	0.1	-0.1	-0.1	0.0	0.0	0.0	-0.1
SFA	0.1	0.0	0.0	0.1	0.1	0.1	0.1	0.0	0.1	0.1	0.0
Total_FA	0.1	0.0	0.0	0.1	0.1	0.1	0.0	0.0	0.0	0.1	0.0
Unsaturation	0.1	-0.1	-0.2	-0.2	0.0	-0.2	-0.2	0.0	0.0	-0.1	0.0
Citrate	0.0	-0.1	0.0	0.0	0.0	-0.1	-0.1	0.0	0.0	0.0	0.0
Glucose	0.0	0.1	-0.1	0.0	-0.1	-0.1	-0.1	0.0	0.0	0.0	0.1
Lactate	0.1	0.0	0.2	0.2	0.2	0.2	0.2	0.0	0.0	-0.1	0.0
Pyruvate	0.0	0.0	-0.1	0.1	0.1	-0.1	0.0	0.0	0.0	-0.1	0.0
Acetate	0.0	0.0	0.0	0.0	-0.1	0.0	0.0	0.1	0.1	0.0	0.0
Acetoacetate	0.0	0.1	0.0	0.1	0.0	0.0	0.0	0.1	0.1	0.0	0.0
Acetone	0.1	0.0	0.0	0.0	-0.1	0.0	0.0	0.1	0.1	0.0	0.0
bOHbutyrate	0.1	0.1	0.0	0.1	0.0	0.0	0.0	0.1	0.0	0.0	0.0
HDL_L	0.1	-0.1	-0.3	-0.1	0.1	-0.3	-0.3	0.1	0.1	0.0	0.0
LDL_L	0.1	0.0	0.1	-0.1	0.1	0.1	0.0	0.0	0.0	0.0	-0.1
Total_L	0.1	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	-0.1
VLDL_L	0.1	0.1	0.2	0.1	0.1	0.2	0.2	-0.1	0.0	0.1	0.0
Clinical_LDL_C	0.1	-0.1	0.0	-0.1	0.1	0.0	0.0	0.0	0.0	0.0	-0.1
HDL_C	0.1	-0.1	-0.4	-0.1	0.1	-0.3	-0.3	0.1	0.1	-0.1	0.0
LDL_C	0.1	-0.1	0.0	-0.1	0.1	0.0	0.0	0.0	0.0	0.0	-0.1
non_HDL_C	0.1	0.0	0.0	-0.1	0.1	0.0	0.0	0.0	0.0	0.0	-0.1
Remnant_C	0.1	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	-0.1
Total_C	0.1	-0.1	-0.1	-0.1	0.1	-0.1	-0.1	0.0	0.0	0.0	-0.1
VLDL_C	0.1	0.0	0.1	0.0	0.1	0.1	0.1	0.0	0.0	0.1	0.0
HDL_FC	0.0	-0.1	-0.4	-0.1	0.1	-0.3	-0.3	0.1	0.1	0.0	0.0
LDL_FC	0.1	-0.1	0.0	-0.1	0.1	0.0	0.0	0.0	0.0	0.0	-0.1
Total_FC	0.1	0.0	-0.1	-0.1	0.1	0.0	0.0	0.0	0.0	0.0	-0.1
VLDL_FC	0.1	0.0	0.2	0.1	0.1	0.2	0.2	0.0	0.0	0.1	0.0
HDL_CE	0.1	-0.1	-0.4	-0.1	0.1	-0.3	-0.3	0.1	0.1	-0.1	0.0
LDL_CE	0.2	0.0	0.1	-0.1	0.1	0.1	0.1	0.0	0.0	0.0	-0.1
Total_CE	0.1	-0.1	-0.1	-0.1	0.1	-0.1	-0.1	0.0	0.0	0.0	-0.1
VLDL_CE	0.1	0.0	0.1	0.0	0.1	0.1	0.1	0.0	0.0	0.0	-0.1
HDL_PL	0.0	-0.1	-0.3	0.0	0.1	-0.3	-0.3	0.1	0.1	0.0	0.0
LDL_PL	0.1	0.0	0.1	-0.1	0.1	0.1	0.1	0.0	0.0	0.0	-0.1
Total_PL	0.1	0.0	-0.2	0.0	0.1	-0.1	-0.1	0.1	0.1	0.0	-0.1
VLDL_PL	0.1	0.1	0.2	0.1	0.1	0.2	0.2	-0.1	0.0	0.1	0.0

Supplementary Tables

Table 10 continued from previous page

	Albumin (g/L)	C-Reactive Protein (mg/L)	Erythrocytes (10 ¹² cells/L)	Leucocytes (10 ⁹ cells/L)	Platelets (10 ⁹ cells/L)	Haemoglobin (g/dL)	Haematocrit (%)	Mean Corpuscular Volume	Mean Corpuscular Haemoglobin (pg)	Mean Corpuscular Haemoglobin (g/dL)	Anthypertensives
HDL_TG	0.0	0.1	0.0	0.1	0.1	0.1	0.1	0.0	0.1	0.1	0.0
LDL_TG	0.0	0.1	0.1	0.1	0.1	0.1	0.1	0.0	0.0	0.1	0.0
Total_TG	0.1	0.1	0.2	0.2	0.0	0.2	0.2	0.0	0.0	0.1	0.0
VLDL_TG	0.1	0.1	0.2	0.1	0.0	0.3	0.2	-0.1	0.0	0.1	0.0
Cholines	0.1	-0.1	-0.2	0.0	0.1	-0.2	-0.2	0.1	0.1	0.0	-0.1
Phosphatidylc	0.1	-0.1	-0.2	0.0	0.1	-0.2	-0.2	0.1	0.1	0.0	-0.1
Phosphoglyc	0.1	0.0	-0.2	0.0	0.1	-0.1	-0.2	0.1	0.1	0.0	-0.1
Sphingomyelins	0.1	0.0	-0.2	-0.1	0.1	-0.2	-0.2	0.1	0.0	0.0	-0.1
HDL_size	0.0	0.0	-0.3	-0.1	0.0	-0.2	-0.2	0.1	0.0	0.0	0.0
LDL_size	0.0	0.0	-0.2	-0.1	0.1	-0.2	-0.2	0.0	0.0	-0.1	0.0
VLDL_size	0.1	0.0	0.2	0.1	0.0	0.2	0.2	-0.1	0.0	0.1	0.0
HDL_P	0.1	-0.1	-0.2	-0.1	0.1	-0.2	-0.2	0.1	0.1	0.0	0.0
LDL_P	0.1	0.0	0.1	0.0	0.1	0.1	0.1	0.0	0.0	0.0	-0.1
Total_P	0.1	-0.1	-0.2	-0.1	0.1	-0.2	-0.2	0.1	0.1	0.0	0.0
VLDL_P	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.0	0.0	0.1	0.0
XXL_VLDL_C	0.1	0.1	0.2	0.2	0.0	0.2	0.2	0.0	0.0	0.1	0.0
XXL_VLDL_CE	0.1	0.1	0.2	0.2	0.0	0.2	0.2	0.0	0.0	0.1	0.0
XXL_VLDL_FC	0.1	0.1	0.2	0.2	0.0	0.2	0.2	0.0	0.0	0.1	0.0
XXL_VLDL_L	0.0	0.1	0.2	0.2	0.0	0.3	0.2	0.0	0.0	0.1	0.0
XXL_VLDL_P	0.0	0.1	0.2	0.2	0.0	0.2	0.2	0.0	0.0	0.1	0.0
XXL_VLDL_PL	0.0	0.1	0.2	0.2	0.0	0.2	0.2	0.0	0.0	0.1	0.0
XXL_VLDL_TG	0.0	0.1	0.2	0.2	0.0	0.2	0.2	0.0	0.0	0.1	0.0
XL_VLDL_C	0.1	0.1	0.2	0.1	0.0	0.2	0.2	-0.1	0.0	0.1	0.0
XL_VLDL_CE	0.1	0.1	0.2	0.1	0.1	0.2	0.2	-0.1	0.0	0.1	0.0
XL_VLDL_FC	0.1	0.1	0.2	0.1	0.0	0.3	0.2	-0.1	0.0	0.1	0.0
XL_VLDL_L	0.1	0.1	0.2	0.1	0.0	0.3	0.2	-0.1	0.0	0.1	0.0
XL_VLDL_P	0.1	0.1	0.2	0.1	0.0	0.3	0.2	-0.1	0.0	0.1	0.0
XL_VLDL_PL	0.1	0.1	0.2	0.1	0.0	0.3	0.2	-0.1	0.0	0.1	0.0
XL_VLDL_TG	0.1	0.1	0.2	0.2	0.0	0.3	0.2	0.0	0.0	0.1	0.0
L_VLDL_C	0.1	0.1	0.2	0.1	0.0	0.2	0.2	-0.1	0.0	0.1	0.0
L_VLDL_CE	0.1	0.0	0.2	0.1	0.1	0.2	0.2	-0.1	0.0	0.1	0.0
L_VLDL_FC	0.1	0.1	0.3	0.1	0.0	0.3	0.2	-0.1	0.0	0.1	0.0
L_VLDL_L	0.1	0.1	0.3	0.1	0.0	0.3	0.2	-0.1	0.0	0.1	0.0
L_VLDL_P	0.1	0.1	0.2	0.1	0.0	0.3	0.2	-0.1	0.0	0.1	0.0
L_VLDL_PL	0.1	0.1	0.2	0.1	0.0	0.3	0.2	-0.1	0.0	0.1	0.0
L_VLDL_TG	0.1	0.0	0.2	0.1	0.0	0.3	0.2	-0.1	0.0	0.1	0.0
M_VLDL_C	0.1	0.0	0.1	-0.1	0.1	0.0	0.0	0.0	0.0	0.0	-0.1
M_VLDL_CE	0.1	0.0	0.0	-0.1	0.1	0.0	0.0	0.0	0.0	0.0	-0.1
M_VLDL_FC	0.1	0.0	0.1	0.0	0.1	0.1	0.1	0.0	0.0	0.0	-0.1
M_VLDL_L	0.1	0.0	0.2	0.1	0.1	0.2	0.2	-0.1	0.0	0.1	0.0
M_VLDL_P	0.1	0.0	0.2	0.0	0.1	0.2	0.1	-0.1	0.0	0.1	0.0
M_VLDL_PL	0.1	0.0	0.1	0.0	0.1	0.1	0.1	0.0	0.0	0.1	0.0
M_VLDL_TG	0.1	0.0	0.2	0.1	0.0	0.2	0.2	-0.1	0.0	0.1	0.0
S_VLDL_C	0.1	0.0	0.1	0.0	0.1	0.1	0.1	0.0	0.0	0.0	0.0
S_VLDL_CE	0.1	0.1	0.1	0.0	0.1	0.1	0.1	0.0	0.0	0.0	0.0
S_VLDL_FC	0.1	0.0	0.1	0.0	0.1	0.1	0.1	0.0	0.0	0.0	-0.1
S_VLDL_L	0.1	0.1	0.2	0.1	0.1	0.2	0.2	0.0	0.0	0.1	0.0
S_VLDL_P	0.1	0.1	0.2	0.1	0.1	0.2	0.2	-0.1	0.0	0.1	0.0
S_VLDL_PL	0.1	0.0	0.1	0.0	0.1	0.1	0.1	0.0	0.0	0.1	0.0
S_VLDL_TG	0.1	0.1	0.2	0.1	0.0	0.2	0.2	0.0	0.0	0.1	0.0
XS_VLDL_C	0.0	0.0	-0.1	0.0	0.1	-0.1	-0.1	0.0	0.0	0.0	-0.1
XS_VLDL_CE	0.0	0.0	-0.1	0.0	0.1	-0.1	-0.1	0.0	0.0	0.0	-0.1
XS_VLDL_FC	0.0	0.1	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	-0.1
XS_VLDL_L	0.0	0.1	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0

3 Metabolomic profiles predict individual multi-disease outcomes

Table 10 continued from previous page

	Albumin (g/L)	C-Reactive Protein (mg/L)	Erythrocytes (10 ¹² cells/L)	Leucocytes (10 ⁹ cells/L)	Platelets (10 ⁹ cells/L)	Haemoglobin (g/dL)	Haematocrit (%)	Mean Corpuscular Volume	Mean Corpuscular Haemoglobin (pg)	Mean Corpuscular Haemoglobin (g/dL)	Anthypertensives
XS_VLDL_P	0.0	0.1	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0
XS_VLDL_PL	0.0	0.1	0.0	0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.0
XS_VLDL_TG	0.0	0.1	0.1	0.1	0.1	0.1	0.1	0.0	0.0	0.1	0.0
L_LDL_C	0.1	-0.1	0.0	-0.1	0.1	0.0	0.0	0.0	0.0	0.0	-0.1
L_LDL_CE	0.1	-0.1	0.0	-0.1	0.1	0.0	0.0	0.0	0.0	0.0	-0.1
L_LDL_FC	0.1	-0.1	0.0	-0.1	0.1	0.0	0.0	0.0	0.0	0.0	-0.1
L_LDL_L	0.1	-0.1	0.0	-0.1	0.1	0.0	0.0	0.0	0.0	0.0	-0.1
L_LDL_P	0.1	0.0	0.1	0.0	0.1	0.0	0.0	0.0	0.0	0.0	-0.1
L_LDL_PL	0.1	-0.1	0.0	-0.1	0.1	0.0	0.0	0.0	0.0	0.0	-0.1
L_LDL_TG	0.0	0.1	0.1	0.1	0.1	0.1	0.1	0.0	0.0	0.1	0.0
M_LDL_C	0.2	0.0	0.1	0.0	0.1	0.1	0.1	0.0	0.0	0.0	-0.1
M_LDL_CE	0.2	0.0	0.1	0.0	0.1	0.1	0.1	0.0	0.0	0.0	-0.1
M_LDL_FC	0.2	-0.1	0.1	-0.1	0.1	0.1	0.1	0.0	0.0	0.0	-0.1
M_LDL_L	0.2	0.0	0.1	0.0	0.1	0.1	0.1	0.0	0.0	0.0	-0.1
M_LDL_P	0.1	0.0	0.1	0.0	0.1	0.1	0.1	0.0	0.0	0.0	-0.1
M_LDL_PL	0.2	0.0	0.1	0.0	0.1	0.1	0.1	0.0	0.0	0.0	-0.1
M_LDL_TG	0.1	0.1	0.1	0.1	0.1	0.2	0.1	0.0	0.0	0.1	0.0
S_LDL_C	0.1	0.0	0.1	0.0	0.1	0.1	0.1	0.0	0.0	0.0	-0.1
S_LDL_CE	0.1	0.0	0.1	0.0	0.1	0.1	0.1	0.0	0.0	0.0	-0.1
S_LDL_FC	0.1	-0.1	0.1	-0.1	0.1	0.0	0.0	0.0	0.0	0.0	-0.1
S_LDL_L	0.1	0.0	0.1	0.0	0.1	0.1	0.1	0.0	0.0	0.0	-0.1
S_LDL_P	0.1	0.0	0.1	0.0	0.1	0.1	0.1	0.0	0.0	0.0	-0.1
S_LDL_PL	0.1	0.0	0.1	0.0	0.1	0.1	0.1	0.0	0.0	0.0	-0.1
S_LDL_TG	0.1	0.1	0.2	0.1	0.0	0.2	0.2	0.0	0.0	0.1	0.0
IDL_C	0.1	-0.1	-0.1	-0.1	0.1	-0.1	-0.1	0.0	0.0	0.0	-0.1
IDL_CE	0.1	-0.1	-0.1	-0.1	0.1	-0.1	-0.1	0.0	0.0	0.0	-0.1
IDL_FC	0.1	0.0	-0.1	-0.1	0.1	-0.1	-0.1	0.0	0.0	0.0	-0.1
IDL_L	0.1	0.0	-0.1	-0.1	0.1	-0.1	-0.1	0.0	0.0	0.0	-0.1
IDL_P	0.1	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	-0.1
IDL_PL	0.1	0.0	-0.1	-0.1	0.1	-0.1	-0.1	0.0	0.0	0.0	-0.1
IDL_TG	0.0	0.1	0.0	0.1	0.1	0.1	0.0	0.0	0.0	0.1	0.0
XL_HDL_C	0.0	-0.1	-0.3	-0.1	0.0	-0.3	-0.3	0.1	0.0	-0.1	-0.1
XL_HDL_CE	0.0	-0.1	-0.3	-0.1	0.0	-0.3	-0.3	0.1	0.0	-0.1	-0.1
XL_HDL_FC	-0.1	0.0	-0.3	-0.1	0.0	-0.3	-0.3	0.1	0.0	-0.1	-0.1
XL_HDL_L	-0.1	-0.1	-0.4	-0.1	0.0	-0.3	-0.3	0.1	0.0	-0.1	0.0
XL_HDL_P	0.0	-0.1	-0.3	-0.1	0.0	-0.3	-0.3	0.1	0.0	0.0	-0.1
XL_HDL_PL	-0.1	-0.1	-0.4	-0.1	0.0	-0.3	-0.3	0.1	0.0	-0.1	0.0
XL_HDL_TG	0.0	0.0	0.0	0.1	0.1	0.0	0.0	0.0	0.0	0.1	0.0
L_HDL_C	0.0	-0.1	-0.4	-0.1	0.0	-0.4	-0.3	0.1	0.1	-0.1	0.0
L_HDL_CE	0.0	-0.1	-0.4	-0.1	0.0	-0.4	-0.3	0.1	0.1	-0.1	0.0
L_HDL_FC	0.0	-0.1	-0.4	-0.1	0.0	-0.4	-0.3	0.1	0.1	-0.1	0.0
L_HDL_L	0.0	-0.1	-0.4	-0.1	0.1	-0.4	-0.4	0.1	0.1	-0.1	0.0
L_HDL_P	0.0	-0.1	-0.4	-0.1	0.0	-0.4	-0.3	0.1	0.1	-0.1	0.0
L_HDL_PL	0.0	-0.1	-0.4	-0.1	0.1	-0.4	-0.3	0.1	0.1	-0.1	0.0
L_HDL_TG	0.0	0.0	-0.2	0.1	0.1	-0.1	-0.1	0.1	0.1	0.0	0.0
M_HDL_C	0.1	-0.1	-0.3	-0.1	0.1	-0.3	-0.3	0.1	0.1	0.0	0.0
M_HDL_CE	0.1	-0.1	-0.3	-0.1	0.1	-0.3	-0.3	0.1	0.1	0.0	0.0
M_HDL_FC	0.1	-0.1	-0.3	-0.1	0.1	-0.3	-0.3	0.1	0.1	0.0	0.0
M_HDL_L	0.1	-0.1	-0.3	0.0	0.1	-0.2	-0.2	0.1	0.1	0.0	0.0
M_HDL_P	0.1	-0.1	-0.3	-0.1	0.1	-0.3	-0.2	0.1	0.1	0.0	0.0
M_HDL_PL	0.1	-0.1	-0.3	0.0	0.1	-0.2	-0.2	0.1	0.1	0.0	0.0
M_HDL_TG	0.0	0.1	0.0	0.1	0.1	0.1	0.0	0.0	0.1	0.1	0.0
S_HDL_C	0.2	-0.1	0.0	0.0	0.1	0.0	0.0	0.1	0.1	0.0	0.0
S_HDL_CE	0.2	-0.1	0.0	0.0	0.1	0.1	0.0	0.0	0.0	0.0	0.0

Supplementary Tables

Table 10 continued from previous page

	Albumin (g/L)	C-Reactive Protein (mg/L)	Erythrocytes (10 ¹² cells/L)	Leucocytes (10 ⁹ cells/L)	Platelets (10 ⁹ cells/L)	Haemoglobin (g/dL)	Haematocrit (%)	Mean Corpuscular Volume	Mean Corpuscular Haemoglobin (pg)	Mean Corpuscular Haemoglobin (g/dL)	Anthypertensives
S_HDL_FC	0.2	0.0	-0.1	0.0	0.1	0.0	0.0	0.1	0.1	0.0	0.0
S_HDL_L	0.2	0.0	0.0	0.0	0.1	0.0	0.0	0.1	0.1	0.0	0.0
S_HDL_P	0.2	-0.1	0.0	0.0	0.1	0.1	0.1	0.0	0.1	0.0	0.0
S_HDL_PL	0.1	0.0	0.0	0.1	0.1	0.0	0.0	0.1	0.1	0.0	0.0
S_HDL_TG	0.0	0.1	0.2	0.2	0.0	0.2	0.2	0.0	0.0	0.1	0.0
ApoA1	0.1	-0.1	-0.3	-0.1	0.1	-0.3	-0.3	0.1	0.1	0.0	0.0
ApoB	0.1	0.0	0.1	0.0	0.1	0.1	0.1	0.0	0.0	0.0	-0.1

3 Metabolomic profiles predict individual multi-disease outcomes

Table 11: Ranked aggregated coefficients of PANEL. Coefficients were aggregated over all 24 endpoints. Rank

Predictor	Mean(coef)	Rank
1	Age	0.518
2	Smoking Status	0.324
3	Weight	0.29
4	Sex	0.281
5	Daily Healthy Food	0.247
6	Body-Mass-Index	0.23
7	Type 2 Diabetes	0.216
8	Standing Height	0.195
9	Antihypertensives	0.19
10	Waist-Hip Ratio	0.119
11	Cystatine C	0.11
12	Erythrocyte Count	0.102
13	LDL Cholesterol	0.098
14	Waist Circumference	0.097
15	Hematocrit	0.088
16	Hemoglobin	0.086
17	Creatinine	0.081
18	Alcohol Intake	0.076
19	Total Cholesterol	0.076
20	FH Type 2 Diabetes	0.07
21	Education	0.064
22	HDL Cholesterol	0.058
23	MCV	0.055
24	HbA1c	0.054
25	ALT	0.054
26	Systolic Blood Pressure	0.054
27	Albumin	0.053
28	Urate	0.051
29	MCH	0.045
30	Leukocyte Count	0.041
31	MCHC	0.039
32	AST	0.038
33	Triglycerides	0.037
34	Urea	0.037
35	Glucose	0.029
36	Platelet Count	0.027
37	C-reactive Protein	0.025
38	AP	0.023
39	Physical Activity	0.021

Supplementary Tables

Table 12: Sensitivity, positive predictive value, and positive likelihood ratios of the metabolomic state addition over false-positive rate thresholds.

Endpoint	Score	5% FPR				10% FPR				15% FPR				20% FPR			
		Threshold	Sensitivity	PPV	LR+												
MACE	MET	0.13	0.13	0.17	2.65	0.11	0.24	0.16	2.37	0.10	0.32	0.15	2.15	0.09	0.40	0.14	2.01
MACE	AgeSex	0.15	0.15	0.20	3.05	0.12	0.26	0.17	2.64	0.10	0.36	0.16	2.38	0.09	0.44	0.15	2.21
MACE	AgeSex+MET	0.15	0.17	0.21	3.35	0.12	0.29	0.19	2.90	0.10	0.39	0.17	2.58	0.09	0.48	0.16	2.38
MACE	ASCVD	0.15	0.17	0.22	3.49	0.12	0.30	0.19	2.95	0.10	0.40	0.17	2.63	0.09	0.49	0.16	2.43
MACE	ASCVD+MET	0.15	0.18	0.22	3.63	0.12	0.31	0.20	3.06	0.10	0.41	0.18	2.73	0.09	0.50	0.16	2.48
MACE	PANEL(noB)+MET	0.15	0.19	0.23	3.74	0.12	0.31	0.20	3.10	0.10	0.41	0.18	2.76	0.09	0.50	0.16	2.48
MACE	PANEL	0.15	0.19	0.23	3.71	0.12	0.31	0.20	3.12	0.10	0.41	0.18	2.74	0.08	0.50	0.17	2.52
MACE	PANEL+MET	0.16	0.19	0.24	3.87	0.12	0.32	0.20	3.18	0.10	0.42	0.18	2.79	0.09	0.51	0.17	2.54
CHD	MET	0.14	0.13	0.19	2.70	0.12	0.23	0.16	2.31	0.11	0.31	0.15	2.10	0.10	0.39	0.14	1.94
CHD	AgeSex	0.14	0.12	0.17	2.46	0.12	0.23	0.16	2.28	0.11	0.32	0.15	2.11	0.09	0.40	0.14	1.97
CHD	AgeSex+MET	0.15	0.15	0.20	2.94	0.13	0.26	0.18	2.57	0.11	0.36	0.17	2.38	0.10	0.44	0.16	2.18
CHD	ASCVD	0.15	0.15	0.20	2.97	0.12	0.26	0.18	2.60	0.10	0.35	0.17	2.33	0.09	0.44	0.16	2.19
CHD	ASCVD+MET	0.15	0.16	0.21	3.14	0.13	0.27	0.19	2.71	0.11	0.37	0.17	2.46	0.09	0.45	0.16	2.26
CHD	PANEL(noB)+MET	0.16	0.16	0.22	3.25	0.13	0.28	0.19	2.76	0.11	0.37	0.17	2.47	0.09	0.45	0.16	2.26
CHD	PANEL	0.15	0.16	0.22	3.28	0.12	0.28	0.19	2.78	0.10	0.37	0.17	2.47	0.09	0.45	0.16	2.27
CHD	PANEL+MET	0.16	0.17	0.22	3.32	0.12	0.28	0.19	2.84	0.11	0.38	0.18	2.51	0.09	0.46	0.16	2.30
Cerebral Stroke	MET	0.04	0.15	0.05	3.04	0.03	0.26	0.05	2.59	0.03	0.35	0.04	2.36	0.02	0.42	0.04	2.11
Cerebral Stroke	AgeSex	0.04	0.16	0.06	3.10	0.03	0.28	0.05	2.82	0.03	0.37	0.05	2.50	0.02	0.46	0.04	2.31
Cerebral Stroke	AgeSex+MET	0.04	0.17	0.06	3.46	0.03	0.30	0.05	2.99	0.03	0.41	0.05	2.71	0.02	0.51	0.05	2.54
Cerebral Stroke	ASCVD	0.04	0.19	0.07	3.84	0.03	0.32	0.06	3.24	0.02	0.45	0.05	2.97	0.02	0.53	0.05	2.66
Cerebral Stroke	ASCVD+MET	0.04	0.19	0.07	3.77	0.03	0.34	0.06	3.41	0.03	0.45	0.05	3.00	0.02	0.54	0.05	2.72
Cerebral Stroke	PANEL(noB)+MET	0.04	0.19	0.07	3.82	0.03	0.34	0.06	3.40	0.03	0.45	0.05	3.01	0.02	0.54	0.05	2.71
Cerebral Stroke	PANEL	0.04	0.20	0.07	4.09	0.03	0.34	0.06	3.39	0.02	0.46	0.05	3.08	0.02	0.55	0.05	2.77
Cerebral Stroke	PANEL+MET	0.04	0.20	0.07	4.04	0.03	0.35	0.06	3.50	0.02	0.47	0.06	3.11	0.02	0.56	0.05	2.80
Dementia	MET	0.04	0.15	0.06	2.95	0.03	0.24	0.05	2.42	0.02	0.32	0.04	2.13	0.02	0.39	0.04	1.97
Dementia	AgeSex	0.03	0.16	0.06	3.27	0.03	0.27	0.05	2.76	0.03	0.40	0.05	2.66	0.02	0.48	0.05	2.40
Dementia	AgeSex+MET	0.04	0.18	0.07	3.58	0.03	0.29	0.06	2.93	0.03	0.40	0.05	2.64	0.02	0.47	0.05	2.36
Dementia	ASCVD	0.03	0.17	0.07	3.43	0.03	0.29	0.06	2.88	0.03	0.40	0.05	2.68	0.02	0.49	0.05	2.47
Dementia	ASCVD+MET	0.04	0.19	0.07	3.73	0.03	0.30	0.06	2.98	0.03	0.39	0.05	2.63	0.02	0.48	0.05	2.42
Dementia	PANEL(noB)+MET	0.04	0.19	0.07	3.84	0.03	0.31	0.06	3.13	0.03	0.41	0.05	2.74	0.02	0.49	0.05	2.42
Dementia	PANEL	0.04	0.19	0.07	3.78	0.03	0.32	0.06	3.23	0.03	0.44	0.06	2.91	0.02	0.51	0.05	2.57
Dementia	PANEL+MET	0.04	0.20	0.08	4.04	0.03	0.33	0.06	3.28	0.03	0.43	0.06	2.86	0.02	0.51	0.05	2.55
Heart Failure	MET	0.08	0.20	0.13	4.08	0.06	0.32	0.11	3.21	0.04	0.41	0.09	2.73	0.04	0.48	0.08	2.39
Heart Failure	AgeSex	0.06	0.15	0.10	2.93	0.05	0.26	0.09	2.58	0.05	0.35	0.08	2.36	0.04	0.43	0.07	2.17
Heart Failure	AgeSex+MET	0.08	0.23	0.14	4.59	0.06	0.35	0.11	3.55	0.05	0.44	0.10	2.93	0.04	0.52	0.09	2.58
Heart Failure	ASCVD	0.07	0.20	0.13	4.06	0.05	0.32	0.10	3.21	0.04	0.42	0.09	2.82	0.04	0.49	0.08	2.47
Heart Failure	ASCVD+MET	0.08	0.24	0.15	4.87	0.06	0.36	0.12	3.62	0.05	0.45	0.10	3.01	0.04	0.53	0.09	2.63
Heart Failure	PANEL(noB)+MET	0.08	0.25	0.15	5.00	0.06	0.37	0.12	3.72	0.05	0.47	0.10	3.11	0.04	0.54	0.09	2.70
Heart Failure	PANEL	0.08	0.26	0.16	5.12	0.05	0.38	0.12	3.78	0.04	0.48	0.10	3.17	0.04	0.55	0.09	2.77
Heart Failure	PANEL+MET	0.08	0.26	0.16	5.22	0.06	0.39	0.13	3.93	0.04	0.48	0.11	3.23	0.04	0.55	0.09	2.76
Atrial Fibrillation	MET	0.10	0.16	0.16	3.10	0.09	0.26	0.13	2.58	0.07	0.34	0.12	2.28	0.06	0.42	0.11	2.08
Atrial Fibrillation	AgeSex	0.11	0.16	0.16	3.21	0.09	0.29	0.15	2.87	0.08	0.39	0.14	2.63	0.07	0.49	0.13	2.43
Atrial Fibrillation	AgeSex+MET	0.12	0.19	0.19	3.83	0.10	0.31	0.16	3.13	0.08	0.42	0.14	2.81	0.07	0.51	0.13	2.54
Atrial Fibrillation	ASCVD	0.12	0.18	0.18	3.63	0.09	0.32	0.16	3.17	0.08	0.42	0.15	2.82	0.07	0.51	0.13	2.53
Atrial Fibrillation	ASCVD+MET	0.12	0.19	0.19	3.82	0.09	0.33	0.17	3.30	0.08	0.43	0.15	2.85	0.07	0.52	0.13	2.58
Atrial Fibrillation	PANEL(noB)+MET	0.13	0.21	0.20	4.22	0.10	0.35	0.17	3.48	0.08	0.45	0.15	3.02	0.07	0.54	0.14	2.72
Atrial Fibrillation	PANEL	0.12	0.22	0.21	4.50	0.09	0.37	0.18	3.66	0.07	0.47	0.16	3.16	0.06	0.56	0.14	2.81
Atrial Fibrillation	PANEL+MET	0.13	0.23	0.22	4.59	0.09	0.37	0.18	3.69	0.08	0.47	0.16	3.17	0.06	0.56	0.14	2.81
T2 Diabetes	MET	0.14	0.36	0.27	7.29	0.09	0.50	0.20	5.00	0.06	0.60	0.17	3.99	0.05	0.67	0.14	3.36
T2 Diabetes	AgeSex	0.07	0.09	0.09	1.87	0.06	0.18	0.08	1.75	0.06	0.25	0.08	1.65	0.05	0.32	0.07	1.60
T2 Diabetes	AgeSex+MET	0.14	0.37	0.27	7.38	0.09	0.50	0.20	5.00	0.06	0.60	0.17	3.99	0.05	0.67	0.14	3.37
T2 Diabetes	ASCVD	0.10	0.18	0.15	3.51	0.08	0.29	0.13	2.94	0.07	0.40	0.12	2.67	0.06	0.49	0.11	2.45
T2 Diabetes	ASCVD+MET	0.14	0.37	0.27	7.34	0.09	0.50	0.20	5.02	0.06	0.60	0.17	4.02	0.05	0.68	0.14	3.39
T2 Diabetes	PANEL(noB)+MET	0.14	0.40	0.29	8.01	0.09	0.56	0.22	5.56	0.06	0.65	0.18	4.35	0.05	0.72	0.15	3.60
T2 Diabetes	PANEL	0.12	0.42	0.30	8.36	0.08	0.57	0.22	5.74	0.06	0.68	0.19	4.55	0.05	0.75	0.16	3.74
T2 Diabetes	PANEL+MET	0.13	0.46	0.31	9.17	0.08	0.60	0.23	6.03	0.06	0.69	0.19	4.61	0.04	0.76	0.16	3.81
Liver Disease	MET	0.09	0.16	0.15	3.14	0.07	0.26	0.12	2.57	0.06	0.33	0.11	2.23	0.05	0.40	0.10	2.00
Liver Disease	AgeSex	0.06	0.07	0.07	1.47	0.06	0.15	0.07	1.46	0.05	0.22	0.07	1.47	0.05	0.28	0.07	1.43
Liver Disease	AgeSex+MET	0.09	0.16	0.15	3.14	0.07	0.26	0.12	2.59	0.06	0.34	0.11	2.24	0.05	0.41	0.10	2.04

3 Metabolomic profiles predict individual multi-disease outcomes

Table 12 continued from previous page

		5% FPR				10% FPR				15% FPR				20% FPR			
Liver Disease	ASCVD	0.07	0.13	0.12	2.54	0.06	0.21	0.10	2.11	0.05	0.29	0.09	1.92	0.05	0.36	0.09	1.78
Liver Disease	ASCVD+MET	0.09	0.16	0.15	3.23	0.07	0.26	0.13	2.64	0.06	0.34	0.11	2.29	0.05	0.41	0.10	2.07
Liver Disease	PANEL(noB)+MET	0.09	0.16	0.15	3.24	0.07	0.27	0.13	2.66	0.06	0.35	0.11	2.31	0.05	0.42	0.10	2.10
Liver Disease	PANEL	0.08	0.18	0.17	3.70	0.07	0.29	0.14	2.93	0.06	0.38	0.12	2.52	0.05	0.45	0.11	2.24
Liver Disease	PANEL+MET	0.09	0.19	0.17	3.83	0.07	0.30	0.14	2.96	0.06	0.38	0.12	2.53	0.05	0.45	0.11	2.24
Renal Disease	MET	0.19	0.19	0.31	3.81	0.15	0.28	0.25	2.83	0.13	0.36	0.23	2.43	0.11	0.43	0.21	2.17
Renal Disease	AgeSex	0.16	0.12	0.22	2.31	0.15	0.22	0.21	2.17	0.13	0.31	0.20	2.05	0.12	0.39	0.19	1.93
Renal Disease	AgeSex+MET	0.20	0.21	0.33	4.11	0.16	0.31	0.27	3.09	0.13	0.39	0.24	2.62	0.12	0.46	0.22	2.31
Renal Disease	ASCVD	0.17	0.16	0.28	3.23	0.14	0.26	0.24	2.60	0.13	0.35	0.22	2.34	0.12	0.43	0.20	2.14
Renal Disease	ASCVD+MET	0.20	0.21	0.33	4.14	0.16	0.32	0.27	3.16	0.13	0.40	0.24	2.64	0.12	0.47	0.22	2.33
Renal Disease	PANEL(noB)+MET	0.20	0.21	0.34	4.28	0.16	0.32	0.28	3.20	0.13	0.40	0.24	2.70	0.12	0.47	0.22	2.37
Renal Disease	PANEL	0.18	0.23	0.36	4.70	0.15	0.34	0.29	3.41	0.13	0.42	0.25	2.80	0.11	0.49	0.23	2.44
Renal Disease	PANEL+MET	0.19	0.23	0.36	4.68	0.15	0.34	0.29	3.40	0.13	0.43	0.25	2.85	0.11	0.49	0.23	2.47
PAD	MET	0.07	0.16	0.10	3.11	0.06	0.27	0.09	2.66	0.05	0.35	0.08	2.31	0.04	0.42	0.07	2.09
PAD	AgeSex	0.06	0.14	0.09	2.74	0.05	0.24	0.08	2.40	0.05	0.33	0.07	2.21	0.04	0.41	0.07	2.03
PAD	AgeSex+MET	0.08	0.17	0.11	3.49	0.06	0.29	0.09	2.88	0.05	0.39	0.09	2.58	0.04	0.46	0.08	2.30
PAD	ASCVD	0.07	0.18	0.12	3.67	0.05	0.30	0.10	2.99	0.05	0.39	0.09	2.59	0.04	0.46	0.08	2.32
PAD	ASCVD+MET	0.08	0.20	0.13	4.02	0.06	0.32	0.11	3.23	0.05	0.41	0.09	2.72	0.04	0.48	0.08	2.42
PAD	PANEL(noB)+MET	0.08	0.21	0.13	4.11	0.06	0.32	0.10	3.22	0.05	0.41	0.09	2.75	0.04	0.49	0.08	2.46
PAD	PANEL	0.07	0.21	0.13	4.24	0.05	0.32	0.10	3.20	0.05	0.42	0.09	2.79	0.04	0.49	0.08	2.44
PAD	PANEL+MET	0.08	0.21	0.14	4.29	0.06	0.33	0.11	3.31	0.05	0.42	0.09	2.81	0.04	0.50	0.08	2.49
Ven. Thrombosis	MET	0.03	0.11	0.04	2.17	0.02	0.19	0.03	1.91	0.02	0.27	0.03	1.79	0.02	0.34	0.03	1.72
Ven. Thrombosis	AgeSex	0.02	0.08	0.03	1.61	0.02	0.16	0.03	1.58	0.02	0.23	0.03	1.52	0.02	0.30	0.03	1.51
Ven. Thrombosis	AgeSex+MET	0.03	0.11	0.04	2.23	0.03	0.21	0.03	2.06	0.02	0.28	0.03	1.89	0.02	0.35	0.03	1.76
Ven. Thrombosis	ASCVD	0.03	0.09	0.03	1.76	0.02	0.18	0.03	1.76	0.02	0.25	0.03	1.69	0.02	0.33	0.03	1.64
Ven. Thrombosis	ASCVD+MET	0.03	0.11	0.04	2.18	0.03	0.20	0.03	1.97	0.02	0.28	0.03	1.88	0.02	0.35	0.03	1.76
Ven. Thrombosis	PANEL(noB)+MET	0.03	0.12	0.04	2.47	0.03	0.23	0.04	2.25	0.02	0.31	0.03	2.09	0.02	0.39	0.03	1.93
Ven. Thrombosis	PANEL	0.03	0.13	0.04	2.61	0.03	0.23	0.04	2.29	0.02	0.32	0.04	2.13	0.02	0.39	0.03	1.94
Ven. Thrombosis	PANEL+MET	0.03	0.13	0.04	2.60	0.03	0.23	0.04	2.27	0.02	0.32	0.04	2.15	0.02	0.39	0.03	1.97
AAA	MET	0.02	0.19	0.03	3.86	0.02	0.32	0.02	3.25	0.01	0.42	0.02	2.83	0.01	0.49	0.02	2.46
AAA	AgeSex	0.02	0.19	0.03	3.85	0.02	0.34	0.03	3.40	0.01	0.47	0.02	3.16	0.01	0.60	0.02	3.00
AAA	AgeSex+MET	0.02	0.24	0.04	4.80	0.02	0.41	0.03	4.06	0.01	0.53	0.03	3.55	0.01	0.63	0.02	3.12
AAA	ASCVD	0.02	0.26	0.04	5.22	0.02	0.41	0.03	4.10	0.01	0.52	0.03	3.49	0.01	0.62	0.02	3.12
AAA	ASCVD+MET	0.02	0.28	0.04	5.50	0.02	0.42	0.03	4.24	0.01	0.55	0.03	3.66	0.01	0.64	0.02	3.21
AAA	PANEL(noB)+MET	0.02	0.29	0.04	5.79	0.02	0.44	0.03	4.40	0.01	0.55	0.03	3.70	0.01	0.63	0.02	3.17
AAA	PANEL	0.02	0.29	0.04	5.75	0.02	0.42	0.03	4.21	0.01	0.54	0.03	3.62	0.01	0.64	0.02	3.19
AAA	PANEL+MET	0.02	0.30	0.04	5.94	0.02	0.44	0.03	4.36	0.01	0.55	0.03	3.66	0.01	0.65	0.02	3.23
COPD	MET	0.15	0.14	0.19	2.80	0.12	0.23	0.17	2.33	0.11	0.31	0.15	2.07	0.10	0.38	0.14	1.92
COPD	AgeSex	0.11	0.09	0.13	1.75	0.11	0.17	0.12	1.66	0.10	0.24	0.12	1.61	0.10	0.31	0.12	1.56
COPD	AgeSex+MET	0.15	0.15	0.20	2.97	0.13	0.24	0.17	2.44	0.11	0.33	0.16	2.19	0.10	0.40	0.15	2.02
COPD	ASCVD	0.14	0.16	0.22	3.28	0.11	0.26	0.18	2.56	0.10	0.33	0.16	2.22	0.09	0.39	0.14	1.97
COPD	ASCVD+MET	0.16	0.18	0.24	3.59	0.12	0.28	0.20	2.84	0.11	0.37	0.17	2.44	0.09	0.43	0.16	2.17
COPD	PANEL(noB)+MET	0.16	0.19	0.25	3.84	0.13	0.30	0.20	2.96	0.11	0.38	0.18	2.57	0.09	0.45	0.16	2.27
COPD	PANEL	0.15	0.19	0.25	3.87	0.12	0.30	0.20	2.99	0.10	0.38	0.18	2.51	0.09	0.45	0.16	2.24
COPD	PANEL+MET	0.16	0.20	0.25	3.97	0.13	0.30	0.21	3.03	0.11	0.39	0.18	2.61	0.09	0.46	0.16	2.28
Asthma	MET	0.06	0.09	0.07	1.84	0.05	0.17	0.06	1.66	0.05	0.24	0.06	1.59	0.04	0.30	0.06	1.50
Asthma	AgeSex	0.04	0.07	0.05	1.35	0.04	0.13	0.05	1.25	0.04	0.19	0.05	1.26	0.04	0.25	0.05	1.23
Asthma	AgeSex+MET	0.06	0.09	0.07	1.82	0.05	0.18	0.07	1.77	0.05	0.25	0.06	1.64	0.04	0.31	0.06	1.55
Asthma	ASCVD	0.05	0.09	0.06	1.74	0.05	0.16	0.06	1.58	0.04	0.23	0.06	1.54	0.04	0.29	0.05	1.47
Asthma	ASCVD+MET	0.06	0.09	0.07	1.90	0.05	0.17	0.06	1.74	0.05	0.25	0.06	1.67	0.04	0.32	0.06	1.59
Asthma	PANEL(noB)+MET	0.06	0.11	0.08	2.24	0.05	0.20	0.07	1.96	0.05	0.27	0.07	1.83	0.04	0.35	0.06	1.75
Asthma	PANEL	0.06	0.11	0.08	2.22	0.05	0.20	0.07	1.99	0.05	0.28	0.07	1.88	0.04	0.35	0.06	1.75
Asthma	PANEL+MET	0.06	0.12	0.08	2.32	0.05	0.20	0.07	2.03	0.05	0.28	0.07	1.88	0.04	0.35	0.06	1.75
Parkinson's	MET	0.01	0.13	0.02	2.52	0.01	0.21	0.01	2.09	0.01	0.29	0.01	1.93	0.01	0.37	0.01	1.86
Parkinson's	AgeSex	0.01	0.18	0.02	3.63	0.01	0.33	0.02	3.28	0.01	0.42	0.02	2.79	0.01	0.50	0.02	2.50
Parkinson's	AgeSex+MET	0.01	0.17	0.02	3.36	0.01	0.29	0.02	2.93	0.01	0.41	0.02	2.75	0.01	0.50	0.01	2.49
Parkinson's	ASCVD	0.01	0.19	0.02	3.75	0.01	0.33	0.02	3.29	0.01	0.45	0.02	2.98	0.01	0.53	0.02	2.65
Parkinson's	ASCVD+MET	0.01	0.19	0.02	3.78	0.01	0.31	0.02	3.07	0.01	0.41	0.02	2.75	0.01	0.50	0.01	2.49
Parkinson's	PANEL(noB)+MET	0.01	0.18	0.02	3.64	0.01	0.30	0.02	2.96	0.01	0.42	0.02	2.77	0.01	0.52	0.02	2.61
Parkinson's	PANEL	0.01	0.18	0.02	3.61	0.01	0.35	0.02	3.47	0.01	0.45	0.02	2.98	0.01	0.53	0.02	2.66
Parkinson's	PANEL+MET	0.01	0.16	0.02	3.22	0.01	0.32	0.02	3.22	0.01	0.43	0.02	2.86	0.01	0.53	0.02	2.64
Cataracts	MET	0.14	0.11	0.19	2.23	0.12	0.20	0.18	2.04	0.11	0.28	0.17	1.88	0.11	0.35	0.16	1.76
Cataracts	AgeSex	0.19	0.17	0.27	3.50	0.16	0.31	0.25	3.06	0.14	0.41	0.23	2.75	0.12	0.50	0.21	2.52
Cataracts	AgeSex+MET	0.19	0.18	0.27	3.52	0.16	0.31	0.25	3.09	0.13	0.41	0.23	2.75	0.12	0.51	0.21	2.53
Cataracts	ASCVD	0.19	0.18	0.28	3.67	0.16	0.31	0.25	3.12	0.13	0.43	0.23	2.84	0.12	0.51	0.22	2.57

Supplementary Tables

Table 12 continued from previous page

		5% FPR				10% FPR				15% FPR				20% FPR			
Cataracts	ASCVD+MET	0.19	0.18	0.28	3.63	0.16	0.31	0.25	3.14	0.13	0.42	0.23	2.83	0.12	0.52	0.22	2.58
Cataracts	PANEL(noB)+MET	0.19	0.18	0.28	3.67	0.16	0.32	0.25	3.16	0.13	0.42	0.23	2.82	0.12	0.52	0.22	2.58
Cataracts	PANEL	0.19	0.19	0.29	3.78	0.16	0.32	0.26	3.23	0.13	0.43	0.23	2.85	0.11	0.52	0.22	2.60
Cataracts	PANEL+MET	0.19	0.18	0.28	3.69	0.16	0.32	0.26	3.23	0.13	0.43	0.23	2.84	0.11	0.52	0.22	2.60
Glaucoma	MET	0.03	0.08	0.04	1.67	0.03	0.14	0.03	1.37	0.03	0.21	0.03	1.38	0.02	0.27	0.03	1.34
Glaucoma	AgeSex	0.04	0.10	0.05	2.06	0.03	0.20	0.04	1.96	0.03	0.29	0.04	1.92	0.03	0.36	0.04	1.83
Glaucoma	AgeSex+MET	0.04	0.10	0.04	1.93	0.03	0.20	0.04	1.97	0.03	0.28	0.04	1.89	0.03	0.37	0.04	1.86
Glaucoma	ASCVD	0.04	0.11	0.05	2.21	0.03	0.20	0.05	2.01	0.03	0.29	0.04	1.91	0.03	0.37	0.04	1.83
Glaucoma	ASCVD+MET	0.04	0.11	0.05	2.17	0.03	0.20	0.05	2.01	0.03	0.29	0.04	1.93	0.03	0.37	0.04	1.83
Glaucoma	PANEL(noB)+MET	0.04	0.10	0.05	2.08	0.03	0.20	0.05	1.99	0.03	0.29	0.04	1.94	0.03	0.37	0.04	1.84
Glaucoma	PANEL	0.04	0.11	0.05	2.13	0.03	0.20	0.05	2.02	0.03	0.29	0.04	1.92	0.03	0.36	0.04	1.82
Glaucoma	PANEL+MET	0.04	0.10	0.05	2.10	0.03	0.20	0.05	2.01	0.03	0.29	0.04	1.95	0.03	0.36	0.04	1.82
Fractures	MET	0.10	0.09	0.13	1.85	0.09	0.16	0.11	1.62	0.08	0.23	0.11	1.52	0.08	0.29	0.10	1.46
Fractures	AgeSex	0.10	0.10	0.14	2.04	0.09	0.18	0.12	1.78	0.09	0.25	0.12	1.66	0.08	0.32	0.11	1.61
Fractures	AgeSex+MET	0.11	0.10	0.14	2.09	0.10	0.19	0.13	1.86	0.09	0.26	0.12	1.71	0.08	0.33	0.12	1.64
Fractures	ASCVD	0.10	0.10	0.14	1.97	0.09	0.18	0.13	1.82	0.09	0.26	0.12	1.70	0.08	0.32	0.11	1.61
Fractures	ASCVD+MET	0.11	0.10	0.14	2.08	0.10	0.19	0.13	1.85	0.09	0.26	0.12	1.72	0.08	0.33	0.12	1.64
Fractures	PANEL(noB)+MET	0.11	0.11	0.15	2.12	0.10	0.19	0.13	1.89	0.09	0.26	0.12	1.74	0.08	0.33	0.12	1.65
Fractures	PANEL	0.11	0.11	0.15	2.21	0.10	0.19	0.13	1.91	0.09	0.27	0.12	1.77	0.08	0.34	0.12	1.68
Fractures	PANEL+MET	0.11	0.11	0.15	2.22	0.10	0.20	0.14	1.95	0.09	0.27	0.13	1.81	0.08	0.34	0.12	1.70
Lung Cancer	MET	0.03	0.20	0.04	4.10	0.02	0.31	0.03	3.14	0.02	0.41	0.03	2.74	0.01	0.49	0.03	2.45
Lung Cancer	AgeSex	0.02	0.12	0.03	2.37	0.02	0.23	0.02	2.26	0.01	0.31	0.02	2.09	0.01	0.41	0.02	2.03
Lung Cancer	AgeSex+MET	0.03	0.21	0.04	4.14	0.02	0.34	0.04	3.43	0.02	0.45	0.03	3.02	0.01	0.53	0.03	2.64
Lung Cancer	ASCVD	0.02	0.33	0.07	6.64	0.01	0.45	0.05	4.47	0.01	0.52	0.04	3.47	0.01	0.59	0.03	2.96
Lung Cancer	ASCVD+MET	0.03	0.33	0.07	6.54	0.02	0.46	0.05	4.63	0.01	0.55	0.04	3.69	0.01	0.61	0.03	3.05
Lung Cancer	PANEL(noB)+MET	0.03	0.34	0.07	6.87	0.02	0.48	0.05	4.78	0.01	0.56	0.04	3.76	0.01	0.62	0.03	3.12
Lung Cancer	PANEL	0.03	0.35	0.07	6.97	0.02	0.48	0.05	4.78	0.01	0.57	0.04	3.81	0.01	0.63	0.03	3.13
Lung Cancer	PANEL+MET	0.03	0.35	0.07	6.98	0.02	0.48	0.05	4.85	0.01	0.56	0.04	3.74	0.01	0.64	0.03	3.19
Skin Cancer	MET	0.06	0.09	0.07	1.78	0.05	0.15	0.06	1.52	0.05	0.22	0.06	1.48	0.04	0.28	0.06	1.40
Skin Cancer	AgeSex	0.07	0.12	0.10	2.42	0.06	0.21	0.09	2.13	0.06	0.31	0.08	2.04	0.05	0.39	0.08	1.93
Skin Cancer	AgeSex+MET	0.07	0.12	0.09	2.35	0.06	0.21	0.09	2.12	0.06	0.31	0.08	2.04	0.05	0.39	0.08	1.93
Skin Cancer	ASCVD	0.07	0.12	0.10	2.42	0.06	0.22	0.09	2.19	0.06	0.31	0.08	2.06	0.05	0.39	0.08	1.96
Skin Cancer	ASCVD+MET	0.07	0.12	0.09	2.35	0.06	0.22	0.09	2.19	0.06	0.31	0.08	2.04	0.05	0.39	0.08	1.95
Skin Cancer	PANEL(noB)+MET	0.07	0.13	0.10	2.51	0.06	0.22	0.09	2.23	0.06	0.31	0.08	2.06	0.05	0.39	0.08	1.96
Skin Cancer	PANEL	0.08	0.13	0.10	2.61	0.06	0.22	0.09	2.20	0.06	0.32	0.09	2.10	0.05	0.40	0.08	1.98
Skin Cancer	PANEL+MET	0.08	0.13	0.10	2.55	0.07	0.23	0.09	2.25	0.06	0.31	0.08	2.10	0.05	0.39	0.08	1.97
Colon Cancer	MET	0.02	0.09	0.02	1.79	0.01	0.18	0.02	1.80	0.01	0.25	0.02	1.69	0.01	0.31	0.02	1.57
Colon Cancer	AgeSex	0.02	0.12	0.02	2.39	0.02	0.22	0.02	2.20	0.01	0.32	0.02	2.10	0.01	0.41	0.02	2.05
Colon Cancer	AgeSex+MET	0.02	0.13	0.03	2.54	0.02	0.23	0.02	2.31	0.01	0.32	0.02	2.12	0.01	0.39	0.02	1.96
Colon Cancer	ASCVD	0.02	0.12	0.02	2.38	0.02	0.22	0.02	2.22	0.01	0.32	0.02	2.13	0.01	0.41	0.02	2.07
Colon Cancer	ASCVD+MET	0.02	0.13	0.03	2.51	0.02	0.23	0.02	2.26	0.01	0.31	0.02	2.06	0.01	0.40	0.02	2.01
Colon Cancer	PANEL(noB)+MET	0.02	0.13	0.03	2.52	0.02	0.23	0.02	2.28	0.01	0.32	0.02	2.13	0.01	0.41	0.02	2.03
Colon Cancer	PANEL	0.02	0.13	0.03	2.56	0.02	0.24	0.02	2.35	0.01	0.32	0.02	2.11	0.01	0.41	0.02	2.04
Colon Cancer	PANEL+MET	0.02	0.13	0.03	2.57	0.02	0.22	0.02	2.24	0.01	0.31	0.02	2.09	0.01	0.39	0.02	1.93
Rectal Cancer	MET	0.01	0.09	0.01	1.88	0.01	0.18	0.01	1.84	0.01	0.26	0.01	1.72	0.01	0.32	0.01	1.61
Rectal Cancer	AgeSex	0.01	0.13	0.02	2.49	0.01	0.26	0.02	2.60	0.01	0.35	0.02	2.35	0.01	0.43	0.01	2.13
Rectal Cancer	AgeSex+MET	0.01	0.12	0.02	2.35	0.01	0.26	0.02	2.56	0.01	0.37	0.02	2.45	0.01	0.44	0.01	2.21
Rectal Cancer	ASCVD	0.01	0.13	0.02	2.61	0.01	0.25	0.02	2.50	0.01	0.36	0.02	2.39	0.01	0.43	0.01	2.14
Rectal Cancer	ASCVD+MET	0.01	0.11	0.01	2.27	0.01	0.25	0.02	2.53	0.01	0.36	0.02	2.38	0.01	0.44	0.01	2.20
Rectal Cancer	PANEL(noB)+MET	0.01	0.12	0.01	2.32	0.01	0.24	0.02	2.39	0.01	0.35	0.01	2.32	0.01	0.45	0.01	2.24
Rectal Cancer	PANEL	0.01	0.13	0.02	2.56	0.01	0.25	0.02	2.45	0.01	0.35	0.01	2.31	0.01	0.42	0.01	2.11
Rectal Cancer	PANEL+MET	0.01	0.12	0.02	2.42	0.01	0.25	0.02	2.45	0.01	0.35	0.02	2.34	0.01	0.43	0.01	2.15
Prostate Cancer	MET	0.05	0.06	0.06	1.20	0.04	0.12	0.06	1.20	0.04	0.17	0.06	1.17	0.04	0.23	0.05	1.14
Prostate Cancer	AgeSex	0.08	0.09	0.09	1.90	0.07	0.19	0.09	1.86	0.06	0.28	0.09	1.83	0.06	0.35	0.08	1.75
Prostate Cancer	AgeSex+MET	0.08	0.10	0.09	1.94	0.07	0.19	0.09	1.88	0.07	0.28	0.09	1.87	0.06	0.35	0.08	1.76
Prostate Cancer	ASCVD	0.08	0.10	0.09	2.05	0.07	0.20	0.09	1.95	0.07	0.28	0.09	1.86	0.06	0.36	0.08	1.80
Prostate Cancer	ASCVD+MET	0.08	0.10	0.09	2.04	0.07	0.20	0.09	1.95	0.07	0.28	0.09	1.89	0.06	0.36	0.08	1.80
Prostate Cancer	PANEL(noB)+MET	0.08	0.10	0.09	1.99	0.07	0.20	0.09	1.96	0.07	0.28	0.09	1.85	0.06	0.37	0.09	1.83
Prostate Cancer	PANEL	0.09	0.10	0.09	1.99	0.07	0.19	0.09	1.95	0.07	0.28	0.09	1.88	0.06	0.37	0.09	1.83
Prostate Cancer	PANEL+MET	0.09	0.10	0.09	1.99	0.07	0.20	0.09	1.97	0.07	0.28	0.09	1.88	0.06	0.37	0.09	1.83
Breast Cancer	MET	0.04	0.06	0.05	1.20	0.04	0.11	0.05	1.14	0.04	0.18	0.05	1.20	0.04	0.23	0.05	1.17
Breast Cancer	AgeSex	0.04	0.05	0.04	1.01	0.04	0.10	0.04	1.04	0.04	0.16	0.04	1.05	0.04	0.22	0.04	1.10
Breast Cancer	AgeSex+MET	0.04	0.06	0.04	1.12	0.04	0.12	0.05	1.16	0.04	0.18	0.05	1.17	0.04	0.23	0.05	1.16
Breast Cancer	ASCVD	0.04	0.05	0.04	1.04	0.04	0.11	0.04	1.07	0.04	0.16	0.04	1.06	0.04	0.22	0.04	1.09
Breast Cancer	ASCVD+MET	0.04	0.05	0.04	1.03	0.04	0.11	0.04	1.10	0.04	0.17	0.05	1.15	0.04	0.22	0.04	1.12

3 Metabolomic profiles predict individual multi-disease outcomes

Table 12 continued from previous page

		5% FPR				10% FPR				15% FPR				20% FPR			
Breast Cancer	PANEL(noB)+MET	0.05	0.07	0.06	1.48	0.04	0.14	0.05	1.36	0.04	0.20	0.05	1.36	0.04	0.26	0.05	1.31
Breast Cancer	PANEL	0.05	0.07	0.05	1.39	0.04	0.14	0.05	1.37	0.04	0.20	0.05	1.33	0.04	0.26	0.05	1.30
Breast Cancer	PANEL+MET	0.05	0.07	0.05	1.40	0.04	0.14	0.05	1.36	0.04	0.20	0.05	1.33	0.04	0.26	0.05	1.31

Table 13: Sensitivity, positive predictive value, and positive likelihood ratios of the metabolomic state addition over false-positive rate thresholds.

Endpoint	Score	25% FPR				30% FPR				35% FPR				40% FPR			
		Threshold	Sensitivity	PPV	LR+												
MACE	MET	0.08	0.48	0.13	1.90	0.07	0.54	0.13	1.81	0.07	0.60	0.12	1.72	0.06	0.65	0.11	1.63
MACE	AgeSex	0.08	0.52	0.14	2.07	0.07	0.58	0.13	1.93	0.06	0.64	0.13	1.84	0.06	0.70	0.12	1.74
MACE	AgeSex+MET	0.08	0.55	0.15	2.21	0.07	0.61	0.14	2.05	0.06	0.68	0.13	1.93	0.05	0.73	0.13	1.82
MACE	ASCVD	0.08	0.56	0.15	2.24	0.07	0.63	0.14	2.09	0.06	0.69	0.14	1.97	0.05	0.74	0.13	1.84
MACE	ASCVD+MET	0.08	0.57	0.15	2.26	0.07	0.64	0.14	2.12	0.06	0.69	0.14	1.98	0.05	0.75	0.13	1.87
MACE	PANEL(noB)+MET	0.08	0.57	0.15	2.26	0.07	0.64	0.14	2.13	0.06	0.69	0.14	1.98	0.05	0.75	0.13	1.86
MACE	PANEL	0.07	0.58	0.16	2.32	0.07	0.64	0.15	2.14	0.06	0.70	0.14	1.99	0.05	0.75	0.13	1.87
MACE	PANEL+MET	0.07	0.58	0.15	2.30	0.07	0.64	0.14	2.13	0.06	0.70	0.14	2.00	0.05	0.75	0.13	1.88
CHD	MET	0.09	0.46	0.14	1.85	0.08	0.53	0.13	1.76	0.07	0.58	0.12	1.66	0.07	0.64	0.12	1.60
CHD	AgeSex	0.08	0.47	0.14	1.86	0.08	0.53	0.13	1.76	0.07	0.59	0.13	1.68	0.06	0.64	0.12	1.60
CHD	AgeSex+MET	0.08	0.51	0.15	2.04	0.08	0.57	0.14	1.91	0.07	0.63	0.13	1.79	0.06	0.68	0.13	1.71
CHD	ASCVD	0.08	0.51	0.15	2.03	0.08	0.57	0.14	1.91	0.07	0.63	0.13	1.80	0.06	0.68	0.13	1.71
CHD	ASCVD+MET	0.08	0.52	0.15	2.09	0.08	0.59	0.14	1.95	0.07	0.64	0.14	1.84	0.06	0.70	0.13	1.74
CHD	PANEL(noB)+MET	0.08	0.53	0.15	2.10	0.07	0.59	0.14	1.96	0.07	0.65	0.14	1.85	0.06	0.70	0.13	1.75
CHD	PANEL	0.08	0.53	0.15	2.11	0.07	0.59	0.14	1.98	0.07	0.65	0.14	1.86	0.06	0.70	0.13	1.75
CHD	PANEL+MET	0.08	0.54	0.15	2.14	0.07	0.60	0.15	1.99	0.07	0.65	0.14	1.87	0.06	0.71	0.13	1.77
Cerebral Stroke	MET	0.02	0.50	0.04	1.99	0.02	0.56	0.03	1.86	0.01	0.62	0.03	1.76	0.01	0.66	0.03	1.66
Cerebral Stroke	AgeSex	0.02	0.54	0.04	2.17	0.02	0.62	0.04	2.06	0.02	0.68	0.04	1.95	0.01	0.73	0.03	1.83
Cerebral Stroke	AgeSex+MET	0.02	0.58	0.04	2.34	0.02	0.65	0.04	2.18	0.01	0.71	0.04	2.04	0.01	0.76	0.03	1.91
Cerebral Stroke	ASCVD	0.02	0.61	0.04	2.43	0.02	0.67	0.04	2.22	0.01	0.72	0.04	2.05	0.01	0.77	0.03	1.92
Cerebral Stroke	ASCVD+MET	0.02	0.62	0.04	2.48	0.02	0.68	0.04	2.27	0.01	0.73	0.04	2.09	0.01	0.78	0.04	1.96
Cerebral Stroke	PANEL(noB)+MET	0.02	0.61	0.04	2.44	0.02	0.68	0.04	2.28	0.01	0.73	0.04	2.10	0.01	0.79	0.04	1.96
Cerebral Stroke	PANEL	0.02	0.63	0.05	2.52	0.02	0.69	0.04	2.30	0.01	0.74	0.04	2.12	0.01	0.78	0.04	1.95
Cerebral Stroke	PANEL+MET	0.02	0.63	0.05	2.52	0.01	0.69	0.04	2.30	0.01	0.74	0.04	2.12	0.01	0.78	0.04	1.95
Dementia	MET	0.02	0.46	0.04	1.82	0.02	0.51	0.03	1.71	0.01	0.56	0.03	1.61	0.01	0.61	0.03	1.53
Dementia	AgeSex	0.02	0.55	0.04	2.20	0.02	0.60	0.04	2.02	0.02	0.65	0.04	1.86	0.02	0.69	0.03	1.72
Dementia	AgeSex+MET	0.02	0.55	0.04	2.21	0.02	0.61	0.04	2.04	0.02	0.66	0.04	1.90	0.01	0.70	0.04	1.76
Dementia	ASCVD	0.02	0.56	0.05	2.24	0.02	0.62	0.04	2.05	0.02	0.67	0.04	1.91	0.02	0.71	0.04	1.77
Dementia	ASCVD+MET	0.02	0.56	0.04	2.22	0.02	0.62	0.04	2.06	0.02	0.67	0.04	1.91	0.01	0.71	0.04	1.78
Dementia	PANEL(noB)+MET	0.02	0.57	0.05	2.28	0.02	0.63	0.04	2.10	0.02	0.68	0.04	1.94	0.01	0.71	0.04	1.79
Dementia	PANEL	0.02	0.58	0.05	2.33	0.02	0.64	0.04	2.12	0.02	0.68	0.04	1.94	0.01	0.72	0.04	1.80
Dementia	PANEL+MET	0.02	0.58	0.05	2.31	0.02	0.63	0.04	2.09	0.02	0.68	0.04	1.93	0.01	0.72	0.04	1.79
Heart Failure	MET	0.03	0.53	0.07	2.14	0.03	0.59	0.07	1.98	0.02	0.64	0.06	1.84	0.02	0.68	0.06	1.71
Heart Failure	AgeSex	0.04	0.51	0.07	2.05	0.04	0.58	0.07	1.94	0.03	0.64	0.06	1.82	0.03	0.69	0.06	1.72
Heart Failure	AgeSex+MET	0.03	0.58	0.08	2.33	0.03	0.63	0.07	2.11	0.03	0.69	0.07	1.96	0.02	0.74	0.06	1.84
Heart Failure	ASCVD	0.03	0.56	0.08	2.23	0.03	0.62	0.07	2.07	0.03	0.67	0.07	1.92	0.03	0.72	0.06	1.79
Heart Failure	ASCVD+MET	0.03	0.59	0.08	2.36	0.03	0.65	0.07	2.16	0.03	0.69	0.07	1.98	0.02	0.74	0.06	1.85
Heart Failure	PANEL(noB)+MET	0.03	0.61	0.08	2.42	0.03	0.66	0.07	2.21	0.03	0.71	0.07	2.02	0.02	0.75	0.06	1.88
Heart Failure	PANEL	0.03	0.62	0.08	2.47	0.03	0.67	0.08	2.24	0.03	0.72	0.07	2.05	0.02	0.76	0.06	1.89
Heart Failure	PANEL+MET	0.03	0.62	0.08	2.46	0.03	0.68	0.08	2.25	0.02	0.72	0.07	2.06	0.02	0.76	0.07	1.91
Atrial Fibrillation	MET	0.06	0.49	0.10	1.94	0.05	0.55	0.10	1.84	0.04	0.61	0.09	1.73	0.04	0.65	0.09	1.63
Atrial Fibrillation	AgeSex	0.06	0.57	0.12	2.27	0.05	0.64	0.11	2.13	0.05	0.69	0.11	1.99	0.04	0.75	0.10	1.87
Atrial Fibrillation	AgeSex+MET	0.06	0.59	0.12	2.34	0.05	0.65	0.12	2.18	0.04	0.71	0.11	2.03	0.04	0.77	0.10	1.91
Atrial Fibrillation	ASCVD	0.06	0.58	0.12	2.34	0.05	0.65	0.12	2.18	0.04	0.71	0.11	2.03	0.04	0.76	0.10	1.90
Atrial Fibrillation	ASCVD+MET	0.06	0.59	0.12	2.37	0.05	0.65	0.12	2.18	0.04	0.72	0.11	2.05	0.04	0.77	0.10	1.92
Atrial Fibrillation	PANEL(noB)+MET	0.06	0.62	0.13	2.49	0.05	0.68	0.12	2.27	0.04	0.74	0.11	2.11	0.04	0.78	0.11	1.96
Atrial Fibrillation	PANEL	0.05	0.63	0.13	2.52	0.05	0.69	0.12	2.31	0.04	0.75	0.11	2.14	0.04	0.79	0.11	1.97
Atrial Fibrillation	PANEL+MET	0.05	0.63	0.13	2.52	0.05	0.69	0.12	2.31	0.04	0.75	0.11	2.13	0.03	0.79	0.11	1.97
T2 Diabetes	MET	0.04	0.73	0.13	2.94	0.03	0.78	0.12	2.61	0.03	0.82	0.11	2.34	0.02	0.85	0.10	2.13
T2 Diabetes	AgeSex	0.05	0.38	0.07	1.53	0.05	0.45	0.07	1.50	0.04	0.51	0.07	1.47	0.04	0.57	0.07	1.42
T2 Diabetes	AgeSex+MET	0.04	0.73	0.13	2.94	0.03	0.78	0.12	2.61	0.02	0.82	0.11	2.35	0.02	0.85	0.10	2.14

Supplementary Tables

Table 13 continued from previous page

		25% FPR			30% FPR			35% FPR			40% FPR						
T2 Diabetes	ASCVD	0.05	0.56	0.10	2.26	0.05	0.63	0.09	2.09	0.04	0.69	0.09	1.96	0.04	0.74	0.08	1.84
T2 Diabetes	ASCVD+MET	0.04	0.74	0.13	2.95	0.03	0.79	0.12	2.63	0.02	0.83	0.11	2.36	0.02	0.86	0.10	2.14
T2 Diabetes	PANEL(noB)+MET	0.04	0.78	0.13	3.10	0.03	0.82	0.12	2.72	0.02	0.86	0.11	2.45	0.02	0.89	0.10	2.21
T2 Diabetes	PANEL	0.04	0.80	0.14	3.19	0.03	0.84	0.12	2.79	0.03	0.87	0.11	2.48	0.02	0.89	0.10	2.24
T2 Diabetes	PANEL+MET	0.04	0.81	0.14	3.24	0.03	0.85	0.12	2.84	0.02	0.88	0.11	2.52	0.02	0.91	0.10	2.26
Liver Disease	MET	0.05	0.46	0.09	1.86	0.04	0.52	0.09	1.73	0.04	0.57	0.08	1.62	0.04	0.61	0.08	1.54
Liver Disease	AgeSex	0.05	0.35	0.07	1.40	0.05	0.41	0.07	1.37	0.04	0.47	0.07	1.33	0.04	0.52	0.07	1.29
Liver Disease	AgeSex+MET	0.05	0.47	0.09	1.87	0.04	0.52	0.09	1.75	0.04	0.58	0.08	1.65	0.04	0.63	0.08	1.56
Liver Disease	ASCVD	0.05	0.42	0.08	1.67	0.04	0.48	0.08	1.59	0.04	0.53	0.08	1.51	0.04	0.58	0.07	1.45
Liver Disease	ASCVD+MET	0.05	0.47	0.09	1.89	0.04	0.53	0.09	1.76	0.04	0.58	0.08	1.67	0.04	0.63	0.08	1.57
Liver Disease	PANEL(noB)+MET	0.05	0.49	0.10	1.95	0.04	0.54	0.09	1.80	0.04	0.60	0.09	1.71	0.04	0.65	0.08	1.62
Liver Disease	PANEL	0.05	0.51	0.10	2.03	0.04	0.56	0.09	1.87	0.04	0.62	0.09	1.76	0.04	0.66	0.08	1.65
Liver Disease	PANEL+MET	0.05	0.51	0.10	2.05	0.04	0.56	0.09	1.88	0.04	0.62	0.09	1.76	0.04	0.66	0.08	1.66
Renal Disease	MET	0.10	0.49	0.19	1.97	0.09	0.55	0.18	1.82	0.08	0.59	0.17	1.70	0.08	0.64	0.16	1.60
Renal Disease	AgeSex	0.11	0.45	0.18	1.80	0.11	0.51	0.17	1.70	0.10	0.57	0.16	1.62	0.10	0.62	0.16	1.55
Renal Disease	AgeSex+MET	0.10	0.53	0.20	2.10	0.09	0.58	0.19	1.94	0.09	0.63	0.18	1.80	0.08	0.68	0.17	1.69
Renal Disease	ASCVD	0.11	0.49	0.19	1.97	0.10	0.55	0.18	1.83	0.09	0.60	0.17	1.72	0.09	0.65	0.16	1.63
Renal Disease	ASCVD+MET	0.10	0.53	0.20	2.13	0.09	0.59	0.19	1.96	0.09	0.64	0.18	1.83	0.08	0.68	0.17	1.71
Renal Disease	PANEL(noB)+MET	0.10	0.54	0.21	2.15	0.09	0.59	0.19	1.98	0.08	0.64	0.18	1.84	0.08	0.69	0.17	1.72
Renal Disease	PANEL	0.10	0.55	0.21	2.22	0.09	0.60	0.19	2.01	0.09	0.65	0.18	1.85	0.08	0.69	0.17	1.74
Renal Disease	PANEL+MET	0.10	0.56	0.21	2.23	0.09	0.61	0.20	2.03	0.08	0.65	0.18	1.87	0.08	0.70	0.17	1.75
PAD	MET	0.03	0.48	0.07	1.91	0.03	0.53	0.06	1.78	0.03	0.59	0.06	1.68	0.02	0.64	0.05	1.59
PAD	AgeSex	0.04	0.48	0.07	1.92	0.03	0.54	0.06	1.82	0.03	0.61	0.06	1.74	0.03	0.66	0.06	1.65
PAD	AgeSex+MET	0.04	0.53	0.07	2.11	0.03	0.59	0.07	1.96	0.03	0.64	0.06	1.83	0.03	0.69	0.06	1.72
PAD	ASCVD	0.04	0.52	0.07	2.10	0.03	0.58	0.07	1.94	0.03	0.63	0.06	1.81	0.03	0.69	0.06	1.71
PAD	ASCVD+MET	0.03	0.55	0.07	2.20	0.03	0.60	0.07	2.01	0.03	0.65	0.06	1.85	0.02	0.69	0.06	1.73
PAD	PANEL(noB)+MET	0.03	0.55	0.07	2.22	0.03	0.61	0.07	2.03	0.03	0.65	0.06	1.86	0.02	0.69	0.06	1.73
PAD	PANEL	0.03	0.55	0.07	2.20	0.03	0.61	0.07	2.02	0.03	0.66	0.06	1.89	0.03	0.71	0.06	1.77
PAD	PANEL+MET	0.03	0.56	0.08	2.24	0.03	0.62	0.07	2.05	0.03	0.66	0.06	1.88	0.02	0.70	0.06	1.76
Ven. Thrombosis	MET	0.02	0.39	0.03	1.57	0.02	0.45	0.03	1.51	0.02	0.51	0.02	1.46	0.01	0.56	0.02	1.40
Ven. Thrombosis	AgeSex	0.02	0.37	0.02	1.48	0.02	0.43	0.02	1.43	0.02	0.50	0.02	1.42	0.02	0.55	0.02	1.37
Ven. Thrombosis	AgeSex+MET	0.02	0.42	0.03	1.67	0.02	0.48	0.03	1.61	0.02	0.54	0.03	1.54	0.01	0.59	0.02	1.47
Ven. Thrombosis	ASCVD	0.02	0.39	0.03	1.58	0.02	0.45	0.03	1.50	0.02	0.51	0.02	1.46	0.02	0.56	0.02	1.40
Ven. Thrombosis	ASCVD+MET	0.02	0.42	0.03	1.68	0.02	0.48	0.03	1.60	0.02	0.55	0.03	1.56	0.01	0.60	0.03	1.49
Ven. Thrombosis	PANEL(noB)+MET	0.02	0.46	0.03	1.84	0.02	0.52	0.03	1.75	0.02	0.57	0.03	1.64	0.01	0.63	0.03	1.57
Ven. Thrombosis	PANEL	0.02	0.45	0.03	1.81	0.02	0.52	0.03	1.72	0.02	0.58	0.03	1.65	0.01	0.64	0.03	1.59
Ven. Thrombosis	PANEL+MET	0.02	0.46	0.03	1.84	0.02	0.53	0.03	1.76	0.02	0.59	0.03	1.67	0.01	0.64	0.03	1.60
AAA	MET	0.01	0.55	0.02	2.21	0.01	0.63	0.02	2.08	0.01	0.70	0.02	2.00	0.00	0.75	0.01	1.86
AAA	AgeSex	0.01	0.69	0.02	2.77	0.01	0.76	0.02	2.51	0.01	0.81	0.02	2.31	0.00	0.84	0.02	2.09
AAA	AgeSex+MET	0.01	0.70	0.02	2.82	0.01	0.76	0.02	2.53	0.00	0.81	0.02	2.32	0.00	0.84	0.02	2.09
AAA	ASCVD	0.01	0.72	0.02	2.88	0.01	0.77	0.02	2.57	0.00	0.82	0.02	2.35	0.00	0.84	0.02	2.11
AAA	ASCVD+MET	0.01	0.71	0.02	2.82	0.01	0.76	0.02	2.54	0.00	0.81	0.02	2.32	0.00	0.85	0.02	2.12
AAA	PANEL(noB)+MET	0.01	0.72	0.02	2.86	0.01	0.77	0.02	2.56	0.00	0.82	0.02	2.33	0.00	0.85	0.02	2.12
AAA	PANEL	0.01	0.72	0.02	2.88	0.01	0.78	0.02	2.59	0.00	0.82	0.02	2.33	0.00	0.85	0.02	2.12
AAA	PANEL+MET	0.01	0.72	0.02	2.87	0.01	0.77	0.02	2.56	0.00	0.81	0.02	2.32	0.00	0.85	0.02	2.12
COPD	MET	0.09	0.45	0.13	1.79	0.08	0.50	0.13	1.68	0.07	0.56	0.12	1.59	0.07	0.61	0.12	1.52
COPD	AgeSex	0.09	0.38	0.11	1.51	0.09	0.44	0.11	1.47	0.08	0.50	0.11	1.43	0.08	0.56	0.11	1.39
COPD	AgeSex+MET	0.09	0.47	0.14	1.86	0.08	0.53	0.13	1.77	0.07	0.58	0.13	1.67	0.07	0.63	0.12	1.57
COPD	ASCVD	0.08	0.46	0.14	1.83	0.08	0.52	0.13	1.73	0.07	0.57	0.12	1.62	0.07	0.62	0.12	1.55
COPD	ASCVD+MET	0.08	0.50	0.15	1.99	0.08	0.55	0.14	1.85	0.07	0.61	0.13	1.74	0.07	0.65	0.12	1.63
COPD	PANEL(noB)+MET	0.08	0.51	0.15	2.04	0.08	0.56	0.14	1.88	0.07	0.61	0.13	1.75	0.07	0.66	0.12	1.64
COPD	PANEL	0.08	0.51	0.15	2.04	0.08	0.56	0.14	1.88	0.07	0.61	0.13	1.75	0.07	0.66	0.12	1.65
COPD	PANEL+MET	0.08	0.52	0.15	2.07	0.08	0.57	0.14	1.90	0.07	0.62	0.13	1.76	0.06	0.67	0.12	1.66
Asthma	MET	0.04	0.36	0.05	1.45	0.04	0.42	0.05	1.41	0.04	0.47	0.05	1.35	0.03	0.53	0.05	1.33
Asthma	AgeSex	0.04	0.31	0.05	1.23	0.04	0.37	0.05	1.22	0.04	0.42	0.05	1.20	0.04	0.48	0.04	1.19
Asthma	AgeSex+MET	0.04	0.37	0.06	1.49	0.04	0.44	0.05	1.45	0.04	0.49	0.05	1.40	0.03	0.53	0.05	1.33
Asthma	ASCVD	0.04	0.35	0.05	1.40	0.04	0.40	0.05	1.35	0.04	0.46	0.05	1.31	0.03	0.50	0.05	1.25
Asthma	ASCVD+MET	0.04	0.38	0.06	1.53	0.04	0.43	0.05	1.44	0.04	0.48	0.05	1.38	0.03	0.54	0.05	1.36
Asthma	PANEL(noB)+MET	0.04	0.41	0.06	1.63	0.04	0.47	0.06	1.55	0.04	0.52	0.06	1.48	0.03	0.56	0.05	1.41
Asthma	PANEL	0.04	0.41	0.06	1.63	0.04	0.46	0.06	1.54	0.03	0.52	0.06	1.49	0.03	0.57	0.05	1.43
Asthma	PANEL+MET	0.04	0.41	0.06	1.65	0.04	0.47	0.06	1.55	0.04	0.52	0.06	1.49	0.03	0.57	0.05	1.42
Parkinson's	MET	0.01	0.43	0.01	1.74	0.00	0.49	0.01	1.63	0.00	0.54	0.01	1.56	0.00	0.58	0.01	1.45
Parkinson's	AgeSex	0.01	0.59	0.01	2.35	0.01	0.67	0.01	2.24	0.00	0.75	0.01	2.14	0.00	0.78	0.01	1.96
Parkinson's	AgeSex+MET	0.01	0.58	0.01	2.32	0.01	0.67	0.01	2.22	0.00	0.72	0.01	2.06	0.00	0.78	0.01	1.96
Parkinson's	ASCVD	0.01	0.59	0.01	2.37	0.01	0.67	0.01	2.22	0.00	0.74	0.01	2.12	0.00	0.79	0.01	1.96

3 Metabolomic profiles predict individual multi-disease outcomes

Table 13 continued from previous page

		25% FPR				30% FPR				35% FPR				40% FPR			
Parkinson's	ASCVD+MET	0.01	0.59	0.01	2.34	0.01	0.68	0.01	2.26	0.00	0.72	0.01	2.07	0.00	0.78	0.01	1.95
Parkinson's	PANEL(noB)+MET	0.01	0.59	0.01	2.38	0.01	0.66	0.01	2.21	0.00	0.72	0.01	2.06	0.00	0.77	0.01	1.92
Parkinson's	PANEL	0.01	0.60	0.01	2.39	0.01	0.66	0.01	2.20	0.00	0.71	0.01	2.04	0.00	0.76	0.01	1.90
Parkinson's	PANEL+MET	0.01	0.60	0.01	2.38	0.00	0.66	0.01	2.18	0.00	0.70	0.01	2.01	0.00	0.75	0.01	1.88
Cataracts	MET	0.10	0.42	0.15	1.70	0.09	0.49	0.15	1.64	0.08	0.55	0.14	1.57	0.08	0.60	0.14	1.50
Cataracts	AgeSex	0.10	0.58	0.20	2.32	0.09	0.64	0.19	2.14	0.08	0.71	0.18	2.02	0.07	0.76	0.17	1.89
Cataracts	AgeSex+MET	0.10	0.58	0.20	2.33	0.09	0.65	0.19	2.16	0.08	0.71	0.18	2.03	0.07	0.77	0.17	1.91
Cataracts	ASCVD	0.10	0.59	0.20	2.37	0.09	0.66	0.19	2.19	0.08	0.72	0.18	2.05	0.07	0.76	0.17	1.91
Cataracts	ASCVD+MET	0.10	0.59	0.20	2.36	0.09	0.66	0.19	2.19	0.08	0.71	0.18	2.04	0.07	0.77	0.17	1.92
Cataracts	PANEL(noB)+MET	0.10	0.59	0.20	2.36	0.09	0.66	0.19	2.19	0.08	0.72	0.18	2.05	0.07	0.77	0.17	1.92
Cataracts	PANEL	0.10	0.59	0.20	2.38	0.09	0.66	0.19	2.20	0.08	0.72	0.18	2.07	0.07	0.77	0.17	1.92
Cataracts	PANEL+MET	0.10	0.59	0.20	2.37	0.09	0.66	0.19	2.20	0.08	0.72	0.18	2.06	0.07	0.77	0.17	1.93
Glaucoma	MET	0.02	0.32	0.03	1.30	0.02	0.38	0.03	1.28	0.02	0.44	0.03	1.26	0.02	0.50	0.03	1.24
Glaucoma	AgeSex	0.03	0.44	0.04	1.77	0.02	0.51	0.04	1.69	0.02	0.56	0.04	1.60	0.02	0.63	0.04	1.57
Glaucoma	AgeSex+MET	0.03	0.44	0.04	1.74	0.02	0.50	0.04	1.66	0.02	0.57	0.04	1.62	0.02	0.62	0.04	1.56
Glaucoma	ASCVD	0.03	0.45	0.04	1.79	0.02	0.52	0.04	1.73	0.02	0.58	0.04	1.65	0.02	0.63	0.04	1.58
Glaucoma	ASCVD+MET	0.03	0.45	0.04	1.78	0.02	0.51	0.04	1.69	0.02	0.57	0.04	1.62	0.02	0.63	0.04	1.57
Glaucoma	PANEL(noB)+MET	0.03	0.45	0.04	1.79	0.02	0.50	0.04	1.68	0.02	0.57	0.04	1.62	0.02	0.63	0.04	1.56
Glaucoma	PANEL	0.03	0.44	0.04	1.75	0.02	0.50	0.04	1.68	0.02	0.57	0.04	1.62	0.02	0.63	0.04	1.57
Glaucoma	PANEL+MET	0.03	0.44	0.04	1.75	0.02	0.50	0.04	1.67	0.02	0.56	0.04	1.61	0.02	0.62	0.04	1.54
Fractures	MET	0.08	0.35	0.10	1.41	0.07	0.41	0.10	1.37	0.07	0.47	0.10	1.34	0.07	0.52	0.09	1.30
Fractures	AgeSex	0.08	0.39	0.11	1.56	0.07	0.44	0.11	1.48	0.07	0.50	0.10	1.42	0.07	0.55	0.10	1.37
Fractures	AgeSex+MET	0.08	0.39	0.11	1.57	0.07	0.46	0.11	1.52	0.07	0.51	0.10	1.45	0.07	0.56	0.10	1.41
Fractures	ASCVD	0.08	0.39	0.11	1.55	0.07	0.45	0.11	1.49	0.07	0.50	0.10	1.43	0.07	0.55	0.10	1.38
Fractures	ASCVD+MET	0.08	0.39	0.11	1.58	0.07	0.45	0.11	1.51	0.07	0.51	0.11	1.46	0.07	0.56	0.10	1.41
Fractures	PANEL(noB)+MET	0.08	0.39	0.11	1.56	0.07	0.46	0.11	1.53	0.07	0.51	0.11	1.46	0.07	0.56	0.10	1.41
Fractures	PANEL	0.08	0.40	0.11	1.61	0.07	0.46	0.11	1.54	0.07	0.52	0.11	1.48	0.07	0.57	0.10	1.43
Fractures	PANEL+MET	0.08	0.40	0.11	1.61	0.07	0.46	0.11	1.54	0.07	0.52	0.11	1.48	0.07	0.57	0.10	1.43
Lung Cancer	MET	0.01	0.54	0.02	2.17	0.01	0.60	0.02	2.00	0.01	0.65	0.02	1.85	0.01	0.69	0.02	1.71
Lung Cancer	AgeSex	0.01	0.48	0.02	1.92	0.01	0.55	0.02	1.84	0.01	0.61	0.02	1.73	0.01	0.68	0.02	1.70
Lung Cancer	AgeSex+MET	0.01	0.59	0.03	2.36	0.01	0.66	0.02	2.19	0.01	0.70	0.02	2.01	0.01	0.77	0.02	1.92
Lung Cancer	ASCVD	0.01	0.65	0.03	2.60	0.01	0.71	0.02	2.35	0.01	0.75	0.02	2.15	0.01	0.79	0.02	1.97
Lung Cancer	ASCVD+MET	0.01	0.67	0.03	2.67	0.01	0.72	0.03	2.39	0.01	0.77	0.02	2.20	0.01	0.81	0.02	2.02
Lung Cancer	PANEL(noB)+MET	0.01	0.69	0.03	2.75	0.01	0.73	0.03	2.45	0.01	0.78	0.02	2.23	0.01	0.81	0.02	2.03
Lung Cancer	PANEL	0.01	0.69	0.03	2.77	0.01	0.74	0.03	2.46	0.01	0.78	0.02	2.24	0.01	0.82	0.02	2.06
Lung Cancer	PANEL+MET	0.01	0.69	0.03	2.76	0.01	0.74	0.03	2.46	0.01	0.78	0.02	2.23	0.01	0.81	0.02	2.04
Skin Cancer	MET	0.04	0.34	0.06	1.36	0.04	0.40	0.06	1.33	0.04	0.46	0.06	1.32	0.04	0.51	0.05	1.28
Skin Cancer	AgeSex	0.05	0.46	0.08	1.85	0.04	0.53	0.07	1.76	0.04	0.59	0.07	1.68	0.04	0.64	0.07	1.60
Skin Cancer	AgeSex+MET	0.05	0.46	0.08	1.84	0.04	0.53	0.07	1.75	0.04	0.59	0.07	1.68	0.04	0.64	0.07	1.60
Skin Cancer	ASCVD	0.05	0.46	0.08	1.84	0.04	0.53	0.07	1.76	0.04	0.59	0.07	1.69	0.04	0.65	0.07	1.62
Skin Cancer	ASCVD+MET	0.05	0.46	0.08	1.84	0.04	0.53	0.07	1.76	0.04	0.59	0.07	1.68	0.04	0.64	0.07	1.61
Skin Cancer	PANEL(noB)+MET	0.05	0.47	0.08	1.88	0.04	0.53	0.07	1.77	0.04	0.59	0.07	1.70	0.04	0.65	0.07	1.63
Skin Cancer	PANEL	0.05	0.47	0.08	1.89	0.04	0.53	0.07	1.77	0.04	0.59	0.07	1.70	0.04	0.65	0.07	1.63
Skin Cancer	PANEL+MET	0.05	0.46	0.08	1.86	0.04	0.53	0.07	1.76	0.04	0.59	0.07	1.70	0.04	0.65	0.07	1.63
Colon Cancer	MET	0.01	0.39	0.02	1.56	0.01	0.45	0.02	1.50	0.01	0.50	0.01	1.42	0.01	0.55	0.01	1.38
Colon Cancer	AgeSex	0.01	0.48	0.02	1.90	0.01	0.55	0.02	1.83	0.01	0.60	0.02	1.72	0.01	0.66	0.02	1.65
Colon Cancer	AgeSex+MET	0.01	0.48	0.02	1.92	0.01	0.55	0.02	1.83	0.01	0.61	0.02	1.73	0.01	0.66	0.02	1.66
Colon Cancer	ASCVD	0.01	0.49	0.02	1.95	0.01	0.55	0.02	1.84	0.01	0.60	0.02	1.72	0.01	0.66	0.02	1.66
Colon Cancer	ASCVD+MET	0.01	0.48	0.02	1.90	0.01	0.55	0.02	1.82	0.01	0.60	0.02	1.72	0.01	0.66	0.02	1.65
Colon Cancer	PANEL(noB)+MET	0.01	0.48	0.02	1.91	0.01	0.55	0.02	1.82	0.01	0.61	0.02	1.73	0.01	0.67	0.02	1.67
Colon Cancer	PANEL	0.01	0.48	0.02	1.92	0.01	0.56	0.02	1.87	0.01	0.62	0.02	1.76	0.01	0.67	0.02	1.67
Colon Cancer	PANEL+MET	0.01	0.47	0.02	1.90	0.01	0.55	0.02	1.83	0.01	0.60	0.02	1.72	0.01	0.66	0.02	1.64
Rectal Cancer	MET	0.01	0.38	0.01	1.54	0.01	0.44	0.01	1.48	0.01	0.51	0.01	1.46	0.01	0.57	0.01	1.42
Rectal Cancer	AgeSex	0.01	0.49	0.01	1.98	0.01	0.56	0.01	1.86	0.01	0.62	0.01	1.78	0.01	0.68	0.01	1.69
Rectal Cancer	AgeSex+MET	0.01	0.52	0.01	2.07	0.01	0.56	0.01	1.88	0.01	0.61	0.01	1.75	0.01	0.67	0.01	1.68
Rectal Cancer	ASCVD	0.01	0.49	0.01	1.96	0.01	0.55	0.01	1.84	0.01	0.63	0.01	1.80	0.01	0.68	0.01	1.70
Rectal Cancer	ASCVD+MET	0.01	0.52	0.01	2.07	0.01	0.56	0.01	1.86	0.01	0.61	0.01	1.75	0.01	0.67	0.01	1.67
Rectal Cancer	PANEL(noB)+MET	0.01	0.51	0.01	2.03	0.01	0.56	0.01	1.88	0.01	0.62	0.01	1.77	0.01	0.67	0.01	1.68
Rectal Cancer	PANEL	0.01	0.49	0.01	1.96	0.01	0.54	0.01	1.80	0.01	0.61	0.01	1.75	0.01	0.68	0.01	1.70
Rectal Cancer	PANEL+MET	0.01	0.49	0.01	1.98	0.01	0.54	0.01	1.82	0.01	0.61	0.01	1.74	0.01	0.67	0.01	1.67
Prostate Cancer	MET	0.04	0.29	0.05	1.14	0.04	0.34	0.06	1.14	0.04	0.41	0.06	1.16	0.04	0.46	0.06	1.16
Prostate Cancer	AgeSex	0.05	0.44	0.08	1.76	0.05	0.52	0.08	1.72	0.05	0.59	0.08	1.69	0.04	0.65	0.08	1.63
Prostate Cancer	AgeSex+MET	0.05	0.44	0.08	1.78	0.05	0.52	0.08	1.72	0.05	0.59	0.08	1.69	0.04	0.65	0.08	1.64
Prostate Cancer	ASCVD	0.05	0.45	0.08	1.78	0.05	0.52	0.08	1.73	0.05	0.59	0.08	1.70	0.04	0.66	0.08	1.64
Prostate Cancer	ASCVD+MET	0.05	0.45	0.08	1.79	0.05	0.52	0.08	1.73	0.05	0.59	0.08	1.70	0.04	0.66	0.08	1.65

Supplementary Tables

Table 13 continued from previous page

		25% FPR				30% FPR				35% FPR				40% FPR			
Prostate Cancer	PANEL(noB)+MET	0.05	0.46	0.08	1.82	0.05	0.53	0.08	1.78	0.05	0.60	0.08	1.70	0.04	0.66	0.08	1.65
Prostate Cancer	PANEL	0.05	0.45	0.08	1.80	0.05	0.53	0.08	1.75	0.04	0.60	0.08	1.72	0.04	0.67	0.08	1.67
Prostate Cancer	PANEL+MET	0.05	0.45	0.08	1.80	0.05	0.53	0.08	1.75	0.04	0.60	0.08	1.72	0.04	0.67	0.08	1.67
Breast Cancer	MET	0.04	0.30	0.05	1.19	0.03	0.34	0.05	1.15	0.03	0.39	0.04	1.12	0.03	0.44	0.04	1.09
Breast Cancer	AgeSex	0.04	0.28	0.04	1.13	0.04	0.34	0.04	1.13	0.04	0.40	0.04	1.13	0.04	0.46	0.05	1.14
Breast Cancer	AgeSex+MET	0.04	0.29	0.05	1.15	0.04	0.34	0.04	1.13	0.04	0.40	0.05	1.14	0.03	0.46	0.05	1.14
Breast Cancer	ASCVD	0.04	0.27	0.04	1.10	0.04	0.34	0.04	1.12	0.04	0.40	0.05	1.14	0.03	0.45	0.05	1.14
Breast Cancer	ASCVD+MET	0.04	0.28	0.04	1.13	0.04	0.34	0.04	1.13	0.04	0.40	0.04	1.13	0.03	0.46	0.05	1.14
Breast Cancer	PANEL(noB)+MET	0.04	0.33	0.05	1.30	0.04	0.38	0.05	1.27	0.04	0.44	0.05	1.25	0.03	0.48	0.05	1.20
Breast Cancer	PANEL	0.04	0.32	0.05	1.28	0.04	0.38	0.05	1.25	0.04	0.43	0.05	1.22	0.03	0.48	0.05	1.21
Breast Cancer	PANEL+MET	0.04	0.32	0.05	1.29	0.04	0.38	0.05	1.25	0.04	0.43	0.05	1.23	0.03	0.48	0.05	1.20

3 Metabolomic profiles predict individual multi-disease outcomes

Table 14: Top 25 metabolites per endpoint.

	MACE	CHD	Cerebral Stroke	Dementia	Heart Failure	Atrial Fibrillation	T2 Diabetes	Liver Disease
1	Creatinine	Creatinine	Creatinine	LA	LA	Creatinine	Glucose	GlycA
2	Albumin	Albumin	LA	Albumin	Albumin	LA	Lactate	LA
3	Gly	Gly	Albumin	Gln	Creatinine	Gly	LA	Albumin
4	LA	LA	Gln	GlycA	GlycA	Albumin	Tyr	MUFA
5	Gln	Gln	Gly	Creatinine	Tyr	Tyr	Albumin	Tyr
6	GlycA	GlycA	XXL-VLDL-CE	Leu	Gly	Gln	Gly	SFA
7	XL-HDL-FC	XXL-VLDL-CE	XL-HDL-FC	Glucose	XXL-VLDL-CE	XXL-VLDL-CE	GlycA	XXL-VLDL-CE
8	Tyr	XL-HDL-FC	SFA	MUFA	MUFA	Acetoacetate	Leu	Leu
9	XXL-VLDL-CE	MUFA	MUFA	SFA	Glucose	SFA	DHA	Gly
10	SFA	SFA	Tyr	XXL-VLDL-CE	SFA	Acetone	L-LDL-TG	XL-HDL-FC
11	Acetoacetate	Tyr	S-LDL-PL	Tyr	Gln	MUFA	S-LDL-PL	XS-VLDL-FC
12	MUFA	Acetoacetate	Acetoacetate	XL-HDL-FC	Leu	Leu	SFA	Glucose
13	S-LDL-PL	S-LDL-PL	GlycA	Total BCAA	XL-HDL-FC	XL-HDL-FC	XL-HDL-FC	L-LDL-TG
14	IDL-FC	Glucose	Glucose	Acetoacetate	XS-VLDL-FC	Lactate	LDL-TG	Total-FA
15	Leu	L-LDL-TG	Acetone	S-LDL-PL	S-LDL-PL	IDL-FC	XXL-VLDL-CE	Lactate
16	Acetone	Total-FA	Leu	XS-VLDL-FC	Acetoacetate	LDL size	MUFA	LDL-TG
17	Lactate	DHA	His	Val	Acetone	S-LDL-PL	XS-VLDL-FC	DHA
18	Glucose	IDL-FC	LDL size	Acetone	DHA	Glucose	Total-FA	Creatinine
19	XS-VLDL-FC	M-VLDL-CE	Lactate	Citrate	L-LDL-TG	Ile	His	Sphingomyelins
20	His	LDL-TG	IDL-FC	boHbutyrate	IDL-FC	His	Unsaturation	M-VLDL-CE
21	DHA	Acetone	M-VLDL-CE	Total-FA	Total-FA	M-VLDL-CE	M-HDL-TG	S-LDL-PL
22	M-VLDL-CE	XS-VLDL-FC	Total-FA	L-LDL-TG	Lactate	XS-VLDL-FC	M-VLDL-CE	IDL-FC
23	Total-FA	Leu	XS-VLDL-FC	IDL-FC	His	Citrate	L-HDL-TG	XS-VLDL-PL
24	Cholines	Lactate	S-LDL-CE	LDL-TG	LDL-TG	Cholines	VLDL-CE	S-LDL-CE
25	L-LDL-TG	His	XXL-VLDL-C	Sphingomyelins	S-LDL-CE	S-LDL-CE	XS-VLDL-PL	Acetoacetate

	Renal Disease	PAD	Ven. Thrombosis	AAA	COPD	Asthma	Parkinson's	Cataracts
1	Creatinine	Creatinine	Albumin	Creatinine	GlycA	GlycA	Creatinine	Albumin
2	Albumin	LA	LA	Gln	Albumin	Albumin	LA	Tyr
3	LA	Gln	Creatinine	Gly	LA	LA	Gln	Gln
4	GlycA	Albumin	GlycA	LA	MUFA	Tyr	Gly	LA
5	MUFA	GlycA	Tyr	Tyr	Gln	MUFA	Acetoacetate	Leu
6	XXL-VLDL-CE	XXL-VLDL-CE	Gln	XXL-VLDL-CE	Tyr	SFA	XXL-VLDL-CE	XXL-VLDL-CE
7	SFA	Gly	Gly	Acetoacetate	DHA	Leu	Citrate	Citrate
8	Leu	Tyr	XXL-VLDL-CE	XL-HDL-FC	SFA	XXL-VLDL-CE	Albumin	Glucose
9	Glucose	Acetoacetate	SFA	Albumin	XXL-VLDL-CE	DHA	XL-HDL-FC	GlycA
10	Tyr	MUFA	MUFA	SFA	S-LDL-PL	Gln	Acetone	Unsaturation
11	His	SFA	His	Acetone	Total-FA	Total-FA	Tyr	IDL-FC
12	XS-VLDL-FC	XL-HDL-FC	XL-HDL-FC	S-LDL-PL	Sphingomyelins	S-LDL-PL	SFA	SFA
13	Total-FA	Leu	IDL-FC	Lactate	XL-HDL-FC	XS-VLDL-FC	Glucose	MUFA
14	XL-HDL-FC	S-LDL-PL	Leu	MUFA	Creatinine	His	LDL size	His
15	IDL-FC	Lactate	DHA	LDL size	XS-VLDL-FC	Glucose	S-LDL-PL	Total-FA
16	Sphingomyelins	XS-VLDL-FC	Acetoacetate	Cholines	Leu	XL-HDL-FC	Leu	Omega-3
17	Citrate	Acetone	LDL size	IDL-FC	Acetoacetate	Sphingomyelins	MUFA	S-LDL-PL
18	S-LDL-CE	S-LDL-CE	XS-VLDL-FC	IDL-P	Lactate	Total BCAA	IDL-FC	XS-VLDL-FC
19	LDL size	XXL-VLDL-C	S-LDL-PL	DHA	L-LDL-TG	L-LDL-TG	IDL-P	XL-HDL-FC
20	M-VLDL-CE	DHA	Total-FA	M-VLDL-CE	IDL-FC	LDL-TG	XS-VLDL-FC	M-VLDL-CE
21	DHA	IDL-P	S-LDL-TG	LDL-TG	XXL-VLDL-C	Ala	Lactate	XXL-VLDL-C
22	Total BCAA	M-VLDL-CE	M-VLDL-CE	M-HDL-PL	LDL-TG	XS-VLDL-PL	S-LDL-CE	LDL size
23	Acetoacetate	Total BCAA	Sphingomyelins	S-VLDL-TG	Total BCAA	IDL-FC	Cholines	Acetoacetate
24	Gln	Total-FA	S-LDL-CE	Leu	Val	Lactate	M-VLDL-CE	Creatinine
25	XXL-VLDL-C	LDL size	Pyruvate	L-LDL-TG	XS-VLDL-PL	Phe	XXL-VLDL-C	Total BCAA

	Glaucoma	Fractures	Lung Cancer	Skin Cancer	Colon Cancer	Rectal Cancer	Prostate Cancer	Breast Cancer
--	-----------------	------------------	--------------------	--------------------	---------------------	----------------------	------------------------	----------------------

Supplementary Tables

Table 14 continued from previous page

	Glaucoma	Fractures	Lung Cancer	Skin Cancer	Colon Cancer	Rectal Cancer	Prostate Cancer	Breast Cancer
1	Albumin	Creatinine	GlycA	Creatinine	LA	Creatinine	LA	LA
2	Glucose	Gly	LA	Albumin	Albumin	LA	Albumin	Albumin
3	Creatinine	Glucose	Gln	LA	Creatinine	Albumin	Creatinine	Creatinine
4	XXL-VLDL-CE	GlycA	Albumin	Gln	XXL-VLDL-CE	Gln	Gln	Gln
5	LA	Albumin	MUFA	XXL-VLDL-CE	Gln	XXL-VLDL-CE	GlycA	GlycA
6	Tyr	Lactate	XXL-VLDL-CE	Gly	GlycA	Gly	XXL-VLDL-CE	XXL-VLDL-CE
7	Gly	Val	XL-HDL-FC	Omega-3	Leu	Tyr	Gly	Gly
8	Gln	LA	SFA	Tyr	Tyr	Acetoacetate	Tyr	Tyr
9	XL-HDL-FC	Tyr	S-LDL-PL	S-LDL-PL	MUFA	XL-HDL-FC	MUFA	MUFA
10	Unsaturation	MUFA	Lactate	XL-HDL-FC	S-LDL-PL	S-LDL-PL	SFA	SFA
11	SFA	Total BCAA	Sphingomyelins	Citrate	SFA	LDL size	Acetoacetate	Acetoacetate
12	Leu	Leu	DHA	Acetone	XL-HDL-FC	SFA	S-LDL-PL	S-LDL-PL
13	Acetoacetate	Unsaturation	L-LDL-TG	LDL size	Acetoacetate	GlycA	XL-HDL-FC	XL-HDL-FC
14	GlycA	XS-VLDL-PL	Val	SFA	LDL size	Leu	XS-VLDL-FC	XS-VLDL-FC
15	XXL-VLDL-C	XS-VLDL-FC	Acetoacetate	Leu	Gly	MUFA	IDL-FC	IDL-FC
16	IDL-FC	bOHbutyrate	XS-VLDL-FC	XXL-VLDL-C	Acetone	Acetone	Lactate	Lactate
17	MUFA	His	Leu	Unsaturation	Total-FA	IDL-FC	S-LDL-CE	S-LDL-CE
18	S-LDL-PL	XS-VLDL-CE	LDL-TG	Acetoacetate	XXL-VLDL-C	Total-FA	His	His
19	Citrate	M-HDL-C	Total BCAA	IDL-FC	Total BCAA	M-VLDL-CE	Leu	Leu
20	Total-FA	SFA	Total-FA	His	Lactate	XXL-VLDL-C	Glucose	Glucose
21	LDL size	Citrate	XXL-VLDL-C	M-VLDL-CE	M-VLDL-CE	His	Total-FA	Total-FA
22	XS-VLDL-FC	XL-HDL-TG	Creatinine	Cholines	IDL-FC	Cholines	Acetone	Acetone
23	M-VLDL-CE	Acetoacetate	XS-VLDL-PL	L-VLDL-TG	XS-VLDL-FC	Lactate	L-LDL-TG	L-LDL-TG
24	Pyruvate	XXL-VLDL-CE	IDL-FC	S-VLDL-TG	Glucose	S-HDL-P	XL-VLDL-CE	XL-VLDL-CE
25	Omega-3	S-HDL-PL	Acetone	IDL-PL	S-LDL-CE	S-LDL-CE	Cholines	Cholines

4 Medical history predicts future health trajectories over the human phenome

Supplementary Tables

Table 15: Incident event counts per 10% Quantile and absolute incident event count for all endpoints.

Endpoint	1	2	3	4	5	6	7	8	9	10	Total
OMOP_4306655	162	309	480	706	1205	1736	2688	4837	8200	17374	37697
phecode_001	13	12	21	23	27	28	45	59	61	88	377
phecode_002	167	226	296	374	408	477	589	698	978	2374	6587
phecode_002-1	133	164	223	262	327	308	390	531	707	1864	4909
phecode_003	394	504	604	574	680	761	838	1060	1414	2858	9687
phecode_004	215	263	297	342	381	419	387	456	689	1398	4847
phecode_004-1	33	29	45	55	63	59	83	94	108	294	863
phecode_004-2	15	17	21	19	22	28	41	26	40	103	332
phecode_004-3	7	16	12	15	22	24	37	45	40	176	394
phecode_005	26	46	36	42	49	66	68	79	119	256	787
phecode_005-1	24	21	22	37	35	43	63	56	81	211	593
phecode_005-2	7	9	11	7	13	10	20	23	32	82	214
phecode_007	21	20	33	38	46	69	56	74	121	449	927
phecode_007-1	20	20	30	33	48	65	58	68	122	443	907
phecode_008	12	41	158	219	242	313	389	437	554	888	3253
phecode_009	41	38	56	58	94	129	180	203	286	990	2075
phecode_010	4	15	32	36	30	34	43	65	87	222	568
phecode_011	28	26	38	55	64	95	141	143	200	539	1329
phecode_012	13	14	16	15	30	54	61	63	134	385	785
phecode_015	66	60	103	109	133	160	154	152	283	619	1839
phecode_015-2	52	45	82	89	111	142	145	148	251	597	1662
phecode_019	4	3	4	1	9	16	13	21	17	82	170
phecode_020	12	14	50	42	69	59	78	102	123	165	714
phecode_020-1	12	14	50	41	68	57	83	98	127	161	711
phecode_024	2	4	26	21	41	32	32	52	53	66	329
phecode_025	22	17	27	31	42	60	72	77	127	273	748
phecode_030	56	99	164	234	282	346	466	532	665	1013	3857
phecode_050	8	13	19	20	31	39	30	27	31	54	272
phecode_050-4	1	2	10	10	20	23	22	31	34	28	181
phecode_052	207	455	2100	2743	3124	3523	3881	4470	4935	6351	31789
phecode_052-1	64	141	286	419	496	579	753	936	1214	1948	6836
phecode_052-3	130	368	1800	2375	2804	3048	3456	3723	4236	5290	27230
phecode_052-31	8	18	45	55	70	67	95	74	73	131	636
phecode_052-32	126	386	1762	2316	2772	3041	3424	3721	4216	5293	27057
phecode_052-4	10	15	10	19	20	19	23	22	38	42	218
phecode_052-5	15	15	16	13	14	16	15	29	36	92	261
phecode_054	45	61	61	60	69	91	74	89	94	268	912
phecode_054-2	6	15	16	11	19	18	35	19	40	138	317
phecode_054-3	6	10	15	12	17	20	19	17	26	138	280
phecode_054-31	5	13	13	14	14	20	15	17	25	147	283
phecode_054-5	8	7	9	10	8	16	8	8	15	15	104
phecode_055	7	9	20	29	38	41	52	64	70	98	428
phecode_055-1	7	8	19	30	33	36	51	60	66	96	406
phecode_056	79	183	1106	1450	1814	1950	2269	2648	3032	4029	18560
phecode_056-1	12	61	332	436	577	650	763	950	1025	1459	6265
phecode_057	0	4	3	7	5	12	21	24	31	90	197
phecode_057-1	0	3	3	8	2	11	19	25	29	90	190
phecode_058	6	15	13	12	15	24	24	20	51	103	283
phecode_058-1	6	12	13	10	16	20	21	21	41	98	258
phecode_059	240	262	370	515	591	748	930	1092	1338	2041	8127
phecode_059-1	226	259	348	491	586	710	898	1092	1288	2013	7911
phecode_060	10	13	7	9	14	9	12	11	19	27	131
phecode_061	147	182	202	246	272	301	372	436	603	1105	3866
phecode_066	2	1	12	11	11	12	14	20	27	42	152
phecode_069	5	9	8	15	9	15	13	20	28	70	192
phecode_070	524	747	1142	1486	1835	2218	2851	3599	4699	8043	27144
phecode_074	7	8	11	7	17	28	27	28	41	237	411
phecode_076	2	5	8	4	9	20	20	18	19	80	185
phecode_084	26	39	108	158	168	199	271	321	334	592	2216
phecode_084-2	6	3	12	10	18	23	23	29	50	132	306
phecode_084-4	0	0	1	1	3	12	13	38	82	229	379
phecode_084-6	13	13	30	44	44	62	67	68	82	119	542
phecode_084-7	7	7	19	32	32	49	56	64	60	83	409
phecode_086	11	33	122	131	157	179	198	265	276	549	1921

4 Medical history predicts future health trajectories over the human phenome

Table 15 continued from previous page

Endpoint	1	2	3	4	5	6	7	8	9	10	Total
phecode_088	21	30	98	117	112	155	222	292	423	730	2200
phecode_089	2368	2827	5851	8417	9854	11100	12058	13483	15011	18377	99346
phecode_089-1	1817	2186	2415	2486	2899	3253	3813	4561	5432	9045	37907
phecode_089-2	988	1289	4107	5970	6867	7830	8476	9624	11040	13726	69917
phecode_089-3	644	1121	3482	4706	5591	6298	7150	8074	9572	12941	59579
phecode_091	12	12	14	10	10	16	25	33	47	119	298
phecode_092	373	541	664	874	987	1273	1679	1946	2641	5573	16551
phecode_092-1	12	21	31	37	39	47	64	80	121	319	771
phecode_092-2	362	534	668	844	973	1247	1657	1920	2601	5484	16290
phecode_092-8	0	1	1	9	6	10	12	17	21	55	132
phecode_095	70	88	88	129	150	152	196	258	374	750	2255
phecode_096	39	75	162	226	257	287	328	378	528	891	3171
phecode_097	3	10	40	61	60	67	118	133	221	514	1227
phecode_097-1	3	7	35	45	54	66	100	118	212	454	1094
phecode_098	71	108	133	150	200	231	265	296	478	1184	3116
phecode_098-2	0	2	9	3	10	8	7	17	28	70	154
phecode_099	43	93	241	358	461	639	751	1044	1429	2338	7397
phecode_100	87	108	123	172	170	182	246	272	313	609	2282
phecode_100-1	24	46	40	54	70	67	75	66	98	206	746
phecode_100-12	21	24	28	34	59	42	43	41	67	122	481
phecode_100-2	15	12	20	16	35	44	52	49	63	128	434
phecode_100-5	7	7	6	10	14	8	9	13	14	17	105
phecode_100-6	2	5	5	11	11	20	29	47	68	163	361
phecode_100-7	4	3	4	4	7	6	7	10	16	67	128
phecode_100-8	5	8	1	8	9	4	10	13	14	29	101
phecode_100-9	17	18	16	22	13	19	21	22	18	37	203
phecode_101	463	616	867	1007	1248	1424	1729	2188	2407	3586	15535
phecode_101-1	15	26	51	62	93	116	148	257	306	580	1654
phecode_101-2	16	46	48	81	83	117	137	206	240	433	1407
phecode_101-21	7	10	13	23	36	48	64	105	123	218	647
phecode_101-3	15	18	31	34	32	60	48	54	67	132	491
phecode_101-4	276	365	522	615	709	822	1002	1105	1193	1442	8051
phecode_101-41	197	265	380	445	537	633	753	829	887	1124	6050
phecode_101-42	110	126	176	209	222	257	296	356	393	511	2656
phecode_101-5	13	24	23	24	34	39	32	40	59	94	382
phecode_101-6	19	23	26	39	52	65	91	146	213	425	1099
phecode_101-61	7	7	10	6	9	12	34	60	78	277	500
phecode_101-62	10	9	26	22	26	25	48	88	127	185	566
phecode_101-7	12	16	19	20	28	47	55	76	96	141	510
phecode_101-71	2	1	7	2	14	14	20	23	37	61	181
phecode_101-8	32	31	57	70	121	123	188	266	411	609	1908
phecode_102	84	116	149	201	308	397	587	941	1229	2798	6810
phecode_102-1	61	84	116	151	232	317	492	810	1119	2450	5832
phecode_102-3	7	6	10	14	21	12	39	41	43	120	313
phecode_102-5	5	5	6	12	15	16	16	18	33	62	188
phecode_103	1044	1271	1583	1940	2334	2838	3502	4041	4966	7522	31041
phecode_103-1	216	309	329	385	396	445	511	576	713	1122	5002
phecode_103-2	31	122	562	868	1099	1501	1824	2243	3084	5306	16640
phecode_103-21	27	100	484	738	951	1329	1596	1980	2759	4781	14745
phecode_103-22	7	25	51	123	130	192	249	363	473	1018	2631
phecode_103-3	56	86	130	204	265	336	511	702	867	1695	4852
phecode_104	112	124	174	173	192	196	250	245	282	417	2165
phecode_104-1	20	25	38	28	41	42	33	47	65	92	431
phecode_104-2	9	11	10	18	22	48	54	87	84	136	479
phecode_104-3	51	50	79	75	79	93	101	121	135	202	986
phecode_104-5	0	1	9	13	18	15	21	23	25	34	159
phecode_105	34	40	44	266	1392	1623	1950	2074	2398	2903	12724
phecode_105-1	149	293	869	936	978	972	1029	1157	1167	1403	8953
phecode_106	340	291	351	395	409	386	422	449	492	680	4215
phecode_106-1	56	68	70	70	74	73	74	103	110	156	854
phecode_106-11	18	18	22	16	23	27	32	29	31	74	290
phecode_106-13	18	29	44	37	50	43	57	65	70	118	531
phecode_106-2	105	110	131	168	178	157	209	223	260	425	1966
phecode_106-21	98	100	123	171	154	152	205	215	244	415	1877
phecode_106-3	83	123	110	121	143	171	201	192	208	294	1646

Supplementary Tables

Table 15 continued from previous page

Endpoint	1	2	3	4	5	6	7	8	9	10	Total
phecode_106-4	5	7	15	17	22	16	21	18	23	28	172
phecode_107	756	855	767	982	1240	1350	1502	1734	2040	2554	13780
phecode_107-1	6	15	11	14	18	31	32	23	48	132	330
phecode_107-2	700	849	736	966	1180	1320	1453	1731	1958	2480	13373
phecode_107-3	10	13	18	14	16	16	16	21	16	24	164
phecode_108	129	200	292	389	472	589	732	840	1119	1902	6664
phecode_108-4	44	73	98	113	165	208	264	276	352	602	2195
phecode_108-41	42	64	96	98	146	192	241	253	320	552	2004
phecode_108-42	2	2	7	6	14	11	25	20	26	66	179
phecode_108-5	75	112	158	199	276	319	367	494	657	1088	3745
phecode_108-7	3	2	5	17	16	19	32	20	39	104	257
phecode_109	43	68	77	77	124	118	160	225	320	441	1653
phecode_109-1	6	16	11	15	18	22	24	21	30	31	194
phecode_109-16	7	8	5	10	7	13	19	17	15	16	117
phecode_109-3	33	39	45	53	85	106	128	166	287	382	1324
phecode_110	39	69	69	61	70	65	86	102	126	223	910
phecode_110-1	25	33	37	46	46	49	53	68	84	132	573
phecode_110-4	6	9	4	11	4	16	9	14	20	47	140
phecode_112	1218	1653	2120	2410	2730	3184	3830	4708	6089	9986	37928
phecode_112-1	7	7	6	6	14	26	41	76	123	225	531
phecode_114	21	38	36	59	58	80	107	123	173	307	1002
phecode_114-4	3	2	4	10	8	13	13	14	29	32	128
phecode_114-6	12	24	29	40	37	52	77	81	115	261	728
phecode_116	595	819	852	1070	1260	1587	1974	2725	3926	6371	21179
phecode_116-1	361	529	556	650	748	895	999	1329	1743	2628	10438
phecode_116-2	111	127	178	182	275	404	549	820	1316	2574	6536
phecode_116-3	75	97	159	160	229	298	406	584	881	1435	4324
phecode_116-4	116	145	187	248	319	483	641	879	1541	2707	7266
phecode_116-5	35	33	39	60	92	124	182	298	477	944	2284
phecode_116-6	98	111	149	201	246	370	505	779	1243	2439	6141
phecode_116-7	12	8	15	20	18	21	35	75	97	175	476
phecode_120	381	491	561	624	695	813	887	993	1058	1630	8133
phecode_120-1	163	221	224	266	277	332	396	463	473	933	3748
phecode_120-11	4	2	5	11	16	12	22	26	31	53	182
phecode_120-12	2	1	6	10	16	12	11	28	25	40	151
phecode_120-13	9	25	24	36	27	33	31	42	50	132	409
phecode_120-2	162	232	284	318	365	441	454	478	495	766	3995
phecode_120-21	134	213	254	280	341	380	437	408	447	692	3586
phecode_120-22	16	11	16	24	32	26	30	47	35	73	310
phecode_121	66	114	142	178	185	201	268	255	301	512	2222
phecode_121-1	18	27	30	48	42	47	58	77	154	265	766
phecode_121-11	4	6	5	7	4	7	7	11	18	35	104
phecode_121-12	15	17	26	39	33	39	47	80	129	236	661
phecode_121-2	40	76	94	118	136	155	167	159	189	272	1406
phecode_121-21	29	58	79	84	114	135	132	134	157	245	1167
phecode_121-22	5	18	11	10	14	19	19	29	22	55	202
phecode_121-23	2	1	3	10	9	17	10	20	20	31	123
phecode_122	142	189	221	253	278	327	366	379	382	637	3174
phecode_122-1	20	18	14	19	22	21	29	25	33	88	289
phecode_122-2	120	171	194	228	264	299	339	363	361	607	2946
phecode_122-21	40	54	53	56	67	75	75	64	77	108	669
phecode_122-22	31	61	64	83	74	113	149	123	167	290	1155
phecode_122-24	8	4	5	10	17	13	16	22	28	35	158
phecode_123	51	54	74	77	115	104	129	173	177	299	1253
phecode_123-1	51	55	70	82	111	95	130	161	181	290	1226
phecode_124	98	129	134	158	158	193	227	236	281	540	2154
phecode_124-3	9	21	24	33	29	31	31	42	42	131	393
phecode_124-5	45	57	49	52	51	76	79	89	117	246	861
phecode_124-6	12	18	32	36	35	49	62	92	95	274	705
phecode_124-7	18	24	29	26	44	34	39	39	54	127	434
phecode_125	21	15	17	35	21	18	35	35	42	83	322
phecode_130	3368	3960	4763	5556	6353	7057	7456	8367	9742	13065	69687
phecode_132	27	59	98	237	346	393	404	494	619	2816	5493
phecode_135	234	287	322	394	453	569	601	708	892	1430	5890
phecode_135-1	67	81	117	121	125	133	157	164	191	258	1414

4 Medical history predicts future health trajectories over the human phenome

Table 15 continued from previous page

Endpoint	1	2	3	4	5	6	7	8	9	10	Total
phecode_135-12	11	10	8	10	17	11	18	14	16	12	127
phecode_135-16	45	40	37	53	57	66	62	58	84	143	645
phecode_135-5	133	177	207	249	289	356	436	493	663	1200	4203
phecode_135-6	3	20	23	15	21	28	32	32	51	86	311
phecode_136	2252	2831	2993	3807	4304	4911	5553	6493	8202	11008	52354
phecode_136-1	6	6	7	13	16	7	17	17	33	41	163
phecode_136-2	255	356	428	477	654	809	1022	1355	1852	4065	11273
phecode_136-3	7	10	10	8	16	18	35	38	40	72	254
phecode_136-4	1951	2480	2598	3300	3684	4170	4683	5482	6808	9010	44166
phecode_136-41	1546	2014	2229	2633	3083	3511	3925	4729	5881	7973	37524
phecode_136-42	759	872	1023	1298	1351	1570	1822	2221	2706	3908	17530
phecode_136-6	2	17	22	14	10	12	17	31	24	34	183
phecode_136-8	6	9	5	13	13	14	21	22	28	20	151
phecode_137	11	10	19	19	16	20	26	42	40	45	248
phecode_137-5	5	3	19	13	14	10	16	21	16	28	145
phecode_138	382	518	1789	2423	2936	3583	4272	4823	5855	7838	34419
phecode_138-1	67	125	191	245	312	403	419	474	569	830	3635
phecode_138-2	200	293	1255	1682	2239	2715	3321	3947	4732	6619	27003
phecode_139	561	805	1402	1619	1796	1972	2223	2315	2828	3333	18854
phecode_139-1	18	18	28	36	28	42	47	48	60	100	425
phecode_139-3	39	37	80	89	84	114	128	138	156	238	1103
phecode_139-4	35	35	30	39	35	44	46	66	77	102	509
phecode_139-5	376	535	952	1198	1327	1445	1515	1730	1961	2419	13458
phecode_139-51	3	8	9	8	10	12	14	15	26	41	146
phecode_139-52	20	40	67	62	83	102	130	131	154	231	1020
phecode_139-53	226	270	340	422	418	477	545	564	684	886	4832
phecode_139-54	63	65	78	100	99	115	136	131	142	195	1124
phecode_139-6	129	175	294	368	351	426	528	542	717	880	4410
phecode_139-61	124	156	286	369	345	423	498	534	690	887	4312
phecode_139-62	6	4	7	9	10	8	5	14	21	25	109
phecode_140	5	1	7	16	59	144	168	266	323	480	1469
phecode_142	90	145	257	308	351	634	3275	4035	5138	6981	21214
phecode_142-1	61	122	230	312	352	639	1543	2193	3216	5003	13671
phecode_142-2	0	0	1	0	0	8	32	32	51	62	186
phecode_142-21	0	0	1	0	0	8	31	33	51	62	186
phecode_144	948	743	787	940	1079	1268	1669	2032	2423	3571	15460
phecode_144-1	155	118	150	175	233	308	383	490	596	757	3365
phecode_144-11	12	9	6	11	8	15	17	20	16	24	138
phecode_144-12	8	12	15	15	22	16	18	27	31	56	220
phecode_144-13	138	91	126	136	208	265	348	476	555	722	3065
phecode_144-2	766	575	645	738	849	1019	1283	1654	2012	3098	12639
phecode_144-21	399	345	386	454	540	667	835	1183	1646	2523	8978
phecode_144-3	127	131	124	125	125	172	179	163	187	232	1565
phecode_146	22	20	29	42	41	72	132	227	1379	4428	6392
phecode_146-2	29	44	37	61	57	126	445	1081	1624	2457	5961
phecode_146-4	7	10	9	18	9	23	23	22	38	33	192
phecode_146-5	5	6	10	7	11	7	15	22	21	79	183
phecode_148	100	141	235	253	331	328	349	434	508	578	3257
phecode_148-1	12	33	107	164	175	211	211	230	251	349	1743
phecode_148-16	3	30	103	135	158	192	188	217	239	334	1599
phecode_148-2	39	46	60	74	59	92	114	150	139	238	1011
phecode_148-3	3	7	2	10	11	8	10	9	12	49	121
phecode_148-5	31	22	41	33	38	48	60	47	63	68	451
phecode_149	86	94	108	109	132	141	169	160	223	401	1623
phecode_149-1	20	16	14	15	19	39	43	38	53	69	326
phecode_149-3	18	37	40	43	44	74	76	109	107	208	756
phecode_149-4	28	39	31	39	36	38	52	63	73	141	540
phecode_153	18	26	52	51	63	68	75	61	75	91	580
phecode_159	0	4	21	18	29	21	41	35	59	126	354
phecode_159-1	0	3	10	25	15	22	26	36	40	94	271
phecode_160	914	1141	1180	1473	1730	2016	2342	3030	4088	8005	25919
phecode_160-1	819	1002	1147	1303	1613	1776	2117	2760	3686	7313	23536
phecode_160-2	108	120	146	146	205	229	293	444	600	1437	3728
phecode_160-4	4	13	20	23	25	29	59	88	124	257	642
phecode_161	22	13	21	20	32	37	33	37	34	133	382

Supplementary Tables

Table 15 continued from previous page

Endpoint	1	2	3	4	5	6	7	8	9	10	Total
phecode_161-1	10	5	10	8	18	14	16	18	16	78	193
phecode_161-2	16	10	15	15	18	22	26	25	28	83	258
phecode_161-21	12	7	10	11	13	17	17	26	18	54	185
phecode_162	24	47	41	62	80	94	174	162	187	509	1380
phecode_162-8	4	3	10	10	21	25	40	39	58	126	336
phecode_164	1762	2267	2469	2785	2914	3496	4070	5142	6732	12171	43808
phecode_164-1	822	1002	1145	1303	1619	1774	2113	2765	3683	7325	23551
phecode_164-2	119	117	151	157	214	234	311	457	644	1534	3938
phecode_164-3	10	12	12	6	17	22	18	26	23	69	215
phecode_164-6	28	39	47	62	63	76	115	200	233	829	1692
phecode_164-62	10	20	14	29	33	49	37	93	161	331	777
phecode_165	24	24	28	39	33	42	59	75	94	353	771
phecode_165-2	14	11	11	16	18	16	20	36	50	160	352
phecode_165-25	11	6	10	11	16	11	15	31	48	161	320
phecode_165-3	6	2	7	5	16	16	14	25	40	177	308
phecode_168	151	176	268	296	396	479	595	801	1031	2317	6510
phecode_168-1	48	61	132	154	218	275	332	470	648	1163	3501
phecode_168-11	8	14	7	18	8	13	20	13	23	44	168
phecode_168-12	0	3	0	2	1	3	7	10	25	114	165
phecode_168-15	0	2	3	2	1	0	0	9	10	76	103
phecode_168-18	8	4	22	27	52	37	68	101	138	259	716
phecode_168-19	25	40	63	105	137	178	216	300	416	791	2271
phecode_168-2	35	42	43	41	48	48	48	69	72	388	834
phecode_168-21	25	19	33	32	42	42	35	40	53	323	644
phecode_168-211	0	2	6	11	9	10	12	13	22	45	130
phecode_168-214	0	3	5	6	4	13	8	14	19	62	134
phecode_168-3	12	18	16	28	21	16	35	33	33	161	373
phecode_168-4	1	9	21	27	35	56	74	83	152	739	1197
phecode_169	164	196	258	278	345	432	513	596	720	1416	4918
phecode_169-1	161	200	256	267	350	433	514	581	719	1409	4890
phecode_169-11	29	29	45	50	49	45	68	60	73	235	683
phecode_169-14	12	7	7	19	15	35	29	26	35	97	282
phecode_170	408	493	569	592	665	772	797	963	980	1886	8125
phecode_170-1	375	492	546	591	632	748	762	930	983	1849	7908
phecode_170-13	0	1	9	18	15	14	45	47	55	87	291
phecode_170-19	5	35	144	177	186	214	260	347	409	647	2424
phecode_170-2	3	5	12	19	19	29	27	29	37	90	270
phecode_171	17	27	50	82	99	135	148	165	240	435	1398
phecode_171-1	3	7	25	40	62	74	85	116	130	198	740
phecode_171-7	17	8	26	19	30	32	31	45	72	180	460
phecode_171-9	1	1	1	4	12	13	12	15	25	64	148
phecode_172	65	89	113	131	162	186	208	291	321	534	2100
phecode_172-2	2	4	7	12	13	14	21	35	44	92	244
phecode_174	79	84	105	127	140	144	180	199	282	920	2260
phecode_174-1	13	16	26	28	22	23	38	46	51	399	662
phecode_174-2	35	44	53	71	75	96	91	127	174	471	1237
phecode_174-6	11	7	4	7	8	5	6	13	16	25	102
phecode_174-7	6	11	14	16	13	25	22	27	33	69	236
phecode_175	29	45	53	60	69	79	89	117	161	340	1042
phecode_175-2	20	37	38	42	45	55	85	97	138	352	909
phecode_176	40	99	135	157	181	232	273	340	460	688	2605
phecode_177	634	726	934	1127	1189	1454	1638	1966	2352	3749	15769
phecode_177-1	13	26	37	48	54	72	67	91	122	175	705
phecode_177-13	2	9	18	27	24	41	34	29	49	82	315
phecode_177-2	493	547	744	881	962	1105	1314	1545	1819	2556	11966
phecode_177-3	17	11	26	19	15	24	38	39	39	96	324
phecode_177-4	68	92	119	129	155	202	259	341	468	1339	3172
phecode_179	26	45	79	100	104	95	112	141	182	704	1588
phecode_179-1	9	16	7	16	14	19	14	24	32	155	306
phecode_179-9	5	15	43	56	64	63	77	87	117	526	1053
phecode_180	99	111	148	205	234	268	339	410	519	1092	3425
phecode_180-3	35	76	73	103	141	173	221	265	369	612	2068
phecode_180-31	29	73	59	96	130	157	207	245	350	580	1926
phecode_180-33	2	6	10	17	10	10	19	26	24	52	176
phecode_181	636	757	932	914	1075	1215	1413	1681	2127	3860	14610

4 Medical history predicts future health trajectories over the human phenome

Table 15 continued from previous page

Endpoint	1	2	3	4	5	6	7	8	9	10	Total
phecode_200	1045	1220	1474	1640	2118	2635	3028	3444	3948	7454	28006
phecode_200-1	762	958	1209	1271	1604	2028	2304	2572	3067	7737	23512
phecode_200-12	6	1	6	5	5	6	11	17	21	68	146
phecode_200-13	20	46	47	68	78	102	108	126	219	1335	2149
phecode_200-14	605	781	1050	1114	1332	1695	1956	2189	2626	7320	20668
phecode_200-2	168	206	249	283	386	538	690	823	1011	1367	5721
phecode_200-21	55	51	56	64	79	102	149	180	255	414	1405
phecode_200-22	53	73	114	147	158	201	313	368	348	612	2387
phecode_200-23	42	50	54	88	113	172	227	260	376	531	1913
phecode_200-3	153	163	181	226	323	359	434	546	662	1297	4344
phecode_200-31	47	49	49	53	67	86	85	116	190	329	1071
phecode_200-4	29	31	31	38	41	77	95	134	129	333	938
phecode_200-41	24	16	21	21	22	45	65	72	81	264	631
phecode_200-7	17	66	224	248	356	394	509	486	595	726	3621
phecode_200-9	11	37	92	100	153	211	209	306	371	710	2200
phecode_202	803	1370	1842	2187	2424	2941	3835	4784	6587	11180	37953
phecode_202-1	33	49	55	80	110	142	219	231	351	2452	3722
phecode_202-2	748	1350	1787	2146	2363	2898	3726	4760	6444	11238	37460
phecode_202-3	0	1	9	9	10	12	21	30	48	57	197
phecode_202-32	0	0	8	7	12	11	19	21	44	50	172
phecode_202-4	151	328	650	1105	1460	2006	2695	3680	5549	20597	38221
phecode_203	1	3	1	6	10	15	18	27	28	82	191
phecode_204	238	566	1677	2316	2937	3596	4296	5396	6800	9368	37190
phecode_204-1	9	40	207	304	374	532	637	796	1202	1910	6011
phecode_204-2	36	114	505	709	958	1228	1427	2007	2501	4102	13587
phecode_204-4	0	0	4	7	5	14	43	74	111	312	570
phecode_205	48	63	89	87	110	137	179	249	334	2736	4032
phecode_205-3	0	1	1	2	2	3	3	3	11	117	143
phecode_206	0	0	3	4	2	7	14	19	29	49	127
phecode_208	77	93	110	144	145	251	278	347	513	995	2953
phecode_208-1	2	6	13	8	14	14	14	10	33	82	196
phecode_208-2	69	89	94	131	153	238	263	342	498	967	2844
phecode_208-21	36	60	74	76	96	148	182	244	317	538	1771
phecode_208-22	13	13	21	22	17	35	41	55	84	386	687
phecode_209	75	66	80	111	110	133	146	171	228	489	1609
phecode_209-1	32	34	40	63	67	71	89	99	124	263	882
phecode_209-12	12	14	25	44	45	53	74	85	106	216	674
phecode_209-13	5	5	7	8	3	2	9	14	13	53	119
phecode_209-2	13	23	24	31	33	50	46	53	89	281	643
phecode_209-21	8	2	7	1	3	6	10	12	13	61	123
phecode_209-22	9	12	18	28	25	19	48	44	73	259	535
phecode_209-23	4	5	3	3	4	5	5	5	6	74	114
phecode_210	2	6	5	5	8	7	14	14	27	41	129
phecode_211	37	46	68	96	82	120	118	152	200	511	1430
phecode_211-1	4	2	3	1	5	6	3	8	16	53	101
phecode_211-2	20	35	32	38	51	45	56	51	110	336	774
phecode_211-21	9	10	7	11	11	13	23	19	29	145	277
phecode_211-22	4	3	4	6	2	4	6	8	11	59	107
phecode_214	5	9	24	30	42	34	52	65	69	149	479
phecode_214-1	5	7	22	31	35	36	50	57	69	143	455
phecode_214-11	3	8	20	29	34	38	50	49	64	127	422
phecode_215	17	25	27	33	36	45	52	76	111	219	641
phecode_215-1	13	22	26	31	38	41	49	62	106	224	612
phecode_229	2	7	7	11	15	16	30	34	41	173	336
phecode_230	877	1197	1273	1545	1942	2279	2759	3257	4209	6718	26056
phecode_230-1	10	19	18	25	37	47	61	84	115	303	719
phecode_230-2	728	984	1051	1273	1548	1856	2227	2658	3381	5386	21092
phecode_230-21	657	883	978	1180	1391	1740	2163	2571	3227	5225	20015
phecode_230-22	13	22	49	73	94	124	153	234	257	458	1477
phecode_230-3	159	205	249	332	436	491	565	830	1092	2083	6442
phecode_230-4	4	10	9	13	13	23	25	27	46	143	313
phecode_230-5	0	0	0	2	5	12	8	15	22	64	128
phecode_232	445	903	1549	1965	2417	2926	3705	4898	6239	10796	35843
phecode_232-1	1	0	1	1	6	8	6	11	32	47	113
phecode_232-2	272	352	505	653	751	970	1285	1676	2294	4731	13489

Supplementary Tables

Table 15 continued from previous page

Endpoint	1	2	3	4	5	6	7	8	9	10	Total
phecode_232-27	74	76	86	113	125	159	202	254	410	1063	2562
phecode_232-29	28	31	47	53	63	68	106	143	216	475	1230
phecode_232-4	187	510	973	1299	1718	2023	2598	3444	4469	7921	25142
phecode_234	21	73	228	286	330	365	517	630	761	1321	4532
phecode_236	875	1584	2500	3217	4001	4694	5712	6942	8966	14095	52586
phecode_236-1	872	1590	2490	3213	3997	4693	5713	6920	8939	14037	52464
phecode_236-11	13	27	29	52	65	103	131	230	375	1643	2668
phecode_236-2	4	2	5	8	4	8	10	7	20	176	244
phecode_237	8	31	111	168	222	267	366	493	649	1195	3510
phecode_239	2217	3262	4317	5570	6539	7883	9146	10627	12630	16394	78585
phecode_239-1	2007	2938	3943	5024	6014	7226	8537	9797	11865	15750	73101
phecode_239-11	2007	2941	3938	4992	5982	7231	8534	9765	11927	15882	73199
phecode_239-12	3	12	35	54	85	129	182	225	307	503	1535
phecode_239-2	33	45	92	103	186	202	266	289	410	749	2375
phecode_239-21	11	13	28	25	50	62	83	102	132	398	904
phecode_239-3	20	33	60	78	114	147	176	213	267	421	1529
phecode_240	6	1	6	9	12	12	15	14	29	64	168
phecode_241	17	12	12	25	45	56	46	67	77	136	493
phecode_242	4	4	9	9	15	22	18	21	28	25	155
phecode_244	52	69	121	136	173	246	277	371	472	942	2859
phecode_244-4	0	3	4	3	1	5	4	8	11	203	242
phecode_247	1316	1715	1783	2048	2381	2692	3249	4020	5329	9647	34180
phecode_247-3	44	76	102	122	113	149	173	203	291	532	1805
phecode_247-4	49	74	92	113	124	211	244	315	513	1264	2999
phecode_247-42	3	16	10	10	19	17	35	41	83	342	576
phecode_247-5	172	288	310	384	446	521	591	750	994	1812	6268
phecode_247-51	1	15	41	70	91	112	131	167	241	424	1293
phecode_247-52	6	20	61	66	108	138	168	246	282	586	1681
phecode_247-7	1012	1088	1378	1628	1894	2076	2510	3130	4250	7902	26868
phecode_247-71	0	15	48	58	66	78	78	90	111	199	743
phecode_247-711	0	1	2	10	10	13	15	41	43	87	222
phecode_247-72	927	985	1313	1504	1803	1973	2382	2984	4056	7827	25754
phecode_248	32	49	62	78	75	103	109	127	139	362	1136
phecode_248-1	0	2	16	20	7	10	11	17	21	28	132
phecode_249	13	19	23	36	48	55	60	61	80	155	550
phecode_249-1	1	9	13	13	23	18	22	29	37	46	211
phecode_251	47	57	74	92	92	129	175	237	282	405	1590
phecode_251-1	37	45	66	84	90	124	160	224	269	400	1499
phecode_252	8	7	10	13	24	28	38	44	53	120	345
phecode_256	1058	1525	1771	2212	2515	2864	3241	4315	6064	10931	36496
phecode_256-1	45	87	90	117	151	163	184	192	343	670	2042
phecode_256-2	522	691	791	903	1063	1245	1475	1792	2423	3736	14641
phecode_256-3	50	93	123	171	188	275	384	507	721	2144	4656
phecode_256-31	45	67	102	130	151	218	308	437	609	1990	4057
phecode_256-32	8	24	27	37	51	52	60	82	126	213	680
phecode_256-4	92	131	188	232	289	355	375	503	801	2452	5418
phecode_256-5	160	271	330	442	572	636	777	1108	1478	2635	8409
phecode_256-6	28	51	70	88	150	166	271	285	449	1515	3073
phecode_256-7	193	249	378	532	593	779	988	1158	1726	3988	10584
phecode_256-71	3	9	15	25	40	79	72	107	152	343	845
phecode_257	17	34	75	82	102	130	160	196	280	471	1547
phecode_280	602	848	997	1319	1515	1808	2163	2621	3403	5431	20707
phecode_280-1	573	791	917	1241	1459	1705	2082	2461	3256	5118	19603
phecode_280-11	260	417	485	544	641	832	1071	1343	1909	3587	11089
phecode_280-12	111	190	222	308	348	454	534	754	1060	2304	6285
phecode_280-13	15	31	33	54	87	90	159	207	248	852	1776
phecode_280-14	2	3	5	10	6	12	13	17	27	120	215
phecode_280-2	1	5	8	2	12	16	12	24	28	140	248
phecode_280-22	0	3	2	2	6	7	13	17	20	111	181
phecode_280-3	3	0	2	6	7	15	16	23	39	275	386
phecode_280-31	2	2	0	8	5	14	18	13	28	218	308
phecode_280-4	3	1	3	2	8	8	14	18	46	175	278
phecode_280-42	2	1	5	1	7	7	11	18	44	168	264
phecode_280-8	12	22	27	34	46	42	57	101	120	443	904
phecode_280-81	3	6	8	6	9	10	15	23	28	127	235

4 Medical history predicts future health trajectories over the human phenome

Table 15 continued from previous page

Endpoint	1	2	3	4	5	6	7	8	9	10	Total
phecode_280-82	12	11	22	34	31	36	41	78	104	341	710
phecode_281	1131	1570	1637	2032	2283	2629	3177	3792	4983	8499	31733
phecode_281-1	264	427	516	552	676	853	1116	1419	1960	3746	11529
phecode_281-2	940	1250	1344	1552	1867	2123	2606	3237	4167	7764	26850
phecode_281-21	1	5	7	8	15	25	25	43	55	256	440
phecode_282-1	709	980	1268	1456	1541	1780	2077	2529	3267	5513	21120
phecode_283	980	1497	4596	6127	7313	8493	9848	11206	13084	16795	79939
phecode_283-3	2	3	8	12	10	11	14	26	49	132	267
phecode_283-4	125	334	737	993	1226	1556	1991	2545	3674	6148	19329
phecode_283-8	841	1354	4142	5552	6628	7629	8790	9898	11695	14744	71273
phecode_284	68	98	122	136	186	229	321	438	745	2910	5253
phecode_284-1	11	31	71	135	137	176	234	351	553	2260	3959
phecode_284-2	25	38	48	37	47	82	82	141	228	1203	1931
phecode_284-29	25	38	45	37	46	81	77	140	226	1199	1914
phecode_286	981	1298	1761	2167	2590	3239	3914	4921	6460	11407	38738
phecode_286-1	17	19	22	30	27	57	71	71	122	612	1048
phecode_286-2	928	1261	1757	2154	2568	3314	3906	4954	6589	11652	39083
phecode_286-21	31	33	46	46	85	93	110	190	330	1043	2007
phecode_286-3	1	3	2	0	4	6	23	37	105	308	489
phecode_286-4	11	4	14	13	18	22	34	48	99	274	537
phecode_287	53	55	68	92	88	127	149	168	213	674	1687
phecode_287-1	3	6	9	15	17	18	25	43	54	355	545
phecode_287-2	6	2	0	2	2	4	7	8	10	135	176
phecode_287-4	16	29	37	35	39	51	62	80	122	332	803
phecode_287-5	5	8	14	14	12	14	15	24	32	107	245
phecode_288	777	1203	1775	2213	2690	3288	3974	5021	6550	10598	38089
phecode_288-2	60	101	139	192	252	273	392	563	805	1443	4220
phecode_288-3	24	100	284	432	540	701	821	1161	1600	2812	8475
phecode_288-4	47	127	178	290	368	413	573	728	996	1537	5257
phecode_288-41	4	4	5	9	6	9	6	14	24	93	174
phecode_289	14	14	18	25	27	36	32	43	70	257	536
phecode_290	58	192	610	935	1184	1518	2002	2550	3410	5821	18280
phecode_290-1	18	11	23	43	51	63	72	83	139	458	961
phecode_291	9	16	16	14	32	42	50	61	106	257	603
phecode_292	24	40	52	49	67	74	85	123	169	326	1009
phecode_293	10	12	21	25	21	40	45	57	76	212	519
phecode_293-1	3	2	9	4	7	8	10	12	10	44	109
phecode_293-4	1	4	6	15	6	16	17	32	51	125	273
phecode_294	113	151	195	261	282	311	1909	4993	7143	11498	26856
phecode_296	3	7	6	12	12	14	14	37	47	342	494
phecode_296-4	1	1	1	5	3	4	12	9	22	202	260
phecode_299	54	82	108	150	161	162	206	331	419	1028	2701
phecode_308	258	605	1315	1756	2256	2913	3541	4534	6304	10738	34220
phecode_308-1	11	13	35	53	53	83	103	138	223	424	1136
phecode_308-3	3	8	32	40	56	66	105	141	222	580	1253
phecode_308-4	1	5	7	16	16	35	59	94	160	419	812
phecode_308-5	6	22	43	85	82	111	134	195	280	522	1480
phecode_308-6	0	3	15	27	26	30	68	105	156	369	799
phecode_308-7	55	63	101	105	132	149	182	231	314	691	2023
phecode_320	16	15	32	41	34	41	53	64	64	122	471
phecode_320-1	14	14	25	29	32	25	31	37	47	68	322
phecode_320-11	5	11	9	16	20	12	24	20	34	33	184
phecode_320-12	2	4	10	10	7	11	18	12	26	40	140
phecode_320-3	5	5	13	7	4	14	16	13	31	51	159
phecode_321	40	53	51	44	54	62	63	70	90	160	687
phecode_321-1	18	20	22	28	22	20	28	33	45	55	291
phecode_321-12	14	7	14	14	32	19	26	24	40	41	231
phecode_321-2	7	7	8	14	9	7	14	14	13	37	130
phecode_321-21	8	5	11	5	8	8	5	13	11	32	106
phecode_322	10	21	14	17	28	31	45	50	52	169	437
phecode_322-4	12	5	12	9	12	18	23	30	27	52	200
phecode_323	24	45	70	75	87	96	111	138	181	301	1128
phecode_323-1	9	20	19	25	22	31	31	46	54	122	379
phecode_323-3	16	23	42	35	39	59	64	82	134	190	684
phecode_323-31	3	4	1	2	1	4	27	41	49	58	190

Supplementary Tables

Table 15 continued from previous page

Endpoint	1	2	3	4	5	6	7	8	9	10	Total
phecode_324	524	758	1109	1357	1569	2011	2384	2821	3714	5643	21890
phecode_324-1	53	158	214	289	381	416	432	541	718	1161	4363
phecode_324-11	55	156	209	278	371	398	418	529	688	1094	4196
phecode_324-12	4	11	7	9	20	22	26	24	57	163	343
phecode_324-2	6	14	25	33	34	39	39	54	69	94	407
phecode_324-21	2	9	8	12	21	16	29	32	38	48	215
phecode_324-3	84	166	340	497	566	665	780	1050	1281	2181	7610
phecode_324-34	17	44	222	312	383	463	531	622	889	1347	4830
phecode_324-36	4	7	15	20	33	32	29	62	66	109	377
phecode_324-4	94	176	356	440	539	732	920	1130	1519	2515	8421
phecode_324-41	35	68	99	134	190	269	281	359	515	841	2791
phecode_324-5	14	13	15	26	27	35	29	41	62	139	401
phecode_324-8	12	33	109	197	235	294	402	516	729	1364	3891
phecode_325	510	720	905	1056	1264	1658	2024	2794	3893	7628	22452
phecode_325-1	32	46	73	92	114	121	131	147	171	352	1279
phecode_325-12	19	24	36	47	54	83	74	99	89	164	689
phecode_325-2	386	572	817	965	1113	1458	1893	2536	3651	7343	20734
phecode_325-21	6	10	12	10	18	26	39	40	56	118	335
phecode_325-23	3	15	52	67	141	204	268	396	613	1257	3016
phecode_325-3	59	86	139	176	219	275	402	469	713	1325	3863
phecode_326	41	45	67	44	64	80	73	83	108	282	887
phecode_326-1	26	35	52	36	46	59	69	71	93	274	761
phecode_327	32	71	101	122	155	235	253	333	485	1026	2813
phecode_328	38	162	277	509	669	843	936	1402	1892	3273	10001
phecode_328-1	22	104	174	276	379	424	597	785	1055	1616	5432
phecode_328-2	2	15	16	13	25	32	39	39	57	95	333
phecode_328-4	3	2	9	14	19	40	58	69	93	263	570
phecode_328-7	1	20	49	69	106	178	194	250	412	1025	2304
phecode_328-8	34	89	138	178	235	385	527	752	1007	1711	5056
phecode_328-9	16	72	130	215	305	448	472	632	993	1902	5185
phecode_329	917	1376	1504	1909	2327	2923	3510	4667	6126	9886	35145
phecode_329-1	122	240	466	704	935	1222	1619	2283	3023	5105	15719
phecode_329-4	6	28	128	193	223	308	335	418	583	1002	3224
phecode_329-41	0	6	25	45	37	76	76	106	170	349	890
phecode_329-42	7	20	95	123	183	202	256	294	419	761	2360
phecode_329-5	8	8	33	84	136	214	306	459	698	1175	3121
phecode_329-6	42	40	66	79	118	142	170	177	247	331	1412
phecode_329-8	3	5	17	22	32	35	60	67	100	222	563
phecode_329-9	43	111	165	239	344	514	691	763	1286	2680	6836
phecode_330	242	304	358	438	480	559	668	739	842	1745	6375
phecode_330-1	125	202	198	249	287	334	376	397	507	1168	3843
phecode_330-11	37	34	49	53	67	61	84	80	134	524	1123
phecode_330-12	33	43	51	53	63	59	71	79	106	346	904
phecode_330-3	166	196	248	291	340	414	442	484	588	1508	4677
phecode_331	821	1322	2211	2926	3580	4288	5067	6294	7794	11567	45870
phecode_331-1	34	72	200	260	369	459	642	788	1005	1808	5637
phecode_331-3	4	12	36	27	49	68	83	107	166	340	892
phecode_331-4	0	5	15	21	35	40	48	61	97	194	516
phecode_331-6	244	350	622	682	900	1116	1433	1717	2419	3846	13329
phecode_331-61	18	57	111	153	190	237	329	412	526	887	2920
phecode_331-62	0	1	3	3	6	7	5	16	16	70	127
phecode_331-7	2	5	4	5	5	16	19	33	26	123	238
phecode_331-8	664	1114	1809	2395	3064	3714	4560	5442	7020	10906	40688
phecode_333	459	851	1561	2103	2519	3179	4006	5006	6888	12899	39471
phecode_333-1	136	252	303	391	538	667	896	1210	1843	3734	9970
phecode_333-11	8	39	77	128	229	310	456	632	1003	2154	5036
phecode_333-2	53	209	727	1014	1205	1565	1915	2546	3248	5163	17645
phecode_333-3	2	10	9	18	23	24	40	40	70	181	417
phecode_333-4	1	8	24	37	42	61	86	150	198	498	1105
phecode_333-43	0	2	9	3	5	5	17	17	24	34	116
phecode_333-5	6	8	26	36	39	47	60	86	117	246	671
phecode_334	198	308	415	494	586	673	758	919	1059	1702	7112
phecode_334-1	47	84	126	177	224	292	356	435	574	1008	3323
phecode_334-11	44	68	103	151	161	244	286	372	484	864	2777
phecode_334-12	2	7	19	31	34	31	52	54	103	165	498

4 Medical history predicts future health trajectories over the human phenome

Table 15 continued from previous page

Endpoint	1	2	3	4	5	6	7	8	9	10	Total
phecode_334-2	88	152	197	221	243	282	299	305	426	637	2850
phecode_334-21	68	131	159	175	198	181	238	249	322	472	2193
phecode_334-23	0	2	20	26	33	43	42	55	75	119	415
phecode_334-24	4	4	11	7	22	25	26	24	25	46	194
phecode_334-4	34	53	56	45	69	65	95	118	154	210	899
phecode_334-41	12	13	20	18	14	20	22	27	47	68	261
phecode_334-42	10	10	10	19	23	32	11	28	35	56	234
phecode_334-44	10	22	31	22	28	36	52	55	80	112	448
phecode_335	266	320	368	520	617	742	892	1148	1602	3066	9541
phecode_335-1	9	5	10	12	19	18	20	23	38	64	218
phecode_335-11	5	4	9	16	13	14	24	13	41	58	197
phecode_335-2	0	5	14	11	22	22	23	34	41	98	270
phecode_335-4	1	3	2	6	4	4	5	4	14	85	128
phecode_336	731	945	1400	1758	2047	2535	2907	3359	4257	6282	26221
phecode_336-1	537	620	927	1132	1300	1584	1897	2173	2796	4172	17138
phecode_336-2	67	110	179	208	267	282	342	423	609	1032	3519
phecode_336-4	8	2	9	18	7	24	31	29	37	88	253
phecode_336-5	99	169	340	486	577	683	808	948	1300	2024	7434
phecode_336-51	3	11	26	38	46	59	81	86	124	165	639
phecode_336-52	5	29	118	147	183	201	260	268	339	476	2026
phecode_336-54	4	1	6	10	15	11	18	13	28	36	142
phecode_336-55	74	89	178	228	272	334	438	563	786	1334	4296
phecode_337	230	330	357	419	534	638	842	1083	1381	3226	9040
phecode_337-1	11	16	21	28	40	48	76	75	103	214	632
phecode_337-11	9	12	15	18	25	32	58	59	75	154	457
phecode_337-2	30	32	35	40	35	39	52	62	57	118	500
phecode_337-21	17	17	22	22	27	16	32	29	37	43	262
phecode_337-3	21	20	24	35	22	32	34	36	67	113	404
phecode_337-31	17	19	26	22	21	22	26	41	43	71	308
phecode_337-8	3	12	9	12	18	23	48	57	98	1122	1402
phecode_338	11	20	22	32	34	30	41	60	67	88	405
phecode_338-1	10	16	18	31	24	24	37	56	60	75	351
phecode_339	8	14	11	13	18	17	16	28	27	70	222
phecode_339-1	6	3	7	14	7	8	10	15	15	36	121
phecode_340	16	15	22	24	25	44	34	51	67	168	466
phecode_341	85	154	181	222	307	365	368	479	721	1601	4483
phecode_341-1	2	1	5	1	7	8	14	11	16	45	110
phecode_341-2	60	114	133	193	249	304	318	362	609	1400	3742
phecode_341-6	15	27	28	45	33	37	56	70	100	202	613
phecode_342	40	79	88	123	136	161	233	294	419	817	2390
phecode_342-1	6	10	15	14	16	21	34	41	49	163	369
phecode_342-2	4	11	10	11	8	10	17	18	22	75	186
phecode_342-4	27	22	43	72	93	126	183	240	368	654	1828
phecode_342-5	6	2	3	8	8	14	12	13	22	66	154
phecode_343	39	70	65	87	104	127	191	228	285	707	1903
phecode_343-1	4	4	9	6	15	14	22	19	38	257	388
phecode_343-3	8	15	21	33	50	56	100	118	197	343	941
phecode_343-5	11	12	13	18	11	18	26	22	25	47	203
phecode_343-6	5	8	7	6	9	14	11	21	14	54	149
phecode_344	60	71	94	81	119	123	142	161	199	298	1348
phecode_344-1	28	44	64	63	88	92	105	134	144	236	998
phecode_344-12	6	6	13	10	9	13	19	17	29	36	158
phecode_344-13	2	1	7	12	14	21	24	31	42	65	219
phecode_344-2	13	6	7	12	14	16	9	24	15	37	153
phecode_344-3	13	15	18	19	19	24	29	17	25	60	239
phecode_345	6	15	18	19	24	27	41	50	51	194	445
phecode_346	63	82	109	106	107	126	161	189	258	330	1531
phecode_346-1	0	2	16	14	28	26	29	48	58	113	334
phecode_346-3	6	7	11	15	15	28	33	37	102	171	425
phecode_346-5	10	14	12	13	17	14	27	29	23	48	207
phecode_346-6	18	27	33	32	46	40	50	77	100	122	545
phecode_347	69	103	160	194	208	219	218	284	354	655	2464
phecode_347-1	27	23	24	27	44	37	44	46	68	93	433
phecode_347-2	6	9	8	14	7	11	16	26	28	49	174
phecode_348	139	196	228	269	295	412	485	550	786	1437	4797

Supplementary Tables

Table 15 continued from previous page

Endpoint	1	2	3	4	5	6	7	8	9	10	Total
phecode_348-2	112	132	186	198	221	323	368	417	616	1159	3732
phecode_348-21	1	4	7	17	21	28	41	35	62	122	338
phecode_348-4	25	34	38	46	47	75	74	116	130	263	848
phecode_349	173	243	286	311	443	516	613	789	1018	1532	5924
phecode_349-1	125	161	187	201	247	305	376	464	613	928	3607
phecode_349-12	0	0	8	7	13	13	8	9	20	76	154
phecode_349-13	52	54	113	124	153	204	287	378	503	824	2692
phecode_349-15	18	15	14	21	23	17	39	30	50	86	313
phecode_349-2	3	10	52	85	119	160	186	244	318	516	1693
phecode_349-3	0	1	3	5	7	7	12	18	36	51	140
phecode_350	844	1143	1399	1782	2189	2548	3150	3981	5248	9171	31455
phecode_350-3	4	6	15	17	20	32	43	39	73	132	381
phecode_350-5	235	405	680	832	941	1188	1692	2292	3410	6970	18645
phecode_351	345	623	1573	2283	2821	3416	3997	4972	6415	11688	38133
phecode_351-1	19	80	307	487	619	778	936	1213	1504	2467	8410
phecode_351-2	11	11	18	43	58	80	141	217	392	2412	3383
phecode_351-3	180	335	865	1174	1503	1797	2105	2553	3229	4818	18559
phecode_351-4	1	4	9	21	18	25	27	46	52	148	351
phecode_352	14	63	216	281	389	419	481	595	702	1038	4198
phecode_352-1	9	28	108	143	204	202	239	271	304	420	1928
phecode_352-2	1	1	22	16	29	23	36	36	53	71	288
phecode_352-3	6	15	62	85	104	132	154	206	231	324	1319
phecode_353	143	253	407	573	716	863	973	1169	1464	2366	8927
phecode_353-1	38	66	71	113	145	155	160	221	313	882	2164
phecode_353-11	9	6	17	10	17	19	25	26	57	191	377
phecode_353-12	16	22	24	43	63	71	70	94	128	374	905
phecode_354	616	1100	2511	3478	4280	5223	6221	7684	9389	13398	53900
phecode_355	152	169	231	300	348	453	541	689	1001	2036	5920
phecode_355-1	73	80	82	103	135	126	169	167	241	480	1656
phecode_355-2	87	97	152	189	215	306	405	514	789	1622	4376
phecode_355-21	0	0	1	1	1	6	10	22	45	80	166
phecode_356	153	213	297	351	383	429	560	754	917	1735	5792
phecode_356-1	27	31	37	49	73	71	92	118	144	356	998
phecode_356-2	63	90	102	144	191	204	224	290	400	666	2374
phecode_356-4	2	5	20	26	26	43	41	63	111	213	550
phecode_360	145	345	1712	2271	2671	3041	3550	4150	5105	7253	30243
phecode_360-1	31	106	609	812	951	1066	1231	1484	1768	2500	10558
phecode_360-11	31	106	609	812	951	1066	1230	1485	1768	2500	10558
phecode_360-12	0	12	106	158	182	238	255	296	399	578	2224
phecode_360-13	1	4	18	17	32	41	43	46	40	82	324
phecode_360-2	46	117	551	603	719	861	937	1077	1179	1640	7730
phecode_360-4	86	212	761	1103	1415	1619	1939	2501	3138	4935	17709
phecode_360-5	2	6	31	38	54	91	143	161	212	368	1106
phecode_360-51	2	5	31	35	46	83	131	152	194	348	1027
phecode_361	246	281	296	372	450	563	689	908	1076	1795	6676
phecode_361-1	40	71	78	84	122	115	191	235	295	475	1706
phecode_361-15	0	4	11	25	25	35	27	47	70	130	374
phecode_361-2	6	4	6	6	4	7	12	14	13	53	125
phecode_361-3	120	125	123	167	235	294	325	430	515	878	3212
phecode_361-4	21	49	49	69	80	102	113	131	193	282	1089
phecode_361-9	49	44	45	73	63	86	132	163	236	445	1336
phecode_362	135	257	449	536	640	741	820	948	1157	1639	7322
phecode_362-1	8	12	42	49	49	69	78	102	108	163	680
phecode_362-5	2	13	98	144	137	180	203	230	259	369	1635
phecode_362-6	0	0	0	3	2	4	16	22	36	68	151
phecode_363	215	406	853	1306	1681	2344	2963	3816	5146	8182	26912
phecode_363-2	29	162	618	914	1284	1811	2330	3013	4254	7126	21541
phecode_363-5	65	110	205	289	357	475	570	725	960	1474	5230
phecode_363-51	0	1	6	5	6	12	18	22	26	47	143
phecode_363-6	21	12	21	31	41	41	48	65	84	139	503
phecode_363-61	4	2	18	20	34	27	34	49	58	99	345
phecode_363-7	49	73	72	113	140	165	201	273	326	527	1939
phecode_365	18	38	42	46	69	76	87	104	115	186	781
phecode_365-2	2	3	14	17	35	30	50	49	55	90	345
phecode_365-3	10	9	19	23	17	30	24	43	36	78	289

4 Medical history predicts future health trajectories over the human phenome

Table 15 continued from previous page

Endpoint	1	2	3	4	5	6	7	8	9	10	Total
phecode_366	51	116	224	300	322	359	395	467	589	858	3681
phecode_366-1	22	22	63	74	81	93	100	115	137	204	911
phecode_366-2	0	4	36	54	60	71	69	103	89	159	645
phecode_366-21	0	4	36	49	55	54	71	76	77	132	554
phecode_366-4	15	43	130	106	169	162	211	274	286	465	1861
phecode_366-42	3	12	65	74	84	108	122	183	182	302	1135
phecode_366-5	1	0	9	12	4	18	9	19	16	15	103
phecode_366-6	0	5	19	30	21	24	47	42	64	75	327
phecode_367	175	338	1912	2585	3064	3505	4079	4637	5672	7611	33578
phecode_367-1	109	200	1519	2099	2521	2912	3384	3943	4899	6643	28229
phecode_367-12	12	39	108	194	248	312	392	469	621	1144	3539
phecode_367-13	0	20	66	112	122	155	179	276	305	593	1828
phecode_367-2	28	70	225	231	289	294	361	400	430	760	3088
phecode_367-21	13	30	67	78	82	104	88	118	169	261	1010
phecode_367-22	0	3	4	12	19	8	17	14	32	43	152
phecode_367-3	4	10	25	22	38	34	41	56	70	129	429
phecode_367-4	5	29	36	35	52	60	58	92	103	117	587
phecode_367-41	0	3	30	25	29	43	49	67	81	87	414
phecode_367-5	42	87	219	222	245	279	322	316	392	759	2883
phecode_367-52	42	86	219	227	237	284	316	316	394	757	2878
phecode_367-6	0	22	80	105	117	125	144	208	257	305	1363
phecode_367-7	5	6	28	30	26	33	41	53	52	91	365
phecode_367-9	9	11	10	14	20	18	25	25	21	64	217
phecode_369	139	203	254	340	423	491	591	737	810	1293	5281
phecode_369-1	22	20	41	48	67	90	76	117	124	269	874
phecode_369-2	9	12	27	11	11	30	37	37	68	172	414
phecode_369-4	42	94	93	113	158	200	258	311	427	573	2269
phecode_369-42	0	6	14	21	27	22	24	35	39	59	247
phecode_369-44	0	1	8	6	9	17	17	20	24	34	136
phecode_369-5	32	43	67	107	114	126	160	154	191	316	1310
phecode_369-51	0	7	28	38	54	60	66	79	98	147	577
phecode_369-6	11	9	9	5	7	17	17	22	22	51	170
phecode_369-62	9	8	7	7	3	12	17	15	19	42	139
phecode_370	28	52	95	97	145	171	194	231	305	577	1895
phecode_370-1	1	10	21	15	22	27	33	37	32	117	315
phecode_370-3	5	14	22	43	50	74	84	80	104	181	657
phecode_370-4	5	11	3	12	13	23	29	27	33	113	269
phecode_371	1047	2091	2941	3476	4218	5319	6648	8408	11053	16283	61484
phecode_371-3	234	447	593	799	1179	1542	2132	2806	3880	5588	19200
phecode_371-31	234	447	593	799	1179	1542	2133	2805	3880	5588	19200
phecode_372	9	21	23	26	26	57	50	68	102	201	583
phecode_372-1	6	3	7	10	12	20	21	27	30	98	234
phecode_372-2	3	1	3	6	6	5	10	13	20	43	110
phecode_373	25	38	57	72	86	88	131	194	230	461	1382
phecode_373-1	12	16	23	27	26	37	42	40	87	241	551
phecode_373-2	6	6	4	7	4	8	11	13	13	35	107
phecode_374	842	1079	1566	2047	2562	3044	3764	4433	5458	13496	38291
phecode_374-1	278	314	392	419	533	523	590	643	747	997	5436
phecode_374-11	150	185	150	195	241	234	274	296	381	535	2641
phecode_374-12	4	3	10	9	17	17	15	20	16	50	161
phecode_374-13	0	0	14	11	19	13	13	16	10	15	111
phecode_374-14	0	2	10	12	12	10	9	21	19	26	121
phecode_374-2	12	15	31	63	63	66	94	86	85	128	643
phecode_374-21	0	3	24	41	50	53	49	77	73	87	457
phecode_374-3	78	125	242	365	447	582	657	851	1086	1759	6192
phecode_374-32	0	1	7	17	21	16	24	28	35	109	258
phecode_374-33	1	5	18	27	30	50	45	78	98	179	531
phecode_374-37	8	23	28	51	43	52	81	92	132	254	764
phecode_374-38	4	32	104	144	224	271	315	391	476	690	2651
phecode_374-39	14	18	37	69	66	91	122	165	179	321	1082
phecode_374-4	87	177	236	354	464	614	848	1136	1815	10684	16415
phecode_374-41	1	2	14	14	16	18	16	23	26	36	166
phecode_374-42	25	68	202	319	431	573	802	1049	1708	10394	15571
phecode_374-5	275	434	503	719	954	1240	1651	2198	2766	4168	14908
phecode_374-51	11	44	108	188	299	453	575	914	1219	1906	5717

Supplementary Tables

Table 15 continued from previous page

Endpoint	1	2	3	4	5	6	7	8	9	10	Total
phecode_374-511	1	8	25	46	77	126	155	235	403	745	1821
phecode_374-512	5	8	13	23	36	63	81	132	276	573	1210
phecode_374-52	3	18	54	100	127	188	251	270	344	464	1819
phecode_374-54	1	9	15	27	40	67	52	58	93	277	639
phecode_374-55	7	17	90	179	251	318	385	522	517	885	3171
phecode_374-6	12	20	46	70	33	58	56	62	93	96	546
phecode_374-61	0	1	11	18	13	23	24	24	25	23	162
phecode_374-7	11	10	24	33	54	39	54	64	78	95	462
phecode_374-8	5	22	73	114	128	155	181	265	316	991	2250
phecode_374-9	2	5	23	38	20	39	42	52	56	68	345
phecode_375	595	767	1087	1417	1653	2029	2280	2644	3025	4332	19829
phecode_375-1	447	591	632	927	1105	1318	1485	1716	2135	3656	14012
phecode_375-11	130	193	208	308	391	523	582	714	810	2369	6228
phecode_375-112	0	5	4	7	7	11	11	15	13	65	138
phecode_375-113	2	9	77	128	171	226	277	301	367	1047	2605
phecode_375-12	68	78	93	111	161	178	206	266	314	636	2111
phecode_375-14	3	4	35	47	65	91	114	131	162	301	953
phecode_375-6	8	57	293	424	482	569	641	751	857	1587	5669
phecode_375-7	6	5	2	5	5	6	11	11	9	54	114
phecode_376	115	279	964	1424	1792	2127	2500	2981	3463	4449	20094
phecode_376-1	16	107	662	1041	1314	1505	1831	2086	2533	3324	14419
phecode_376-2	113	272	961	1408	1778	2143	2466	2968	3446	4421	19976
phecode_376-21	49	95	249	372	477	610	695	815	987	1366	5715
phecode_376-4	1	4	12	21	35	36	56	82	82	117	446
phecode_377	123	248	808	1020	1276	1453	1674	1971	2344	3272	14189
phecode_377-2	55	130	658	800	987	1176	1309	1607	1898	2598	11218
phecode_377-4	17	40	65	99	165	201	195	236	297	460	1775
phecode_377-5	51	62	75	75	99	104	128	150	166	434	1344
phecode_377-8	1	5	17	8	11	20	20	16	19	53	170
phecode_379	26	70	176	187	269	221	265	322	395	542	2473
phecode_379-2	28	60	166	198	231	224	254	306	359	495	2321
phecode_379-21	23	44	89	115	105	117	134	161	182	287	1257
phecode_380	64	114	176	185	243	244	338	366	442	709	2881
phecode_380-1	35	34	45	44	60	63	76	105	112	202	776
phecode_380-11	15	13	20	15	10	21	23	17	29	99	262
phecode_380-12	2	0	13	14	24	26	42	31	63	73	288
phecode_380-2	35	59	106	144	155	168	213	263	289	476	1908
phecode_380-21	8	8	8	8	9	15	13	15	16	33	133
phecode_380-22	0	1	22	13	16	21	19	22	22	24	160
phecode_380-3	6	9	16	12	27	26	30	47	48	136	357
phecode_381	113	129	165	162	181	272	291	323	388	657	2681
phecode_381-1	39	69	72	94	82	93	125	171	194	282	1221
phecode_381-11	6	13	11	16	22	23	34	40	51	66	282
phecode_381-3	22	13	21	25	29	32	48	30	57	97	374
phecode_381-4	13	27	25	29	38	44	45	66	90	164	541
phecode_381-6	24	22	21	19	37	48	59	66	87	139	522
phecode_381-8	4	7	20	28	19	45	36	52	62	135	408
phecode_381-81	2	1	5	6	8	11	7	14	23	28	105
phecode_381-82	0	1	12	8	13	11	23	24	39	74	205
phecode_382	2	4	11	15	33	27	45	57	53	94	341
phecode_383	33	40	67	64	82	104	127	147	168	316	1148
phecode_383-1	1	8	39	40	45	64	76	93	122	201	689
phecode_384	23	35	54	76	89	82	112	124	159	281	1035
phecode_384-1	1	2	23	24	12	17	21	41	31	36	208
phecode_384-3	1	3	4	11	21	16	26	31	38	81	232
phecode_384-4	0	0	9	8	7	6	17	11	22	22	102
phecode_385	1	16	60	64	115	138	196	228	301	729	1848
phecode_386	350	589	977	1254	1607	1811	2219	2645	3137	4849	19438
phecode_386-1	76	79	118	153	174	255	325	329	472	784	2765
phecode_386-2	78	138	136	214	264	296	341	448	544	920	3379
phecode_386-3	2	3	7	5	9	14	18	21	39	57	175
phecode_386-4	86	103	135	190	177	276	295	337	441	672	2712
phecode_386-41	0	5	25	35	39	34	58	92	94	181	563
phecode_386-42	1	13	24	24	35	52	71	73	107	172	572
phecode_386-8	1	0	10	12	14	22	18	31	43	60	211

4 Medical history predicts future health trajectories over the human phenome

Table 15 continued from previous page

Endpoint	1	2	3	4	5	6	7	8	9	10	Total
phecode_386-9	1	9	45	59	93	133	140	177	237	429	1323
phecode_387	370	458	484	657	849	1030	1257	1536	1848	2350	10839
phecode_387-1	20	32	40	49	75	100	122	146	217	264	1065
phecode_387-2	210	227	225	327	387	457	498	550	629	1161	4671
phecode_387-21	16	18	17	27	24	33	41	43	52	157	428
phecode_387-3	102	169	182	237	330	433	553	767	918	1150	4841
phecode_387-4	2	4	18	21	13	31	22	28	41	41	221
phecode_387-5	2	7	8	11	14	14	15	24	34	49	178
phecode_388	99	110	169	244	346	438	571	823	1153	2610	6563
phecode_388-1	8	11	10	9	21	20	25	38	60	409	611
phecode_389	104	343	2007	2531	3110	3685	4554	5352	6746	9901	38333
phecode_389-1	19	56	138	191	215	268	364	448	577	1075	3351
phecode_390	267	415	4602	6422	7361	8260	9065	10035	11672	14005	72104
phecode_390-1	95	306	2121	2645	3092	3485	4005	4677	5550	7977	33953
phecode_390-4	128	309	3322	4411	5346	5854	6769	7742	9134	11763	54778
phecode_390-5	3	4	4	6	5	4	11	9	9	49	104
phecode_390-6	23	39	73	102	154	196	226	280	372	611	2076
phecode_391	244	443	2186	2706	3135	3651	4123	4657	5321	7424	33890
phecode_391-1	122	266	1028	1303	1492	1768	2096	2447	2977	4715	18214
phecode_391-11	6	15	69	89	113	144	174	191	301	501	1603
phecode_391-12	37	67	128	181	181	232	288	344	446	874	2778
phecode_391-2	34	147	858	1115	1335	1558	1881	2099	2535	3625	15187
phecode_391-21	2	13	68	95	127	148	164	178	242	409	1446
phecode_391-22	2	5	26	28	32	47	48	50	92	124	454
phecode_391-4	5	4	13	6	13	10	23	26	29	90	219
phecode_391-6	20	27	24	19	28	35	54	39	81	222	549
phecode_391-7	59	94	141	164	203	222	271	327	363	897	2741
phecode_391-8	19	13	20	20	30	32	34	39	69	77	353
phecode_391-9	22	45	112	114	149	191	226	276	414	1016	2565
phecode_392	64	208	948	1358	1733	2162	2579	3248	4196	6664	23160
phecode_393	17	10	12	14	24	41	36	40	55	125	374
phecode_394	141	278	623	848	1202	1425	1869	2229	3002	4440	16057
phecode_394-1	52	50	82	88	107	143	157	212	273	504	1668
phecode_394-2	79	205	450	715	918	1179	1496	1906	2613	3918	13479
phecode_394-21	64	142	290	476	721	944	1212	1668	2314	3610	11441
phecode_394-22	17	49	132	156	187	231	264	324	420	591	2371
phecode_394-4	2	8	30	56	72	90	101	148	199	277	983
phecode_395	98	193	692	971	1178	1381	1683	1949	2556	3512	14213
phecode_395-1	88	192	641	913	1112	1293	1588	1805	2435	3335	13402
phecode_395-3	0	1	3	2	3	9	10	23	36	37	124
phecode_396	431	1067	2562	3569	4463	5349	6505	7732	9371	12847	53896
phecode_396-1	26	24	87	75	115	120	158	171	227	495	1498
phecode_396-11	3	5	1	15	10	6	12	15	21	94	182
phecode_396-2	87	223	448	635	827	1068	1331	1645	2041	3009	11314
phecode_396-21	26	69	106	161	244	328	441	565	745	1203	3888
phecode_396-22	23	64	80	142	214	296	373	479	588	876	3135
phecode_396-3	13	9	21	33	35	52	60	75	108	279	685
phecode_396-5	1	4	13	8	5	10	19	22	18	40	140
phecode_397	94	270	1122	1515	1746	1989	2211	2642	3113	4176	18878
phecode_397-1	93	264	1103	1488	1727	1977	2161	2597	3060	4122	18592
phecode_397-3	0	2	5	6	12	12	17	17	23	35	129
phecode_398	23	87	449	549	660	778	869	1014	1383	2016	7828
phecode_398-1	3	15	77	111	159	173	221	277	366	566	1968
phecode_400	120	247	359	434	489	609	731	991	1532	3385	8897
phecode_400-2	118	250	352	439	485	605	731	995	1539	3391	8905
phecode_401	3740	5528	6740	7235	8473	9658	10872	12392	14880	20391	99909
phecode_401-1	3742	5532	6722	7184	8492	9617	10849	12421	14832	20408	99799
phecode_401-2	4	12	17	32	38	66	77	128	175	495	1044
phecode_401-3	8	8	17	18	31	39	74	108	163	1077	1543
phecode_401-6	11	53	184	226	240	311	365	420	467	718	2995
phecode_402	214	378	2032	2708	3050	3379	3808	3986	4524	5459	29538
phecode_403	288	650	1000	1263	1269	1726	2187	2827	3633	7102	21945
phecode_403-1	13	6	7	9	15	27	37	27	53	118	312
phecode_404	469	1021	1547	2189	2746	3254	3594	4781	6144	10795	36540
phecode_404-1	186	402	673	981	1178	1507	1654	2126	2846	6237	17790

Supplementary Tables

Table 15 continued from previous page

Endpoint	1	2	3	4	5	6	7	8	9	10	Total
phecode_404-11	153	312	558	749	995	1211	1385	1699	2147	3670	12879
phecode_404-2	286	668	1161	1606	2067	2376	2673	3635	4647	8190	27309
phecode_406	24	52	64	104	116	182	240	264	469	1406	2921
phecode_406-1	23	54	57	98	120	158	204	244	405	1209	2572
phecode_406-11	4	10	11	14	29	33	37	50	73	311	572
phecode_406-13	0	2	0	4	1	10	11	14	19	81	142
phecode_408	5	8	6	4	14	6	4	14	21	36	118
phecode_410	116	161	175	223	315	333	443	514	641	1274	4195
phecode_410-1	20	30	35	45	57	77	99	97	138	190	788
phecode_410-2	77	92	121	131	178	255	298	408	501	1054	3115
phecode_410-3	17	28	29	36	35	33	35	44	61	77	395
phecode_411	71	103	148	177	223	283	294	336	464	838	2937
phecode_411-1	0	6	2	4	7	15	23	25	59	65	206
phecode_411-2	47	64	97	93	169	204	209	239	324	632	2078
phecode_413	402	747	1019	1202	1272	1617	2154	2753	3671	6711	21548
phecode_413-1	277	466	624	703	828	1011	1281	1570	2208	4413	13381
phecode_413-11	199	297	377	425	512	663	794	1004	1357	2715	8343
phecode_413-12	30	49	50	61	64	95	103	121	131	248	952
phecode_413-13	6	2	9	13	17	17	30	29	53	186	362
phecode_413-2	153	365	472	620	634	850	1166	1570	2227	3884	11941
phecode_413-21	41	106	217	274	338	433	519	822	1187	2275	6212
phecode_413-22	73	150	184	210	288	343	480	605	753	1337	4423
phecode_413-3	103	183	282	337	375	468	582	741	1137	2731	6939
phecode_413-32	37	43	66	79	104	117	143	217	302	764	1872
phecode_413-4	25	34	38	27	46	72	94	114	172	342	964
phecode_413-42	22	35	31	34	37	61	94	110	162	337	923
phecode_413-6	23	52	70	134	149	161	195	258	368	1501	2911
phecode_414	82	158	198	238	247	251	298	383	506	1388	3749
phecode_414-1	17	34	33	43	52	48	59	62	90	208	646
phecode_414-11	7	12	23	27	19	25	22	36	41	136	348
phecode_414-2	30	40	53	82	88	94	108	142	171	468	1276
phecode_414-5	3	4	6	22	12	35	37	37	64	603	823
phecode_414-9	0	1	3	2	2	12	18	29	33	67	167
phecode_416	1222	2138	2905	3361	3585	4518	5508	6849	8817	13734	52637
phecode_416-1	231	283	389	359	468	509	593	754	854	1756	6196
phecode_416-11	179	215	258	286	349	408	454	552	666	1192	4559
phecode_416-12	26	55	70	102	121	133	161	186	280	793	1927
phecode_416-2	412	944	1506	1989	2357	2575	3314	4307	5742	9765	32911
phecode_416-21	249	427	589	965	1380	1739	2324	3001	4103	7354	22131
phecode_416-211	166	288	348	473	657	731	1023	1204	1681	3012	9583
phecode_416-212	16	37	58	63	87	115	135	163	240	657	1571
phecode_416-213	3	7	14	17	17	34	35	48	90	351	616
phecode_416-214	0	0	1	3	5	6	11	13	17	70	126
phecode_416-22	25	42	69	144	218	306	376	471	679	1485	3815
phecode_416-221	4	6	9	6	15	14	24	25	40	81	224
phecode_416-222	1	4	7	4	4	5	10	5	17	45	102
phecode_416-3	11	20	29	39	60	71	79	84	118	316	827
phecode_416-31	0	1	8	4	8	9	15	22	27	72	166
phecode_416-4	289	626	808	911	1071	1393	1662	2271	3164	5992	18187
phecode_416-41	88	208	306	315	394	498	682	944	1430	2871	7736
phecode_416-42	91	215	259	337	353	457	529	718	983	2022	5964
phecode_416-43	107	222	269	291	391	473	565	807	1097	2083	6305
phecode_416-5	106	211	306	463	511	634	772	924	1187	1735	6849
phecode_416-51	19	35	52	90	123	139	174	220	227	388	1467
phecode_416-52	72	123	180	272	298	330	454	539	668	1012	3948
phecode_416-6	2	5	8	7	8	6	21	18	23	27	125
phecode_416-7	14	34	35	40	64	91	111	137	180	474	1180
phecode_416-71	14	34	35	40	64	90	112	137	180	474	1180
phecode_416-8	8	11	12	19	13	15	13	21	17	29	158
phecode_417	1149	2057	3383	4359	5088	5562	6766	7707	9309	13440	58820
phecode_417-1	337	631	1249	1699	2139	2579	3139	3779	4755	7006	27313
phecode_417-2	294	385	424	526	603	625	783	1000	1248	2286	8174
phecode_417-3	310	400	499	660	836	977	1166	1486	2022	3566	11922
phecode_418	120	262	644	1054	1327	1691	2092	2676	3502	5902	19270
phecode_418-1	25	102	486	737	978	1247	1485	1939	2565	4400	13964

4 Medical history predicts future health trajectories over the human phenome

Table 15 continued from previous page

Endpoint	1	2	3	4	5	6	7	8	9	10	Total
phecode_419	59	128	182	236	292	367	432	682	1023	3123	6524
phecode_419-2	0	0	1	5	4	4	3	9	13	97	136
phecode_420	29	65	94	112	137	164	269	362	460	1238	2930
phecode_423	127	179	262	328	435	556	640	798	1139	1752	6216
phecode_423-1	137	146	227	272	392	462	533	665	989	1471	5294
phecode_424	167	439	650	842	1056	1327	1497	2188	3156	8110	19432
phecode_424-1	76	220	307	430	535	637	719	981	1509	3999	9413
phecode_424-2	5	21	61	107	145	214	264	348	491	1518	3174
phecode_424-3	3	13	36	51	60	108	134	181	292	620	1498
phecode_424-5	0	0	0	3	5	5	3	13	15	62	106
phecode_424-6	2	6	3	7	6	10	24	24	38	162	282
phecode_425	144	332	467	578	660	793	1001	1346	1979	4340	11640
phecode_426	146	327	354	439	601	742	1004	1282	1841	4648	11384
phecode_430	133	191	223	287	348	428	474	477	621	1161	4343
phecode_430-1	94	100	105	134	143	177	199	229	237	406	1824
phecode_430-2	35	69	112	127	178	187	222	269	331	637	2167
phecode_430-3	21	28	47	57	69	76	102	104	176	339	1019
phecode_431	434	839	1159	1322	1494	1781	2308	2738	3621	5205	20901
phecode_431-1	284	516	688	847	1127	1209	1363	1777	2293	3683	13787
phecode_431-11	190	376	532	662	825	873	1222	1500	2018	3236	11434
phecode_431-12	112	154	183	230	286	309	369	392	459	840	3334
phecode_431-14	3	5	10	15	29	39	52	63	66	109	391
phecode_431-15	6	11	9	27	39	31	74	103	156	278	734
phecode_431-2	181	272	381	505	724	886	1144	1389	1837	3162	10481
phecode_433	171	350	506	682	789	944	1298	1701	2510	4364	13315
phecode_433-1	39	94	107	140	166	205	269	336	499	816	2671

Supplementary Tables

Table 16: Incident Event Ratio per 10% Quantile and Rate Ratio for all endpoints.

Endpoint	1	2	3	4	5	6	7	8	9	10	Rate Ratio
OMOP_4306655	0.32%	0.61%	0.96%	1.41%	2.40%	3.46%	5.35%	9.63%	16.32%	34.58%	107.25
phecode_001	0.03%	0.02%	0.04%	0.05%	0.05%	0.06%	0.09%	0.12%	0.12%	0.18%	6.77
phecode_002	0.33%	0.45%	0.59%	0.75%	0.82%	0.95%	1.18%	1.40%	1.96%	4.75%	14.22
phecode_002-1	0.27%	0.33%	0.45%	0.52%	0.65%	0.62%	0.78%	1.06%	1.41%	3.73%	14.02
phecode_003	0.79%	1.01%	1.21%	1.15%	1.36%	1.52%	1.68%	2.12%	2.83%	5.73%	7.25
phecode_004	0.43%	0.53%	0.60%	0.69%	0.77%	0.84%	0.78%	0.92%	1.39%	2.81%	6.50
phecode_004-1	0.07%	0.06%	0.09%	0.11%	0.13%	0.12%	0.17%	0.19%	0.22%	0.59%	8.91
phecode_004-2	0.03%	0.03%	0.04%	0.04%	0.04%	0.06%	0.08%	0.05%	0.08%	0.21%	6.87
phecode_004-3	0.01%	0.03%	0.02%	0.03%	0.04%	0.05%	0.07%	0.09%	0.08%	0.35%	25.14
phecode_005	0.05%	0.09%	0.07%	0.08%	0.10%	0.13%	0.14%	0.16%	0.24%	0.51%	9.85
phecode_005-1	0.05%	0.04%	0.04%	0.07%	0.07%	0.09%	0.13%	0.11%	0.16%	0.42%	8.79
phecode_005-2	0.01%	0.02%	0.02%	0.01%	0.03%	0.02%	0.04%	0.05%	0.06%	0.16%	11.71
phecode_007	0.04%	0.04%	0.07%	0.08%	0.09%	0.14%	0.11%	0.15%	0.24%	0.89%	21.38
phecode_007-1	0.04%	0.04%	0.06%	0.07%	0.10%	0.13%	0.12%	0.14%	0.24%	0.88%	22.15
phecode_008	0.02%	0.08%	0.32%	0.44%	0.49%	0.63%	0.78%	0.88%	1.11%	1.78%	74.00
phecode_009	0.08%	0.08%	0.11%	0.12%	0.19%	0.26%	0.36%	0.40%	0.57%	1.97%	24.15
phecode_010	0.01%	0.03%	0.06%	0.07%	0.06%	0.07%	0.09%	0.13%	0.17%	0.44%	55.50
phecode_011	0.06%	0.05%	0.08%	0.11%	0.13%	0.19%	0.28%	0.28%	0.40%	1.07%	19.25
phecode_012	0.03%	0.03%	0.03%	0.03%	0.06%	0.11%	0.12%	0.13%	0.27%	0.77%	29.62
phecode_015	0.13%	0.12%	0.21%	0.22%	0.27%	0.32%	0.31%	0.30%	0.57%	1.24%	9.38
phecode_015-2	0.10%	0.09%	0.16%	0.18%	0.22%	0.28%	0.29%	0.30%	0.50%	1.19%	11.48
phecode_019	0.01%	0.01%	0.01%	0.00%	0.02%	0.03%	0.03%	0.04%	0.03%	0.16%	20.50
phecode_020	0.02%	0.03%	0.10%	0.08%	0.14%	0.12%	0.16%	0.20%	0.24%	0.33%	13.75
phecode_020-1	0.02%	0.03%	0.10%	0.08%	0.14%	0.11%	0.17%	0.20%	0.25%	0.32%	13.42
phecode_024	0.00%	0.01%	0.05%	0.04%	0.08%	0.06%	0.06%	0.10%	0.11%	0.13%	33.00
phecode_025	0.04%	0.03%	0.05%	0.06%	0.08%	0.12%	0.14%	0.15%	0.25%	0.54%	12.41
phecode_030	0.11%	0.20%	0.33%	0.47%	0.56%	0.69%	0.93%	1.07%	1.33%	2.03%	18.09
phecode_050	0.02%	0.03%	0.04%	0.04%	0.06%	0.08%	0.06%	0.05%	0.06%	0.11%	6.75
phecode_050-4	0.00%	0.00%	0.02%	0.02%	0.04%	0.05%	0.04%	0.06%	0.07%	0.06%	28.00
phecode_052	0.45%	0.98%	4.53%	5.92%	6.74%	7.60%	8.37%	9.64%	10.64%	13.70%	30.68
phecode_052-1	0.13%	0.29%	0.58%	0.85%	1.01%	1.18%	1.53%	1.90%	2.47%	3.96%	30.44
phecode_052-3	0.27%	0.77%	3.77%	4.97%	5.87%	6.38%	7.24%	7.80%	8.87%	11.08%	40.69
phecode_052-31	0.02%	0.04%	0.09%	0.11%	0.14%	0.14%	0.19%	0.15%	0.15%	0.26%	16.38
phecode_052-32	0.26%	0.80%	3.64%	4.78%	5.72%	6.28%	7.07%	7.68%	8.71%	10.93%	42.01
phecode_052-4	0.02%	0.03%	0.02%	0.04%	0.04%	0.04%	0.05%	0.04%	0.08%	0.08%	4.20
phecode_052-5	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.03%	0.06%	0.07%	0.18%	6.13
phecode_054	0.09%	0.12%	0.12%	0.12%	0.14%	0.18%	0.15%	0.18%	0.19%	0.54%	5.96
phecode_054-2	0.01%	0.03%	0.03%	0.02%	0.04%	0.04%	0.07%	0.04%	0.08%	0.28%	23.00
phecode_054-3	0.01%	0.02%	0.03%	0.02%	0.03%	0.04%	0.04%	0.03%	0.05%	0.27%	23.00
phecode_054-31	0.01%	0.03%	0.03%	0.03%	0.03%	0.04%	0.03%	0.03%	0.05%	0.29%	29.40
phecode_054-5	0.02%	0.01%	0.02%	0.02%	0.02%	0.03%	0.02%	0.02%	0.03%	0.03%	1.88
phecode_055	0.01%	0.02%	0.04%	0.06%	0.08%	0.08%	0.10%	0.13%	0.14%	0.20%	14.00
phecode_055-1	0.01%	0.02%	0.04%	0.06%	0.07%	0.07%	0.10%	0.12%	0.13%	0.19%	13.71
phecode_056	0.17%	0.39%	2.34%	3.07%	3.84%	4.13%	4.80%	5.60%	6.42%	8.52%	51.00
phecode_056-1	0.02%	0.12%	0.67%	0.88%	1.17%	1.32%	1.55%	1.93%	2.08%	2.96%	121.59
phecode_057	0.00%	0.01%	0.01%	0.01%	0.01%	0.02%	0.04%	0.05%	0.06%	0.18%	Inf
phecode_057-1	0.00%	0.01%	0.01%	0.02%	0.00%	0.02%	0.04%	0.05%	0.06%	0.18%	Inf
phecode_058	0.01%	0.03%	0.03%	0.02%	0.03%	0.05%	0.05%	0.04%	0.10%	0.20%	17.17
phecode_058-1	0.01%	0.02%	0.03%	0.02%	0.03%	0.04%	0.04%	0.04%	0.08%	0.20%	16.33
phecode_059	0.48%	0.52%	0.74%	1.02%	1.18%	1.49%	1.85%	2.17%	2.66%	4.06%	8.50
phecode_059-1	0.45%	0.52%	0.69%	0.98%	1.17%	1.41%	1.79%	2.17%	2.56%	4.01%	8.91
phecode_060	0.02%	0.03%	0.01%	0.02%	0.03%	0.02%	0.02%	0.02%	0.04%	0.05%	2.70
phecode_061	0.30%	0.37%	0.41%	0.50%	0.55%	0.61%	0.76%	0.89%	1.23%	2.25%	7.52
phecode_066	0.00%	0.00%	0.02%	0.02%	0.02%	0.02%	0.03%	0.04%	0.05%	0.08%	21.00
phecode_069	0.01%	0.02%	0.02%	0.03%	0.02%	0.03%	0.03%	0.04%	0.06%	0.14%	14.00
phecode_070	1.13%	1.60%	2.45%	3.19%	3.94%	4.76%	6.12%	7.73%	10.09%	17.27%	15.35
phecode_074	0.01%	0.02%	0.02%	0.01%	0.03%	0.06%	0.05%	0.06%	0.08%	0.47%	33.86
phecode_076	0.00%	0.01%	0.02%	0.01%	0.02%	0.04%	0.04%	0.04%	0.04%	0.16%	40.00
phecode_084	0.05%	0.08%	0.22%	0.32%	0.34%	0.40%	0.55%	0.65%	0.67%	1.19%	22.77
phecode_084-2	0.01%	0.01%	0.02%	0.02%	0.04%	0.05%	0.05%	0.06%	0.10%	0.26%	22.00
phecode_084-4	0.00%	0.00%	0.00%	0.00%	0.01%	0.02%	0.03%	0.08%	0.16%	0.46%	Inf
phecode_084-6	0.03%	0.03%	0.06%	0.09%	0.09%	0.12%	0.13%	0.14%	0.16%	0.24%	9.15
phecode_084-7	0.01%	0.01%	0.04%	0.06%	0.06%	0.10%	0.11%	0.13%	0.12%	0.17%	11.86
phecode_086	0.02%	0.07%	0.24%	0.26%	0.32%	0.36%	0.40%	0.53%	0.55%	1.10%	49.91

4 Medical history predicts future health trajectories over the human phenome

Table 16 continued from previous page

Endpoint	1	2	3	4	5	6	7	8	9	10	Rate Ratio
phecode_088	0.04%	0.06%	0.20%	0.24%	0.23%	0.31%	0.45%	0.59%	0.86%	1.48%	34.76
phecode_089	7.29%	8.70%	18.01%	25.90%	30.32%	34.16%	37.11%	41.49%	46.20%	56.56%	7.76
phecode_089-1	3.91%	4.71%	5.20%	5.35%	6.24%	7.00%	8.21%	9.82%	11.70%	19.48%	4.98
phecode_089-2	2.49%	3.25%	10.37%	15.07%	17.33%	19.77%	21.40%	24.29%	27.87%	34.65%	13.89
phecode_089-3	1.52%	2.65%	8.22%	11.12%	13.21%	14.88%	16.89%	19.07%	22.61%	30.57%	20.09
phecode_091	0.02%	0.02%	0.03%	0.02%	0.02%	0.03%	0.05%	0.07%	0.09%	0.24%	9.92
phecode_092	0.75%	1.08%	1.33%	1.75%	1.97%	2.54%	3.35%	3.89%	5.28%	11.13%	14.94
phecode_092-1	0.02%	0.04%	0.06%	0.07%	0.08%	0.09%	0.13%	0.16%	0.24%	0.63%	26.58
phecode_092-2	0.72%	1.07%	1.33%	1.69%	1.94%	2.49%	3.31%	3.84%	5.20%	10.96%	15.15
phecode_092-8	0.00%	0.00%	0.00%	0.02%	0.01%	0.02%	0.02%	0.03%	0.04%	0.11%	Inf
phecode_095	0.14%	0.18%	0.18%	0.26%	0.30%	0.30%	0.39%	0.52%	0.75%	1.50%	10.71
phecode_096	0.08%	0.15%	0.32%	0.45%	0.51%	0.57%	0.66%	0.76%	1.06%	1.78%	22.85
phecode_097	0.01%	0.02%	0.08%	0.12%	0.12%	0.13%	0.24%	0.26%	0.44%	1.02%	171.34
phecode_097-1	0.01%	0.01%	0.07%	0.09%	0.11%	0.13%	0.20%	0.24%	0.42%	0.90%	151.34
phecode_098	0.14%	0.22%	0.27%	0.30%	0.40%	0.46%	0.53%	0.59%	0.95%	2.36%	16.68
phecode_098-2	0.00%	0.00%	0.02%	0.01%	0.02%	0.02%	0.01%	0.03%	0.06%	0.14%	Inf
phecode_099	0.09%	0.19%	0.49%	0.72%	0.93%	1.29%	1.51%	2.10%	2.88%	4.71%	54.37
phecode_100	0.17%	0.22%	0.25%	0.34%	0.34%	0.36%	0.49%	0.54%	0.62%	1.21%	7.00
phecode_100-1	0.05%	0.09%	0.08%	0.11%	0.14%	0.13%	0.15%	0.13%	0.20%	0.41%	8.58
phecode_100-12	0.04%	0.05%	0.06%	0.07%	0.12%	0.08%	0.09%	0.08%	0.13%	0.24%	5.81
phecode_100-2	0.03%	0.02%	0.04%	0.03%	0.07%	0.09%	0.10%	0.10%	0.13%	0.25%	8.53
phecode_100-5	0.01%	0.01%	0.01%	0.02%	0.03%	0.02%	0.02%	0.03%	0.03%	0.03%	2.43
phecode_100-6	0.00%	0.01%	0.01%	0.02%	0.02%	0.04%	0.06%	0.09%	0.14%	0.32%	81.50
phecode_100-7	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.02%	0.03%	0.13%	16.75
phecode_100-8	0.01%	0.02%	0.00%	0.02%	0.02%	0.01%	0.02%	0.03%	0.03%	0.06%	5.80
phecode_100-9	0.03%	0.04%	0.03%	0.04%	0.03%	0.04%	0.04%	0.04%	0.04%	0.07%	2.18
phecode_101	0.93%	1.24%	1.74%	2.02%	2.50%	2.86%	3.47%	4.39%	4.83%	7.20%	7.75
phecode_101-1	0.03%	0.05%	0.10%	0.12%	0.19%	0.23%	0.29%	0.51%	0.61%	1.15%	38.67
phecode_101-2	0.03%	0.09%	0.10%	0.16%	0.17%	0.23%	0.27%	0.41%	0.48%	0.86%	27.06
phecode_101-21	0.01%	0.02%	0.03%	0.05%	0.07%	0.10%	0.13%	0.21%	0.24%	0.43%	31.14
phecode_101-3	0.03%	0.04%	0.06%	0.07%	0.06%	0.12%	0.10%	0.11%	0.13%	0.26%	8.80
phecode_101-4	0.55%	0.73%	1.04%	1.23%	1.42%	1.65%	2.01%	2.21%	2.39%	2.89%	5.22
phecode_101-41	0.39%	0.53%	0.76%	0.89%	1.07%	1.26%	1.50%	1.66%	1.77%	2.25%	5.71
phecode_101-42	0.22%	0.25%	0.35%	0.42%	0.44%	0.51%	0.59%	0.71%	0.78%	1.02%	4.65
phecode_101-5	0.03%	0.05%	0.05%	0.05%	0.07%	0.08%	0.06%	0.08%	0.12%	0.19%	7.23
phecode_101-6	0.04%	0.05%	0.05%	0.08%	0.10%	0.13%	0.18%	0.29%	0.42%	0.85%	22.37
phecode_101-61	0.01%	0.01%	0.02%	0.01%	0.02%	0.02%	0.07%	0.12%	0.16%	0.55%	39.57
phecode_101-62	0.02%	0.02%	0.05%	0.04%	0.05%	0.05%	0.10%	0.18%	0.25%	0.37%	18.50
phecode_101-7	0.02%	0.03%	0.04%	0.04%	0.06%	0.09%	0.11%	0.15%	0.19%	0.28%	11.75
phecode_101-71	0.00%	0.00%	0.01%	0.00%	0.03%	0.03%	0.04%	0.05%	0.07%	0.12%	30.50
phecode_101-8	0.06%	0.06%	0.11%	0.14%	0.24%	0.24%	0.37%	0.53%	0.82%	1.21%	19.03
phecode_102	0.17%	0.23%	0.30%	0.40%	0.61%	0.79%	1.17%	1.88%	2.45%	5.58%	33.31
phecode_102-1	0.12%	0.17%	0.23%	0.30%	0.46%	0.63%	0.98%	1.61%	2.23%	4.88%	40.16
phecode_102-3	0.01%	0.01%	0.02%	0.03%	0.04%	0.02%	0.08%	0.08%	0.09%	0.24%	17.14
phecode_102-5	0.01%	0.01%	0.01%	0.02%	0.03%	0.03%	0.03%	0.04%	0.07%	0.12%	12.40
phecode_103	2.14%	2.60%	3.24%	3.97%	4.78%	5.81%	7.17%	8.27%	10.16%	15.39%	7.21
phecode_103-1	0.43%	0.62%	0.66%	0.77%	0.79%	0.89%	1.02%	1.15%	1.43%	2.25%	5.19
phecode_103-2	0.06%	0.25%	1.13%	1.75%	2.21%	3.02%	3.68%	4.52%	6.22%	10.69%	171.16
phecode_103-21	0.05%	0.20%	0.98%	1.49%	1.92%	2.68%	3.22%	3.99%	5.56%	9.63%	177.08
phecode_103-22	0.01%	0.05%	0.10%	0.24%	0.26%	0.38%	0.50%	0.72%	0.94%	2.03%	145.43
phecode_103-3	0.11%	0.17%	0.26%	0.41%	0.53%	0.67%	1.02%	1.40%	1.73%	3.38%	30.27
phecode_104	0.22%	0.25%	0.35%	0.34%	0.38%	0.39%	0.50%	0.49%	0.56%	0.83%	3.72
phecode_104-1	0.04%	0.05%	0.08%	0.06%	0.08%	0.08%	0.07%	0.09%	0.13%	0.18%	4.60
phecode_104-2	0.02%	0.02%	0.02%	0.04%	0.04%	0.10%	0.11%	0.17%	0.17%	0.27%	15.11
phecode_104-3	0.10%	0.10%	0.16%	0.15%	0.16%	0.19%	0.20%	0.24%	0.27%	0.40%	3.96
phecode_104-5	0.00%	0.00%	0.02%	0.03%	0.04%	0.03%	0.04%	0.05%	0.05%	0.07%	Inf
phecode_105	0.07%	0.08%	0.09%	0.54%	2.83%	3.30%	3.97%	4.22%	4.88%	5.91%	85.38
phecode_105-1	0.56%	1.10%	3.27%	3.53%	3.68%	3.66%	3.88%	4.36%	4.40%	5.29%	9.42
phecode_106	1.27%	1.09%	1.31%	1.48%	1.53%	1.45%	1.58%	1.68%	1.84%	2.55%	2.00
phecode_106-1	0.21%	0.25%	0.26%	0.26%	0.28%	0.27%	0.28%	0.38%	0.41%	0.58%	2.79
phecode_106-11	0.07%	0.07%	0.08%	0.06%	0.08%	0.10%	0.12%	0.11%	0.11%	0.27%	4.11
phecode_106-13	0.07%	0.11%	0.16%	0.14%	0.19%	0.16%	0.21%	0.24%	0.26%	0.44%	6.56
phecode_106-2	0.39%	0.40%	0.48%	0.62%	0.65%	0.58%	0.77%	0.82%	0.95%	1.56%	4.05
phecode_106-21	0.36%	0.37%	0.45%	0.63%	0.57%	0.56%	0.75%	0.79%	0.90%	1.52%	4.23
phecode_106-3	0.30%	0.45%	0.40%	0.44%	0.52%	0.63%	0.74%	0.70%	0.76%	1.08%	3.54

Supplementary Tables

Table 16 continued from previous page

Endpoint	1	2	3	4	5	6	7	8	9	10	Rate Ratio
phcode_106-4	0.02%	0.03%	0.05%	0.06%	0.08%	0.06%	0.08%	0.07%	0.08%	0.10%	5.60
phcode_107	3.36%	3.80%	3.41%	4.37%	5.51%	6.00%	6.68%	7.71%	9.07%	11.36%	3.38
phcode_107-1	0.03%	0.07%	0.05%	0.06%	0.08%	0.14%	0.14%	0.10%	0.21%	0.58%	22.00
phcode_107-2	3.10%	3.76%	3.26%	4.28%	5.23%	5.85%	6.44%	7.67%	8.67%	10.99%	3.54
phcode_107-3	0.04%	0.06%	0.08%	0.06%	0.07%	0.07%	0.07%	0.09%	0.07%	0.11%	2.40
phcode_108	0.26%	0.40%	0.58%	0.78%	0.94%	1.18%	1.46%	1.68%	2.24%	3.80%	14.74
phcode_108-4	0.09%	0.15%	0.20%	0.23%	0.33%	0.41%	0.53%	0.55%	0.70%	1.20%	13.68
phcode_108-41	0.08%	0.13%	0.19%	0.20%	0.29%	0.38%	0.48%	0.50%	0.64%	1.10%	13.14
phcode_108-42	0.00%	0.00%	0.01%	0.01%	0.03%	0.02%	0.05%	0.04%	0.05%	0.13%	33.00
phcode_108-5	0.15%	0.22%	0.32%	0.40%	0.55%	0.64%	0.73%	0.99%	1.31%	2.17%	14.51
phcode_108-7	0.01%	0.00%	0.01%	0.03%	0.03%	0.04%	0.06%	0.04%	0.08%	0.21%	34.67
phcode_109	0.09%	0.14%	0.15%	0.15%	0.25%	0.24%	0.32%	0.45%	0.64%	0.88%	10.26
phcode_109-1	0.01%	0.03%	0.02%	0.03%	0.04%	0.04%	0.05%	0.04%	0.06%	0.06%	5.17
phcode_109-16	0.01%	0.02%	0.01%	0.02%	0.01%	0.03%	0.04%	0.03%	0.03%	0.03%	2.29
phcode_109-3	0.07%	0.08%	0.09%	0.11%	0.17%	0.21%	0.25%	0.33%	0.57%	0.76%	11.58
phcode_110	0.08%	0.14%	0.14%	0.12%	0.14%	0.13%	0.17%	0.20%	0.25%	0.44%	5.72
phcode_110-1	0.05%	0.07%	0.07%	0.09%	0.09%	0.10%	0.11%	0.14%	0.17%	0.26%	5.28
phcode_110-4	0.01%	0.02%	0.01%	0.02%	0.01%	0.03%	0.02%	0.03%	0.04%	0.09%	7.83
phcode_112	2.48%	3.36%	4.31%	4.90%	5.55%	6.48%	7.79%	9.58%	12.38%	20.31%	8.20
phcode_112-1	0.01%	0.01%	0.01%	0.01%	0.03%	0.05%	0.08%	0.15%	0.24%	0.45%	32.14
phcode_114	0.04%	0.08%	0.07%	0.12%	0.12%	0.16%	0.21%	0.24%	0.34%	0.61%	14.62
phcode_114-4	0.01%	0.00%	0.01%	0.02%	0.02%	0.03%	0.03%	0.03%	0.06%	0.06%	10.67
phcode_114-6	0.02%	0.05%	0.06%	0.08%	0.07%	0.10%	0.15%	0.16%	0.23%	0.52%	21.75
phcode_116	1.19%	1.64%	1.71%	2.14%	2.52%	3.18%	3.95%	5.46%	7.86%	12.76%	10.71
phcode_116-1	0.72%	1.06%	1.11%	1.30%	1.50%	1.79%	2.00%	2.66%	3.48%	5.25%	7.28
phcode_116-2	0.22%	0.25%	0.35%	0.36%	0.55%	0.80%	1.09%	1.63%	2.62%	5.12%	23.19
phcode_116-3	0.15%	0.19%	0.32%	0.32%	0.46%	0.59%	0.81%	1.16%	1.75%	2.86%	19.13
phcode_116-4	0.23%	0.29%	0.37%	0.49%	0.64%	0.96%	1.28%	1.75%	3.07%	5.39%	23.34
phcode_116-5	0.07%	0.07%	0.08%	0.12%	0.18%	0.25%	0.36%	0.59%	0.95%	1.88%	26.97
phcode_116-6	0.20%	0.22%	0.30%	0.40%	0.49%	0.74%	1.01%	1.55%	2.48%	4.86%	24.89
phcode_116-7	0.02%	0.02%	0.03%	0.04%	0.04%	0.04%	0.07%	0.15%	0.19%	0.35%	14.58
phcode_120	0.76%	0.98%	1.12%	1.25%	1.39%	1.63%	1.78%	1.99%	2.12%	3.27%	4.28
phcode_120-1	0.33%	0.44%	0.45%	0.53%	0.55%	0.66%	0.79%	0.92%	0.94%	1.86%	5.72
phcode_120-11	0.01%	0.00%	0.01%	0.02%	0.03%	0.02%	0.04%	0.05%	0.06%	0.11%	13.25
phcode_120-12	0.00%	0.00%	0.01%	0.02%	0.03%	0.02%	0.02%	0.06%	0.05%	0.08%	20.00
phcode_120-13	0.02%	0.05%	0.05%	0.07%	0.05%	0.07%	0.06%	0.08%	0.10%	0.26%	14.67
phcode_120-2	0.32%	0.46%	0.57%	0.63%	0.73%	0.88%	0.91%	0.95%	0.99%	1.53%	4.73
phcode_120-21	0.27%	0.42%	0.51%	0.56%	0.68%	0.76%	0.87%	0.81%	0.89%	1.38%	5.16
phcode_120-22	0.03%	0.02%	0.03%	0.05%	0.06%	0.05%	0.06%	0.09%	0.07%	0.15%	4.56
phcode_121	0.13%	0.23%	0.28%	0.35%	0.37%	0.40%	0.53%	0.51%	0.60%	1.02%	7.76
phcode_121-1	0.04%	0.05%	0.06%	0.10%	0.08%	0.09%	0.12%	0.15%	0.31%	0.53%	14.72
phcode_121-11	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.02%	0.04%	0.07%	8.75
phcode_121-12	0.03%	0.03%	0.05%	0.08%	0.07%	0.08%	0.09%	0.16%	0.26%	0.47%	15.73
phcode_121-2	0.08%	0.15%	0.19%	0.24%	0.27%	0.31%	0.33%	0.32%	0.38%	0.54%	6.80
phcode_121-21	0.06%	0.12%	0.16%	0.17%	0.23%	0.27%	0.26%	0.27%	0.31%	0.49%	8.45
phcode_121-22	0.01%	0.04%	0.02%	0.02%	0.03%	0.04%	0.04%	0.06%	0.04%	0.11%	11.00
phcode_121-23	0.00%	0.00%	0.01%	0.02%	0.02%	0.03%	0.02%	0.04%	0.04%	0.06%	15.50
phcode_122	0.28%	0.38%	0.44%	0.51%	0.55%	0.65%	0.73%	0.76%	0.76%	1.27%	4.49
phcode_122-1	0.04%	0.04%	0.03%	0.04%	0.04%	0.04%	0.06%	0.05%	0.07%	0.18%	4.40
phcode_122-2	0.24%	0.34%	0.39%	0.45%	0.53%	0.60%	0.68%	0.72%	0.72%	1.21%	5.06
phcode_122-21	0.08%	0.11%	0.11%	0.11%	0.13%	0.15%	0.15%	0.13%	0.15%	0.22%	2.70
phcode_122-22	0.06%	0.12%	0.13%	0.17%	0.15%	0.22%	0.30%	0.24%	0.33%	0.58%	9.36
phcode_122-24	0.02%	0.01%	0.01%	0.02%	0.03%	0.03%	0.03%	0.04%	0.06%	0.07%	4.38
phcode_123	0.10%	0.11%	0.15%	0.15%	0.23%	0.21%	0.26%	0.34%	0.35%	0.60%	5.86
phcode_123-1	0.10%	0.11%	0.14%	0.16%	0.22%	0.19%	0.26%	0.32%	0.36%	0.58%	5.69
phcode_124	0.20%	0.26%	0.27%	0.31%	0.31%	0.38%	0.45%	0.47%	0.56%	1.08%	5.51
phcode_124-3	0.02%	0.04%	0.05%	0.07%	0.06%	0.06%	0.06%	0.08%	0.08%	0.26%	14.56
phcode_124-5	0.09%	0.11%	0.10%	0.10%	0.10%	0.15%	0.16%	0.18%	0.23%	0.49%	5.47
phcode_124-6	0.02%	0.04%	0.06%	0.07%	0.07%	0.10%	0.12%	0.18%	0.19%	0.55%	22.83
phcode_124-7	0.04%	0.05%	0.06%	0.05%	0.09%	0.07%	0.08%	0.08%	0.11%	0.25%	7.06
phcode_125	0.04%	0.03%	0.03%	0.07%	0.04%	0.04%	0.07%	0.07%	0.08%	0.17%	3.95
phcode_130	7.21%	8.48%	10.20%	11.89%	13.60%	15.11%	15.96%	17.91%	20.85%	27.97%	3.88
phcode_132	0.05%	0.12%	0.20%	0.47%	0.69%	0.78%	0.81%	0.99%	1.24%	5.62%	104.30
phcode_135	0.47%	0.58%	0.65%	0.80%	0.91%	1.15%	1.21%	1.43%	1.80%	2.89%	6.11
phcode_135-1	0.13%	0.16%	0.23%	0.24%	0.25%	0.27%	0.31%	0.33%	0.38%	0.51%	3.85

4 Medical history predicts future health trajectories over the human phenome

Table 16 continued from previous page

Endpoint	1	2	3	4	5	6	7	8	9	10	Rate Ratio
phecode_135-12	0.02%	0.02%	0.02%	0.02%	0.03%	0.02%	0.04%	0.03%	0.03%	0.02%	1.09
phecode_135-16	0.09%	0.08%	0.07%	0.11%	0.11%	0.13%	0.12%	0.12%	0.17%	0.28%	3.18
phecode_135-5	0.27%	0.36%	0.42%	0.50%	0.58%	0.72%	0.88%	0.99%	1.33%	2.42%	9.02
phecode_135-6	0.01%	0.04%	0.05%	0.03%	0.04%	0.06%	0.06%	0.06%	0.10%	0.17%	28.67
phecode_136	4.62%	5.81%	6.14%	7.81%	8.83%	10.07%	11.39%	13.32%	16.82%	22.58%	4.89
phecode_136-1	0.01%	0.01%	0.01%	0.03%	0.03%	0.01%	0.03%	0.03%	0.07%	0.08%	6.83
phecode_136-2	0.51%	0.71%	0.86%	0.95%	1.31%	1.62%	2.04%	2.71%	3.70%	8.13%	15.94
phecode_136-3	0.01%	0.02%	0.02%	0.02%	0.03%	0.04%	0.07%	0.08%	0.08%	0.14%	10.29
phecode_136-4	3.98%	5.06%	5.30%	6.74%	7.52%	8.51%	9.56%	11.19%	13.90%	18.40%	4.62
phecode_136-41	3.13%	4.08%	4.51%	5.33%	6.24%	7.11%	7.95%	9.58%	11.91%	16.15%	5.16
phecode_136-42	1.53%	1.76%	2.06%	2.61%	2.72%	3.16%	3.67%	4.47%	5.45%	7.87%	5.15
phecode_136-6	0.00%	0.03%	0.04%	0.03%	0.02%	0.02%	0.03%	0.06%	0.05%	0.07%	17.00
phecode_136-8	0.01%	0.02%	0.01%	0.03%	0.03%	0.03%	0.04%	0.04%	0.06%	0.04%	3.33
phecode_137	0.02%	0.02%	0.04%	0.04%	0.03%	0.04%	0.05%	0.08%	0.08%	0.09%	4.09
phecode_137-5	0.01%	0.01%	0.04%	0.03%	0.03%	0.02%	0.03%	0.04%	0.03%	0.06%	5.60
phecode_138	0.84%	1.14%	3.93%	5.32%	6.45%	7.87%	9.38%	10.59%	12.86%	17.21%	20.52
phecode_138-1	0.13%	0.25%	0.38%	0.49%	0.62%	0.81%	0.84%	0.95%	1.14%	1.66%	12.39
phecode_138-2	0.42%	0.62%	2.65%	3.55%	4.73%	5.73%	7.01%	8.33%	9.99%	13.97%	33.10
phecode_139	1.16%	1.67%	2.91%	3.36%	3.73%	4.09%	4.61%	4.80%	5.87%	6.91%	5.94
phecode_139-1	0.04%	0.04%	0.06%	0.07%	0.06%	0.08%	0.09%	0.10%	0.12%	0.20%	5.56
phecode_139-3	0.08%	0.07%	0.16%	0.18%	0.17%	0.23%	0.26%	0.28%	0.31%	0.48%	6.10
phecode_139-4	0.07%	0.07%	0.06%	0.08%	0.07%	0.09%	0.09%	0.13%	0.15%	0.20%	2.91
phecode_139-5	0.77%	1.10%	1.95%	2.46%	2.72%	2.96%	3.11%	3.55%	4.02%	4.96%	6.43
phecode_139-51	0.01%	0.02%	0.02%	0.02%	0.02%	0.02%	0.03%	0.03%	0.05%	0.08%	13.67
phecode_139-52	0.04%	0.08%	0.13%	0.12%	0.17%	0.20%	0.26%	0.26%	0.31%	0.46%	11.55
phecode_139-53	0.46%	0.54%	0.68%	0.85%	0.84%	0.96%	1.10%	1.14%	1.38%	1.78%	3.92
phecode_139-54	0.28%	0.28%	0.34%	0.44%	0.43%	0.50%	0.59%	0.57%	0.62%	0.85%	3.10
phecode_139-6	0.26%	0.35%	0.59%	0.74%	0.70%	0.85%	1.06%	1.09%	1.44%	1.76%	6.82
phecode_139-61	0.25%	0.31%	0.57%	0.74%	0.69%	0.85%	1.00%	1.07%	1.38%	1.78%	7.15
phecode_139-62	0.01%	0.01%	0.01%	0.02%	0.02%	0.02%	0.01%	0.03%	0.04%	0.05%	4.17
phecode_140	0.01%	0.00%	0.01%	0.03%	0.12%	0.29%	0.34%	0.54%	0.65%	0.97%	96.00
phecode_142	0.19%	0.31%	0.55%	0.66%	0.75%	1.36%	7.03%	8.66%	11.03%	14.99%	77.57
phecode_142-1	0.13%	0.26%	0.49%	0.66%	0.74%	1.35%	3.26%	4.64%	6.80%	10.58%	82.02
phecode_142-2	0.00%	0.00%	0.00%	0.00%	0.00%	0.02%	0.06%	0.06%	0.10%	0.12%	Inf
phecode_142-21	0.00%	0.00%	0.00%	0.00%	0.00%	0.02%	0.06%	0.07%	0.10%	0.12%	Inf
phecode_144	3.92%	3.07%	3.26%	3.89%	4.46%	5.24%	6.90%	8.40%	10.02%	14.77%	3.77
phecode_144-1	0.58%	0.45%	0.57%	0.66%	0.88%	1.16%	1.44%	1.85%	2.25%	2.86%	4.88
phecode_144-11	0.04%	0.03%	0.02%	0.04%	0.03%	0.05%	0.06%	0.07%	0.06%	0.09%	2.00
phecode_144-12	0.03%	0.04%	0.05%	0.05%	0.08%	0.06%	0.07%	0.10%	0.11%	0.21%	7.00
phecode_144-13	0.52%	0.34%	0.47%	0.51%	0.78%	1.00%	1.31%	1.79%	2.09%	2.72%	5.23
phecode_144-2	3.07%	2.31%	2.59%	2.96%	3.41%	4.09%	5.15%	6.64%	8.07%	12.43%	4.04
phecode_144-21	1.57%	1.36%	1.52%	1.79%	2.12%	2.62%	3.29%	4.65%	6.48%	9.93%	6.32
phecode_144-3	0.47%	0.48%	0.46%	0.46%	0.46%	0.63%	0.66%	0.60%	0.69%	0.85%	1.83
phecode_146	0.04%	0.04%	0.06%	0.08%	0.08%	0.14%	0.26%	0.45%	2.76%	8.85%	201.28
phecode_146-2	0.13%	0.19%	0.16%	0.27%	0.25%	0.55%	1.96%	4.76%	7.15%	10.82%	84.73
phecode_146-4	0.01%	0.02%	0.02%	0.04%	0.02%	0.05%	0.05%	0.04%	0.08%	0.07%	4.71
phecode_146-5	0.01%	0.01%	0.02%	0.01%	0.02%	0.01%	0.03%	0.04%	0.04%	0.16%	15.80
phecode_148	0.20%	0.28%	0.47%	0.50%	0.66%	0.65%	0.70%	0.87%	1.01%	1.15%	5.78
phecode_148-1	0.02%	0.07%	0.21%	0.33%	0.35%	0.42%	0.42%	0.46%	0.50%	0.70%	29.08
phecode_148-16	0.01%	0.06%	0.21%	0.27%	0.31%	0.38%	0.37%	0.43%	0.48%	0.67%	111.34
phecode_148-2	0.08%	0.09%	0.12%	0.15%	0.12%	0.18%	0.23%	0.30%	0.28%	0.47%	6.10
phecode_148-3	0.01%	0.01%	0.00%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.10%	16.33
phecode_148-5	0.06%	0.04%	0.08%	0.07%	0.08%	0.10%	0.12%	0.09%	0.13%	0.14%	2.19
phecode_149	0.17%	0.19%	0.22%	0.22%	0.26%	0.28%	0.34%	0.32%	0.44%	0.80%	4.66
phecode_149-1	0.04%	0.03%	0.03%	0.03%	0.04%	0.08%	0.09%	0.08%	0.11%	0.14%	3.45
phecode_149-3	0.04%	0.07%	0.08%	0.09%	0.09%	0.15%	0.15%	0.22%	0.21%	0.41%	11.56
phecode_149-4	0.06%	0.08%	0.06%	0.08%	0.07%	0.08%	0.10%	0.13%	0.15%	0.28%	5.04
phecode_153	0.04%	0.05%	0.10%	0.10%	0.13%	0.14%	0.15%	0.12%	0.15%	0.18%	5.06
phecode_159	0.00%	0.01%	0.04%	0.04%	0.06%	0.04%	0.08%	0.07%	0.12%	0.25%	Inf
phecode_159-1	0.00%	0.01%	0.02%	0.05%	0.03%	0.04%	0.05%	0.07%	0.08%	0.19%	Inf
phecode_160	1.88%	2.35%	2.43%	3.04%	3.57%	4.16%	4.83%	6.25%	8.43%	16.50%	8.76
phecode_160-1	1.68%	2.06%	2.36%	2.68%	3.31%	3.65%	4.35%	5.67%	7.57%	15.02%	8.93
phecode_160-2	0.22%	0.24%	0.29%	0.29%	0.41%	0.46%	0.59%	0.89%	1.20%	2.87%	13.31
phecode_160-4	0.01%	0.03%	0.04%	0.05%	0.05%	0.06%	0.12%	0.18%	0.25%	0.51%	64.25
phecode_161	0.04%	0.03%	0.04%	0.04%	0.06%	0.07%	0.07%	0.07%	0.07%	0.26%	6.05

Supplementary Tables

Table 16 continued from previous page

Endpoint	1	2	3	4	5	6	7	8	9	10	Rate Ratio
phecode_161-1	0.02%	0.01%	0.02%	0.02%	0.04%	0.03%	0.03%	0.04%	0.03%	0.16%	7.80
phecode_161-2	0.03%	0.02%	0.03%	0.03%	0.04%	0.04%	0.05%	0.05%	0.06%	0.17%	5.19
phecode_161-21	0.02%	0.01%	0.02%	0.02%	0.03%	0.03%	0.03%	0.05%	0.04%	0.11%	4.50
phecode_162	0.05%	0.09%	0.08%	0.12%	0.16%	0.19%	0.35%	0.32%	0.37%	1.01%	21.21
phecode_162-8	0.01%	0.01%	0.02%	0.02%	0.04%	0.05%	0.08%	0.08%	0.12%	0.25%	31.50
phecode_164	3.72%	4.78%	5.21%	5.88%	6.15%	7.38%	8.59%	10.85%	14.20%	25.68%	6.91
phecode_164-1	1.69%	2.06%	2.35%	2.68%	3.32%	3.64%	4.34%	5.68%	7.56%	15.04%	8.91
phecode_164-2	0.24%	0.23%	0.30%	0.31%	0.43%	0.47%	0.62%	0.91%	1.29%	3.07%	12.89
phecode_164-3	0.02%	0.02%	0.02%	0.01%	0.03%	0.04%	0.04%	0.05%	0.05%	0.14%	6.90
phecode_164-6	0.06%	0.08%	0.09%	0.12%	0.13%	0.15%	0.23%	0.40%	0.47%	1.66%	29.61
phecode_164-62	0.02%	0.04%	0.03%	0.06%	0.07%	0.10%	0.07%	0.19%	0.32%	0.66%	33.10
phecode_165	0.05%	0.05%	0.06%	0.08%	0.07%	0.08%	0.12%	0.15%	0.19%	0.70%	14.71
phecode_165-2	0.03%	0.02%	0.02%	0.03%	0.04%	0.03%	0.04%	0.07%	0.10%	0.32%	11.43
phecode_165-25	0.02%	0.01%	0.02%	0.02%	0.03%	0.02%	0.03%	0.06%	0.10%	0.32%	14.64
phecode_165-3	0.01%	0.00%	0.01%	0.01%	0.03%	0.03%	0.03%	0.05%	0.08%	0.35%	29.50
phecode_168	0.30%	0.35%	0.54%	0.59%	0.79%	0.96%	1.19%	1.61%	2.07%	4.65%	15.34
phecode_168-1	0.10%	0.12%	0.26%	0.31%	0.44%	0.55%	0.66%	0.94%	1.30%	2.33%	24.23
phecode_168-11	0.02%	0.03%	0.01%	0.04%	0.02%	0.03%	0.04%	0.03%	0.05%	0.09%	5.50
phecode_168-12	0.00%	0.01%	0.00%	0.00%	0.00%	0.01%	0.01%	0.02%	0.05%	0.23%	Inf
phecode_168-15	0.00%	0.00%	0.01%	0.00%	0.00%	0.00%	0.00%	0.02%	0.02%	0.15%	Inf
phecode_168-18	0.02%	0.01%	0.04%	0.05%	0.10%	0.07%	0.14%	0.20%	0.27%	0.52%	32.38
phecode_168-19	0.05%	0.08%	0.13%	0.21%	0.27%	0.36%	0.43%	0.60%	0.83%	1.58%	31.64
phecode_168-2	0.07%	0.08%	0.09%	0.08%	0.10%	0.10%	0.10%	0.14%	0.14%	0.77%	11.09
phecode_168-21	0.05%	0.04%	0.07%	0.06%	0.08%	0.08%	0.07%	0.08%	0.11%	0.64%	12.92
phecode_168-211	0.00%	0.00%	0.01%	0.02%	0.02%	0.02%	0.02%	0.03%	0.04%	0.09%	Inf
phecode_168-214	0.00%	0.01%	0.01%	0.01%	0.01%	0.03%	0.02%	0.03%	0.04%	0.12%	Inf
phecode_168-3	0.02%	0.04%	0.03%	0.06%	0.04%	0.03%	0.07%	0.07%	0.07%	0.32%	13.42
phecode_168-4	0.00%	0.02%	0.04%	0.05%	0.07%	0.11%	0.15%	0.17%	0.30%	1.47%	739.01
phecode_169	0.33%	0.39%	0.52%	0.56%	0.69%	0.86%	1.02%	1.19%	1.44%	2.83%	8.63
phecode_169-1	0.32%	0.40%	0.51%	0.53%	0.70%	0.86%	1.03%	1.16%	1.44%	2.81%	8.75
phecode_169-11	0.06%	0.06%	0.09%	0.10%	0.10%	0.09%	0.14%	0.12%	0.15%	0.47%	8.10
phecode_169-14	0.02%	0.01%	0.01%	0.04%	0.03%	0.07%	0.06%	0.05%	0.07%	0.19%	8.08
phecode_170	0.82%	0.99%	1.14%	1.18%	1.33%	1.54%	1.59%	1.93%	1.96%	3.77%	4.62
phecode_170-1	0.75%	0.98%	1.09%	1.18%	1.26%	1.50%	1.52%	1.86%	1.97%	3.70%	4.93
phecode_170-13	0.00%	0.00%	0.02%	0.04%	0.03%	0.03%	0.09%	0.09%	0.11%	0.17%	Inf
phecode_170-19	0.01%	0.07%	0.29%	0.35%	0.37%	0.43%	0.52%	0.69%	0.82%	1.29%	129.40
phecode_170-2	0.01%	0.01%	0.02%	0.04%	0.04%	0.06%	0.05%	0.06%	0.07%	0.18%	30.00
phecode_171	0.03%	0.05%	0.10%	0.16%	0.20%	0.27%	0.30%	0.33%	0.48%	0.87%	25.59
phecode_171-1	0.01%	0.01%	0.05%	0.08%	0.12%	0.15%	0.17%	0.23%	0.26%	0.39%	66.00
phecode_171-7	0.03%	0.02%	0.05%	0.04%	0.06%	0.06%	0.06%	0.09%	0.14%	0.36%	10.59
phecode_171-9	0.00%	0.00%	0.00%	0.01%	0.02%	0.03%	0.02%	0.03%	0.05%	0.13%	64.00
phecode_172	0.13%	0.18%	0.23%	0.26%	0.32%	0.37%	0.42%	0.58%	0.64%	1.07%	8.22
phecode_172-2	0.00%	0.01%	0.01%	0.02%	0.03%	0.03%	0.04%	0.07%	0.09%	0.18%	46.00
phecode_174	0.16%	0.17%	0.21%	0.25%	0.28%	0.29%	0.36%	0.40%	0.56%	1.83%	11.65
phecode_174-1	0.03%	0.03%	0.05%	0.06%	0.04%	0.05%	0.08%	0.09%	0.10%	0.79%	30.69
phecode_174-2	0.07%	0.09%	0.11%	0.14%	0.15%	0.19%	0.18%	0.25%	0.35%	0.94%	13.46
phecode_174-6	0.02%	0.01%	0.01%	0.01%	0.02%	0.01%	0.01%	0.03%	0.03%	0.05%	2.27
phecode_174-7	0.01%	0.02%	0.03%	0.03%	0.03%	0.05%	0.04%	0.05%	0.07%	0.14%	11.50
phecode_175	0.06%	0.09%	0.11%	0.12%	0.14%	0.16%	0.18%	0.23%	0.32%	0.68%	11.72
phecode_175-2	0.04%	0.07%	0.08%	0.08%	0.09%	0.11%	0.17%	0.19%	0.27%	0.70%	17.60
phecode_176	0.08%	0.20%	0.27%	0.31%	0.36%	0.46%	0.55%	0.68%	0.92%	1.38%	17.20
phecode_177	1.29%	1.48%	1.90%	2.30%	2.42%	2.96%	3.34%	4.01%	4.80%	7.64%	5.91
phecode_177-1	0.03%	0.05%	0.07%	0.10%	0.11%	0.14%	0.13%	0.18%	0.24%	0.35%	13.46
phecode_177-13	0.00%	0.02%	0.04%	0.05%	0.05%	0.08%	0.07%	0.06%	0.10%	0.16%	41.00
phecode_177-2	1.00%	1.11%	1.51%	1.78%	1.95%	2.24%	2.66%	3.13%	3.68%	5.17%	5.18
phecode_177-3	0.03%	0.02%	0.05%	0.04%	0.03%	0.05%	0.08%	0.08%	0.08%	0.19%	5.65
phecode_177-4	0.14%	0.18%	0.24%	0.26%	0.31%	0.40%	0.52%	0.68%	0.93%	2.67%	19.69
phecode_179	0.05%	0.09%	0.16%	0.20%	0.21%	0.19%	0.22%	0.28%	0.36%	1.40%	27.08
phecode_179-1	0.02%	0.03%	0.01%	0.03%	0.03%	0.04%	0.03%	0.05%	0.06%	0.31%	17.22
phecode_179-9	0.01%	0.03%	0.09%	0.11%	0.13%	0.13%	0.15%	0.17%	0.23%	1.05%	105.20
phecode_180	0.20%	0.22%	0.29%	0.41%	0.47%	0.53%	0.68%	0.82%	1.03%	2.18%	11.03
phecode_180-3	0.07%	0.15%	0.15%	0.21%	0.28%	0.34%	0.44%	0.53%	0.73%	1.22%	17.49
phecode_180-31	0.06%	0.15%	0.12%	0.19%	0.26%	0.31%	0.41%	0.49%	0.70%	1.15%	20.00
phecode_180-33	0.00%	0.01%	0.02%	0.03%	0.02%	0.02%	0.04%	0.05%	0.05%	0.10%	26.00
phecode_181	1.31%	1.56%	1.92%	1.88%	2.21%	2.50%	2.90%	3.45%	4.37%	7.93%	6.07

4 Medical history predicts future health trajectories over the human phenome

Table 16 continued from previous page

Endpoint	1	2	3	4	5	6	7	8	9	10	Rate Ratio
phecode_200	2.20%	2.57%	3.11%	3.46%	4.46%	5.55%	6.38%	7.26%	8.32%	15.71%	7.13
phecode_200-1	1.58%	1.98%	2.50%	2.63%	3.32%	4.19%	4.76%	5.32%	6.34%	16.00%	10.15
phecode_200-12	0.01%	0.00%	0.01%	0.01%	0.01%	0.01%	0.02%	0.03%	0.04%	0.14%	11.33
phecode_200-13	0.04%	0.09%	0.09%	0.14%	0.16%	0.20%	0.22%	0.25%	0.44%	2.66%	66.75
phecode_200-14	1.25%	1.61%	2.17%	2.30%	2.75%	3.50%	4.04%	4.52%	5.42%	15.11%	12.10
phecode_200-2	0.34%	0.42%	0.50%	0.57%	0.78%	1.08%	1.39%	1.66%	2.04%	2.75%	8.14
phecode_200-21	0.11%	0.10%	0.11%	0.13%	0.16%	0.20%	0.30%	0.36%	0.51%	0.83%	7.53
phecode_200-22	0.11%	0.15%	0.23%	0.29%	0.32%	0.40%	0.62%	0.73%	0.69%	1.22%	11.55
phecode_200-23	0.08%	0.10%	0.11%	0.18%	0.23%	0.34%	0.45%	0.52%	0.75%	1.06%	12.64
phecode_200-3	0.31%	0.33%	0.36%	0.45%	0.65%	0.72%	0.87%	1.10%	1.33%	2.61%	8.48
phecode_200-31	0.09%	0.10%	0.10%	0.11%	0.13%	0.17%	0.17%	0.23%	0.38%	0.66%	7.00
phecode_200-4	0.06%	0.06%	0.06%	0.08%	0.08%	0.15%	0.19%	0.27%	0.26%	0.66%	11.48
phecode_200-41	0.05%	0.03%	0.04%	0.04%	0.04%	0.09%	0.13%	0.14%	0.16%	0.53%	11.00
phecode_200-7	0.03%	0.13%	0.45%	0.49%	0.71%	0.78%	1.01%	0.97%	1.18%	1.45%	42.71
phecode_200-9	0.02%	0.07%	0.18%	0.20%	0.31%	0.42%	0.42%	0.61%	0.74%	1.42%	64.55
phecode_202	1.68%	2.87%	3.86%	4.58%	5.07%	6.16%	8.03%	10.02%	13.79%	23.41%	13.92
phecode_202-1	0.07%	0.10%	0.11%	0.16%	0.22%	0.28%	0.44%	0.46%	0.70%	4.91%	74.30
phecode_202-2	1.56%	2.82%	3.73%	4.48%	4.94%	6.06%	7.79%	9.95%	13.47%	23.48%	15.02
phecode_202-3	0.00%	0.00%	0.02%	0.02%	0.02%	0.02%	0.04%	0.06%	0.10%	0.11%	Inf
phecode_202-32	0.00%	0.00%	0.02%	0.01%	0.02%	0.02%	0.04%	0.04%	0.09%	0.10%	Inf
phecode_202-4	0.30%	0.66%	1.31%	2.23%	2.94%	4.04%	5.43%	7.41%	11.18%	41.50%	136.41
phecode_203	0.00%	0.01%	0.00%	0.01%	0.02%	0.03%	0.04%	0.05%	0.06%	0.16%	82.00
phecode_204	0.49%	1.16%	3.43%	4.74%	6.00%	7.35%	8.78%	11.03%	13.90%	19.15%	39.36
phecode_204-1	0.02%	0.08%	0.41%	0.61%	0.75%	1.07%	1.28%	1.59%	2.41%	3.83%	212.22
phecode_204-2	0.07%	0.23%	1.01%	1.42%	1.92%	2.46%	2.86%	4.03%	5.02%	8.23%	113.94
phecode_204-4	0.00%	0.00%	0.01%	0.01%	0.01%	0.03%	0.09%	0.15%	0.22%	0.62%	Inf
phecode_205	0.10%	0.13%	0.18%	0.17%	0.22%	0.27%	0.36%	0.50%	0.67%	5.46%	57.00
phecode_205-3	0.00%	0.00%	0.00%	0.00%	0.00%	0.01%	0.01%	0.01%	0.02%	0.23%	Inf
phecode_206	0.00%	0.00%	0.01%	0.01%	0.00%	0.01%	0.03%	0.04%	0.06%	0.10%	Inf
phecode_208	0.15%	0.19%	0.22%	0.29%	0.29%	0.50%	0.55%	0.69%	1.02%	1.98%	12.92
phecode_208-1	0.00%	0.01%	0.03%	0.02%	0.03%	0.03%	0.03%	0.02%	0.07%	0.16%	41.00
phecode_208-2	0.14%	0.18%	0.19%	0.26%	0.30%	0.47%	0.52%	0.68%	0.99%	1.93%	14.01
phecode_208-21	0.07%	0.12%	0.15%	0.15%	0.19%	0.29%	0.36%	0.49%	0.63%	1.07%	14.94
phecode_208-22	0.03%	0.03%	0.04%	0.04%	0.03%	0.07%	0.08%	0.11%	0.17%	0.77%	29.69
phecode_209	0.15%	0.13%	0.16%	0.22%	0.22%	0.27%	0.29%	0.34%	0.45%	0.98%	6.52
phecode_209-1	0.06%	0.07%	0.08%	0.13%	0.13%	0.14%	0.18%	0.20%	0.25%	0.52%	8.22
phecode_209-12	0.02%	0.03%	0.05%	0.09%	0.09%	0.11%	0.15%	0.17%	0.21%	0.43%	18.00
phecode_209-13	0.01%	0.01%	0.01%	0.02%	0.01%	0.00%	0.02%	0.03%	0.03%	0.11%	10.60
phecode_209-2	0.03%	0.05%	0.05%	0.06%	0.07%	0.10%	0.09%	0.11%	0.18%	0.56%	21.62
phecode_209-21	0.02%	0.00%	0.01%	0.00%	0.01%	0.01%	0.02%	0.02%	0.03%	0.12%	7.63
phecode_209-22	0.02%	0.02%	0.04%	0.06%	0.05%	0.04%	0.10%	0.09%	0.15%	0.52%	28.78
phecode_209-23	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.15%	18.50
phecode_210	0.00%	0.01%	0.01%	0.01%	0.02%	0.01%	0.03%	0.03%	0.05%	0.08%	20.50
phecode_211	0.07%	0.09%	0.14%	0.19%	0.16%	0.24%	0.24%	0.30%	0.40%	1.02%	13.81
phecode_211-1	0.01%	0.00%	0.01%	0.00%	0.01%	0.01%	0.01%	0.02%	0.03%	0.11%	13.25
phecode_211-2	0.04%	0.07%	0.06%	0.08%	0.10%	0.09%	0.11%	0.10%	0.22%	0.67%	16.80
phecode_211-21	0.02%	0.02%	0.01%	0.02%	0.02%	0.03%	0.05%	0.04%	0.06%	0.29%	16.11
phecode_211-22	0.01%	0.01%	0.01%	0.01%	0.00%	0.01%	0.01%	0.02%	0.02%	0.12%	14.75
phecode_214	0.02%	0.03%	0.09%	0.11%	0.15%	0.12%	0.19%	0.24%	0.25%	0.55%	29.80
phecode_214-1	0.02%	0.03%	0.08%	0.11%	0.13%	0.13%	0.18%	0.21%	0.25%	0.53%	28.60
phecode_214-11	0.01%	0.03%	0.07%	0.11%	0.12%	0.14%	0.18%	0.18%	0.23%	0.47%	42.33
phecode_215	0.07%	0.11%	0.12%	0.14%	0.16%	0.20%	0.23%	0.33%	0.49%	0.96%	12.88
phecode_215-1	0.06%	0.10%	0.11%	0.14%	0.17%	0.18%	0.21%	0.27%	0.46%	0.98%	17.23
phecode_229	0.00%	0.01%	0.01%	0.02%	0.03%	0.03%	0.06%	0.07%	0.08%	0.34%	86.50
phecode_230	1.77%	2.42%	2.57%	3.12%	3.92%	4.60%	5.57%	6.58%	8.50%	13.56%	7.66
phecode_230-1	0.02%	0.04%	0.04%	0.05%	0.07%	0.09%	0.12%	0.17%	0.23%	0.60%	30.30
phecode_230-2	1.47%	1.98%	2.12%	2.56%	3.12%	3.74%	4.49%	5.35%	6.81%	10.85%	7.40
phecode_230-21	1.32%	1.78%	1.97%	2.37%	2.80%	3.50%	4.35%	5.17%	6.49%	10.51%	7.95
phecode_230-22	0.03%	0.04%	0.10%	0.15%	0.19%	0.25%	0.30%	0.47%	0.51%	0.91%	35.23
phecode_230-3	0.32%	0.41%	0.50%	0.66%	0.87%	0.98%	1.13%	1.66%	2.18%	4.16%	13.10
phecode_230-4	0.01%	0.02%	0.02%	0.03%	0.03%	0.05%	0.05%	0.05%	0.09%	0.28%	35.75
phecode_230-5	0.00%	0.00%	0.00%	0.00%	0.01%	0.02%	0.02%	0.03%	0.04%	0.13%	Inf
phecode_232	0.90%	1.82%	3.12%	3.95%	4.86%	5.89%	7.45%	9.85%	12.55%	21.72%	24.26
phecode_232-1	0.00%	0.00%	0.00%	0.00%	0.01%	0.02%	0.01%	0.02%	0.06%	0.09%	47.00
phecode_232-2	0.55%	0.71%	1.01%	1.31%	1.51%	1.95%	2.58%	3.36%	4.60%	9.49%	17.39

Supplementary Tables

Table 16 continued from previous page

Endpoint	1	2	3	4	5	6	7	8	9	10	Rate Ratio
phecode_232-27	0.15%	0.15%	0.17%	0.23%	0.25%	0.32%	0.40%	0.51%	0.82%	2.12%	14.36
phecode_232-29	0.06%	0.06%	0.09%	0.11%	0.13%	0.14%	0.21%	0.28%	0.43%	0.95%	16.96
phecode_232-4	0.37%	1.02%	1.94%	2.59%	3.43%	4.03%	5.18%	6.87%	8.91%	15.79%	42.36
phecode_234	0.04%	0.15%	0.46%	0.57%	0.66%	0.73%	1.03%	1.26%	1.52%	2.64%	62.91
phecode_236	1.86%	3.36%	5.30%	6.82%	8.48%	9.95%	12.11%	14.72%	19.01%	29.88%	16.11
phecode_236-1	1.85%	3.37%	5.28%	6.81%	8.47%	9.95%	12.11%	14.67%	18.95%	29.75%	16.10
phecode_236-11	0.03%	0.05%	0.06%	0.10%	0.13%	0.21%	0.26%	0.46%	0.75%	3.28%	126.39
phecode_236-2	0.01%	0.00%	0.01%	0.02%	0.01%	0.02%	0.02%	0.01%	0.04%	0.35%	44.00
phecode_237	0.02%	0.06%	0.22%	0.34%	0.45%	0.54%	0.73%	0.99%	1.30%	2.40%	149.38
phecode_239	5.01%	7.38%	9.76%	12.60%	14.79%	17.83%	20.69%	24.04%	28.57%	37.08%	7.39
phecode_239-1	4.44%	6.51%	8.73%	11.12%	13.32%	16.00%	18.90%	21.69%	26.27%	34.87%	7.85
phecode_239-11	4.42%	6.48%	8.68%	11.00%	13.19%	15.94%	18.81%	21.52%	26.29%	35.01%	7.91
phecode_239-12	0.01%	0.02%	0.07%	0.11%	0.17%	0.26%	0.36%	0.45%	0.61%	1.00%	167.67
phecode_239-2	0.07%	0.09%	0.18%	0.21%	0.37%	0.40%	0.53%	0.58%	0.82%	1.50%	22.70
phecode_239-21	0.02%	0.03%	0.06%	0.05%	0.10%	0.12%	0.17%	0.20%	0.26%	0.79%	36.18
phecode_239-3	0.04%	0.07%	0.12%	0.16%	0.23%	0.29%	0.35%	0.43%	0.53%	0.84%	21.05
phecode_240	0.01%	0.00%	0.01%	0.02%	0.02%	0.02%	0.03%	0.03%	0.06%	0.13%	10.67
phecode_241	0.03%	0.02%	0.02%	0.05%	0.09%	0.11%	0.09%	0.13%	0.15%	0.27%	8.00
phecode_242	0.01%	0.01%	0.02%	0.02%	0.03%	0.04%	0.04%	0.04%	0.06%	0.05%	6.25
phecode_244	0.10%	0.14%	0.24%	0.27%	0.35%	0.49%	0.55%	0.74%	0.94%	1.88%	18.12
phecode_244-4	0.00%	0.01%	0.01%	0.01%	0.00%	0.01%	0.01%	0.02%	0.02%	0.40%	Inf
phecode_247	2.72%	3.54%	3.68%	4.23%	4.92%	5.56%	6.71%	8.30%	11.00%	19.92%	7.33
phecode_247-3	0.09%	0.15%	0.20%	0.24%	0.22%	0.30%	0.34%	0.40%	0.58%	1.06%	12.09
phecode_247-4	0.10%	0.15%	0.18%	0.22%	0.25%	0.42%	0.49%	0.63%	1.02%	2.52%	25.80
phecode_247-42	0.01%	0.03%	0.02%	0.02%	0.04%	0.03%	0.07%	0.08%	0.17%	0.68%	114.00
phecode_247-5	0.34%	0.57%	0.62%	0.77%	0.89%	1.04%	1.18%	1.50%	1.98%	3.62%	10.54
phecode_247-51	0.00%	0.03%	0.08%	0.14%	0.18%	0.22%	0.26%	0.33%	0.48%	0.84%	424.01
phecode_247-52	0.01%	0.04%	0.12%	0.13%	0.22%	0.27%	0.33%	0.49%	0.56%	1.17%	97.67
phecode_247-7	2.08%	2.24%	2.84%	3.35%	3.90%	4.27%	5.17%	6.45%	8.75%	16.27%	7.81
phecode_247-71	0.00%	0.03%	0.10%	0.12%	0.13%	0.16%	0.16%	0.18%	0.22%	0.40%	Inf
phecode_247-711	0.00%	0.00%	0.00%	0.02%	0.02%	0.03%	0.03%	0.08%	0.09%	0.17%	Inf
phecode_247-72	1.91%	2.03%	2.70%	3.09%	3.71%	4.06%	4.90%	6.14%	8.34%	16.10%	8.44
phecode_248	0.06%	0.10%	0.12%	0.16%	0.15%	0.21%	0.22%	0.25%	0.28%	0.72%	11.31
phecode_248-1	0.00%	0.00%	0.03%	0.04%	0.01%	0.02%	0.02%	0.03%	0.04%	0.06%	Inf
phecode_249	0.03%	0.04%	0.05%	0.07%	0.10%	0.11%	0.12%	0.12%	0.16%	0.31%	11.92
phecode_249-1	0.00%	0.02%	0.03%	0.03%	0.05%	0.04%	0.04%	0.06%	0.07%	0.09%	46.00
phecode_251	0.09%	0.11%	0.15%	0.18%	0.18%	0.26%	0.35%	0.47%	0.56%	0.81%	8.62
phecode_251-1	0.07%	0.09%	0.13%	0.17%	0.18%	0.25%	0.32%	0.45%	0.54%	0.80%	10.81
phecode_252	0.02%	0.01%	0.02%	0.03%	0.05%	0.06%	0.08%	0.09%	0.11%	0.24%	15.00
phecode_256	2.13%	3.08%	3.57%	4.46%	5.07%	5.78%	6.54%	8.70%	12.23%	22.05%	10.33
phecode_256-1	0.09%	0.17%	0.18%	0.23%	0.30%	0.32%	0.37%	0.38%	0.68%	1.33%	14.89
phecode_256-2	1.04%	1.38%	1.58%	1.81%	2.13%	2.49%	2.95%	3.58%	4.85%	7.47%	7.16
phecode_256-3	0.10%	0.19%	0.24%	0.34%	0.37%	0.55%	0.76%	1.01%	1.44%	4.27%	42.88
phecode_256-31	0.09%	0.13%	0.20%	0.26%	0.30%	0.43%	0.61%	0.87%	1.21%	3.96%	44.22
phecode_256-32	0.02%	0.05%	0.05%	0.07%	0.10%	0.10%	0.12%	0.16%	0.25%	0.42%	26.63
phecode_256-4	0.18%	0.26%	0.37%	0.46%	0.58%	0.71%	0.75%	1.00%	1.60%	4.89%	26.65
phecode_256-5	0.32%	0.54%	0.66%	0.88%	1.14%	1.27%	1.55%	2.21%	2.95%	5.26%	16.47
phecode_256-6	0.06%	0.10%	0.14%	0.18%	0.30%	0.33%	0.54%	0.57%	0.89%	3.02%	54.11
phecode_256-7	0.38%	0.50%	0.75%	1.06%	1.18%	1.55%	1.97%	2.31%	3.44%	7.95%	20.66
phecode_256-71	0.01%	0.02%	0.03%	0.05%	0.08%	0.16%	0.14%	0.21%	0.30%	0.68%	114.34
phecode_257	0.03%	0.07%	0.15%	0.16%	0.20%	0.26%	0.32%	0.39%	0.56%	0.94%	27.71
phecode_280	1.23%	1.73%	2.04%	2.69%	3.09%	3.69%	4.42%	5.35%	6.95%	11.09%	9.02
phecode_280-1	1.17%	1.61%	1.87%	2.53%	2.97%	3.47%	4.24%	5.01%	6.63%	10.43%	8.93
phecode_280-11	0.52%	0.84%	0.98%	1.09%	1.29%	1.67%	2.15%	2.70%	3.84%	7.21%	13.80
phecode_280-12	0.22%	0.38%	0.45%	0.62%	0.70%	0.91%	1.07%	1.52%	2.13%	4.63%	20.76
phecode_280-13	0.03%	0.06%	0.07%	0.11%	0.17%	0.18%	0.32%	0.41%	0.49%	1.70%	56.80
phecode_280-14	0.00%	0.01%	0.01%	0.02%	0.01%	0.02%	0.03%	0.03%	0.05%	0.24%	60.00
phecode_280-2	0.00%	0.01%	0.02%	0.00%	0.02%	0.03%	0.02%	0.05%	0.06%	0.28%	140.00
phecode_280-22	0.00%	0.01%	0.00%	0.00%	0.01%	0.01%	0.03%	0.03%	0.04%	0.22%	Inf
phecode_280-3	0.01%	0.00%	0.00%	0.01%	0.01%	0.03%	0.03%	0.05%	0.08%	0.55%	91.67
phecode_280-31	0.00%	0.00%	0.00%	0.02%	0.01%	0.03%	0.04%	0.03%	0.06%	0.43%	109.00
phecode_280-4	0.01%	0.00%	0.01%	0.00%	0.02%	0.02%	0.03%	0.04%	0.09%	0.35%	58.33
phecode_280-42	0.00%	0.00%	0.01%	0.00%	0.01%	0.01%	0.02%	0.04%	0.09%	0.33%	84.00
phecode_280-8	0.02%	0.04%	0.05%	0.07%	0.09%	0.08%	0.11%	0.20%	0.24%	0.88%	36.92
phecode_280-81	0.01%	0.01%	0.02%	0.01%	0.02%	0.02%	0.03%	0.05%	0.06%	0.25%	42.33

4 Medical history predicts future health trajectories over the human phenome

Table 16 continued from previous page

Endpoint	1	2	3	4	5	6	7	8	9	10	Rate Ratio
phecode_280-82	0.02%	0.02%	0.04%	0.07%	0.06%	0.07%	0.08%	0.16%	0.21%	0.68%	28.42
phecode_281	2.82%	3.92%	4.08%	5.07%	5.69%	6.56%	7.92%	9.46%	12.43%	21.20%	7.51
phecode_281-1	0.53%	0.86%	1.04%	1.11%	1.36%	1.72%	2.25%	2.86%	3.94%	7.54%	14.19
phecode_281-2	2.33%	3.10%	3.33%	3.85%	4.63%	5.26%	6.46%	8.02%	10.33%	19.24%	8.26
phecode_281-21	0.00%	0.01%	0.01%	0.02%	0.03%	0.05%	0.05%	0.09%	0.11%	0.51%	256.01
phecode_282-1	1.65%	2.29%	2.96%	3.39%	3.59%	4.15%	4.84%	5.90%	7.62%	12.85%	7.78
phecode_283	2.58%	3.94%	12.11%	16.15%	19.27%	22.38%	25.95%	29.53%	34.48%	44.26%	17.14
phecode_283-3	0.00%	0.01%	0.02%	0.02%	0.02%	0.02%	0.03%	0.05%	0.10%	0.26%	66.00
phecode_283-4	0.25%	0.67%	1.48%	2.00%	2.47%	3.13%	4.01%	5.12%	7.39%	12.37%	49.18
phecode_283-8	2.16%	3.48%	10.65%	14.28%	17.04%	19.62%	22.60%	25.45%	30.07%	37.91%	17.53
phecode_284	0.14%	0.20%	0.24%	0.27%	0.37%	0.46%	0.64%	0.88%	1.49%	5.83%	42.79
phecode_284-1	0.02%	0.06%	0.14%	0.27%	0.27%	0.35%	0.47%	0.70%	1.11%	4.52%	205.46
phecode_284-2	0.05%	0.08%	0.10%	0.07%	0.09%	0.16%	0.16%	0.28%	0.45%	2.40%	48.12
phecode_284-29	0.05%	0.08%	0.09%	0.07%	0.09%	0.16%	0.15%	0.28%	0.45%	2.39%	47.96
phecode_286	2.29%	3.03%	4.11%	5.06%	6.05%	7.56%	9.14%	11.49%	15.09%	26.64%	11.63
phecode_286-1	0.03%	0.04%	0.04%	0.06%	0.05%	0.11%	0.14%	0.24%	0.24%	1.22%	36.00
phecode_286-2	2.13%	2.90%	4.03%	4.95%	5.90%	7.61%	8.97%	11.38%	15.13%	26.76%	12.56
phecode_286-21	0.06%	0.07%	0.09%	0.09%	0.17%	0.19%	0.22%	0.38%	0.66%	2.10%	33.65
phecode_286-3	0.00%	0.01%	0.01%	0.00%	0.02%	0.02%	0.09%	0.14%	0.39%	1.16%	308.01
phecode_286-4	0.02%	0.01%	0.03%	0.03%	0.04%	0.04%	0.07%	0.10%	0.20%	0.55%	24.91
phecode_287	0.11%	0.11%	0.14%	0.18%	0.18%	0.25%	0.30%	0.34%	0.43%	1.35%	12.72
phecode_287-1	0.01%	0.01%	0.02%	0.03%	0.03%	0.04%	0.05%	0.09%	0.11%	0.71%	118.34
phecode_287-2	0.01%	0.00%	0.00%	0.00%	0.00%	0.01%	0.01%	0.02%	0.02%	0.27%	22.50
phecode_287-4	0.03%	0.06%	0.07%	0.07%	0.08%	0.10%	0.12%	0.16%	0.24%	0.66%	20.75
phecode_287-5	0.01%	0.02%	0.03%	0.03%	0.02%	0.03%	0.03%	0.05%	0.06%	0.21%	21.40
phecode_288	1.72%	2.67%	3.94%	4.91%	5.97%	7.30%	8.82%	11.14%	14.53%	23.51%	13.64
phecode_288-2	0.12%	0.20%	0.28%	0.39%	0.51%	0.55%	0.79%	1.14%	1.62%	2.91%	24.05
phecode_288-3	0.05%	0.20%	0.58%	0.88%	1.11%	1.44%	1.68%	2.38%	3.28%	5.76%	117.17
phecode_288-4	0.09%	0.26%	0.36%	0.58%	0.74%	0.83%	1.15%	1.46%	2.00%	3.09%	32.70
phecode_288-41	0.01%	0.01%	0.01%	0.02%	0.01%	0.02%	0.01%	0.03%	0.05%	0.19%	23.25
phecode_289	0.03%	0.03%	0.04%	0.05%	0.05%	0.07%	0.06%	0.09%	0.14%	0.51%	18.36
phecode_290	0.12%	0.40%	1.28%	1.95%	2.47%	3.17%	4.18%	5.33%	7.13%	12.17%	100.36
phecode_290-1	0.04%	0.02%	0.05%	0.09%	0.10%	0.13%	0.14%	0.17%	0.28%	0.91%	25.44
phecode_291	0.02%	0.03%	0.03%	0.03%	0.06%	0.08%	0.10%	0.12%	0.21%	0.51%	28.56
phecode_292	0.05%	0.08%	0.10%	0.10%	0.13%	0.15%	0.17%	0.25%	0.34%	0.65%	13.58
phecode_293	0.02%	0.02%	0.04%	0.05%	0.04%	0.08%	0.09%	0.11%	0.15%	0.42%	21.20
phecode_293-1	0.01%	0.00%	0.02%	0.01%	0.01%	0.02%	0.02%	0.02%	0.02%	0.09%	14.67
phecode_293-4	0.00%	0.01%	0.01%	0.03%	0.01%	0.03%	0.03%	0.06%	0.10%	0.25%	125.00
phecode_294	0.24%	0.31%	0.41%	0.54%	0.59%	0.65%	3.98%	10.41%	14.89%	23.97%	101.75
phecode_296	0.01%	0.01%	0.01%	0.02%	0.02%	0.03%	0.03%	0.07%	0.09%	0.68%	114.00
phecode_296-4	0.00%	0.00%	0.00%	0.01%	0.01%	0.01%	0.02%	0.02%	0.04%	0.40%	202.00
phecode_299	0.11%	0.16%	0.22%	0.30%	0.32%	0.33%	0.41%	0.66%	0.84%	2.06%	19.04
phecode_308	0.55%	1.29%	2.81%	3.76%	4.83%	6.23%	7.58%	9.70%	13.49%	22.97%	41.62
phecode_308-1	0.02%	0.03%	0.07%	0.11%	0.11%	0.17%	0.21%	0.28%	0.44%	0.85%	38.55
phecode_308-3	0.01%	0.02%	0.06%	0.08%	0.11%	0.13%	0.21%	0.28%	0.44%	1.16%	193.34
phecode_308-4	0.00%	0.01%	0.01%	0.03%	0.03%	0.07%	0.12%	0.19%	0.32%	0.83%	419.01
phecode_308-5	0.01%	0.04%	0.09%	0.17%	0.16%	0.22%	0.27%	0.39%	0.56%	1.04%	87.00
phecode_308-6	0.00%	0.01%	0.03%	0.05%	0.05%	0.06%	0.14%	0.21%	0.31%	0.74%	Inf
phecode_308-7	0.11%	0.13%	0.20%	0.21%	0.26%	0.30%	0.36%	0.46%	0.63%	1.38%	12.56
phecode_320	0.03%	0.03%	0.06%	0.08%	0.07%	0.08%	0.11%	0.11%	0.13%	0.24%	7.63
phecode_320-1	0.03%	0.03%	0.05%	0.06%	0.06%	0.05%	0.06%	0.07%	0.09%	0.14%	4.86
phecode_320-11	0.01%	0.02%	0.02%	0.03%	0.04%	0.02%	0.05%	0.04%	0.07%	0.07%	6.60
phecode_320-12	0.00%	0.01%	0.02%	0.02%	0.01%	0.02%	0.04%	0.02%	0.05%	0.08%	20.00
phecode_320-3	0.01%	0.01%	0.03%	0.01%	0.01%	0.03%	0.03%	0.03%	0.06%	0.10%	10.20
phecode_321	0.08%	0.11%	0.10%	0.09%	0.11%	0.12%	0.13%	0.14%	0.18%	0.32%	4.00
phecode_321-1	0.04%	0.04%	0.04%	0.06%	0.04%	0.04%	0.06%	0.07%	0.09%	0.11%	3.06
phecode_321-12	0.03%	0.01%	0.03%	0.03%	0.06%	0.04%	0.05%	0.05%	0.08%	0.08%	2.93
phecode_321-2	0.01%	0.01%	0.02%	0.03%	0.02%	0.01%	0.03%	0.03%	0.03%	0.07%	5.29
phecode_321-21	0.02%	0.01%	0.02%	0.01%	0.02%	0.02%	0.01%	0.03%	0.02%	0.06%	4.00
phecode_322	0.02%	0.04%	0.03%	0.03%	0.06%	0.06%	0.09%	0.10%	0.10%	0.34%	16.90
phecode_322-4	0.02%	0.01%	0.02%	0.02%	0.02%	0.04%	0.05%	0.06%	0.05%	0.10%	4.33
phecode_323	0.05%	0.09%	0.14%	0.15%	0.17%	0.19%	0.22%	0.27%	0.36%	0.60%	12.54
phecode_323-1	0.02%	0.04%	0.04%	0.05%	0.04%	0.06%	0.06%	0.09%	0.11%	0.24%	13.56
phecode_323-3	0.03%	0.05%	0.08%	0.07%	0.08%	0.12%	0.13%	0.16%	0.27%	0.38%	11.88
phecode_323-31	0.01%	0.01%	0.00%	0.00%	0.00%	0.01%	0.05%	0.08%	0.10%	0.12%	19.33

Supplementary Tables

Table 16 continued from previous page

Endpoint	1	2	3	4	5	6	7	8	9	10	Rate Ratio
phecode_324	1.07%	1.55%	2.27%	2.78%	3.22%	4.12%	4.89%	5.78%	7.62%	11.57%	10.77
phecode_324-1	0.11%	0.31%	0.43%	0.58%	0.76%	0.83%	0.86%	1.08%	1.43%	2.31%	21.91
phecode_324-11	0.11%	0.31%	0.42%	0.55%	0.74%	0.79%	0.83%	1.05%	1.37%	2.18%	19.89
phecode_324-12	0.01%	0.02%	0.01%	0.02%	0.04%	0.04%	0.05%	0.05%	0.11%	0.32%	40.75
phecode_324-2	0.01%	0.03%	0.05%	0.07%	0.07%	0.07%	0.08%	0.08%	0.11%	0.14%	15.67
phecode_324-21	0.00%	0.02%	0.02%	0.02%	0.04%	0.03%	0.06%	0.06%	0.08%	0.10%	24.00
phecode_324-3	0.17%	0.34%	0.69%	1.01%	1.15%	1.35%	1.58%	2.13%	2.60%	4.42%	25.96
phecode_324-34	0.03%	0.09%	0.45%	0.63%	0.77%	0.94%	1.07%	1.26%	1.80%	2.72%	79.24
phecode_324-36	0.01%	0.01%	0.03%	0.04%	0.07%	0.06%	0.06%	0.12%	0.13%	0.22%	27.25
phecode_324-4	0.19%	0.35%	0.71%	0.88%	1.08%	1.47%	1.84%	2.26%	3.04%	5.03%	26.76
phecode_324-41	0.07%	0.14%	0.20%	0.27%	0.38%	0.54%	0.56%	0.72%	1.03%	1.68%	24.03
phecode_324-5	0.03%	0.03%	0.03%	0.05%	0.05%	0.07%	0.06%	0.08%	0.12%	0.28%	9.93
phecode_324-8	0.02%	0.07%	0.22%	0.39%	0.47%	0.59%	0.80%	1.03%	1.46%	2.73%	113.67
phecode_325	1.03%	1.45%	1.82%	2.13%	2.55%	3.34%	4.08%	5.63%	7.84%	15.37%	14.96
phecode_325-1	0.06%	0.09%	0.15%	0.18%	0.23%	0.24%	0.26%	0.29%	0.34%	0.70%	11.00
phecode_325-12	0.04%	0.05%	0.07%	0.09%	0.11%	0.17%	0.15%	0.20%	0.18%	0.33%	8.63
phecode_325-2	0.77%	1.14%	1.63%	1.93%	2.23%	2.92%	3.79%	5.07%	7.30%	14.69%	19.02
phecode_325-21	0.01%	0.02%	0.02%	0.02%	0.04%	0.05%	0.08%	0.08%	0.11%	0.23%	19.67
phecode_325-23	0.01%	0.03%	0.10%	0.13%	0.28%	0.41%	0.53%	0.79%	1.22%	2.51%	419.01
phecode_325-3	0.12%	0.17%	0.28%	0.35%	0.44%	0.55%	0.81%	0.94%	1.43%	2.66%	22.46
phecode_326	0.08%	0.09%	0.13%	0.09%	0.13%	0.16%	0.15%	0.17%	0.22%	0.56%	6.88
phecode_326-1	0.05%	0.07%	0.10%	0.07%	0.09%	0.12%	0.14%	0.14%	0.19%	0.55%	10.54
phecode_327	0.06%	0.14%	0.20%	0.24%	0.31%	0.47%	0.50%	0.66%	0.97%	2.04%	32.06
phecode_328	0.08%	0.32%	0.55%	1.01%	1.33%	1.68%	1.86%	2.79%	3.77%	6.52%	86.13
phecode_328-1	0.04%	0.21%	0.35%	0.55%	0.75%	0.84%	1.19%	1.56%	2.10%	3.22%	73.46
phecode_328-2	0.00%	0.03%	0.03%	0.03%	0.05%	0.06%	0.08%	0.08%	0.11%	0.19%	47.50
phecode_328-4	0.01%	0.00%	0.02%	0.03%	0.04%	0.08%	0.12%	0.14%	0.19%	0.52%	87.67
phecode_328-7	0.00%	0.04%	0.10%	0.14%	0.21%	0.35%	0.39%	0.50%	0.82%	2.04%	1025.02
phecode_328-8	0.07%	0.18%	0.27%	0.35%	0.47%	0.77%	1.05%	1.50%	2.00%	3.41%	50.32
phecode_328-9	0.03%	0.14%	0.26%	0.43%	0.61%	0.89%	0.94%	1.26%	1.98%	3.79%	118.88
phecode_329	1.86%	2.78%	3.04%	3.86%	4.71%	5.92%	7.10%	9.44%	12.40%	20.01%	10.78
phecode_329-1	0.24%	0.48%	0.93%	1.41%	1.87%	2.45%	3.24%	4.58%	6.06%	10.23%	41.85
phecode_329-4	0.01%	0.06%	0.26%	0.39%	0.45%	0.62%	0.67%	0.84%	1.17%	2.00%	167.00
phecode_329-41	0.00%	0.01%	0.05%	0.09%	0.07%	0.15%	0.15%	0.21%	0.34%	0.70%	Inf
phecode_329-42	0.01%	0.04%	0.19%	0.25%	0.37%	0.40%	0.51%	0.59%	0.84%	1.52%	108.72
phecode_329-5	0.02%	0.02%	0.07%	0.17%	0.27%	0.43%	0.61%	0.91%	1.39%	2.34%	146.88
phecode_329-6	0.08%	0.08%	0.13%	0.16%	0.24%	0.28%	0.34%	0.35%	0.49%	0.66%	7.88
phecode_329-8	0.01%	0.01%	0.03%	0.04%	0.06%	0.07%	0.12%	0.13%	0.20%	0.44%	74.00
phecode_329-9	0.09%	0.22%	0.33%	0.48%	0.68%	1.02%	1.38%	1.52%	2.56%	5.34%	62.33
phecode_330	0.49%	0.61%	0.72%	0.89%	0.97%	1.13%	1.35%	1.49%	1.70%	3.53%	7.21
phecode_330-1	0.25%	0.41%	0.40%	0.50%	0.58%	0.67%	0.76%	0.80%	1.02%	2.35%	9.34
phecode_330-11	0.07%	0.07%	0.10%	0.11%	0.13%	0.12%	0.17%	0.16%	0.27%	1.05%	14.16
phecode_330-12	0.07%	0.09%	0.10%	0.11%	0.13%	0.12%	0.14%	0.16%	0.21%	0.69%	10.48
phecode_330-3	0.33%	0.39%	0.50%	0.58%	0.68%	0.83%	0.89%	0.97%	1.18%	3.02%	9.08
phecode_331	1.90%	3.06%	5.12%	6.78%	8.29%	9.93%	11.73%	14.57%	18.05%	26.78%	14.09
phecode_331-1	0.07%	0.15%	0.40%	0.52%	0.74%	0.93%	1.30%	1.59%	2.03%	3.65%	53.18
phecode_331-3	0.01%	0.02%	0.07%	0.05%	0.10%	0.14%	0.17%	0.21%	0.33%	0.68%	85.00
phecode_331-4	0.00%	0.01%	0.03%	0.04%	0.07%	0.08%	0.10%	0.12%	0.19%	0.39%	Inf
phecode_331-6	0.52%	0.74%	1.31%	1.44%	1.90%	2.36%	3.03%	3.63%	5.11%	8.12%	15.76
phecode_331-61	0.04%	0.11%	0.22%	0.31%	0.38%	0.47%	0.66%	0.82%	1.05%	1.78%	49.28
phecode_331-62	0.00%	0.00%	0.01%	0.01%	0.01%	0.01%	0.01%	0.03%	0.03%	0.14%	Inf
phecode_331-7	0.00%	0.01%	0.01%	0.01%	0.01%	0.03%	0.04%	0.07%	0.05%	0.24%	61.50
phecode_331-8	1.45%	2.44%	3.96%	5.25%	6.71%	8.14%	9.99%	11.92%	15.38%	23.90%	16.43
phecode_333	0.98%	1.81%	3.32%	4.48%	5.36%	6.77%	8.53%	10.66%	14.67%	27.47%	28.10
phecode_333-1	0.27%	0.51%	0.61%	0.79%	1.08%	1.34%	1.80%	2.43%	3.70%	7.50%	27.46
phecode_333-11	0.02%	0.08%	0.15%	0.26%	0.46%	0.62%	0.91%	1.26%	2.00%	4.30%	269.26
phecode_333-2	0.11%	0.43%	1.50%	2.09%	2.49%	3.23%	3.95%	5.25%	6.70%	10.65%	97.42
phecode_333-3	0.00%	0.02%	0.02%	0.04%	0.05%	0.05%	0.08%	0.08%	0.14%	0.36%	90.50
phecode_333-4	0.00%	0.02%	0.05%	0.07%	0.08%	0.12%	0.17%	0.30%	0.40%	0.99%	498.00
phecode_333-43	0.00%	0.00%	0.02%	0.01%	0.01%	0.01%	0.03%	0.03%	0.05%	0.07%	Inf
phecode_333-5	0.01%	0.02%	0.05%	0.07%	0.08%	0.09%	0.12%	0.17%	0.23%	0.49%	41.00
phecode_334	0.40%	0.62%	0.84%	1.00%	1.18%	1.36%	1.53%	1.85%	2.13%	3.43%	8.60
phecode_334-1	0.09%	0.17%	0.25%	0.35%	0.45%	0.58%	0.71%	0.87%	1.15%	2.02%	21.45
phecode_334-11	0.09%	0.14%	0.21%	0.30%	0.32%	0.49%	0.57%	0.74%	0.97%	1.73%	19.64
phecode_334-12	0.00%	0.01%	0.04%	0.06%	0.07%	0.06%	0.10%	0.11%	0.21%	0.33%	82.50

4 Medical history predicts future health trajectories over the human phenome

Table 16 continued from previous page

Endpoint	1	2	3	4	5	6	7	8	9	10	Rate Ratio
phecode_334-2	0.18%	0.30%	0.39%	0.44%	0.49%	0.56%	0.60%	0.61%	0.85%	1.28%	7.24
phecode_334-21	0.14%	0.26%	0.32%	0.35%	0.40%	0.36%	0.48%	0.50%	0.64%	0.94%	6.94
phecode_334-23	0.00%	0.00%	0.04%	0.05%	0.07%	0.09%	0.08%	0.11%	0.15%	0.24%	Inf
phecode_334-24	0.01%	0.01%	0.02%	0.01%	0.04%	0.05%	0.05%	0.05%	0.05%	0.09%	11.50
phecode_334-4	0.07%	0.11%	0.11%	0.09%	0.14%	0.13%	0.19%	0.24%	0.31%	0.42%	6.18
phecode_334-41	0.02%	0.03%	0.04%	0.04%	0.03%	0.04%	0.04%	0.05%	0.09%	0.14%	5.67
phecode_334-42	0.02%	0.02%	0.02%	0.04%	0.05%	0.06%	0.02%	0.06%	0.07%	0.11%	5.60
phecode_334-44	0.02%	0.04%	0.06%	0.04%	0.06%	0.07%	0.10%	0.11%	0.16%	0.22%	11.20
phecode_335	0.53%	0.64%	0.74%	1.04%	1.24%	1.49%	1.79%	2.30%	3.21%	6.14%	11.53
phecode_335-1	0.02%	0.01%	0.02%	0.02%	0.04%	0.04%	0.04%	0.05%	0.08%	0.13%	7.11
phecode_335-11	0.01%	0.01%	0.02%	0.03%	0.03%	0.03%	0.05%	0.03%	0.08%	0.12%	11.60
phecode_335-2	0.00%	0.01%	0.03%	0.02%	0.04%	0.04%	0.05%	0.07%	0.08%	0.20%	Inf
phecode_335-4	0.00%	0.01%	0.00%	0.01%	0.01%	0.01%	0.01%	0.01%	0.03%	0.17%	85.00
phecode_336	1.53%	1.97%	2.92%	3.67%	4.27%	5.29%	6.07%	7.01%	8.88%	13.11%	8.59
phecode_336-1	1.11%	1.28%	1.91%	2.33%	2.68%	3.27%	3.91%	4.48%	5.77%	8.60%	7.77
phecode_336-2	0.13%	0.22%	0.36%	0.42%	0.53%	0.56%	0.68%	0.85%	1.22%	2.06%	15.40
phecode_336-4	0.02%	0.00%	0.02%	0.04%	0.01%	0.05%	0.06%	0.06%	0.07%	0.18%	11.00
phecode_336-5	0.20%	0.34%	0.68%	0.97%	1.16%	1.37%	1.62%	1.90%	2.61%	4.06%	20.44
phecode_336-51	0.01%	0.02%	0.05%	0.08%	0.09%	0.12%	0.16%	0.17%	0.25%	0.33%	55.00
phecode_336-52	0.01%	0.06%	0.24%	0.29%	0.36%	0.40%	0.52%	0.53%	0.68%	0.95%	95.20
phecode_336-54	0.01%	0.00%	0.01%	0.02%	0.03%	0.02%	0.04%	0.03%	0.06%	0.07%	9.00
phecode_336-55	0.15%	0.18%	0.36%	0.46%	0.54%	0.67%	0.88%	1.13%	1.57%	2.67%	18.03
phecode_337	0.46%	0.66%	0.71%	0.84%	1.07%	1.28%	1.69%	2.17%	2.76%	6.46%	14.03
phecode_337-1	0.02%	0.03%	0.04%	0.06%	0.08%	0.10%	0.15%	0.15%	0.21%	0.43%	19.45
phecode_337-11	0.02%	0.02%	0.03%	0.04%	0.05%	0.06%	0.12%	0.12%	0.15%	0.31%	17.11
phecode_337-2	0.06%	0.06%	0.07%	0.08%	0.07%	0.08%	0.10%	0.12%	0.11%	0.24%	3.93
phecode_337-21	0.03%	0.03%	0.04%	0.04%	0.05%	0.03%	0.06%	0.06%	0.07%	0.09%	2.53
phecode_337-3	0.04%	0.04%	0.05%	0.07%	0.04%	0.06%	0.07%	0.07%	0.13%	0.22%	5.38
phecode_337-31	0.03%	0.04%	0.05%	0.04%	0.04%	0.04%	0.05%	0.08%	0.09%	0.14%	4.18
phecode_337-8	0.01%	0.02%	0.02%	0.02%	0.04%	0.05%	0.10%	0.11%	0.20%	2.23%	374.01
phecode_338	0.02%	0.04%	0.04%	0.06%	0.07%	0.06%	0.08%	0.12%	0.13%	0.18%	8.00
phecode_338-1	0.02%	0.03%	0.04%	0.06%	0.05%	0.05%	0.07%	0.11%	0.12%	0.15%	7.50
phecode_339	0.02%	0.03%	0.02%	0.03%	0.04%	0.03%	0.03%	0.06%	0.05%	0.14%	8.75
phecode_339-1	0.01%	0.01%	0.01%	0.03%	0.01%	0.02%	0.02%	0.03%	0.03%	0.07%	6.00
phecode_340	0.03%	0.03%	0.04%	0.05%	0.05%	0.09%	0.07%	0.10%	0.13%	0.33%	10.50
phecode_341	0.17%	0.31%	0.36%	0.44%	0.61%	0.73%	0.74%	0.96%	1.44%	3.20%	18.84
phecode_341-1	0.00%	0.00%	0.01%	0.00%	0.01%	0.02%	0.03%	0.02%	0.03%	0.09%	22.50
phecode_341-2	0.12%	0.23%	0.27%	0.39%	0.50%	0.61%	0.63%	0.72%	1.22%	2.79%	23.33
phecode_341-6	0.03%	0.05%	0.06%	0.09%	0.07%	0.07%	0.11%	0.14%	0.20%	0.40%	13.47
phecode_342	0.08%	0.16%	0.18%	0.25%	0.27%	0.32%	0.46%	0.59%	0.84%	1.63%	20.43
phecode_342-1	0.01%	0.02%	0.03%	0.03%	0.03%	0.04%	0.07%	0.08%	0.10%	0.32%	27.17
phecode_342-2	0.01%	0.02%	0.02%	0.02%	0.02%	0.02%	0.03%	0.04%	0.04%	0.15%	18.75
phecode_342-4	0.05%	0.04%	0.09%	0.14%	0.19%	0.25%	0.36%	0.48%	0.73%	1.30%	24.22
phecode_342-5	0.01%	0.00%	0.01%	0.02%	0.02%	0.03%	0.02%	0.03%	0.04%	0.13%	11.00
phecode_343	0.08%	0.14%	0.13%	0.17%	0.21%	0.25%	0.38%	0.45%	0.57%	1.41%	18.13
phecode_343-1	0.01%	0.01%	0.02%	0.01%	0.03%	0.03%	0.04%	0.04%	0.08%	0.51%	64.25
phecode_343-3	0.02%	0.03%	0.04%	0.07%	0.10%	0.11%	0.20%	0.24%	0.39%	0.68%	42.88
phecode_343-5	0.02%	0.02%	0.03%	0.04%	0.02%	0.04%	0.05%	0.04%	0.05%	0.09%	4.27
phecode_343-6	0.01%	0.02%	0.01%	0.01%	0.02%	0.03%	0.02%	0.04%	0.03%	0.11%	10.80
phecode_344	0.12%	0.14%	0.19%	0.16%	0.24%	0.25%	0.28%	0.32%	0.40%	0.59%	4.97
phecode_344-1	0.06%	0.09%	0.13%	0.13%	0.18%	0.18%	0.21%	0.27%	0.29%	0.47%	8.43
phecode_344-12	0.01%	0.01%	0.03%	0.02%	0.02%	0.03%	0.04%	0.03%	0.06%	0.07%	6.00
phecode_344-13	0.00%	0.00%	0.01%	0.02%	0.03%	0.04%	0.05%	0.06%	0.08%	0.13%	32.50
phecode_344-2	0.03%	0.01%	0.01%	0.02%	0.03%	0.03%	0.02%	0.05%	0.03%	0.07%	2.85
phecode_344-3	0.03%	0.03%	0.04%	0.04%	0.04%	0.05%	0.06%	0.03%	0.05%	0.12%	4.62
phecode_345	0.01%	0.03%	0.04%	0.04%	0.05%	0.05%	0.08%	0.10%	0.10%	0.39%	32.33
phecode_346	0.13%	0.16%	0.22%	0.21%	0.21%	0.25%	0.32%	0.38%	0.51%	0.66%	5.24
phecode_346-1	0.00%	0.00%	0.03%	0.03%	0.06%	0.05%	0.06%	0.10%	0.12%	0.22%	Inf
phecode_346-3	0.01%	0.01%	0.02%	0.03%	0.03%	0.06%	0.07%	0.07%	0.20%	0.34%	28.50
phecode_346-5	0.02%	0.03%	0.02%	0.03%	0.03%	0.03%	0.05%	0.06%	0.05%	0.10%	4.80
phecode_346-6	0.04%	0.05%	0.07%	0.06%	0.09%	0.08%	0.10%	0.15%	0.20%	0.24%	6.78
phecode_347	0.14%	0.21%	0.32%	0.39%	0.41%	0.44%	0.43%	0.57%	0.71%	1.31%	9.49
phecode_347-1	0.05%	0.05%	0.05%	0.05%	0.09%	0.07%	0.09%	0.09%	0.14%	0.19%	3.44
phecode_347-2	0.01%	0.02%	0.02%	0.03%	0.01%	0.02%	0.03%	0.05%	0.06%	0.10%	8.17
phecode_348	0.28%	0.39%	0.46%	0.54%	0.59%	0.82%	0.97%	1.10%	1.57%	2.88%	10.34

Supplementary Tables

Table 16 continued from previous page

Endpoint	1	2	3	4	5	6	7	8	9	10	Rate Ratio
phecode_348-2	0.22%	0.26%	0.37%	0.40%	0.44%	0.65%	0.74%	0.83%	1.23%	2.32%	10.35
phecode_348-21	0.00%	0.01%	0.01%	0.03%	0.04%	0.06%	0.08%	0.07%	0.12%	0.24%	122.00
phecode_348-4	0.05%	0.07%	0.08%	0.09%	0.09%	0.15%	0.15%	0.23%	0.26%	0.52%	10.52
phecode_349	0.35%	0.49%	0.57%	0.62%	0.89%	1.04%	1.23%	1.58%	2.04%	3.08%	8.86
phecode_349-1	0.25%	0.32%	0.37%	0.40%	0.49%	0.61%	0.75%	0.93%	1.22%	1.85%	7.42
phecode_349-12	0.00%	0.00%	0.02%	0.01%	0.03%	0.03%	0.02%	0.02%	0.04%	0.15%	Inf
phecode_349-13	0.10%	0.11%	0.23%	0.25%	0.31%	0.41%	0.57%	0.75%	1.00%	1.64%	15.85
phecode_349-15	0.04%	0.03%	0.03%	0.04%	0.05%	0.03%	0.08%	0.06%	0.10%	0.17%	4.78
phecode_349-2	0.01%	0.02%	0.10%	0.17%	0.24%	0.32%	0.37%	0.49%	0.63%	1.03%	172.00
phecode_349-3	0.00%	0.00%	0.01%	0.01%	0.01%	0.01%	0.02%	0.04%	0.07%	0.10%	Inf
phecode_350	1.73%	2.34%	2.87%	3.65%	4.48%	5.22%	6.45%	8.15%	10.75%	18.78%	10.87
phecode_350-3	0.01%	0.01%	0.03%	0.03%	0.04%	0.06%	0.09%	0.08%	0.15%	0.26%	33.00
phecode_350-5	0.47%	0.81%	1.36%	1.66%	1.88%	2.37%	3.38%	4.58%	6.81%	13.92%	29.66
phecode_351	0.73%	1.31%	3.31%	4.80%	5.93%	7.18%	8.40%	10.45%	13.49%	24.58%	33.88
phecode_351-1	0.04%	0.16%	0.62%	0.98%	1.25%	1.57%	1.89%	2.45%	3.03%	4.98%	129.84
phecode_351-2	0.02%	0.02%	0.04%	0.09%	0.12%	0.16%	0.28%	0.43%	0.78%	4.81%	219.28
phecode_351-3	0.37%	0.69%	1.77%	2.41%	3.08%	3.69%	4.32%	5.24%	6.62%	9.88%	26.77
phecode_351-4	0.00%	0.01%	0.02%	0.04%	0.04%	0.05%	0.05%	0.09%	0.10%	0.29%	148.00
phecode_352	0.03%	0.13%	0.43%	0.56%	0.78%	0.84%	0.96%	1.19%	1.40%	2.07%	74.14
phecode_352-1	0.02%	0.06%	0.22%	0.29%	0.41%	0.40%	0.48%	0.54%	0.61%	0.84%	46.67
phecode_352-2	0.00%	0.00%	0.04%	0.03%	0.06%	0.05%	0.07%	0.07%	0.11%	0.14%	71.00
phecode_352-3	0.01%	0.03%	0.12%	0.17%	0.21%	0.26%	0.31%	0.41%	0.46%	0.65%	54.00
phecode_353	0.29%	0.51%	0.81%	1.15%	1.43%	1.73%	1.95%	2.34%	2.93%	4.74%	16.55
phecode_353-1	0.08%	0.13%	0.14%	0.23%	0.29%	0.31%	0.32%	0.44%	0.62%	1.76%	23.21
phecode_353-11	0.02%	0.01%	0.03%	0.02%	0.03%	0.04%	0.05%	0.05%	0.11%	0.38%	21.22
phecode_353-12	0.03%	0.04%	0.05%	0.09%	0.13%	0.14%	0.14%	0.19%	0.25%	0.74%	23.38
phecode_354	1.33%	2.37%	5.41%	7.49%	9.22%	11.25%	13.40%	16.55%	20.22%	28.86%	21.75
phecode_355	0.30%	0.34%	0.46%	0.60%	0.70%	0.91%	1.08%	1.38%	2.00%	4.07%	13.39
phecode_355-1	0.15%	0.16%	0.16%	0.21%	0.27%	0.25%	0.34%	0.33%	0.48%	0.96%	6.58
phecode_355-2	0.17%	0.19%	0.30%	0.38%	0.43%	0.61%	0.81%	1.03%	1.58%	3.24%	18.64
phecode_355-21	0.00%	0.00%	0.00%	0.00%	0.00%	0.01%	0.02%	0.04%	0.09%	0.16%	Inf
phecode_356	0.31%	0.43%	0.59%	0.70%	0.76%	0.86%	1.12%	1.51%	1.83%	3.46%	11.34
phecode_356-1	0.05%	0.06%	0.07%	0.10%	0.15%	0.14%	0.18%	0.23%	0.29%	0.71%	13.19
phecode_356-2	0.13%	0.18%	0.20%	0.29%	0.38%	0.41%	0.45%	0.58%	0.80%	1.33%	10.57
phecode_356-4	0.00%	0.01%	0.04%	0.05%	0.05%	0.09%	0.08%	0.13%	0.22%	0.42%	106.50
phecode_360	0.30%	0.72%	3.58%	4.75%	5.59%	6.37%	7.43%	8.69%	10.69%	15.19%	50.02
phecode_360-1	0.06%	0.21%	1.23%	1.64%	1.92%	2.16%	2.49%	3.00%	3.58%	5.06%	80.65
phecode_360-11	0.06%	0.21%	1.23%	1.64%	1.92%	2.16%	2.49%	3.00%	3.58%	5.06%	80.65
phecode_360-12	0.00%	0.02%	0.21%	0.32%	0.36%	0.48%	0.51%	0.59%	0.80%	1.15%	Inf
phecode_360-13	0.00%	0.01%	0.04%	0.03%	0.06%	0.08%	0.09%	0.09%	0.08%	0.16%	82.00
phecode_360-2	0.09%	0.24%	1.11%	1.22%	1.45%	1.74%	1.89%	2.18%	2.38%	3.31%	35.65
phecode_360-4	0.18%	0.43%	1.55%	2.25%	2.88%	3.30%	3.95%	5.09%	6.39%	10.05%	57.38
phecode_360-5	0.00%	0.01%	0.06%	0.08%	0.11%	0.18%	0.29%	0.32%	0.42%	0.73%	184.00
phecode_360-51	0.00%	0.01%	0.06%	0.07%	0.09%	0.17%	0.26%	0.30%	0.39%	0.69%	174.00
phecode_361	0.49%	0.56%	0.59%	0.74%	0.90%	1.13%	1.38%	1.82%	2.15%	3.59%	7.30
phecode_361-1	0.08%	0.14%	0.16%	0.17%	0.24%	0.23%	0.38%	0.47%	0.59%	0.95%	11.88
phecode_361-15	0.00%	0.01%	0.02%	0.05%	0.05%	0.07%	0.05%	0.09%	0.14%	0.26%	Inf
phecode_361-2	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.02%	0.03%	0.03%	0.11%	8.83
phecode_361-3	0.24%	0.25%	0.25%	0.33%	0.47%	0.59%	0.65%	0.86%	1.03%	1.75%	7.32
phecode_361-4	0.04%	0.10%	0.10%	0.14%	0.16%	0.20%	0.23%	0.26%	0.38%	0.56%	13.43
phecode_361-9	0.10%	0.09%	0.09%	0.15%	0.13%	0.17%	0.26%	0.32%	0.47%	0.89%	9.08
phecode_362	0.27%	0.52%	0.90%	1.08%	1.29%	1.49%	1.65%	1.91%	2.33%	3.30%	12.14
phecode_362-1	0.02%	0.02%	0.08%	0.10%	0.10%	0.14%	0.16%	0.20%	0.22%	0.33%	20.38
phecode_362-5	0.00%	0.03%	0.20%	0.29%	0.27%	0.36%	0.41%	0.46%	0.52%	0.74%	184.50
phecode_362-6	0.00%	0.00%	0.00%	0.01%	0.00%	0.01%	0.03%	0.04%	0.07%	0.14%	Inf
phecode_363	0.44%	0.83%	1.74%	2.67%	3.43%	4.79%	6.05%	7.79%	10.51%	16.71%	38.06
phecode_363-2	0.06%	0.33%	1.25%	1.85%	2.60%	3.67%	4.72%	6.10%	8.61%	14.43%	245.73
phecode_363-5	0.13%	0.22%	0.41%	0.58%	0.71%	0.95%	1.14%	1.45%	1.92%	2.95%	22.68
phecode_363-51	0.00%	0.00%	0.01%	0.01%	0.01%	0.02%	0.04%	0.04%	0.05%	0.09%	Inf
phecode_363-6	0.04%	0.02%	0.04%	0.06%	0.08%	0.08%	0.10%	0.13%	0.17%	0.28%	6.62
phecode_363-61	0.01%	0.00%	0.04%	0.04%	0.07%	0.05%	0.07%	0.10%	0.12%	0.20%	24.75
phecode_363-7	0.10%	0.15%	0.14%	0.23%	0.28%	0.33%	0.40%	0.54%	0.65%	1.05%	10.76
phecode_365	0.04%	0.08%	0.08%	0.09%	0.14%	0.15%	0.17%	0.21%	0.23%	0.37%	10.33
phecode_365-2	0.00%	0.01%	0.03%	0.03%	0.07%	0.06%	0.10%	0.10%	0.11%	0.18%	45.00
phecode_365-3	0.02%	0.02%	0.04%	0.05%	0.03%	0.06%	0.05%	0.09%	0.07%	0.16%	7.80

4 Medical history predicts future health trajectories over the human phenome

Table 16 continued from previous page

Endpoint	1	2	3	4	5	6	7	8	9	10	Rate Ratio
phecode_366	0.10%	0.23%	0.45%	0.60%	0.65%	0.72%	0.79%	0.94%	1.18%	1.72%	16.82
phecode_366-1	0.04%	0.04%	0.13%	0.15%	0.16%	0.19%	0.20%	0.23%	0.27%	0.41%	9.27
phecode_366-2	0.00%	0.01%	0.07%	0.11%	0.12%	0.14%	0.14%	0.21%	0.18%	0.32%	Inf
phecode_366-21	0.00%	0.01%	0.07%	0.10%	0.11%	0.11%	0.14%	0.15%	0.15%	0.26%	Inf
phecode_366-4	0.03%	0.09%	0.26%	0.21%	0.34%	0.32%	0.42%	0.55%	0.57%	0.93%	31.00
phecode_366-42	0.01%	0.02%	0.13%	0.15%	0.17%	0.22%	0.24%	0.37%	0.36%	0.60%	100.67
phecode_366-5	0.00%	0.00%	0.02%	0.02%	0.01%	0.04%	0.02%	0.04%	0.03%	0.03%	15.00
phecode_366-6	0.00%	0.01%	0.04%	0.06%	0.04%	0.05%	0.09%	0.08%	0.13%	0.15%	Inf
phecode_367	0.39%	0.75%	4.22%	5.70%	6.76%	7.73%	9.00%	10.23%	12.52%	16.80%	43.49
phecode_367-1	0.24%	0.43%	3.29%	4.55%	5.47%	6.31%	7.34%	8.55%	10.62%	14.40%	60.94
phecode_367-12	0.02%	0.08%	0.22%	0.39%	0.50%	0.63%	0.79%	0.94%	1.25%	2.30%	95.34
phecode_367-13	0.00%	0.04%	0.13%	0.22%	0.24%	0.31%	0.36%	0.55%	0.61%	1.18%	Inf
phecode_367-2	0.06%	0.14%	0.45%	0.46%	0.58%	0.59%	0.72%	0.80%	0.86%	1.52%	27.14
phecode_367-21	0.03%	0.06%	0.13%	0.16%	0.16%	0.21%	0.18%	0.24%	0.34%	0.52%	20.08
phecode_367-22	0.00%	0.01%	0.01%	0.02%	0.04%	0.02%	0.03%	0.03%	0.06%	0.09%	Inf
phecode_367-3	0.01%	0.02%	0.05%	0.04%	0.08%	0.07%	0.08%	0.11%	0.14%	0.26%	32.25
phecode_367-4	0.01%	0.06%	0.07%	0.07%	0.10%	0.12%	0.12%	0.18%	0.21%	0.23%	23.40
phecode_367-41	0.00%	0.01%	0.06%	0.05%	0.06%	0.09%	0.10%	0.13%	0.16%	0.17%	Inf
phecode_367-5	0.08%	0.17%	0.44%	0.45%	0.49%	0.56%	0.65%	0.63%	0.79%	1.52%	18.07
phecode_367-52	0.08%	0.17%	0.44%	0.46%	0.48%	0.57%	0.63%	0.63%	0.79%	1.52%	18.02
phecode_367-6	0.00%	0.04%	0.16%	0.21%	0.23%	0.25%	0.29%	0.42%	0.51%	0.61%	Inf
phecode_367-7	0.01%	0.01%	0.06%	0.06%	0.05%	0.07%	0.08%	0.11%	0.10%	0.18%	18.20
phecode_367-9	0.02%	0.02%	0.02%	0.03%	0.04%	0.04%	0.05%	0.05%	0.04%	0.13%	7.11
phecode_369	0.28%	0.41%	0.51%	0.68%	0.85%	0.98%	1.18%	1.47%	1.62%	2.59%	9.30
phecode_369-1	0.04%	0.04%	0.08%	0.10%	0.13%	0.18%	0.15%	0.23%	0.25%	0.54%	12.23
phecode_369-2	0.02%	0.02%	0.05%	0.02%	0.02%	0.06%	0.07%	0.07%	0.14%	0.34%	19.11
phecode_369-4	0.08%	0.19%	0.19%	0.23%	0.31%	0.40%	0.51%	0.62%	0.85%	1.14%	13.64
phecode_369-42	0.00%	0.01%	0.03%	0.04%	0.05%	0.04%	0.05%	0.07%	0.08%	0.12%	Inf
phecode_369-44	0.00%	0.00%	0.02%	0.01%	0.02%	0.03%	0.03%	0.04%	0.05%	0.07%	Inf
phecode_369-5	0.06%	0.09%	0.13%	0.21%	0.23%	0.25%	0.32%	0.31%	0.38%	0.63%	9.88
phecode_369-51	0.00%	0.01%	0.06%	0.08%	0.11%	0.12%	0.13%	0.16%	0.20%	0.29%	Inf
phecode_369-6	0.02%	0.02%	0.02%	0.01%	0.01%	0.03%	0.03%	0.04%	0.04%	0.10%	4.64
phecode_369-62	0.02%	0.02%	0.01%	0.01%	0.01%	0.02%	0.03%	0.03%	0.04%	0.08%	4.67
phecode_370	0.06%	0.10%	0.19%	0.19%	0.29%	0.34%	0.39%	0.46%	0.61%	1.15%	20.61
phecode_370-1	0.00%	0.02%	0.04%	0.03%	0.04%	0.05%	0.07%	0.07%	0.06%	0.23%	117.00
phecode_370-3	0.01%	0.03%	0.04%	0.09%	0.10%	0.15%	0.17%	0.16%	0.21%	0.36%	36.20
phecode_370-4	0.01%	0.02%	0.01%	0.02%	0.03%	0.05%	0.06%	0.05%	0.07%	0.22%	22.60
phecode_371	2.14%	4.27%	6.00%	7.09%	8.61%	10.85%	13.57%	17.16%	22.55%	33.23%	15.55
phecode_371-3	0.47%	0.89%	1.18%	1.59%	2.35%	3.08%	4.25%	5.60%	7.74%	11.15%	23.88
phecode_371-31	0.47%	0.89%	1.18%	1.59%	2.35%	3.08%	4.26%	5.60%	7.74%	11.15%	23.88
phecode_372	0.02%	0.04%	0.05%	0.05%	0.05%	0.11%	0.10%	0.14%	0.20%	0.40%	22.33
phecode_372-1	0.01%	0.01%	0.01%	0.02%	0.02%	0.04%	0.04%	0.05%	0.06%	0.20%	16.33
phecode_372-2	0.01%	0.00%	0.01%	0.01%	0.01%	0.01%	0.02%	0.03%	0.04%	0.09%	14.33
phecode_373	0.05%	0.08%	0.11%	0.14%	0.17%	0.18%	0.26%	0.39%	0.46%	0.92%	18.44
phecode_373-1	0.02%	0.03%	0.05%	0.05%	0.05%	0.07%	0.08%	0.08%	0.17%	0.48%	20.08
phecode_373-2	0.01%	0.01%	0.01%	0.01%	0.01%	0.02%	0.02%	0.03%	0.03%	0.07%	5.83
phecode_374	1.73%	2.22%	3.22%	4.20%	5.26%	6.25%	7.73%	9.11%	11.21%	27.72%	16.03
phecode_374-1	0.56%	0.63%	0.79%	0.84%	1.07%	1.05%	1.18%	1.29%	1.50%	2.00%	3.59
phecode_374-11	0.30%	0.37%	0.30%	0.39%	0.48%	0.47%	0.55%	0.59%	0.76%	1.07%	3.57
phecode_374-12	0.01%	0.01%	0.02%	0.02%	0.03%	0.03%	0.03%	0.04%	0.03%	0.10%	12.50
phecode_374-13	0.00%	0.00%	0.03%	0.02%	0.04%	0.03%	0.03%	0.03%	0.02%	0.03%	Inf
phecode_374-14	0.00%	0.00%	0.02%	0.02%	0.02%	0.02%	0.02%	0.04%	0.04%	0.05%	Inf
phecode_374-2	0.02%	0.03%	0.06%	0.13%	0.13%	0.13%	0.19%	0.17%	0.17%	0.25%	10.67
phecode_374-21	0.00%	0.01%	0.05%	0.08%	0.10%	0.11%	0.10%	0.15%	0.15%	0.17%	Inf
phecode_374-3	0.16%	0.25%	0.48%	0.73%	0.89%	1.17%	1.32%	1.70%	2.17%	3.52%	22.55
phecode_374-32	0.00%	0.00%	0.01%	0.03%	0.04%	0.03%	0.05%	0.06%	0.07%	0.22%	Inf
phecode_374-33	0.00%	0.01%	0.04%	0.05%	0.06%	0.10%	0.09%	0.16%	0.20%	0.36%	179.00
phecode_374-37	0.02%	0.05%	0.06%	0.10%	0.09%	0.10%	0.16%	0.18%	0.26%	0.51%	31.75
phecode_374-38	0.01%	0.06%	0.21%	0.29%	0.45%	0.54%	0.63%	0.78%	0.95%	1.38%	172.50
phecode_374-39	0.03%	0.04%	0.07%	0.14%	0.13%	0.18%	0.24%	0.33%	0.36%	0.64%	22.93
phecode_374-4	0.18%	0.36%	0.48%	0.71%	0.93%	1.24%	1.71%	2.29%	3.65%	21.51%	122.81
phecode_374-41	0.00%	0.00%	0.03%	0.03%	0.03%	0.04%	0.03%	0.05%	0.05%	0.07%	36.00
phecode_374-42	0.05%	0.14%	0.41%	0.64%	0.87%	1.15%	1.61%	2.11%	3.43%	20.90%	415.77
phecode_374-5	0.55%	0.87%	1.01%	1.44%	1.91%	2.48%	3.30%	4.40%	5.53%	8.33%	15.16
phecode_374-51	0.02%	0.09%	0.22%	0.38%	0.60%	0.90%	1.15%	1.82%	2.43%	3.80%	173.28

Supplementary Tables

Table 16 continued from previous page

Endpoint	1	2	3	4	5	6	7	8	9	10	Rate Ratio
phecode_374-511	0.00%	0.02%	0.05%	0.09%	0.15%	0.25%	0.31%	0.47%	0.80%	1.48%	745.01
phecode_374-512	0.01%	0.02%	0.03%	0.05%	0.07%	0.13%	0.16%	0.26%	0.55%	1.14%	114.60
phecode_374-52	0.01%	0.04%	0.11%	0.20%	0.25%	0.37%	0.50%	0.54%	0.69%	0.92%	154.67
phecode_374-54	0.00%	0.02%	0.03%	0.05%	0.08%	0.13%	0.10%	0.12%	0.19%	0.55%	277.01
phecode_374-55	0.01%	0.03%	0.18%	0.36%	0.50%	0.63%	0.77%	1.04%	1.03%	1.76%	126.43
phecode_374-6	0.02%	0.04%	0.09%	0.14%	0.07%	0.12%	0.11%	0.12%	0.19%	0.19%	8.00
phecode_374-61	0.00%	0.00%	0.02%	0.04%	0.03%	0.05%	0.05%	0.05%	0.05%	0.05%	Inf
phecode_374-7	0.02%	0.02%	0.05%	0.07%	0.11%	0.08%	0.11%	0.13%	0.16%	0.19%	8.64
phecode_374-8	0.01%	0.04%	0.15%	0.23%	0.25%	0.31%	0.36%	0.53%	0.63%	1.97%	198.20
phecode_374-9	0.00%	0.01%	0.05%	0.08%	0.04%	0.08%	0.08%	0.10%	0.11%	0.14%	34.00
phecode_375	1.20%	1.55%	2.20%	2.87%	3.34%	4.10%	4.61%	5.35%	6.12%	8.76%	7.28
phecode_375-1	0.90%	1.19%	1.27%	1.87%	2.22%	2.65%	2.99%	3.45%	4.30%	7.36%	8.18
phecode_375-11	0.26%	0.39%	0.42%	0.61%	0.78%	1.04%	1.16%	1.43%	1.62%	4.73%	18.22
phecode_375-112	0.00%	0.01%	0.01%	0.01%	0.01%	0.02%	0.02%	0.03%	0.03%	0.13%	Inf
phecode_375-113	0.00%	0.02%	0.15%	0.26%	0.34%	0.45%	0.55%	0.60%	0.73%	2.09%	523.51
phecode_375-12	0.14%	0.16%	0.19%	0.22%	0.32%	0.35%	0.41%	0.53%	0.63%	1.27%	9.35
phecode_375-14	0.01%	0.01%	0.07%	0.09%	0.13%	0.18%	0.23%	0.26%	0.32%	0.60%	100.34
phecode_375-6	0.02%	0.11%	0.59%	0.85%	0.96%	1.14%	1.28%	1.50%	1.71%	3.17%	198.38
phecode_375-7	0.01%	0.01%	0.00%	0.01%	0.01%	0.01%	0.02%	0.02%	0.02%	0.11%	9.00
phecode_376	0.23%	0.57%	1.95%	2.89%	3.63%	4.31%	5.07%	6.04%	7.02%	9.02%	38.69
phecode_376-1	0.03%	0.22%	1.33%	2.10%	2.64%	3.03%	3.69%	4.20%	5.10%	6.69%	207.75
phecode_376-2	0.23%	0.55%	1.95%	2.85%	3.60%	4.34%	5.00%	6.01%	6.98%	8.96%	39.12
phecode_376-21	0.10%	0.19%	0.50%	0.75%	0.96%	1.22%	1.39%	1.63%	1.98%	2.74%	27.88
phecode_376-4	0.00%	0.01%	0.02%	0.04%	0.07%	0.07%	0.11%	0.16%	0.16%	0.23%	117.00
phecode_377	0.25%	0.50%	1.64%	2.07%	2.58%	2.94%	3.39%	3.99%	4.75%	6.63%	26.60
phecode_377-2	0.11%	0.26%	1.33%	1.61%	1.99%	2.37%	2.64%	3.24%	3.83%	5.24%	47.24
phecode_377-4	0.03%	0.08%	0.13%	0.20%	0.33%	0.40%	0.39%	0.47%	0.59%	0.92%	27.06
phecode_377-5	0.10%	0.12%	0.15%	0.15%	0.20%	0.21%	0.26%	0.30%	0.33%	0.86%	8.51
phecode_377-8	0.00%	0.01%	0.03%	0.02%	0.02%	0.04%	0.04%	0.03%	0.04%	0.11%	53.00
phecode_379	0.05%	0.14%	0.35%	0.37%	0.54%	0.44%	0.53%	0.64%	0.79%	1.08%	20.85
phecode_379-2	0.06%	0.12%	0.33%	0.40%	0.46%	0.45%	0.51%	0.61%	0.72%	0.99%	17.68
phecode_379-21	0.05%	0.09%	0.18%	0.23%	0.21%	0.23%	0.27%	0.32%	0.36%	0.57%	12.48
phecode_380	0.13%	0.23%	0.35%	0.37%	0.49%	0.49%	0.68%	0.73%	0.88%	1.42%	11.08
phecode_380-1	0.07%	0.07%	0.09%	0.09%	0.12%	0.13%	0.15%	0.21%	0.22%	0.40%	5.77
phecode_380-11	0.03%	0.03%	0.04%	0.03%	0.02%	0.04%	0.05%	0.03%	0.06%	0.20%	6.60
phecode_380-12	0.00%	0.00%	0.03%	0.03%	0.05%	0.05%	0.08%	0.06%	0.13%	0.15%	36.50
phecode_380-2	0.07%	0.12%	0.21%	0.29%	0.31%	0.34%	0.42%	0.52%	0.58%	0.95%	13.60
phecode_380-21	0.02%	0.02%	0.02%	0.02%	0.02%	0.03%	0.03%	0.03%	0.03%	0.07%	4.13
phecode_380-22	0.00%	0.00%	0.04%	0.03%	0.03%	0.04%	0.04%	0.04%	0.04%	0.05%	Inf
phecode_380-3	0.01%	0.02%	0.03%	0.02%	0.05%	0.05%	0.06%	0.09%	0.10%	0.27%	22.67
phecode_381	0.23%	0.26%	0.33%	0.33%	0.36%	0.55%	0.58%	0.65%	0.78%	1.32%	5.81
phecode_381-1	0.08%	0.14%	0.14%	0.19%	0.16%	0.19%	0.25%	0.34%	0.39%	0.56%	7.23
phecode_381-11	0.01%	0.03%	0.02%	0.03%	0.04%	0.05%	0.07%	0.08%	0.10%	0.13%	11.00
phecode_381-3	0.04%	0.03%	0.04%	0.05%	0.06%	0.06%	0.10%	0.06%	0.11%	0.19%	4.41
phecode_381-4	0.03%	0.05%	0.05%	0.06%	0.08%	0.09%	0.09%	0.13%	0.18%	0.33%	12.62
phecode_381-6	0.05%	0.04%	0.04%	0.04%	0.07%	0.10%	0.12%	0.13%	0.17%	0.28%	5.79
phecode_381-8	0.01%	0.01%	0.04%	0.06%	0.04%	0.09%	0.07%	0.10%	0.12%	0.27%	33.75
phecode_381-81	0.00%	0.00%	0.01%	0.01%	0.02%	0.02%	0.01%	0.03%	0.05%	0.06%	14.00
phecode_381-82	0.00%	0.00%	0.02%	0.02%	0.03%	0.02%	0.05%	0.05%	0.08%	0.15%	Inf
phecode_382	0.00%	0.01%	0.02%	0.03%	0.07%	0.05%	0.09%	0.11%	0.11%	0.19%	47.00
phecode_383	0.07%	0.08%	0.13%	0.13%	0.16%	0.21%	0.25%	0.29%	0.33%	0.63%	9.58
phecode_383-1	0.00%	0.02%	0.08%	0.08%	0.09%	0.13%	0.15%	0.19%	0.24%	0.40%	201.00
phecode_384	0.05%	0.07%	0.11%	0.15%	0.18%	0.16%	0.22%	0.25%	0.32%	0.56%	12.22
phecode_384-1	0.00%	0.00%	0.05%	0.05%	0.02%	0.03%	0.04%	0.08%	0.06%	0.07%	36.00
phecode_384-3	0.00%	0.01%	0.01%	0.02%	0.04%	0.03%	0.05%	0.06%	0.08%	0.16%	81.00
phecode_384-4	0.00%	0.00%	0.02%	0.02%	0.01%	0.01%	0.03%	0.02%	0.04%	0.04%	Inf
phecode_385	0.00%	0.03%	0.12%	0.13%	0.23%	0.27%	0.39%	0.45%	0.60%	1.45%	729.01
phecode_386	0.71%	1.20%	1.99%	2.56%	3.28%	3.70%	4.53%	5.40%	6.40%	9.89%	13.85
phecode_386-1	0.15%	0.16%	0.24%	0.31%	0.35%	0.51%	0.65%	0.66%	0.94%	1.56%	10.32
phecode_386-2	0.16%	0.28%	0.27%	0.43%	0.53%	0.59%	0.68%	0.90%	1.09%	1.84%	11.80
phecode_386-3	0.00%	0.01%	0.01%	0.01%	0.02%	0.03%	0.04%	0.04%	0.08%	0.11%	28.50
phecode_386-4	0.17%	0.21%	0.27%	0.38%	0.35%	0.55%	0.59%	0.67%	0.88%	1.34%	7.81
phecode_386-41	0.00%	0.01%	0.05%	0.07%	0.08%	0.07%	0.12%	0.18%	0.19%	0.36%	Inf
phecode_386-42	0.00%	0.03%	0.05%	0.05%	0.07%	0.10%	0.14%	0.15%	0.21%	0.34%	172.00
phecode_386-8	0.00%	0.00%	0.02%	0.02%	0.03%	0.04%	0.04%	0.06%	0.09%	0.12%	60.00

4 Medical history predicts future health trajectories over the human phenome

Table 16 continued from previous page

Endpoint	1	2	3	4	5	6	7	8	9	10	Rate Ratio
phecode_386-9	0.00%	0.02%	0.09%	0.12%	0.19%	0.27%	0.28%	0.35%	0.47%	0.86%	429.01
phecode_387	0.74%	0.92%	0.97%	1.32%	1.70%	2.07%	2.52%	3.08%	3.71%	4.72%	6.35
phecode_387-1	0.04%	0.06%	0.08%	0.10%	0.15%	0.20%	0.24%	0.29%	0.43%	0.53%	13.20
phecode_387-2	0.42%	0.45%	0.45%	0.65%	0.77%	0.91%	1.00%	1.10%	1.26%	2.32%	5.53
phecode_387-21	0.03%	0.04%	0.03%	0.05%	0.05%	0.07%	0.08%	0.09%	0.10%	0.31%	9.81
phecode_387-3	0.20%	0.34%	0.36%	0.47%	0.66%	0.86%	1.10%	1.53%	1.83%	2.29%	11.27
phecode_387-4	0.00%	0.01%	0.04%	0.04%	0.03%	0.06%	0.04%	0.06%	0.08%	0.08%	20.50
phecode_387-5	0.00%	0.01%	0.02%	0.02%	0.03%	0.03%	0.03%	0.05%	0.07%	0.10%	24.50
phecode_388	0.20%	0.22%	0.34%	0.49%	0.69%	0.88%	1.14%	1.64%	2.30%	5.21%	26.36
phecode_388-1	0.02%	0.02%	0.02%	0.02%	0.04%	0.04%	0.05%	0.08%	0.12%	0.81%	51.13
phecode_389	0.22%	0.73%	4.25%	5.35%	6.58%	7.80%	9.64%	11.32%	14.27%	20.95%	95.20
phecode_389-1	0.04%	0.11%	0.28%	0.38%	0.43%	0.54%	0.73%	0.90%	1.16%	2.15%	56.58
phecode_390	0.62%	0.96%	10.68%	14.91%	17.09%	19.17%	21.04%	23.29%	27.09%	32.51%	52.45
phecode_390-1	0.20%	0.66%	4.56%	5.68%	6.64%	7.49%	8.60%	10.05%	11.92%	17.14%	83.97
phecode_390-4	0.28%	0.67%	7.21%	9.58%	11.61%	12.71%	14.70%	16.81%	19.84%	25.55%	91.90
phecode_390-5	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%	0.02%	0.02%	0.02%	0.10%	16.33
phecode_390-6	0.05%	0.08%	0.15%	0.20%	0.31%	0.39%	0.45%	0.56%	0.74%	1.22%	26.57
phecode_391	0.53%	0.96%	4.73%	5.85%	6.78%	7.89%	8.91%	10.07%	11.50%	16.05%	30.43
phecode_391-1	0.26%	0.56%	2.15%	2.73%	3.13%	3.70%	4.39%	5.13%	6.24%	9.88%	38.65
phecode_391-11	0.01%	0.03%	0.14%	0.18%	0.23%	0.29%	0.35%	0.38%	0.61%	1.01%	83.50
phecode_391-12	0.07%	0.13%	0.26%	0.36%	0.36%	0.47%	0.58%	0.69%	0.90%	1.76%	23.62
phecode_391-2	0.07%	0.30%	1.75%	2.28%	2.73%	3.19%	3.85%	4.29%	5.19%	7.41%	106.62
phecode_391-21	0.00%	0.03%	0.14%	0.19%	0.25%	0.30%	0.33%	0.36%	0.48%	0.82%	204.50
phecode_391-22	0.00%	0.01%	0.05%	0.06%	0.06%	0.09%	0.10%	0.10%	0.18%	0.25%	62.00
phecode_391-4	0.01%	0.01%	0.03%	0.01%	0.03%	0.02%	0.05%	0.05%	0.06%	0.18%	18.00
phecode_391-6	0.04%	0.05%	0.05%	0.04%	0.06%	0.07%	0.11%	0.08%	0.16%	0.44%	11.10
phecode_391-7	0.12%	0.19%	0.28%	0.33%	0.41%	0.45%	0.54%	0.66%	0.73%	1.80%	15.20
phecode_391-8	0.04%	0.03%	0.04%	0.04%	0.06%	0.06%	0.07%	0.08%	0.14%	0.15%	4.05
phecode_391-9	0.04%	0.09%	0.22%	0.23%	0.30%	0.38%	0.45%	0.55%	0.83%	2.03%	46.18
phecode_392	0.13%	0.43%	1.97%	2.83%	3.61%	4.50%	5.37%	6.76%	8.73%	13.87%	104.13
phecode_393	0.03%	0.02%	0.02%	0.03%	0.05%	0.08%	0.07%	0.08%	0.11%	0.25%	7.35
phecode_394	0.29%	0.56%	1.26%	1.72%	2.44%	2.89%	3.79%	4.52%	6.09%	9.01%	31.49
phecode_394-1	0.10%	0.10%	0.16%	0.18%	0.21%	0.29%	0.31%	0.42%	0.55%	1.01%	9.69
phecode_394-2	0.16%	0.41%	0.91%	1.44%	1.85%	2.37%	3.01%	3.84%	5.26%	7.89%	49.60
phecode_394-21	0.13%	0.29%	0.58%	0.96%	1.45%	1.90%	2.43%	3.35%	4.65%	7.25%	56.41
phecode_394-22	0.03%	0.10%	0.26%	0.31%	0.37%	0.46%	0.53%	0.65%	0.84%	1.18%	34.77
phecode_394-4	0.00%	0.02%	0.06%	0.11%	0.14%	0.18%	0.20%	0.29%	0.40%	0.55%	138.50
phecode_395	0.20%	0.39%	1.42%	1.99%	2.41%	2.83%	3.44%	3.99%	5.23%	7.19%	35.84
phecode_395-1	0.18%	0.39%	1.31%	1.86%	2.27%	2.64%	3.24%	3.68%	4.96%	6.80%	37.90
phecode_395-3	0.00%	0.00%	0.01%	0.00%	0.01%	0.02%	0.02%	0.05%	0.07%	0.07%	Inf
phecode_396	0.92%	2.27%	5.45%	7.60%	9.50%	11.38%	13.84%	16.46%	19.94%	27.34%	29.81
phecode_396-1	0.05%	0.05%	0.17%	0.15%	0.23%	0.24%	0.32%	0.34%	0.45%	0.99%	19.04
phecode_396-11	0.01%	0.01%	0.00%	0.03%	0.02%	0.01%	0.02%	0.03%	0.04%	0.19%	31.33
phecode_396-2	0.18%	0.45%	0.90%	1.28%	1.67%	2.15%	2.68%	3.32%	4.12%	6.07%	34.59
phecode_396-21	0.05%	0.14%	0.21%	0.32%	0.49%	0.65%	0.88%	1.13%	1.49%	2.40%	46.27
phecode_396-22	0.05%	0.13%	0.16%	0.28%	0.43%	0.59%	0.74%	0.96%	1.17%	1.75%	38.09
phecode_396-3	0.03%	0.02%	0.04%	0.07%	0.07%	0.10%	0.12%	0.15%	0.22%	0.56%	21.46
phecode_396-5	0.00%	0.01%	0.03%	0.02%	0.01%	0.02%	0.04%	0.04%	0.04%	0.08%	40.00
phecode_397	0.19%	0.55%	2.30%	3.11%	3.58%	4.08%	4.53%	5.42%	6.38%	8.56%	44.43
phecode_397-1	0.19%	0.54%	2.26%	3.05%	3.54%	4.05%	4.43%	5.32%	6.27%	8.45%	44.32
phecode_397-3	0.00%	0.00%	0.01%	0.01%	0.02%	0.02%	0.03%	0.03%	0.05%	0.07%	Inf
phecode_398	0.05%	0.18%	0.90%	1.10%	1.33%	1.57%	1.75%	2.04%	2.78%	4.06%	87.65
phecode_398-1	0.01%	0.03%	0.15%	0.22%	0.32%	0.35%	0.44%	0.55%	0.73%	1.13%	188.67
phecode_400	0.24%	0.49%	0.72%	0.87%	0.98%	1.22%	1.46%	1.98%	3.06%	6.76%	28.21
phecode_400-2	0.24%	0.50%	0.70%	0.88%	0.97%	1.21%	1.46%	1.99%	3.07%	6.77%	28.74
phecode_401	9.18%	13.57%	16.55%	17.76%	20.80%	23.71%	26.69%	30.43%	36.54%	50.07%	5.45
phecode_401-1	9.18%	13.58%	16.50%	17.63%	20.84%	23.60%	26.62%	30.48%	36.40%	50.08%	5.45
phecode_401-2	0.01%	0.02%	0.03%	0.06%	0.08%	0.13%	0.15%	0.25%	0.35%	0.99%	123.75
phecode_401-3	0.02%	0.02%	0.03%	0.04%	0.06%	0.08%	0.15%	0.22%	0.32%	2.15%	134.63
phecode_401-6	0.02%	0.11%	0.37%	0.45%	0.48%	0.62%	0.73%	0.84%	0.93%	1.43%	65.27
phecode_402	0.45%	0.80%	4.31%	5.74%	6.46%	7.16%	8.07%	8.45%	9.59%	11.57%	25.51
phecode_403	0.59%	1.34%	2.06%	2.61%	2.62%	3.56%	4.52%	5.84%	7.50%	14.66%	24.66
phecode_403-1	0.03%	0.01%	0.01%	0.02%	0.03%	0.05%	0.07%	0.05%	0.11%	0.23%	9.08
phecode_404	0.97%	2.12%	3.21%	4.55%	5.70%	6.76%	7.46%	9.93%	12.76%	22.42%	23.02
phecode_404-1	0.38%	0.81%	1.36%	1.99%	2.39%	3.05%	3.35%	4.31%	5.77%	12.64%	33.53

Supplementary Tables

Table 16 continued from previous page

Endpoint	1	2	3	4	5	6	7	8	9	10	Rate Ratio
phcode_404-11	0.31%	0.63%	1.13%	1.51%	2.01%	2.45%	2.80%	3.44%	4.34%	7.42%	23.99
phcode_404-2	0.59%	1.37%	2.38%	3.29%	4.24%	4.87%	5.48%	7.46%	9.53%	16.80%	28.64
phcode_406	0.05%	0.10%	0.13%	0.21%	0.23%	0.36%	0.48%	0.53%	0.93%	2.80%	58.58
phcode_406-1	0.05%	0.11%	0.11%	0.20%	0.24%	0.31%	0.41%	0.49%	0.81%	2.41%	52.57
phcode_406-11	0.01%	0.02%	0.02%	0.03%	0.06%	0.07%	0.07%	0.10%	0.15%	0.62%	77.75
phcode_406-13	0.00%	0.00%	0.00%	0.01%	0.00%	0.02%	0.02%	0.03%	0.04%	0.16%	Inf
phcode_408	0.01%	0.02%	0.01%	0.01%	0.03%	0.01%	0.01%	0.03%	0.04%	0.07%	7.20
phcode_410	0.23%	0.32%	0.35%	0.45%	0.63%	0.67%	0.89%	1.03%	1.29%	2.56%	10.98
phcode_410-1	0.04%	0.06%	0.07%	0.09%	0.11%	0.15%	0.20%	0.19%	0.28%	0.38%	9.50
phcode_410-2	0.15%	0.18%	0.24%	0.26%	0.36%	0.51%	0.60%	0.82%	1.00%	2.11%	13.69
phcode_410-3	0.03%	0.06%	0.06%	0.07%	0.07%	0.07%	0.07%	0.09%	0.12%	0.15%	4.53
phcode_411	0.14%	0.21%	0.29%	0.35%	0.44%	0.56%	0.59%	0.67%	0.92%	1.67%	11.80
phcode_411-1	0.00%	0.01%	0.00%	0.01%	0.01%	0.03%	0.05%	0.05%	0.12%	0.13%	Inf
phcode_411-2	0.09%	0.13%	0.19%	0.19%	0.34%	0.41%	0.42%	0.48%	0.65%	1.26%	13.45
phcode_413	0.81%	1.50%	2.05%	2.42%	2.56%	3.26%	4.34%	5.54%	7.39%	13.51%	16.69
phcode_413-1	0.55%	0.93%	1.25%	1.41%	1.66%	2.03%	2.57%	3.15%	4.42%	8.84%	15.93
phcode_413-11	0.40%	0.59%	0.75%	0.85%	1.02%	1.33%	1.59%	2.01%	2.71%	5.43%	13.64
phcode_413-12	0.06%	0.10%	0.10%	0.12%	0.13%	0.19%	0.21%	0.24%	0.26%	0.49%	8.27
phcode_413-13	0.01%	0.00%	0.02%	0.03%	0.03%	0.03%	0.06%	0.06%	0.11%	0.37%	31.00
phcode_413-2	0.31%	0.73%	0.94%	1.24%	1.27%	1.70%	2.33%	3.14%	4.46%	7.77%	25.39
phcode_413-21	0.08%	0.21%	0.43%	0.55%	0.67%	0.86%	1.04%	1.64%	2.37%	4.54%	55.49
phcode_413-22	0.15%	0.30%	0.37%	0.42%	0.57%	0.68%	0.96%	1.21%	1.50%	2.67%	18.32
phcode_413-3	0.21%	0.36%	0.56%	0.67%	0.75%	0.93%	1.16%	1.48%	2.26%	5.44%	26.52
phcode_413-32	0.07%	0.09%	0.13%	0.16%	0.21%	0.23%	0.28%	0.43%	0.60%	1.52%	20.65
phcode_413-4	0.05%	0.07%	0.08%	0.05%	0.09%	0.14%	0.19%	0.23%	0.34%	0.68%	13.68
phcode_413-42	0.04%	0.07%	0.06%	0.07%	0.07%	0.12%	0.19%	0.22%	0.32%	0.67%	15.32
phcode_413-6	0.05%	0.10%	0.14%	0.27%	0.30%	0.32%	0.39%	0.51%	0.73%	2.99%	65.26
phcode_414	0.16%	0.32%	0.39%	0.47%	0.49%	0.50%	0.59%	0.76%	1.01%	2.77%	16.93
phcode_414-1	0.03%	0.07%	0.07%	0.09%	0.10%	0.10%	0.12%	0.12%	0.18%	0.41%	12.24
phcode_414-11	0.01%	0.02%	0.05%	0.05%	0.04%	0.05%	0.04%	0.07%	0.08%	0.27%	19.43
phcode_414-2	0.06%	0.08%	0.11%	0.16%	0.18%	0.19%	0.22%	0.28%	0.34%	0.93%	15.60
phcode_414-5	0.01%	0.01%	0.01%	0.04%	0.02%	0.07%	0.07%	0.07%	0.13%	1.20%	201.00
phcode_414-9	0.00%	0.00%	0.01%	0.00%	0.00%	0.02%	0.04%	0.06%	0.07%	0.13%	Inf
phcode_416	2.52%	4.41%	5.99%	6.93%	7.39%	9.32%	11.36%	14.12%	18.18%	28.32%	11.24
phcode_416-1	0.46%	0.57%	0.78%	0.72%	0.94%	1.02%	1.19%	1.51%	1.71%	3.52%	7.60
phcode_416-11	0.36%	0.43%	0.52%	0.57%	0.70%	0.82%	0.91%	1.11%	1.33%	2.39%	6.66
phcode_416-12	0.05%	0.11%	0.14%	0.20%	0.24%	0.27%	0.32%	0.37%	0.56%	1.58%	30.50
phcode_416-2	0.83%	1.91%	3.05%	4.02%	4.77%	5.21%	6.71%	8.71%	11.62%	19.76%	23.70
phcode_416-21	0.50%	0.86%	1.18%	1.94%	2.78%	3.50%	4.67%	6.03%	8.25%	14.79%	29.53
phcode_416-211	0.33%	0.58%	0.70%	0.94%	1.31%	1.46%	2.04%	2.41%	3.36%	6.02%	18.14
phcode_416-212	0.03%	0.07%	0.12%	0.13%	0.17%	0.23%	0.27%	0.32%	0.48%	1.31%	41.06
phcode_416-213	0.01%	0.01%	0.03%	0.03%	0.03%	0.07%	0.07%	0.10%	0.18%	0.70%	117.00
phcode_416-214	0.00%	0.00%	0.00%	0.01%	0.01%	0.01%	0.02%	0.03%	0.03%	0.14%	Inf
phcode_416-22	0.05%	0.08%	0.14%	0.29%	0.43%	0.61%	0.75%	0.94%	1.35%	2.96%	59.40
phcode_416-221	0.01%	0.01%	0.02%	0.01%	0.03%	0.03%	0.05%	0.05%	0.08%	0.16%	20.25
phcode_416-222	0.00%	0.01%	0.01%	0.01%	0.01%	0.01%	0.02%	0.01%	0.03%	0.09%	45.00
phcode_416-3	0.02%	0.04%	0.06%	0.08%	0.12%	0.14%	0.16%	0.17%	0.23%	0.63%	28.73
phcode_416-31	0.00%	0.00%	0.02%	0.01%	0.02%	0.02%	0.03%	0.04%	0.05%	0.14%	Inf
phcode_416-4	0.58%	1.25%	1.62%	1.82%	2.14%	2.79%	3.33%	4.55%	6.34%	12.00%	20.73
phcode_416-41	0.18%	0.41%	0.61%	0.63%	0.79%	0.99%	1.36%	1.88%	2.85%	5.72%	32.63
phcode_416-42	0.18%	0.43%	0.52%	0.67%	0.70%	0.91%	1.06%	1.43%	1.96%	4.03%	22.22
phcode_416-43	0.21%	0.44%	0.54%	0.58%	0.78%	0.94%	1.13%	1.61%	2.19%	4.16%	19.47
phcode_416-5	0.21%	0.42%	0.61%	0.93%	1.02%	1.27%	1.54%	1.85%	2.38%	3.47%	16.37
phcode_416-51	0.04%	0.07%	0.10%	0.18%	0.24%	0.28%	0.35%	0.44%	0.45%	0.77%	20.42
phcode_416-52	0.14%	0.25%	0.36%	0.54%	0.59%	0.66%	0.91%	1.08%	1.33%	2.02%	14.06
phcode_416-6	0.00%	0.01%	0.02%	0.01%	0.02%	0.01%	0.04%	0.04%	0.05%	0.05%	13.50
phcode_416-7	0.03%	0.07%	0.07%	0.08%	0.13%	0.18%	0.22%	0.27%	0.36%	0.94%	33.86
phcode_416-71	0.03%	0.07%	0.07%	0.08%	0.13%	0.18%	0.22%	0.27%	0.36%	0.94%	33.86
phcode_416-8	0.02%	0.02%	0.02%	0.04%	0.03%	0.03%	0.03%	0.04%	0.03%	0.06%	3.63
phcode_417	2.42%	4.33%	7.12%	9.17%	10.71%	11.70%	14.24%	16.22%	19.59%	28.28%	11.70
phcode_417-1	0.70%	1.31%	2.60%	3.54%	4.46%	5.37%	6.54%	7.87%	9.91%	14.60%	20.79
phcode_417-2	0.59%	0.77%	0.85%	1.05%	1.21%	1.25%	1.56%	2.00%	2.49%	4.57%	7.78
phcode_417-3	0.62%	0.80%	1.00%	1.32%	1.67%	1.95%	2.33%	2.97%	4.04%	7.13%	11.50
phcode_418	0.24%	0.53%	1.31%	2.14%	2.70%	3.44%	4.25%	5.44%	7.12%	12.00%	49.18
phcode_418-1	0.05%	0.21%	0.98%	1.48%	1.97%	2.51%	2.98%	3.90%	5.16%	8.84%	176.00

4 Medical history predicts future health trajectories over the human phenome

Table 16 continued from previous page

Endpoint	1	2	3	4	5	6	7	8	9	10	Rate Ratio
phecode_419	0.12%	0.26%	0.36%	0.47%	0.58%	0.73%	0.86%	1.36%	2.04%	6.23%	52.93
phecode_419-2	0.00%	0.00%	0.00%	0.01%	0.01%	0.01%	0.01%	0.02%	0.03%	0.19%	Inf
phecode_420	0.06%	0.13%	0.19%	0.22%	0.27%	0.33%	0.54%	0.72%	0.92%	2.47%	42.69
phecode_423	0.26%	0.36%	0.53%	0.66%	0.88%	1.12%	1.29%	1.61%	2.29%	3.53%	13.80
phecode_423-1	0.28%	0.29%	0.46%	0.55%	0.79%	0.93%	1.07%	1.34%	1.99%	2.96%	10.74
phecode_424	0.33%	0.88%	1.30%	1.69%	2.12%	2.66%	3.00%	4.38%	6.32%	16.25%	48.56
phecode_424-1	0.15%	0.44%	0.61%	0.86%	1.07%	1.27%	1.44%	1.96%	3.02%	7.99%	52.62
phecode_424-2	0.01%	0.04%	0.12%	0.21%	0.29%	0.43%	0.53%	0.69%	0.98%	3.02%	303.61
phecode_424-3	0.01%	0.03%	0.07%	0.10%	0.12%	0.21%	0.27%	0.36%	0.58%	1.23%	206.67
phecode_424-5	0.00%	0.00%	0.00%	0.01%	0.01%	0.01%	0.01%	0.03%	0.03%	0.12%	Inf
phecode_424-6	0.00%	0.01%	0.01%	0.01%	0.01%	0.02%	0.05%	0.05%	0.08%	0.32%	81.00
phecode_425	0.29%	0.66%	0.93%	1.16%	1.32%	1.59%	2.00%	2.69%	3.96%	8.68%	30.14
phecode_426	0.29%	0.65%	0.71%	0.88%	1.20%	1.48%	2.00%	2.56%	3.68%	9.28%	31.84
phecode_430	0.27%	0.38%	0.45%	0.57%	0.69%	0.85%	0.95%	0.95%	1.24%	2.32%	8.73
phecode_430-1	0.19%	0.20%	0.21%	0.27%	0.29%	0.35%	0.40%	0.46%	0.47%	0.81%	4.32
phecode_430-2	0.07%	0.14%	0.22%	0.25%	0.35%	0.37%	0.44%	0.54%	0.66%	1.27%	18.20
phecode_430-3	0.04%	0.06%	0.09%	0.11%	0.14%	0.15%	0.20%	0.21%	0.35%	0.67%	16.14
phecode_431	0.88%	1.70%	2.34%	2.67%	3.02%	3.60%	4.67%	5.54%	7.33%	10.53%	11.99
phecode_431-1	0.57%	1.04%	1.38%	1.70%	2.26%	2.43%	2.74%	3.57%	4.61%	7.40%	12.97
phecode_431-11	0.38%	0.75%	1.07%	1.33%	1.65%	1.75%	2.45%	3.01%	4.04%	6.48%	17.03
phecode_431-12	0.22%	0.31%	0.37%	0.46%	0.57%	0.62%	0.74%	0.78%	0.92%	1.68%	7.50
phecode_431-14	0.01%	0.01%	0.02%	0.03%	0.06%	0.08%	0.10%	0.13%	0.13%	0.22%	36.33
phecode_431-15	0.01%	0.02%	0.02%	0.05%	0.08%	0.06%	0.15%	0.21%	0.31%	0.55%	46.33
phecode_431-2	0.36%	0.55%	0.76%	1.01%	1.45%	1.78%	2.29%	2.78%	3.68%	6.34%	17.47
phecode_433	0.34%	0.70%	1.01%	1.36%	1.58%	1.89%	2.60%	3.40%	5.02%	8.73%	25.52
phecode_433-1	0.08%	0.19%	0.21%	0.28%	0.33%	0.41%	0.54%	0.67%	0.99%	1.63%	20.92

Supplementary Tables

Table 17: Relative and absolute discriminatory performances for all endpoints of the medical history model. Displayed are absolute discriminatory performances for CPH models trained on Age+Sex and Age+Sex+MedicalHistory as well as the relative performance delta. Reported are median values and 95 % quantile borders of distributions bootstrapped with 100 iterations.

Endpoint	PheCode string	C-Index Age+Sex (IQR)	C-Index Age+Sex+MedicalHistory (IQR)	Difference in C-Index (IQR)
OMOP_4306655	All-Cause Death	0.701 (0.699, 0.703)	0.878 (0.877, 0.88)	0.177 (0.175, 0.179)
phecode_001	Salmonella	0.493 (0.469, 0.518)	0.676 (0.653, 0.7)	0.183 (0.152, 0.217)
phecode_002	Staphylococcus	0.612 (0.606, 0.617)	0.728 (0.722, 0.734)	0.116 (0.11, 0.123)
phecode_002-1	Staphylococcus aureus	0.607 (0.6, 0.613)	0.732 (0.725, 0.739)	0.125 (0.118, 0.133)
phecode_003	Escherichia coli	0.602 (0.597, 0.607)	0.675 (0.67, 0.68)	0.073 (0.068, 0.078)
phecode_004	Streptococcus	0.57 (0.563, 0.576)	0.656 (0.649, 0.664)	0.087 (0.079, 0.095)
phecode_004-1	Streptococcus pneumoniae	0.582 (0.566, 0.597)	0.689 (0.673, 0.705)	0.107 (0.089, 0.126)
phecode_004-2	Group A Streptococcus	0.555 (0.531, 0.581)	0.671 (0.646, 0.696)	0.115 (0.083, 0.145)
phecode_004-3	Group B Streptococcus	0.593 (0.57, 0.618)	0.74 (0.716, 0.764)	0.146 (0.114, 0.178)
phecode_005	Mycobacteria	0.519 (0.502, 0.535)	0.694 (0.676, 0.712)	0.174 (0.152, 0.2)
phecode_005-1	Mycobacterium tuberculosis	0.526 (0.507, 0.546)	0.708 (0.689, 0.728)	0.183 (0.155, 0.208)
phecode_005-2	Nontuberculous mycobacteria	0.608 (0.574, 0.641)	0.742 (0.711, 0.777)	0.133 (0.102, 0.166)
phecode_007	Hemophilus infection	0.633 (0.619, 0.647)	0.79 (0.777, 0.803)	0.157 (0.14, 0.173)
phecode_007-1	Hemophilus influenzae	0.631 (0.619, 0.647)	0.793 (0.778, 0.807)	0.161 (0.143, 0.178)
phecode_008	Helicobacter [H. pylori]	0.518 (0.51, 0.525)	0.72 (0.713, 0.726)	0.202 (0.192, 0.212)
phecode_009	Pseudomonas	0.666 (0.656, 0.674)	0.809 (0.8, 0.817)	0.143 (0.133, 0.154)
phecode_010	Corynebacterium	0.677 (0.657, 0.696)	0.747 (0.731, 0.763)	0.07 (0.054, 0.087)
phecode_011	Klebsiella	0.665 (0.653, 0.679)	0.785 (0.774, 0.797)	0.12 (0.106, 0.135)
phecode_012	Proteus	0.654 (0.635, 0.67)	0.817 (0.801, 0.83)	0.163 (0.144, 0.184)
phecode_015	Clostridium	0.628 (0.617, 0.64)	0.711 (0.7, 0.723)	0.083 (0.071, 0.094)
phecode_015-2	Clostridium difficile	0.648 (0.637, 0.659)	0.738 (0.727, 0.749)	0.09 (0.078, 0.101)
phecode_019	Treponema	0.604 (0.571, 0.636)	0.791 (0.76, 0.823)	0.188 (0.148, 0.226)
phecode_020	Borrelia	0.501 (0.482, 0.52)	0.687 (0.671, 0.703)	0.186 (0.163, 0.207)
phecode_020-1	Lyme disease	0.502 (0.486, 0.521)	0.687 (0.672, 0.702)	0.184 (0.163, 0.207)
phecode_024	Pertussis	0.578 (0.553, 0.6)	0.667 (0.646, 0.688)	0.09 (0.064, 0.114)
phecode_025	Enterococcus	0.69 (0.67, 0.708)	0.772 (0.755, 0.791)	0.083 (0.061, 0.102)
phecode_030	Campylobacter	0.543 (0.536, 0.549)	0.706 (0.698, 0.712)	0.163 (0.154, 0.171)
phecode_050	Enterovirus	0.515 (0.484, 0.544)	0.616 (0.589, 0.643)	0.101 (0.067, 0.14)
phecode_050-4	Hand, foot, and mouth disease	0.576 (0.545, 0.608)	0.673 (0.645, 0.698)	0.096 (0.058, 0.131)
phecode_052	Herpesvirus	0.561 (0.559, 0.564)	0.674 (0.672, 0.676)	0.112 (0.11, 0.115)
phecode_052-1	Herpes simplex	0.604 (0.6, 0.61)	0.719 (0.714, 0.723)	0.114 (0.109, 0.12)
phecode_052-3	Varicella zoster virus	0.571 (0.568, 0.573)	0.673 (0.671, 0.675)	0.102 (0.1, 0.105)
phecode_052-31	Varicella [chickenpox]	0.608 (0.59, 0.626)	0.663 (0.646, 0.678)	0.055 (0.039, 0.069)
phecode_052-32	Herpes zoster	0.575 (0.572, 0.578)	0.675 (0.673, 0.678)	0.1 (0.098, 0.102)
phecode_052-4	Infectious mononucleosis	0.575 (0.546, 0.608)	0.63 (0.6, 0.658)	0.055 (0.018, 0.086)
phecode_052-5	Cytomegalovirus [CMV]	0.488 (0.459, 0.519)	0.691 (0.658, 0.722)	0.201 (0.161, 0.246)
phecode_054	Hepatovirus	0.579 (0.562, 0.594)	0.487 (0.467, 0.504)	-0.092 (-0.115, -0.068)
phecode_054-2	Hepatitis B	0.596 (0.566, 0.623)	0.385 (0.357, 0.414)	-0.21 (-0.25, -0.175)
phecode_054-3	Hepatitis C	0.645 (0.616, 0.666)	0.745 (0.717, 0.772)	0.1 (0.066, 0.133)
phecode_054-31	Chronic hepatitis C	0.652 (0.625, 0.678)	0.764 (0.738, 0.79)	0.113 (0.078, 0.146)
phecode_054-5	Hepatitis E	0.508 (0.462, 0.56)	0.561 (0.503, 0.613)	0.051 (-0.021, 0.126)
phecode_055	Poxvirus	0.624 (0.602, 0.646)	0.687 (0.669, 0.704)	0.063 (0.043, 0.084)
phecode_055-1	Molluscum contagiosum	0.627 (0.606, 0.649)	0.691 (0.671, 0.711)	0.064 (0.042, 0.087)
phecode_056	Human papillomavirus	0.511 (0.508, 0.515)	0.685 (0.682, 0.688)	0.174 (0.17, 0.178)
phecode_056-1	Plantar wart	0.541 (0.535, 0.546)	0.697 (0.692, 0.701)	0.156 (0.149, 0.162)
phecode_057	Retrovirus	0.786 (0.763, 0.808)	0.815 (0.795, 0.837)	0.029 (0.004, 0.054)
phecode_057-1	Human immunodeficiency virus	0.788 (0.765, 0.813)	0.82 (0.799, 0.839)	0.032 (0.008, 0.054)
phecode_058	Pneumoviridae	0.621 (0.592, 0.653)	0.748 (0.714, 0.78)	0.126 (0.086, 0.164)
phecode_058-1	Respiratory syncytial virus	0.622 (0.591, 0.654)	0.75 (0.718, 0.785)	0.128 (0.085, 0.169)
phecode_059	Coronavirus	0.574 (0.541, 0.607)	0.673 (0.643, 0.704)	0.099 (0.058, 0.143)
phecode_059-1	COVID-19*	0.584 (0.548, 0.627)	0.679 (0.648, 0.715)	0.097 (0.044, 0.141)
phecode_060	Adenovirus	0.535 (0.492, 0.577)	0.608 (0.564, 0.652)	0.075 (0.015, 0.133)
phecode_061	Influenza virus	0.537 (0.529, 0.544)	0.681 (0.674, 0.689)	0.144 (0.134, 0.154)
phecode_066	Orthorubulavirus [Mumps]	0.656 (0.623, 0.693)	0.714 (0.681, 0.745)	0.059 (0.028, 0.088)
phecode_069	Other specified viral infections	0.431 (0.401, 0.461)	0.711 (0.675, 0.747)	0.28 (0.232, 0.326)
phecode_070	Candidiasis	0.585 (0.583, 0.588)	0.723 (0.721, 0.726)	0.138 (0.135, 0.141)

4 Medical history predicts future health trajectories over the human phenome

Table 17 continued from previous page

Endpoint	PheCode string	C-Index Age+Sex (IQR)	C-Index Age+Sex+MedicalHistory (IQR)	Difference in C-Index (IQR)
phecode_074	Aspergillosis	0.595 (0.572, 0.617)	0.814 (0.795, 0.834)	0.22 (0.186, 0.25)
phecode_076	Pneumocystosis	0.669 (0.638, 0.702)	0.783 (0.754, 0.813)	0.114 (0.077, 0.149)
phecode_084	Parasites	0.594 (0.584, 0.604)	0.702 (0.694, 0.71)	0.108 (0.098, 0.117)
phecode_084-2	Malaria [Plasmodium]	0.554 (0.531, 0.579)	0.77 (0.749, 0.792)	0.215 (0.187, 0.246)
phecode_084-4	Trichomoniasis	0.794 (0.781, 0.808)	0.892 (0.883, 0.901)	0.098 (0.083, 0.112)
phecode_084-6	Enterobiasis	0.591 (0.576, 0.608)	0.664 (0.647, 0.681)	0.071 (0.055, 0.089)
phecode_084-7	Giardiasis	0.499 (0.473, 0.524)	0.678 (0.658, 0.698)	0.179 (0.153, 0.209)
phecode_086	Pediculosis, acariasis and other infestations	0.558 (0.548, 0.57)	0.709 (0.699, 0.717)	0.15 (0.138, 0.163)
phecode_088	Sexually transmitted disease	0.653 (0.644, 0.662)	0.749 (0.741, 0.756)	0.096 (0.087, 0.104)
phecode_089	Infections	0.522 (0.521, 0.524)	0.679 (0.678, 0.68)	0.157 (0.155, 0.159)
phecode_089-1	Bacterial infections	0.558 (0.556, 0.561)	0.649 (0.647, 0.651)	0.091 (0.088, 0.094)
phecode_089-2	Viral infections	0.529 (0.528, 0.531)	0.682 (0.681, 0.684)	0.153 (0.151, 0.155)
phecode_089-3	Fungal infections	0.516 (0.514, 0.518)	0.688 (0.686, 0.689)	0.171 (0.169, 0.174)
phecode_091	Gangrene	0.637 (0.609, 0.664)	0.731 (0.7, 0.761)	0.094 (0.064, 0.125)
phecode_092	Bacteremia, Sepsis, and SIRS	0.654 (0.65, 0.657)	0.739 (0.735, 0.742)	0.085 (0.081, 0.089)
phecode_092-1	Systemic inflammatory response syndrome	0.678 (0.662, 0.694)	0.78 (0.765, 0.794)	0.102 (0.085, 0.117)
phecode_092-2	Sepsis	0.653 (0.65, 0.658)	0.739 (0.735, 0.742)	0.086 (0.082, 0.089)
phecode_092-8	Bacteremia	0.665 (0.612, 0.711)	0.809 (0.771, 0.852)	0.146 (0.094, 0.202)
phecode_095	Sequela of infection	0.6 (0.59, 0.61)	0.715 (0.707, 0.725)	0.115 (0.103, 0.128)
phecode_096	Contact or exposure to infectious agent	0.456 (0.443, 0.467)	0.679 (0.67, 0.69)	0.223 (0.208, 0.239)
phecode_097	Drug resistant microorganisms	0.581 (0.569, 0.595)	0.783 (0.773, 0.794)	0.202 (0.189, 0.215)
phecode_097-1	Methicillin resistant Staphylococcus aureus	0.582 (0.567, 0.597)	0.788 (0.777, 0.8)	0.206 (0.19, 0.222)
phecode_098	Carrier or suspected carrier of infectious diseases	0.601 (0.592, 0.61)	0.748 (0.74, 0.756)	0.147 (0.137, 0.157)
phecode_098-2	Carrier or suspected carrier of Staphylococcus aureus	0.527 (0.489, 0.567)	0.791 (0.763, 0.82)	0.264 (0.226, 0.302)
phecode_099	Lab findings related to infections	0.641 (0.636, 0.646)	0.753 (0.749, 0.756)	0.112 (0.108, 0.116)
phecode_100	Malignant neoplasm of the head and neck	0.627 (0.617, 0.637)	0.681 (0.671, 0.692)	0.053 (0.045, 0.063)
phecode_100-1	Malignant neoplasm of the oral cavity	0.606 (0.589, 0.624)	0.664 (0.647, 0.684)	0.059 (0.037, 0.077)
phecode_100-12	Malignant neoplasm of the tongue	0.6 (0.582, 0.622)	0.629 (0.606, 0.65)	0.028 (0.003, 0.052)
phecode_100-2	Malignant neoplasm of the oropharynx	0.594 (0.571, 0.618)	0.712 (0.69, 0.734)	0.119 (0.093, 0.141)
phecode_100-5	Malignant neoplasm of nasal cavities, middle ear, and accessory sinuses	0.565 (0.522, 0.611)	0.577 (0.525, 0.628)	0.011 (-0.034, 0.055)
phecode_100-6	Malignant neoplasm of the larynx	0.76 (0.739, 0.781)	0.835 (0.818, 0.853)	0.075 (0.056, 0.092)
phecode_100-7	Malignant neoplasm of the pharynx	0.663 (0.628, 0.709)	0.797 (0.762, 0.838)	0.134 (0.092, 0.175)
phecode_100-8	Malignant neoplasm of the lip	0.602 (0.556, 0.65)	0.675 (0.627, 0.72)	0.072 (0.023, 0.122)
phecode_100-9	Malignant neoplasm of the salivary glands	0.507 (0.474, 0.535)	0.546 (0.508, 0.585)	0.04 (-0.001, 0.088)
phecode_101	Malignant neoplasm of the digestive organs	0.668 (0.665, 0.672)	0.702 (0.699, 0.706)	0.034 (0.031, 0.037)
phecode_101-1	Malignant neoplasm of the esophagus	0.738 (0.728, 0.748)	0.8 (0.791, 0.809)	0.062 (0.054, 0.07)
phecode_101-2	Malignant neoplasm of stomach	0.707 (0.696, 0.718)	0.76 (0.749, 0.771)	0.054 (0.044, 0.063)
phecode_101-21	Malignant neoplasm of cardia	0.755 (0.739, 0.77)	0.799 (0.786, 0.813)	0.044 (0.031, 0.058)
phecode_101-3	Malignant neoplasm of the small intestine	0.644 (0.624, 0.664)	0.704 (0.683, 0.722)	0.059 (0.042, 0.08)
phecode_101-4	Malignant neoplasm of the colon and rectum	0.666 (0.661, 0.671)	0.675 (0.669, 0.68)	0.009 (0.005, 0.011)
phecode_101-41	Malignant neoplasm of the colon	0.666 (0.66, 0.672)	0.677 (0.671, 0.683)	0.011 (0.007, 0.015)
phecode_101-42	Malignant neoplasm of the rectum	0.667 (0.659, 0.676)	0.668 (0.659, 0.677)	0.001 (-0.005, 0.006)
phecode_101-5	Malignant neoplasm of the anus and anal canal	0.6 (0.575, 0.624)	0.662 (0.637, 0.688)	0.062 (0.033, 0.093)
phecode_101-6	Malignant neoplasm of the liver and intrahepatic bile ducts	0.693 (0.68, 0.707)	0.812 (0.802, 0.824)	0.119 (0.107, 0.133)
phecode_101-61	Malignant neoplasm of the liver	0.733 (0.714, 0.751)	0.87 (0.854, 0.886)	0.137 (0.118, 0.154)
phecode_101-62	Malignant neoplasm of the intrahepatic bile ducts	0.683 (0.667, 0.701)	0.807 (0.792, 0.821)	0.123 (0.108, 0.141)
phecode_101-7	Malignant neoplasm of the gallbladder and extrahepatic bile ducts	0.675 (0.655, 0.693)	0.767 (0.75, 0.785)	0.092 (0.075, 0.109)
phecode_101-71	Malignant neoplasm of the gallbladder	0.682 (0.652, 0.712)	0.782 (0.755, 0.812)	0.102 (0.071, 0.131)
phecode_101-8	Malignant neoplasm of the pancreas	0.684 (0.674, 0.693)	0.804 (0.797, 0.812)	0.121 (0.111, 0.129)
phecode_102	Malignant neoplasm of the thoracic and respiratory organs	0.685 (0.68, 0.69)	0.814 (0.809, 0.818)	0.129 (0.123, 0.133)

Supplementary Tables

Table 17 continued from previous page

Endpoint	PheCode string	C-Index Age+Sex (IQR)	C-Index Age+Sex+MedicalHistory (IQR)	Difference in C-Index (IQR)
phecode_102-1	Malignant neoplasm of the of bronchus and lung	0.69 (0.685, 0.696)	0.826 (0.821, 0.83)	0.136 (0.131, 0.141)
phecode_102-3	Malignant neoplasm of the trachea	0.678 (0.655, 0.704)	0.77 (0.746, 0.795)	0.092 (0.07, 0.113)
phecode_102-5	Malignant neoplasm of the heart, mediastinum, thymus, and pleura	0.642 (0.61, 0.678)	0.755 (0.726, 0.786)	0.114 (0.077, 0.147)
phecode_103	Malignant neoplasm of the skin	0.658 (0.655, 0.66)	0.69 (0.688, 0.692)	0.032 (0.031, 0.034)
phecode_103-1	Melanomas of skin	0.606 (0.6, 0.612)	0.638 (0.631, 0.645)	0.032 (0.026, 0.037)
phecode_103-2	Keratinocyte carcinoma	0.656 (0.653, 0.659)	0.758 (0.755, 0.761)	0.102 (0.1, 0.105)
phecode_103-21	Basal cell carcinoma	0.648 (0.644, 0.651)	0.759 (0.756, 0.762)	0.111 (0.108, 0.114)
phecode_103-22	Squamous cell carcinoma of the skin	0.735 (0.727, 0.743)	0.798 (0.791, 0.806)	0.063 (0.058, 0.069)
phecode_103-3	Carcinoma in situ of skin	0.712 (0.707, 0.718)	0.777 (0.772, 0.782)	0.065 (0.06, 0.069)
phecode_104	Malignant sarcoma-related cancers	0.588 (0.579, 0.598)	0.614 (0.603, 0.624)	0.026 (0.017, 0.036)
phecode_104-1	Malignant neoplasm of the bone and/or cartilage	0.572 (0.55, 0.595)	0.627 (0.602, 0.653)	0.054 (0.032, 0.08)
phecode_104-2	Malignant neoplasm of retroperitoneum and peritoneum	0.732 (0.711, 0.751)	0.774 (0.757, 0.792)	0.043 (0.025, 0.056)
phecode_104-3	Malignant neoplasm of other connective and soft tissue	0.583 (0.567, 0.599)	0.618 (0.601, 0.635)	0.036 (0.02, 0.051)
phecode_104-5	Gastrointestinal stromal tumor*	0.634 (0.589, 0.673)	0.695 (0.663, 0.723)	0.062 (0.025, 0.098)
phecode_105	Malignant neoplasm of the breast	0.752 (0.75, 0.754)	0.75 (0.748, 0.753)	-0.002 (-0.004, 0.001)
phecode_105-1	Malignant neoplasm of the breast, female	0.534 (0.529, 0.539)	0.617 (0.612, 0.621)	0.082 (0.077, 0.087)
phecode_106	Gynecological malignant neoplasms	0.569 (0.563, 0.577)	0.573 (0.565, 0.58)	0.003 (-0.006, 0.011)
phecode_106-1	Malignant neoplasm of external female genital organs and cervix	0.556 (0.538, 0.572)	0.583 (0.565, 0.599)	0.027 (0.007, 0.047)
phecode_106-11	Malignant neoplasm of the vulva	0.594 (0.569, 0.622)	0.631 (0.603, 0.658)	0.036 (0.001, 0.073)
phecode_106-13	Malignant neoplasm of the cervix	0.65 (0.63, 0.668)	0.636 (0.614, 0.656)	-0.014 (-0.033, 0.006)
phecode_106-2	Malignant neoplasm of the uterus	0.602 (0.593, 0.611)	0.624 (0.613, 0.635)	0.022 (0.011, 0.035)
phecode_106-21	Malignant neoplasm of endometrium	0.606 (0.596, 0.616)	0.629 (0.618, 0.64)	0.023 (0.01, 0.035)
phecode_106-3	Malignant neoplasm of the ovary	0.603 (0.592, 0.614)	0.628 (0.616, 0.64)	0.025 (0.011, 0.036)
phecode_106-4	Malignant neoplasm of the fallopian tube and uterine adnexa	0.609 (0.572, 0.649)	0.63 (0.595, 0.667)	0.022 (-0.017, 0.061)
phecode_107	Malignant neoplasm of male genitalia	0.682 (0.679, 0.685)	0.688 (0.685, 0.692)	0.006 (0.005, 0.008)
phecode_107-1	Malignant neoplasm of the penis	0.665 (0.646, 0.69)	0.763 (0.742, 0.787)	0.097 (0.075, 0.12)
phecode_107-2	Malignant neoplasm of the prostate	0.686 (0.683, 0.69)	0.692 (0.689, 0.696)	0.006 (0.004, 0.007)
phecode_107-3	Malignant neoplasm of the testis	0.595 (0.562, 0.635)	0.562 (0.525, 0.596)	-0.035 (-0.077, 0.011)
phecode_108	Malignant neoplasm of the urinary tract	0.727 (0.723, 0.733)	0.742 (0.737, 0.747)	0.015 (0.011, 0.018)
phecode_108-4	Malignant neoplasm of the kidney	0.684 (0.674, 0.693)	0.72 (0.71, 0.729)	0.036 (0.028, 0.044)
phecode_108-41	Malignant neoplasm of kidney, except pelvis	0.685 (0.676, 0.694)	0.721 (0.713, 0.731)	0.037 (0.028, 0.045)
phecode_108-42	Malignant neoplasm of renal pelvis	0.743 (0.715, 0.773)	0.773 (0.747, 0.804)	0.028 (0.006, 0.055)
phecode_108-5	Malignant neoplasm of the bladder	0.75 (0.743, 0.756)	0.755 (0.748, 0.761)	0.005 (0.002, 0.008)
phecode_108-7	Malignant neoplasm of ureter	0.769 (0.749, 0.79)	0.801 (0.78, 0.824)	0.031 (0.012, 0.052)
phecode_109	Malignant neoplasm of the eye, brain and other parts of central nervous system	0.611 (0.599, 0.622)	0.734 (0.724, 0.744)	0.123 (0.112, 0.136)
phecode_109-1	Malignant neoplasm of eye	0.593 (0.563, 0.626)	0.614 (0.582, 0.646)	0.02 (-0.013, 0.054)
phecode_109-16	Malignant neoplasm of choroid	0.582 (0.541, 0.622)	0.613 (0.574, 0.651)	0.032 (-0.013, 0.072)
phecode_109-3	Malignant neoplasm of brain	0.615 (0.603, 0.627)	0.763 (0.753, 0.774)	0.148 (0.136, 0.161)
phecode_110	Malignant neoplasm of the endocrine glands	0.557 (0.541, 0.572)	0.639 (0.622, 0.655)	0.082 (0.061, 0.103)
phecode_110-1	Malignant neoplasm of the thyroid	0.57 (0.554, 0.589)	0.644 (0.626, 0.662)	0.073 (0.05, 0.095)
phecode_110-4	Malignant neoplasm of the pituitary gland and craniopharyngeal duct	0.552 (0.514, 0.593)	0.417 (0.366, 0.461)	-0.136 (-0.213, -0.067)
phecode_112	Malignant neoplasm of other and ill-defined sites	0.662 (0.66, 0.665)	0.703 (0.7, 0.705)	0.04 (0.038, 0.042)
phecode_112-1	Mesothelioma*	0.811 (0.796, 0.827)	0.872 (0.859, 0.885)	0.061 (0.05, 0.071)
phecode_114	Neuroendocrine tumors	0.593 (0.578, 0.608)	0.725 (0.71, 0.74)	0.132 (0.115, 0.151)
phecode_114-4	Carcinoid tumors	0.635 (0.598, 0.679)	0.726 (0.694, 0.767)	0.091 (0.05, 0.134)
phecode_114-6	Pheochromocytoma (including adrenal gland neoplasms)	0.585 (0.566, 0.602)	0.748 (0.729, 0.765)	0.163 (0.141, 0.186)

4 Medical history predicts future health trajectories over the human phenome

Table 17 continued from previous page

Endpoint	PheCode string	C-Index Age+Sex (IQR)	C-Index Age+Sex+MedicalHistory (IQR)	Difference in C-Index (IQR)
phecode_116	Secondary malignant neoplasm	0.642 (0.639, 0.645)	0.754 (0.751, 0.757)	0.112 (0.109, 0.115)
phecode_116-1	Secondary malignancy of lymph nodes	0.604 (0.6, 0.609)	0.701 (0.696, 0.705)	0.096 (0.091, 0.101)
phecode_116-2	Secondary malignancy of respiratory organs	0.641 (0.635, 0.647)	0.825 (0.821, 0.83)	0.185 (0.179, 0.19)
phecode_116-3	Secondary malignant neoplasm of digestive systems	0.643 (0.636, 0.65)	0.799 (0.793, 0.805)	0.156 (0.149, 0.163)
phecode_116-4	Secondary malignant neoplasm of liver	0.647 (0.643, 0.653)	0.821 (0.817, 0.825)	0.174 (0.169, 0.179)
phecode_116-5	Secondary malignancy of brain/spine	0.622 (0.613, 0.631)	0.848 (0.842, 0.853)	0.226 (0.216, 0.235)
phecode_116-6	Secondary malignancy of bone	0.671 (0.665, 0.677)	0.831 (0.827, 0.835)	0.16 (0.155, 0.166)
phecode_116-7	Secondary malignant neoplasm of skin	0.608 (0.586, 0.628)	0.815 (0.797, 0.833)	0.207 (0.184, 0.23)
phecode_120	Hemo onc - by cell of origin	0.654 (0.649, 0.659)	0.65 (0.645, 0.655)	-0.004 (-0.009, 0)
phecode_120-1	Myeloid	0.658 (0.651, 0.666)	0.685 (0.677, 0.693)	0.026 (0.018, 0.034)
phecode_120-11	Plasma cell	0.68 (0.653, 0.71)	0.72 (0.691, 0.753)	0.04 (0.01, 0.072)
phecode_120-12	Monocyte	0.752 (0.724, 0.786)	0.786 (0.758, 0.819)	0.035 (0.007, 0.063)
phecode_120-13	Erythroid	0.649 (0.626, 0.672)	0.571 (0.54, 0.603)	-0.079 (-0.115, -0.042)
phecode_120-2	Lymphoid	0.66 (0.653, 0.667)	0.647 (0.641, 0.655)	-0.013 (-0.019, -0.006)
phecode_120-21	Mature B-cell	0.665 (0.658, 0.673)	0.65 (0.643, 0.659)	-0.015 (-0.021, -0.008)
phecode_120-22	Mature T-Cell	0.635 (0.609, 0.664)	0.675 (0.649, 0.701)	0.039 (0.017, 0.061)
phecode_121	Leukemia	0.675 (0.666, 0.684)	0.68 (0.671, 0.688)	0.005 (-0.003, 0.013)
phecode_121-1	Acute leukemia	0.668 (0.651, 0.684)	0.774 (0.76, 0.79)	0.106 (0.089, 0.124)
phecode_121-11	Acute lymphoid leukemia	0.569 (0.528, 0.613)	0.708 (0.654, 0.753)	0.138 (0.077, 0.195)
phecode_121-12	Acute myeloid leukemia	0.674 (0.657, 0.692)	0.784 (0.768, 0.801)	0.11 (0.093, 0.127)
phecode_121-2	Chronic leukemia	0.684 (0.673, 0.695)	0.658 (0.648, 0.67)	-0.026 (-0.035, -0.017)
phecode_121-21	Chronic lymphoid leukemia	0.691 (0.68, 0.701)	0.667 (0.656, 0.679)	-0.023 (-0.033, -0.013)
phecode_121-22	Chronic myloid leukemia	0.637 (0.605, 0.67)	0.687 (0.651, 0.722)	0.05 (0.018, 0.08)
phecode_121-23	Chronic myelomonocytic (monocytic) leukemia	0.745 (0.711, 0.784)	0.774 (0.74, 0.813)	0.029 (0.001, 0.058)
phecode_122	Lymphoma	0.652 (0.643, 0.66)	0.654 (0.645, 0.662)	0.002 (-0.005, 0.008)
phecode_122-1	Hodgkin lymphoma	0.585 (0.556, 0.614)	0.652 (0.621, 0.684)	0.067 (0.033, 0.1)
phecode_122-2	Non-Hodgkin lymphoma	0.654 (0.646, 0.662)	0.66 (0.652, 0.67)	0.006 (-0.001, 0.014)
phecode_122-21	Follicular lymphoma	0.62 (0.6, 0.636)	0.606 (0.587, 0.625)	-0.013 (-0.029, 0.003)
phecode_122-22	Diffuse large B-cell lymphoma*	0.676 (0.664, 0.689)	0.711 (0.698, 0.726)	0.035 (0.023, 0.048)
phecode_122-24	T-cell lymphoma	0.648 (0.611, 0.685)	0.688 (0.654, 0.721)	0.04 (0.008, 0.076)
phecode_123	Multiple myeloma and malignant plasma cell neoplasms	0.678 (0.666, 0.69)	0.697 (0.683, 0.71)	0.018 (0.008, 0.029)
phecode_123-1	Multiple myeloma	0.68 (0.668, 0.691)	0.697 (0.683, 0.71)	0.017 (0.005, 0.029)
phecode_124	Myeloproliferative disorder	0.65 (0.64, 0.66)	0.681 (0.671, 0.692)	0.031 (0.02, 0.04)
phecode_124-3	Polycythemia vera	0.647 (0.626, 0.67)	0.61 (0.58, 0.638)	-0.037 (-0.071, -0.005)
phecode_124-5	Essential thrombocythemia	0.619 (0.601, 0.637)	0.673 (0.653, 0.692)	0.054 (0.034, 0.073)
phecode_124-6	Myelodysplastic syndrome	0.737 (0.721, 0.751)	0.79 (0.775, 0.806)	0.053 (0.038, 0.068)
phecode_124-7	Chronic myeloproliferative disease	0.657 (0.637, 0.678)	0.386 (0.36, 0.41)	-0.271 (-0.308, -0.236)
phecode_125	Other malignant neoplasms of lymphoid, hematopoietic and related tissue	0.582 (0.554, 0.609)	0.644 (0.615, 0.674)	0.063 (0.032, 0.093)
phecode_130	Cancer (solid tumor, excluding BCC)	0.64 (0.638, 0.642)	0.649 (0.648, 0.651)	0.009 (0.008, 0.011)
phecode_132	Sequelae of cancer	0.726 (0.721, 0.73)	0.82 (0.815, 0.825)	0.094 (0.09, 0.099)
phecode_135	Benign neoplasm of the head and neck	0.562 (0.556, 0.569)	0.661 (0.655, 0.667)	0.099 (0.092, 0.106)
phecode_135-1	Benign neoplasm of the oral cavity	0.511 (0.5, 0.525)	0.611 (0.599, 0.623)	0.1 (0.084, 0.114)
phecode_135-12	Benign neoplasm of the tongue	0.394 (0.352, 0.44)	0.574 (0.533, 0.613)	0.178 (0.125, 0.237)
phecode_135-16	Benign neoplasm of the salivary glands	0.533 (0.516, 0.552)	0.614 (0.596, 0.632)	0.081 (0.059, 0.102)
phecode_135-5	Benign neoplasm of the paranasal sinus and nasal cavity	0.578 (0.571, 0.585)	0.691 (0.684, 0.697)	0.112 (0.103, 0.121)
phecode_135-6	Benign neoplasm of vocal cord or larynx	0.577 (0.551, 0.6)	0.683 (0.659, 0.707)	0.105 (0.078, 0.137)

Supplementary Tables

Table 17 continued from previous page

Endpoint	PheCode string	C-Index Age+Sex (IQR)	C-Index Age+Sex+MedicalHistory (IQR)	Difference in C-Index (IQR)
phecode_136	Benign neoplasm of the digestive organs	0.594 (0.592, 0.596)	0.655 (0.653, 0.658)	0.061 (0.059, 0.063)
phecode_136-1	Benign neoplasm of the esophagus	0.57 (0.532, 0.607)	0.672 (0.636, 0.713)	0.102 (0.058, 0.146)
phecode_136-2	Benign neoplasm of stomach	0.625 (0.621, 0.63)	0.745 (0.741, 0.749)	0.12 (0.115, 0.124)
phecode_136-3	Benign neoplasm of the small intestine	0.646 (0.62, 0.675)	0.724 (0.7, 0.752)	0.078 (0.051, 0.109)
phecode_136-4	Benign neoplasm of colon, rectum, anus and anal canal	0.598 (0.596, 0.6)	0.648 (0.646, 0.651)	0.05 (0.048, 0.052)
phecode_136-41	Benign neoplasm of the colon	0.612 (0.61, 0.615)	0.658 (0.656, 0.66)	0.046 (0.044, 0.048)
phecode_136-42	Benign neoplasm of rectum and anus	0.593 (0.589, 0.597)	0.652 (0.648, 0.655)	0.059 (0.056, 0.063)
phecode_136-6	Benign neoplasm of the liver and intrahepatic bile ducts	0.568 (0.536, 0.6)	0.622 (0.588, 0.656)	0.054 (0.013, 0.099)
phecode_136-8	Benign neoplasm of the pancreas	0.646 (0.607, 0.683)	0.625 (0.584, 0.666)	-0.022 (-0.066, 0.023)
phecode_137	Benign neoplasm of the thoracic and respiratory organs	0.606 (0.576, 0.634)	0.637 (0.61, 0.668)	0.032 (0.002, 0.063)
phecode_137-5	Benign neoplasm of the heart, mediastinum, thymus, and pleura	0.59 (0.547, 0.633)	0.608 (0.567, 0.649)	0.018 (-0.035, 0.074)
phecode_138	Benign neoplasm of the skin	0.539 (0.536, 0.541)	0.691 (0.689, 0.693)	0.152 (0.149, 0.155)
phecode_138-1	Nevus, non-neoplastic	0.547 (0.54, 0.555)	0.673 (0.666, 0.68)	0.126 (0.116, 0.134)
phecode_138-2	Melanocytic nevi*	0.552 (0.55, 0.555)	0.706 (0.703, 0.708)	0.153 (0.151, 0.156)
phecode_139	Benign sarcoma-related cancers	0.506 (0.502, 0.509)	0.626 (0.623, 0.629)	0.12 (0.115, 0.124)
phecode_139-1	Benign neoplasms of the bone and/or cartilage	0.483 (0.46, 0.505)	0.647 (0.623, 0.667)	0.164 (0.136, 0.191)
phecode_139-3	Benign neoplasm of other connective and soft tissue	0.553 (0.538, 0.568)	0.641 (0.628, 0.654)	0.088 (0.072, 0.106)
phecode_139-4	Benign neoplasm of peripheral nerves*	0.42 (0.398, 0.441)	0.624 (0.602, 0.648)	0.204 (0.175, 0.236)
phecode_139-5	Lipoma	0.526 (0.522, 0.53)	0.628 (0.625, 0.632)	0.103 (0.098, 0.108)
phecode_139-51	Lipomatosis*	0.446 (0.409, 0.488)	0.695 (0.658, 0.732)	0.249 (0.197, 0.297)
phecode_139-52	Lipoma of intrathoracic organs	0.523 (0.507, 0.539)	0.656 (0.643, 0.671)	0.133 (0.115, 0.151)
phecode_139-53	Lipoma of other skin subcutaneous tissue	0.53 (0.523, 0.537)	0.612 (0.606, 0.619)	0.083 (0.074, 0.091)
phecode_139-54	Testicular lipoma	0.593 (0.58, 0.608)	0.613 (0.599, 0.627)	0.02 (0.007, 0.033)
phecode_139-6	Hemangioma and lymphangioma	0.533 (0.526, 0.54)	0.642 (0.636, 0.649)	0.109 (0.101, 0.118)
phecode_139-61	Hemangioma	0.533 (0.526, 0.54)	0.645 (0.638, 0.652)	0.112 (0.103, 0.121)
phecode_139-62	Lymphangioma	0.608 (0.554, 0.664)	0.642 (0.591, 0.69)	0.033 (-0.008, 0.082)
phecode_140	Benign neoplasm of the breast	0.794 (0.786, 0.8)	0.803 (0.796, 0.81)	0.01 (0.005, 0.014)
phecode_142	Lump or mass in breast or nonspecific abnormal breast exam	0.76 (0.758, 0.762)	0.81 (0.809, 0.812)	0.051 (0.049, 0.052)
phecode_142-1	Lump or mass in breast	0.747 (0.744, 0.75)	0.809 (0.807, 0.811)	0.062 (0.059, 0.064)
phecode_142-2	Abnormal mammogram	0.77 (0.75, 0.788)	0.82 (0.803, 0.837)	0.051 (0.03, 0.071)
phecode_142-21	Mammographic microcalcification	0.768 (0.752, 0.789)	0.82 (0.805, 0.835)	0.051 (0.03, 0.071)
phecode_144	Gynecological benign neoplasms	0.639 (0.635, 0.642)	0.66 (0.656, 0.664)	0.021 (0.019, 0.024)
phecode_144-1	Benign neoplasms of external female genital organs and cervix	0.65 (0.643, 0.657)	0.673 (0.665, 0.68)	0.022 (0.016, 0.028)
phecode_144-11	Benign neoplasms of the vulva	0.56 (0.523, 0.601)	0.594 (0.558, 0.634)	0.033 (-0.013, 0.08)
phecode_144-12	Benign neoplasms of the vagina	0.525 (0.495, 0.554)	0.656 (0.625, 0.686)	0.131 (0.087, 0.172)
phecode_144-13	Benign neoplasms of the cervix	0.66 (0.653, 0.667)	0.683 (0.676, 0.691)	0.023 (0.018, 0.03)
phecode_144-2	Benign neoplasms of the uterus	0.647 (0.643, 0.651)	0.669 (0.665, 0.673)	0.022 (0.019, 0.025)
phecode_144-21	Leiomyoma of uterus	0.689 (0.684, 0.694)	0.711 (0.707, 0.715)	0.022 (0.019, 0.025)
phecode_144-3	Benign neoplasms of the ovary	0.475 (0.464, 0.487)	0.556 (0.545, 0.569)	0.081 (0.064, 0.096)
phecode_146	Benign neoplasm of the genitourinary system	0.827 (0.824, 0.829)	0.915 (0.913, 0.918)	0.088 (0.086, 0.091)
phecode_146-2	Benign neoplasm of the prostate	0.638 (0.633, 0.643)	0.851 (0.848, 0.854)	0.213 (0.208, 0.218)
phecode_146-4	Benign neoplasm of the kidney	0.622 (0.592, 0.658)	0.642 (0.607, 0.676)	0.02 (-0.016, 0.053)
phecode_146-5	Benign neoplasm of the bladder	0.73 (0.701, 0.758)	0.742 (0.707, 0.774)	0.012 (-0.018, 0.038)
phecode_148	Benign neoplasm of the eye, brain and other parts of central nervous system	0.587 (0.579, 0.597)	0.628 (0.62, 0.637)	0.041 (0.032, 0.049)

4 Medical history predicts future health trajectories over the human phenome

Table 17 continued from previous page

Endpoint	PheCode string	C-Index Age+Sex (IQR)	C-Index Age+Sex+MedicalHistory (IQR)	Difference in C-Index (IQR)
phecode_148-1	Benign neoplasm of eye	0.572 (0.561, 0.583)	0.659 (0.648, 0.668)	0.087 (0.076, 0.098)
phecode_148-16	Benign neoplasm of choroid	0.577 (0.565, 0.588)	0.669 (0.66, 0.679)	0.091 (0.081, 0.103)
phecode_148-2	Benign neoplasm of meninges (Meningioma)	0.642 (0.626, 0.657)	0.656 (0.64, 0.672)	0.015 (0.002, 0.028)
phecode_148-3	Benign neoplasm of brain	0.553 (0.503, 0.598)	0.749 (0.704, 0.79)	0.195 (0.138, 0.258)
phecode_148-5	Benign neoplasm of cranial nerve	0.554 (0.532, 0.577)	0.587 (0.565, 0.61)	0.033 (0.008, 0.057)
phecode_149	Benign neoplasm of the endocrine glands	0.594 (0.582, 0.604)	0.646 (0.634, 0.657)	0.053 (0.041, 0.064)
phecode_149-1	Benign neoplasm of the thyroid	0.57 (0.547, 0.595)	0.649 (0.622, 0.675)	0.079 (0.051, 0.106)
phecode_149-3	Benign neoplasm of the parathyroid gland	0.668 (0.653, 0.686)	0.694 (0.678, 0.709)	0.025 (0.009, 0.042)
phecode_149-4	Benign neoplasm of the pituitary gland and craniopharyngeal duct	0.538 (0.517, 0.561)	0.469 (0.447, 0.495)	-0.069 (-0.101, -0.037)
phecode_153	Benign neoplasm of other or unspecified sites	0.536 (0.516, 0.555)	0.602 (0.585, 0.62)	0.067 (0.048, 0.089)
phecode_159	Genetic susceptibility to malignant neoplasm	0.659 (0.634, 0.686)	0.766 (0.745, 0.789)	0.107 (0.085, 0.128)
phecode_159-1	Genetic susceptibility to malignant neoplasm of breast	0.678 (0.65, 0.705)	0.767 (0.744, 0.79)	0.089 (0.066, 0.112)
phecode_160	Nutritional anemias	0.565 (0.562, 0.568)	0.7 (0.697, 0.702)	0.135 (0.131, 0.139)
phecode_160-1	Iron deficiency anemia	0.562 (0.558, 0.565)	0.698 (0.695, 0.701)	0.136 (0.132, 0.14)
phecode_160-2	Megaloblastic anemia	0.594 (0.586, 0.602)	0.751 (0.744, 0.758)	0.157 (0.147, 0.166)
phecode_160-4	Other deficiency anemia	0.63 (0.611, 0.65)	0.79 (0.775, 0.807)	0.161 (0.14, 0.181)
phecode_161	Hemolytic anemias	0.614 (0.593, 0.638)	0.688 (0.662, 0.716)	0.074 (0.044, 0.1)
phecode_161-1	Intrinsic (hereditary) hemolytic anemias	0.575 (0.54, 0.609)	0.701 (0.665, 0.735)	0.124 (0.076, 0.168)
phecode_161-2	Extrinsic (acquired) hemolytic anemias	0.649 (0.619, 0.676)	0.693 (0.663, 0.725)	0.045 (0.013, 0.076)
phecode_161-21	Autoimmune hemolytic anemias [AIHA]	0.671 (0.638, 0.701)	0.697 (0.663, 0.739)	0.028 (-0.005, 0.057)
phecode_162	Aplastic anemia	0.662 (0.65, 0.675)	0.772 (0.761, 0.785)	0.11 (0.097, 0.123)
phecode_162-8	Pancytopenia	0.635 (0.609, 0.664)	0.778 (0.756, 0.8)	0.143 (0.111, 0.172)
phecode_164	Anemia	0.593 (0.59, 0.595)	0.681 (0.679, 0.683)	0.088 (0.086, 0.091)
phecode_164-1	Microcytic anemia	0.562 (0.558, 0.565)	0.698 (0.695, 0.701)	0.136 (0.132, 0.14)
phecode_164-2	Macrocytic anemia	0.593 (0.585, 0.601)	0.753 (0.746, 0.76)	0.16 (0.151, 0.169)
phecode_164-3	Acute posthemorrhagic anemia	0.574 (0.539, 0.611)	0.668 (0.639, 0.7)	0.094 (0.055, 0.132)
phecode_164-6	Anemia secondary to chronic diseases and conditions	0.652 (0.641, 0.663)	0.815 (0.805, 0.825)	0.162 (0.151, 0.173)
phecode_164-62	Anemia in neoplastic disease	0.665 (0.648, 0.68)	0.83 (0.818, 0.843)	0.165 (0.147, 0.183)
phecode_165	Hemoglobinopathies	0.588 (0.571, 0.603)	0.757 (0.742, 0.773)	0.169 (0.148, 0.19)
phecode_165-2	Thalassemia	0.574 (0.551, 0.602)	0.75 (0.727, 0.774)	0.176 (0.142, 0.21)
phecode_165-25	Thalassemia minor	0.575 (0.549, 0.601)	0.782 (0.757, 0.806)	0.207 (0.171, 0.242)
phecode_165-3	Hemoglobin C trait [Sickle-cell trait]	0.626 (0.603, 0.65)	0.812 (0.794, 0.834)	0.187 (0.16, 0.217)
phecode_168	Coagulation defects, purpura and other hemorrhagic conditions	0.605 (0.599, 0.611)	0.739 (0.734, 0.745)	0.134 (0.128, 0.14)
phecode_168-1	Hypo-coagulability	0.617 (0.609, 0.624)	0.748 (0.741, 0.755)	0.131 (0.124, 0.138)
phecode_168-11	Hereditary hypo-coagulability	0.47 (0.436, 0.505)	0.638 (0.602, 0.674)	0.166 (0.128, 0.211)
phecode_168-12	Hemorrhagic disorder due to intrinsic circulating anticoagulants	0.72 (0.687, 0.756)	0.912 (0.898, 0.929)	0.193 (0.152, 0.226)
phecode_168-15	Acquired coagulation factor deficiency	0.693 (0.649, 0.746)	0.911 (0.885, 0.941)	0.22 (0.165, 0.263)
phecode_168-18	Other nonthrombocytopenic purpura	0.671 (0.657, 0.686)	0.773 (0.76, 0.785)	0.102 (0.088, 0.115)
phecode_168-19	Spontaneous ecchymoses	0.606 (0.596, 0.615)	0.758 (0.75, 0.766)	0.152 (0.141, 0.162)
phecode_168-2	Hyper-coagulability	0.49 (0.474, 0.506)	0.343 (0.325, 0.363)	-0.147 (-0.169, -0.123)
phecode_168-21	Primary hypercoagulable state [Primary thrombophilia]	0.526 (0.508, 0.546)	0.319 (0.298, 0.337)	-0.207 (-0.235, -0.18)
phecode_168-211	Activated protein C resistance*	0.578 (0.542, 0.62)	0.737 (0.705, 0.772)	0.161 (0.118, 0.201)
phecode_168-214	Antiphospholipid syndrome*	0.451 (0.412, 0.494)	0.797 (0.765, 0.831)	0.347 (0.289, 0.398)
phecode_168-3	Hereditary deficiency of other clotting factors	0.504 (0.48, 0.53)	0.651 (0.623, 0.68)	0.146 (0.107, 0.182)
phecode_168-4	Abnormal coagulation profile	0.715 (0.704, 0.727)	0.866 (0.857, 0.875)	0.15 (0.14, 0.163)
phecode_169	Platelet defects	0.646 (0.639, 0.652)	0.705 (0.699, 0.712)	0.059 (0.053, 0.066)
phecode_169-1	Thrombocytopenia	0.646 (0.639, 0.654)	0.705 (0.698, 0.712)	0.059 (0.052, 0.067)
phecode_169-11	Immune thrombocytopenic purpura [ITP]	0.613 (0.595, 0.633)	0.706 (0.687, 0.725)	0.093 (0.072, 0.113)

Supplementary Tables

Table 17 continued from previous page

Endpoint	PheCode string	C-Index Age+Sex (IQR)	C-Index Age+Sex+MedicalHistory (IQR)	Difference in C-Index (IQR)
phecode_169-14	Secondary thrombocytopenia	0.632 (0.6, 0.658)	0.721 (0.693, 0.748)	0.089 (0.053, 0.122)
phecode_170	Decreased white blood cell count	0.574 (0.569, 0.579)	0.643 (0.638, 0.649)	0.069 (0.063, 0.075)
phecode_170-1	Neutropenia	0.575 (0.57, 0.58)	0.645 (0.64, 0.65)	0.07 (0.065, 0.076)
phecode_170-13	Neutropenia due to infection	0.578 (0.549, 0.609)	0.735 (0.706, 0.759)	0.156 (0.123, 0.188)
phecode_170-19	Neutropenia NOS	0.57 (0.561, 0.579)	0.714 (0.706, 0.722)	0.144 (0.134, 0.152)
phecode_170-2	Lymphocytopenia	0.454 (0.422, 0.482)	0.727 (0.704, 0.751)	0.273 (0.235, 0.313)
phecode_171	Increased white blood cell count	0.593 (0.58, 0.605)	0.721 (0.711, 0.733)	0.129 (0.113, 0.145)
phecode_171-1	Lymphocytosis (symptomatic)	0.634 (0.618, 0.649)	0.722 (0.708, 0.735)	0.088 (0.071, 0.105)
phecode_171-7	Eosinophilia	0.54 (0.516, 0.564)	0.733 (0.711, 0.754)	0.192 (0.165, 0.222)
phecode_171-9	Elevated white blood cell count [Leukocytosis] NOS	0.517 (0.479, 0.56)	0.799 (0.77, 0.827)	0.282 (0.227, 0.327)
phecode_172	Other disorders of white blood cells	0.596 (0.585, 0.608)	0.678 (0.667, 0.689)	0.081 (0.068, 0.094)
phecode_172-2	Genetic anomalies of leukocytes	0.562 (0.528, 0.594)	0.764 (0.737, 0.791)	0.202 (0.164, 0.242)
phecode_174	Diseases of spleen	0.611 (0.601, 0.622)	0.743 (0.733, 0.754)	0.132 (0.121, 0.144)
phecode_174-1	Hyposplenism*	0.577 (0.559, 0.594)	0.424 (0.401, 0.446)	-0.154 (-0.187, -0.12)
phecode_174-2	Splenomegaly	0.636 (0.621, 0.649)	0.738 (0.724, 0.751)	0.102 (0.087, 0.118)
phecode_174-6	Cyst of spleen*	0.534 (0.479, 0.594)	0.634 (0.571, 0.699)	0.1 (0.026, 0.177)
phecode_174-7	Infarction of spleen*	0.612 (0.577, 0.647)	0.708 (0.678, 0.742)	0.096 (0.061, 0.132)
phecode_175	Polycythemias	0.635 (0.623, 0.648)	0.697 (0.683, 0.713)	0.061 (0.045, 0.079)
phecode_175-2	Secondary polycythemia	0.639 (0.624, 0.654)	0.64 (0.622, 0.657)	0 (-0.02, 0.025)
phecode_176	Other diseases of blood and blood-forming organs	0.572 (0.562, 0.582)	0.697 (0.688, 0.707)	0.125 (0.113, 0.137)
phecode_177	Abnormality of the lymph nodes	0.545 (0.541, 0.548)	0.655 (0.651, 0.659)	0.11 (0.106, 0.114)
phecode_177-1	Lymphadenitis	0.55 (0.532, 0.567)	0.691 (0.675, 0.705)	0.142 (0.122, 0.16)
phecode_177-13	Acute lymphadenitis	0.56 (0.537, 0.585)	0.685 (0.664, 0.707)	0.124 (0.095, 0.152)
phecode_177-2	Enlargement of lymph nodes [Lymphadenopathy]	0.523 (0.519, 0.527)	0.643 (0.639, 0.648)	0.12 (0.115, 0.125)
phecode_177-3	Lymphangitis	0.476 (0.448, 0.503)	0.683 (0.655, 0.713)	0.208 (0.175, 0.247)
phecode_177-4	Lymphedema	0.642 (0.634, 0.651)	0.762 (0.755, 0.77)	0.12 (0.112, 0.13)
phecode_179	Immunodeficiencies	0.52 (0.508, 0.533)	0.75 (0.738, 0.761)	0.23 (0.212, 0.246)
phecode_179-1	Hypogammaglobulinemia NOS	0.598 (0.571, 0.622)	0.75 (0.724, 0.779)	0.153 (0.119, 0.189)
phecode_179-9	Immunodeficiency NOS	0.523 (0.508, 0.54)	0.785 (0.771, 0.799)	0.262 (0.243, 0.282)
phecode_180	Other disorders involving the immune mechanism	0.606 (0.597, 0.614)	0.711 (0.704, 0.72)	0.105 (0.096, 0.117)
phecode_180-3	Paraproteinemias	0.687 (0.676, 0.696)	0.726 (0.716, 0.736)	0.04 (0.031, 0.049)
phecode_180-31	Monoclonal gammopathy	0.684 (0.673, 0.695)	0.729 (0.717, 0.74)	0.045 (0.035, 0.055)
phecode_180-33	Macroglobulinemia [Waldenstrom macroglobulinemia]	0.722 (0.69, 0.755)	0.75 (0.716, 0.783)	0.028 (0.001, 0.056)
phecode_181	Autoimmune disease	0.568 (0.564, 0.573)	0.667 (0.662, 0.67)	0.098 (0.094, 0.103)
phecode_200	Disorders of thyroid gland	0.638 (0.636, 0.641)	0.679 (0.677, 0.682)	0.041 (0.039, 0.044)
phecode_200-1	Hypothyroidism	0.644 (0.641, 0.647)	0.706 (0.703, 0.709)	0.062 (0.059, 0.065)
phecode_200-12	Hypothyroidism due to drugs or iatrogenic causes	0.55 (0.512, 0.589)	0.8 (0.765, 0.836)	0.251 (0.205, 0.297)
phecode_200-13	Postprocedural hypothyroidism	0.682 (0.672, 0.69)	0.844 (0.836, 0.853)	0.162 (0.152, 0.174)
phecode_200-14	Hypothyroidism, not specified as secondary	0.648 (0.645, 0.651)	0.717 (0.715, 0.721)	0.07 (0.066, 0.072)
phecode_200-2	Goiter	0.64 (0.634, 0.645)	0.687 (0.682, 0.693)	0.048 (0.042, 0.053)
phecode_200-21	Diffuse goiter	0.659 (0.648, 0.67)	0.709 (0.697, 0.721)	0.049 (0.039, 0.061)
phecode_200-22	Uninodular goiter [single thyroid nodule]	0.622 (0.613, 0.631)	0.689 (0.68, 0.697)	0.067 (0.058, 0.076)
phecode_200-23	Multinodular goiter	0.68 (0.672, 0.689)	0.723 (0.714, 0.732)	0.043 (0.034, 0.05)
phecode_200-3	Thyrotoxicosis [hyperthyroidism]	0.628 (0.621, 0.634)	0.698 (0.691, 0.705)	0.07 (0.064, 0.076)
phecode_200-31	Graves' disease [Toxic diffuse goiter]	0.653 (0.639, 0.665)	0.695 (0.682, 0.709)	0.043 (0.029, 0.057)
phecode_200-4	Thyroiditis	0.65 (0.635, 0.664)	0.724 (0.71, 0.739)	0.076 (0.059, 0.088)

4 Medical history predicts future health trajectories over the human phenome

Table 17 continued from previous page

Endpoint	PheCode string	C-Index Age+Sex (IQR)	C-Index Age+Sex+MedicalHistory (IQR)	Difference in C-Index (IQR)
phecode_200-41	Hashimoto thyroiditis [Chronic lymphocytic thyroiditis]	0.656 (0.638, 0.674)	0.744 (0.726, 0.761)	0.088 (0.069, 0.106)
phecode_200-7	Iodine-deficiency related thyroid disorders*	0.593 (0.586, 0.601)	0.674 (0.668, 0.68)	0.08 (0.074, 0.086)
phecode_200-9	Abnormal thyroid function studies	0.602 (0.593, 0.612)	0.739 (0.732, 0.747)	0.137 (0.126, 0.147)
phecode_202	Diabetes mellitus	0.604 (0.601, 0.606)	0.721 (0.719, 0.723)	0.117 (0.115, 0.12)
phecode_202-1	Type 1 diabetes	0.614 (0.605, 0.622)	0.873 (0.868, 0.879)	0.259 (0.25, 0.269)
phecode_202-2	Type 2 diabetes	0.603 (0.601, 0.606)	0.724 (0.722, 0.726)	0.121 (0.118, 0.123)
phecode_202-3	Secondary diabetes	0.592 (0.561, 0.625)	0.76 (0.731, 0.79)	0.169 (0.127, 0.211)
phecode_202-32	Drug or chemical induced diabetes mellitus*	0.587 (0.549, 0.623)	0.766 (0.734, 0.798)	0.18 (0.134, 0.224)
phecode_202-4	Other specified diabetes*	0.608 (0.606, 0.61)	0.863 (0.861, 0.864)	0.255 (0.253, 0.258)
phecode_203	Metabolic syndrome [Dysmetabolic syndrome X]	0.528 (0.494, 0.562)	0.802 (0.78, 0.824)	0.274 (0.237, 0.313)
phecode_204	Elevated blood glucose level	0.586 (0.583, 0.588)	0.728 (0.725, 0.73)	0.142 (0.139, 0.145)
phecode_204-1	Impaired fasting glucose	0.608 (0.602, 0.613)	0.753 (0.749, 0.757)	0.146 (0.141, 0.151)
phecode_204-2	Impaired glucose tolerance (oral)	0.592 (0.588, 0.596)	0.747 (0.743, 0.749)	0.155 (0.151, 0.159)
phecode_204-4	Prediabetes*	0.596 (0.579, 0.613)	0.865 (0.856, 0.873)	0.269 (0.252, 0.285)
phecode_205	Hypoglycemia	0.629 (0.622, 0.637)	0.869 (0.863, 0.875)	0.24 (0.231, 0.248)
phecode_205-3	Drug-induced hypoglycemia*	0.654 (0.614, 0.696)	0.925 (0.904, 0.949)	0.273 (0.225, 0.315)
phecode_206	Disorders of pancreatic internal secretion (excl. DM)	0.567 (0.533, 0.599)	0.816 (0.795, 0.842)	0.25 (0.216, 0.281)
phecode_208	Disorders of parathyroid gland	0.67 (0.662, 0.678)	0.74 (0.732, 0.747)	0.07 (0.061, 0.078)
phecode_208-1	Hypoparathyroidism	0.57 (0.535, 0.607)	0.763 (0.731, 0.795)	0.193 (0.154, 0.232)
phecode_208-2	Hyperparathyroidism	0.675 (0.666, 0.683)	0.745 (0.736, 0.753)	0.07 (0.062, 0.079)
phecode_208-21	Primary hyperparathyroidism	0.708 (0.698, 0.718)	0.734 (0.724, 0.744)	0.026 (0.017, 0.034)
phecode_208-22	Secondary hyperparathyroidism	0.6 (0.58, 0.618)	0.833 (0.818, 0.849)	0.233 (0.211, 0.256)
phecode_209	Disorders of the pituitary gland and its hypothalamic control	0.607 (0.596, 0.62)	0.699 (0.687, 0.712)	0.092 (0.079, 0.107)
phecode_209-1	Pituitary hyperfunction	0.634 (0.616, 0.651)	0.715 (0.699, 0.732)	0.082 (0.064, 0.101)
phecode_209-12	Syndrome of inappropriate secretion of antidiuretic hormone	0.704 (0.688, 0.722)	0.771 (0.754, 0.788)	0.067 (0.047, 0.084)
phecode_209-13	Hyperprolactinemia*	0.57 (0.523, 0.618)	0.744 (0.701, 0.789)	0.173 (0.121, 0.225)
phecode_209-2	Hypopituitarism	0.635 (0.62, 0.651)	0.361 (0.342, 0.382)	-0.274 (-0.301, -0.246)
phecode_209-21	Diabetes insipidus	0.48 (0.441, 0.522)	0.63 (0.576, 0.686)	0.149 (0.083, 0.228)
phecode_209-22	Hypopituitarism NOS	0.669 (0.652, 0.686)	0.354 (0.33, 0.376)	-0.314 (-0.345, -0.282)
phecode_209-23	Iatrogenic hypopituitarism*	0.501 (0.461, 0.547)	0.823 (0.782, 0.866)	0.321 (0.263, 0.377)
phecode_210	Cushing's syndrome	0.573 (0.534, 0.614)	0.732 (0.691, 0.767)	0.157 (0.114, 0.2)
phecode_211	Disorders of adrenal glands	0.566 (0.553, 0.579)	0.744 (0.733, 0.757)	0.178 (0.162, 0.195)
phecode_211-1	Hyperaldosteronism	0.593 (0.543, 0.642)	0.803 (0.763, 0.849)	0.211 (0.162, 0.267)
phecode_211-2	Adrenocortical insufficiency	0.548 (0.528, 0.565)	0.763 (0.745, 0.782)	0.215 (0.194, 0.237)
phecode_211-21	Primary adrenocortical insufficiency	0.507 (0.476, 0.536)	0.645 (0.611, 0.678)	0.139 (0.095, 0.186)
phecode_211-22	Drug-induced adrenocortical insufficiency*	0.563 (0.51, 0.617)	0.847 (0.806, 0.891)	0.286 (0.223, 0.341)
phecode_214	Ovarian dysfunction	0.71 (0.692, 0.73)	0.705 (0.689, 0.723)	-0.004 (-0.022, 0.015)
phecode_214-1	Primary ovarian failure	0.712 (0.692, 0.731)	0.706 (0.69, 0.724)	-0.005 (-0.024, 0.013)
phecode_214-11	Premature menopause	0.706 (0.687, 0.728)	0.704 (0.684, 0.724)	-0.003 (-0.021, 0.016)
phecode_215	Testicular dysfunction	0.476 (0.459, 0.494)	0.737 (0.722, 0.754)	0.262 (0.238, 0.286)
phecode_215-1	Testicular hypofunction	0.475 (0.456, 0.494)	0.748 (0.731, 0.766)	0.273 (0.249, 0.297)
phecode_229	Other endocrine disorders	0.518 (0.492, 0.542)	0.809 (0.786, 0.831)	0.291 (0.26, 0.321)
phecode_230	Malnutrition and underweight	0.603 (0.6, 0.606)	0.681 (0.678, 0.684)	0.078 (0.074, 0.081)
phecode_230-1	Protein-calorie malnutrition	0.614 (0.595, 0.636)	0.789 (0.772, 0.804)	0.174 (0.153, 0.195)
phecode_230-2	Abnormal loss of weight and underweight	0.605 (0.602, 0.609)	0.677 (0.673, 0.68)	0.071 (0.067, 0.075)
phecode_230-21	Abnormal weight loss	0.611 (0.607, 0.614)	0.682 (0.679, 0.685)	0.072 (0.068, 0.075)
phecode_230-22	Underweight	0.595 (0.581, 0.608)	0.739 (0.728, 0.751)	0.145 (0.13, 0.158)
phecode_230-3	Anorexia	0.609 (0.603, 0.616)	0.723 (0.717, 0.73)	0.115 (0.108, 0.121)
phecode_230-4	Cachexia	0.669 (0.643, 0.699)	0.81 (0.786, 0.834)	0.141 (0.105, 0.172)

Supplementary Tables

Table 17 continued from previous page

Endpoint	PheCode string	C-Index Age+Sex (IQR)	C-Index Age+Sex+MedicalHistory (IQR)	Difference in C-Index (IQR)
phecode_230-5	Early satiety	0.593 (0.539, 0.643)	0.829 (0.795, 0.861)	0.234 (0.184, 0.286)
phecode_232	Vitamin deficiencies	0.55 (0.548, 0.552)	0.735 (0.733, 0.737)	0.185 (0.182, 0.188)
phecode_232-1	Vitamin A deficiency	0.637 (0.599, 0.676)	0.828 (0.804, 0.855)	0.192 (0.151, 0.23)
phecode_232-2	Vitamin B group deficiency	0.572 (0.567, 0.576)	0.745 (0.741, 0.748)	0.173 (0.168, 0.178)
phecode_232-27	Vitamin B12 deficiency	0.583 (0.574, 0.592)	0.761 (0.753, 0.769)	0.178 (0.167, 0.189)
phecode_232-29	Folate deficiency [Vitamin B9]	0.646 (0.632, 0.66)	0.757 (0.743, 0.77)	0.111 (0.096, 0.124)
phecode_232-4	Vitamin D deficiency	0.562 (0.559, 0.565)	0.75 (0.747, 0.752)	0.188 (0.184, 0.191)
phecode_234	Other nutritional deficiencies	0.524 (0.516, 0.53)	0.718 (0.713, 0.723)	0.195 (0.186, 0.203)
phecode_236	Overweight and obesity	0.515 (0.513, 0.518)	0.711 (0.71, 0.713)	0.196 (0.193, 0.199)
phecode_236-1	Obesity	0.515 (0.513, 0.518)	0.711 (0.709, 0.713)	0.196 (0.193, 0.199)
phecode_236-11	Morbid obesity	0.578 (0.568, 0.587)	0.87 (0.864, 0.876)	0.291 (0.281, 0.303)
phecode_236-2	Localized adiposity	0.575 (0.545, 0.604)	0.879 (0.854, 0.906)	0.304 (0.27, 0.337)
phecode_237	Abnormal weight gain	0.553 (0.545, 0.561)	0.764 (0.759, 0.77)	0.211 (0.203, 0.22)
phecode_239	Hyperlipidemia	0.611 (0.609, 0.612)	0.677 (0.675, 0.678)	0.066 (0.064, 0.067)
phecode_239-1	Hypercholesterolemia	0.615 (0.613, 0.617)	0.681 (0.679, 0.683)	0.066 (0.064, 0.068)
phecode_239-11	Pure hypercholesterolemia	0.617 (0.615, 0.618)	0.682 (0.68, 0.683)	0.066 (0.064, 0.067)
phecode_239-12	Familial hypercholesterolemia*	0.594 (0.583, 0.603)	0.768 (0.76, 0.776)	0.174 (0.163, 0.185)
phecode_239-2	Hyperglyceridemia	0.556 (0.546, 0.566)	0.713 (0.705, 0.721)	0.157 (0.146, 0.168)
phecode_239-21	Pure hyperglyceridemia	0.618 (0.601, 0.634)	0.734 (0.719, 0.748)	0.116 (0.093, 0.136)
phecode_239-3	Mixed hyperlipidemia	0.539 (0.527, 0.551)	0.72 (0.71, 0.729)	0.181 (0.166, 0.195)
phecode_240	Disorders of amino-acid transport and metabolism	0.51 (0.476, 0.545)	0.728 (0.687, 0.768)	0.218 (0.169, 0.26)
phecode_241	Disorders of carbohydrate metabolism	0.571 (0.552, 0.591)	0.693 (0.677, 0.711)	0.123 (0.099, 0.142)
phecode_242	Lipid storage disorders	0.526 (0.487, 0.568)	0.643 (0.608, 0.678)	0.116 (0.07, 0.166)
phecode_244	Disorders of lipoprotein metabolism and other lipidemias	0.585 (0.576, 0.595)	0.728 (0.721, 0.736)	0.144 (0.132, 0.155)
phecode_244-4	Lipodystrophy, not elsewhere classified	0.593 (0.566, 0.624)	0.931 (0.914, 0.949)	0.339 (0.304, 0.365)
phecode_247	Disorders of mineral metabolism and mineral deficiencies	0.565 (0.562, 0.567)	0.681 (0.678, 0.684)	0.116 (0.113, 0.119)
phecode_247-3	Disorder of phosphorus metabolism	0.635 (0.623, 0.649)	0.7 (0.687, 0.714)	0.065 (0.05, 0.08)
phecode_247-4	Disorders of magnesium metabolism	0.648 (0.638, 0.657)	0.796 (0.786, 0.804)	0.148 (0.136, 0.159)
phecode_247-42	Hypomagnesemia*	0.663 (0.639, 0.686)	0.834 (0.815, 0.858)	0.171 (0.147, 0.2)
phecode_247-5	Disorders of calcium metabolism	0.608 (0.601, 0.614)	0.695 (0.689, 0.701)	0.087 (0.081, 0.094)
phecode_247-51	Hypocalcemia	0.521 (0.508, 0.535)	0.751 (0.741, 0.761)	0.23 (0.214, 0.246)
phecode_247-52	Hypercalcemia	0.663 (0.652, 0.674)	0.76 (0.751, 0.77)	0.097 (0.087, 0.107)
phecode_247-7	Disorders of iron metabolism	0.554 (0.551, 0.557)	0.689 (0.687, 0.692)	0.136 (0.132, 0.14)
phecode_247-71	Hemochromatosis	0.564 (0.548, 0.583)	0.691 (0.676, 0.705)	0.127 (0.106, 0.146)
phecode_247-711	Hereditary hemochromatosis	0.525 (0.492, 0.559)	0.819 (0.801, 0.837)	0.294 (0.262, 0.333)
phecode_247-72	Iron deficiency	0.554 (0.551, 0.558)	0.695 (0.692, 0.698)	0.141 (0.137, 0.144)
phecode_248	Disorders of plasma-protein metabolism, NEC	0.619 (0.605, 0.634)	0.706 (0.692, 0.721)	0.087 (0.07, 0.103)
phecode_248-1	Alpha-1-antitrypsin deficiency	0.484 (0.44, 0.53)	0.638 (0.602, 0.677)	0.155 (0.092, 0.206)
phecode_249	Amyloidosis	0.707 (0.688, 0.726)	0.712 (0.692, 0.732)	0.005 (-0.011, 0.02)
phecode_249-1	Cerebral amyloid angiopathy*	0.787 (0.761, 0.812)	0.748 (0.72, 0.78)	-0.038 (-0.059, -0.013)
phecode_251	Disorders of bilirubin excretion	0.616 (0.606, 0.628)	0.691 (0.679, 0.701)	0.074 (0.064, 0.084)
phecode_251-1	Gilbert syndrome*	0.618 (0.608, 0.629)	0.698 (0.687, 0.708)	0.08 (0.07, 0.09)
phecode_252	Other and unspecified disorders of metabolism	0.497 (0.469, 0.52)	0.75 (0.731, 0.769)	0.254 (0.221, 0.285)
phecode_256	Disorders of fluid, electrolyte and acid-base balance	0.634 (0.632, 0.637)	0.708 (0.705, 0.71)	0.073 (0.071, 0.076)
phecode_256-1	Hyperosmolality and/or hyponatremia	0.604 (0.593, 0.617)	0.695 (0.685, 0.708)	0.091 (0.08, 0.102)
phecode_256-2	Hyposmolality and/or hyponatremia	0.607 (0.603, 0.611)	0.666 (0.662, 0.67)	0.059 (0.055, 0.062)
phecode_256-3	Mixed disorder of acid-base balance	0.661 (0.654, 0.67)	0.817 (0.811, 0.823)	0.155 (0.148, 0.162)
phecode_256-31	Acidosis	0.663 (0.655, 0.672)	0.825 (0.819, 0.831)	0.162 (0.154, 0.17)

4 Medical history predicts future health trajectories over the human phenome

Table 17 continued from previous page

Endpoint	PheCode string	C-Index Age+Sex (IQR)	C-Index Age+Sex+MedicalHistory (IQR)	Difference in C-Index (IQR)
phecode_256-32	Alkalosis	0.652 (0.632, 0.671)	0.76 (0.739, 0.777)	0.109 (0.083, 0.132)
phecode_256-4	Hyperkalemia [Hyperpotassemia]	0.667 (0.661, 0.674)	0.787 (0.781, 0.793)	0.119 (0.113, 0.126)
phecode_256-5	Hypokalemia [Hypopotassemia]	0.647 (0.641, 0.653)	0.735 (0.73, 0.74)	0.089 (0.083, 0.095)
phecode_256-6	Fluid overload	0.696 (0.688, 0.706)	0.831 (0.824, 0.838)	0.135 (0.125, 0.144)
phecode_256-7	Volume depletion	0.667 (0.662, 0.671)	0.77 (0.766, 0.774)	0.104 (0.099, 0.108)
phecode_256-71	Dehydration	0.618 (0.6, 0.636)	0.781 (0.766, 0.794)	0.163 (0.144, 0.18)
phecode_257	Polydipsia	0.558 (0.546, 0.569)	0.729 (0.719, 0.738)	0.171 (0.158, 0.184)
phecode_280	Substance related disorders	0.583 (0.58, 0.586)	0.69 (0.687, 0.693)	0.107 (0.103, 0.11)
phecode_280-1	Alcohol use disorders	0.588 (0.585, 0.591)	0.69 (0.686, 0.693)	0.102 (0.098, 0.105)
phecode_280-11	Alcohol abuse	0.635 (0.631, 0.639)	0.725 (0.721, 0.729)	0.089 (0.085, 0.094)
phecode_280-12	Alcohol dependence	0.566 (0.559, 0.572)	0.75 (0.745, 0.755)	0.185 (0.176, 0.192)
phecode_280-13	Alcoholic liver disease	0.647 (0.638, 0.658)	0.827 (0.819, 0.835)	0.179 (0.169, 0.19)
phecode_280-14	Alcoholic pancreatitis	0.704 (0.676, 0.731)	0.838 (0.815, 0.863)	0.134 (0.108, 0.159)
phecode_280-2	Opioid related disorders	0.621 (0.591, 0.656)	0.827 (0.804, 0.853)	0.206 (0.166, 0.242)
phecode_280-22	Opioid dependence	0.654 (0.617, 0.691)	0.862 (0.839, 0.89)	0.21 (0.169, 0.251)
phecode_280-3	Cannabis related disorders	0.754 (0.736, 0.776)	0.9 (0.887, 0.915)	0.146 (0.125, 0.166)
phecode_280-31	Cannabis abuse	0.75 (0.726, 0.772)	0.89 (0.874, 0.909)	0.142 (0.119, 0.161)
phecode_280-4	Sedative, hypnotic, or anxiolytic related disorders	0.53 (0.499, 0.558)	0.89 (0.875, 0.906)	0.36 (0.328, 0.395)
phecode_280-42	Sedative, hypnotic or anxiolytic-related dependence	0.531 (0.503, 0.566)	0.891 (0.874, 0.908)	0.36 (0.324, 0.393)
phecode_280-8	Other psychoactive substance related disorders	0.602 (0.584, 0.619)	0.791 (0.778, 0.806)	0.188 (0.169, 0.208)
phecode_280-81	Other psychoactive substance abuse	0.665 (0.632, 0.699)	0.8 (0.773, 0.829)	0.136 (0.097, 0.172)
phecode_280-82	Other psychoactive substance dependence	0.594 (0.576, 0.613)	0.793 (0.778, 0.807)	0.199 (0.178, 0.221)
phecode_281	Substance abuse, dependence, and withdrawal	0.565 (0.562, 0.567)	0.686 (0.684, 0.689)	0.122 (0.119, 0.125)
phecode_281-1	Substance abuse	0.636 (0.633, 0.64)	0.726 (0.721, 0.729)	0.089 (0.085, 0.093)
phecode_281-2	Substance dependence	0.554 (0.551, 0.557)	0.695 (0.693, 0.698)	0.141 (0.138, 0.145)
phecode_281-21	Substance withdrawal	0.607 (0.58, 0.631)	0.85 (0.833, 0.868)	0.243 (0.213, 0.274)
phecode_282-1	Current tobacco use and nicotine dependence	0.563 (0.56, 0.566)	0.677 (0.674, 0.68)	0.114 (0.11, 0.117)
phecode_283	Other behavioral problems	0.52 (0.518, 0.522)	0.692 (0.69, 0.693)	0.172 (0.17, 0.174)
phecode_283-3	High risk sexual behavior	0.769 (0.748, 0.794)	0.818 (0.797, 0.837)	0.049 (0.03, 0.066)
phecode_283-4	Patient's noncompliance with medical treatment and regimen	0.513 (0.509, 0.517)	0.743 (0.74, 0.746)	0.23 (0.226, 0.235)
phecode_283-8	Other problems related to lifestyle	0.524 (0.522, 0.526)	0.687 (0.686, 0.689)	0.164 (0.161, 0.166)
phecode_284	Suicide ideation and attempt or self harm	0.608 (0.602, 0.615)	0.831 (0.826, 0.837)	0.223 (0.216, 0.231)
phecode_284-1	Suicidal ideations	0.621 (0.614, 0.629)	0.849 (0.844, 0.854)	0.228 (0.219, 0.236)
phecode_284-2	Suicide and self-inflicted harm	0.598 (0.588, 0.608)	0.856 (0.848, 0.864)	0.259 (0.245, 0.271)
phecode_284-29	Intentional self-harm*	0.596 (0.584, 0.607)	0.859 (0.851, 0.867)	0.263 (0.249, 0.275)
phecode_286	Mood [affective] disorders	0.576 (0.573, 0.578)	0.72 (0.718, 0.722)	0.144 (0.142, 0.147)
phecode_286-1	Bipolar disorder	0.529 (0.514, 0.543)	0.841 (0.831, 0.853)	0.313 (0.294, 0.331)
phecode_286-2	Major depressive disorder	0.576 (0.574, 0.579)	0.722 (0.721, 0.725)	0.146 (0.144, 0.149)
phecode_286-21	Major depressive disorder, recurrent	0.583 (0.572, 0.591)	0.817 (0.81, 0.826)	0.235 (0.224, 0.248)
phecode_286-3	Premenstrual dysphoric disorder	0.902 (0.896, 0.909)	0.922 (0.916, 0.928)	0.019 (0.015, 0.024)
phecode_286-4	Dysthymic disorder	0.588 (0.569, 0.607)	0.807 (0.792, 0.822)	0.218 (0.194, 0.242)
phecode_287	Psychotic disorder	0.481 (0.469, 0.494)	0.754 (0.743, 0.765)	0.272 (0.255, 0.287)
phecode_287-1	Schizophrenia	0.561 (0.537, 0.586)	0.877 (0.864, 0.892)	0.317 (0.293, 0.341)
phecode_287-2	Schizoaffective disorder	0.569 (0.533, 0.604)	0.915 (0.888, 0.942)	0.345 (0.302, 0.392)
phecode_287-4	Delusional disorders	0.477 (0.459, 0.495)	0.758 (0.741, 0.774)	0.28 (0.256, 0.304)
phecode_287-5	Drug-induced psychotic disorder	0.555 (0.526, 0.589)	0.763 (0.734, 0.79)	0.207 (0.174, 0.243)
phecode_288	Anxiety disorders	0.588 (0.585, 0.59)	0.714 (0.712, 0.716)	0.127 (0.124, 0.129)
phecode_288-2	Panic disorder [episodic paroxysmal anxiety]	0.582 (0.575, 0.589)	0.749 (0.743, 0.756)	0.168 (0.16, 0.175)

Supplementary Tables

Table 17 continued from previous page

Endpoint	PheCode string	C-Index Age+Sex (IQR)	C-Index Age+Sex+MedicalHistory (IQR)	Difference in C-Index (IQR)
phecode_288-3	Generalized anxiety disorder	0.591 (0.586, 0.596)	0.764 (0.76, 0.768)	0.172 (0.168, 0.178)
phecode_288-4	Phobic disorders	0.598 (0.591, 0.604)	0.73 (0.724, 0.734)	0.132 (0.126, 0.138)
phecode_288-41	Agoraphobia	0.638 (0.601, 0.672)	0.815 (0.783, 0.845)	0.177 (0.136, 0.218)
phecode_289	Obsessive-compulsive disorder	0.591 (0.572, 0.611)	0.768 (0.75, 0.785)	0.176 (0.155, 0.2)
phecode_290	Reaction to severe stress, and adjustment disorders	0.656 (0.653, 0.659)	0.758 (0.756, 0.761)	0.102 (0.099, 0.105)
phecode_290-1	Posttraumatic stress disorder	0.661 (0.646, 0.676)	0.787 (0.773, 0.8)	0.125 (0.111, 0.14)
phecode_291	Dissociative, conversion and factitious disorders	0.522 (0.503, 0.542)	0.779 (0.762, 0.796)	0.257 (0.231, 0.281)
phecode_292	Somatoform disorders	0.569 (0.554, 0.583)	0.716 (0.702, 0.728)	0.147 (0.131, 0.164)
phecode_293	Eating disorders	0.632 (0.611, 0.656)	0.767 (0.748, 0.784)	0.134 (0.113, 0.158)
phecode_293-1	Anorexia nervosa	0.573 (0.529, 0.615)	0.71 (0.668, 0.757)	0.137 (0.083, 0.193)
phecode_293-4	Polyphagia	0.628 (0.598, 0.655)	0.813 (0.793, 0.837)	0.184 (0.155, 0.221)
phecode_294	Sexual dysfunction and disorders	0.784 (0.782, 0.785)	0.857 (0.855, 0.858)	0.073 (0.072, 0.075)
phecode_296	Specific personality disorders	0.659 (0.639, 0.681)	0.884 (0.871, 0.899)	0.226 (0.205, 0.248)
phecode_296-4	Borderline personality disorder	0.75 (0.727, 0.774)	0.929 (0.914, 0.944)	0.179 (0.157, 0.202)
phecode_299	Mental disorder, not otherwise specified	0.541 (0.53, 0.551)	0.761 (0.752, 0.769)	0.22 (0.207, 0.232)
phecode_308	Signs and symptoms involving emotional state	0.573 (0.57, 0.576)	0.75 (0.747, 0.752)	0.176 (0.174, 0.179)
phecode_308-1	Irritability	0.631 (0.618, 0.644)	0.779 (0.769, 0.789)	0.148 (0.134, 0.162)
phecode_308-3	Emotional lability	0.697 (0.684, 0.711)	0.821 (0.811, 0.829)	0.124 (0.111, 0.135)
phecode_308-4	Demoralization and apathy	0.541 (0.523, 0.557)	0.853 (0.844, 0.862)	0.312 (0.294, 0.329)
phecode_308-5	Nervousness	0.561 (0.547, 0.574)	0.764 (0.754, 0.774)	0.203 (0.189, 0.218)
phecode_308-6	Excessive crying of child, adolescent, or adult	0.699 (0.684, 0.713)	0.819 (0.808, 0.829)	0.119 (0.108, 0.132)
phecode_308-7	Restlessness and agitation*	0.592 (0.58, 0.604)	0.729 (0.719, 0.74)	0.137 (0.124, 0.15)
phecode_320	Meningitis	0.531 (0.511, 0.551)	0.658 (0.634, 0.679)	0.126 (0.098, 0.156)
phecode_320-1	Infective meningitis	0.512 (0.489, 0.54)	0.618 (0.594, 0.645)	0.106 (0.074, 0.14)
phecode_320-11	Bacterial meningitis	0.564 (0.531, 0.592)	0.63 (0.597, 0.664)	0.069 (0.032, 0.108)
phecode_320-12	Viral meningitis	0.54 (0.493, 0.582)	0.699 (0.663, 0.737)	0.161 (0.115, 0.211)
phecode_320-3	Meningitis NOS	0.52 (0.485, 0.557)	0.712 (0.676, 0.747)	0.192 (0.143, 0.237)
phecode_321	Encephalitis, myelitis and encephalomyelitis	0.524 (0.502, 0.544)	0.622 (0.602, 0.639)	0.098 (0.073, 0.124)
phecode_321-1	Encephalitis	0.55 (0.515, 0.581)	0.595 (0.566, 0.625)	0.045 (0.006, 0.085)
phecode_321-12	Viral encephalitis	0.562 (0.529, 0.598)	0.605 (0.576, 0.636)	0.044 (0.001, 0.082)
phecode_321-2	Myelitis	0.577 (0.537, 0.618)	0.641 (0.597, 0.687)	0.062 (0.012, 0.118)
phecode_321-21	Acute (transverse) myelitis	0.565 (0.519, 0.607)	0.636 (0.584, 0.685)	0.071 (0.008, 0.132)
phecode_322	Other CNS infection	0.627 (0.604, 0.647)	0.75 (0.729, 0.77)	0.124 (0.096, 0.143)
phecode_322-4	Intracranial and intraspinal abscess	0.638 (0.608, 0.673)	0.693 (0.658, 0.726)	0.053 (0.021, 0.088)
phecode_323	Systemic atrophies primarily affecting the central nervous system	0.633 (0.62, 0.647)	0.722 (0.709, 0.736)	0.089 (0.075, 0.103)
phecode_323-1	Hereditary ataxia	0.589 (0.567, 0.614)	0.699 (0.673, 0.724)	0.11 (0.082, 0.139)
phecode_323-3	Motor neuron disease	0.662 (0.645, 0.677)	0.766 (0.751, 0.781)	0.105 (0.088, 0.123)
phecode_323-31	Amyotrophic lateral sclerosis [ALS]	0.639 (0.606, 0.672)	0.816 (0.796, 0.839)	0.177 (0.144, 0.207)
phecode_324	Extrapyramidal and movement disorders	0.595 (0.592, 0.599)	0.692 (0.689, 0.695)	0.097 (0.093, 0.1)
phecode_324-1	Parkinsonism	0.736 (0.73, 0.743)	0.737 (0.731, 0.744)	0.001 (-0.003, 0.005)
phecode_324-11	Parkinson's disease (Primary)	0.738 (0.732, 0.745)	0.737 (0.731, 0.743)	-0.001 (-0.005, 0.002)
phecode_324-12	Secondary parkinsonism	0.706 (0.678, 0.734)	0.815 (0.792, 0.84)	0.11 (0.077, 0.137)
phecode_324-2	Degenerative diseases of the basal ganglia (excluding parkinsons)	0.718 (0.697, 0.737)	0.726 (0.705, 0.748)	0.009 (-0.01, 0.03)
phecode_324-21	Progressive supranuclear ophthalmoplegia [Steele-Richardson-Olszewski]*	0.773 (0.746, 0.8)	0.77 (0.746, 0.797)	-0.003 (-0.029, 0.023)
phecode_324-3	Dystonia	0.545 (0.539, 0.55)	0.708 (0.703, 0.713)	0.163 (0.157, 0.17)
phecode_324-34	Torticollis	0.528 (0.521, 0.535)	0.717 (0.711, 0.722)	0.189 (0.181, 0.196)
phecode_324-36	Blepharospasm	0.555 (0.534, 0.576)	0.709 (0.685, 0.726)	0.152 (0.128, 0.178)
phecode_324-4	Tremor	0.628 (0.623, 0.633)	0.732 (0.728, 0.737)	0.104 (0.099, 0.108)
phecode_324-41	Essential tremor*	0.655 (0.649, 0.664)	0.736 (0.728, 0.743)	0.08 (0.072, 0.088)

4 Medical history predicts future health trajectories over the human phenome

Table 17 continued from previous page

Endpoint	PheCode string	C-Index Age+Sex (IQR)	C-Index Age+Sex+MedicalHistory (IQR)	Difference in C-Index (IQR)
phecode_324-5	Myoclonus	0.562 (0.539, 0.588)	0.707 (0.684, 0.731)	0.145 (0.114, 0.175)
phecode_324-8	Restless legs syndrome	0.605 (0.597, 0.612)	0.763 (0.758, 0.768)	0.158 (0.151, 0.165)
phecode_325	Symptoms and signs related to movement disorders	0.647 (0.643, 0.65)	0.737 (0.734, 0.74)	0.09 (0.087, 0.094)
phecode_325-1	Abnormal involuntary movements	0.506 (0.492, 0.52)	0.682 (0.669, 0.694)	0.176 (0.157, 0.192)
phecode_325-12	Fasciculation*	0.514 (0.497, 0.533)	0.676 (0.66, 0.691)	0.161 (0.14, 0.183)
phecode_325-2	Abnormality of gait and mobility	0.66 (0.656, 0.663)	0.751 (0.748, 0.755)	0.091 (0.088, 0.095)
phecode_325-21	Ataxic gait*	0.63 (0.601, 0.66)	0.745 (0.72, 0.772)	0.113 (0.08, 0.147)
phecode_325-23	Unsteadiness on feet*	0.665 (0.655, 0.673)	0.797 (0.791, 0.804)	0.133 (0.125, 0.141)
phecode_325-3	Lack of coordination	0.622 (0.614, 0.629)	0.74 (0.733, 0.747)	0.118 (0.111, 0.126)
phecode_326	Demyelinating diseases of the central nervous system	0.612 (0.597, 0.628)	0.683 (0.668, 0.7)	0.071 (0.055, 0.089)
phecode_326-1	Multiple sclerosis	0.636 (0.619, 0.652)	0.714 (0.7, 0.73)	0.079 (0.061, 0.098)
phecode_327	Other degenerative diseases of nervous system	0.716 (0.707, 0.725)	0.782 (0.775, 0.791)	0.066 (0.059, 0.075)
phecode_328	Dementias and cerebral degeneration	0.796 (0.792, 0.799)	0.811 (0.808, 0.815)	0.016 (0.013, 0.018)
phecode_328-1	Alzheimer's disease	0.806 (0.801, 0.81)	0.81 (0.805, 0.815)	0.004 (0.002, 0.007)
phecode_328-2	Frontotemporal dementia	0.687 (0.662, 0.71)	0.746 (0.723, 0.769)	0.059 (0.035, 0.082)
phecode_328-4	Dementia with Lewy bodies	0.817 (0.803, 0.831)	0.853 (0.839, 0.868)	0.036 (0.024, 0.049)
phecode_328-7	Vascular dementia	0.814 (0.807, 0.821)	0.862 (0.857, 0.869)	0.049 (0.041, 0.055)
phecode_328-8	Dementia in conditions classified elsewhere	0.799 (0.794, 0.804)	0.813 (0.808, 0.818)	0.014 (0.011, 0.017)
phecode_328-9	Dementia NOS	0.795 (0.79, 0.8)	0.824 (0.819, 0.83)	0.03 (0.026, 0.034)
phecode_329	Symptoms and signs involving cognitive functions and awareness	0.653 (0.65, 0.655)	0.717 (0.715, 0.719)	0.064 (0.062, 0.067)
phecode_329-1	Memory loss	0.653 (0.649, 0.657)	0.759 (0.756, 0.762)	0.106 (0.103, 0.109)
phecode_329-4	Other specified cognitive deficit	0.539 (0.53, 0.547)	0.745 (0.738, 0.751)	0.206 (0.196, 0.216)
phecode_329-41	Attention and concentration deficit	0.617 (0.601, 0.632)	0.79 (0.779, 0.802)	0.174 (0.157, 0.19)
phecode_329-42	Cognitive communication deficit	0.574 (0.566, 0.583)	0.747 (0.739, 0.754)	0.173 (0.163, 0.182)
phecode_329-5	Mild cognitive impairment, so stated	0.761 (0.755, 0.769)	0.824 (0.819, 0.83)	0.063 (0.057, 0.067)
phecode_329-6	Transient global amnesia	0.643 (0.633, 0.653)	0.698 (0.687, 0.709)	0.055 (0.044, 0.065)
phecode_329-8	Altered mental status, unspecified	0.57 (0.541, 0.597)	0.687 (0.666, 0.71)	0.117 (0.088, 0.145)
phecode_329-9	Delirium	0.768 (0.762, 0.773)	0.828 (0.823, 0.834)	0.06 (0.055, 0.066)
phecode_330	Epilepsy, recurrent seizures, convulsions	0.582 (0.577, 0.589)	0.68 (0.675, 0.687)	0.098 (0.091, 0.105)
phecode_330-1	Epilepsy	0.58 (0.572, 0.588)	0.692 (0.684, 0.699)	0.112 (0.103, 0.12)
phecode_330-11	Generalized epilepsy	0.561 (0.546, 0.575)	0.766 (0.754, 0.779)	0.205 (0.188, 0.222)
phecode_330-12	Partial epilepsy	0.554 (0.536, 0.571)	0.721 (0.705, 0.736)	0.168 (0.146, 0.188)
phecode_330-3	Convulsions	0.576 (0.569, 0.583)	0.707 (0.701, 0.714)	0.131 (0.122, 0.139)
phecode_331	Headache	0.568 (0.565, 0.57)	0.704 (0.702, 0.706)	0.137 (0.134, 0.139)
phecode_331-1	Tension headache	0.599 (0.592, 0.605)	0.75 (0.745, 0.755)	0.151 (0.144, 0.158)
phecode_331-3	Headache syndromes, non migraine	0.569 (0.553, 0.587)	0.774 (0.761, 0.785)	0.205 (0.185, 0.224)
phecode_331-4	Cluster headaches	0.598 (0.577, 0.619)	0.765 (0.751, 0.78)	0.167 (0.146, 0.191)
phecode_331-6	Migraine	0.636 (0.632, 0.64)	0.722 (0.718, 0.725)	0.086 (0.082, 0.089)
phecode_331-61	Migraine with aura	0.595 (0.587, 0.604)	0.735 (0.727, 0.742)	0.14 (0.131, 0.149)
phecode_331-62	Hemiplegic migraine	0.669 (0.628, 0.711)	0.831 (0.803, 0.863)	0.162 (0.118, 0.206)
phecode_331-7	Drug induced headache	0.627 (0.596, 0.66)	0.804 (0.782, 0.83)	0.178 (0.143, 0.212)
phecode_331-8	Headache NOS	0.558 (0.556, 0.561)	0.713 (0.711, 0.715)	0.155 (0.152, 0.157)
phecode_333	Sleep disorders	0.519 (0.516, 0.521)	0.747 (0.745, 0.749)	0.228 (0.225, 0.231)
phecode_333-1	Sleep apnea	0.617 (0.613, 0.622)	0.765 (0.761, 0.769)	0.148 (0.143, 0.152)
phecode_333-11	Obstructive sleep apnea	0.626 (0.62, 0.632)	0.814 (0.81, 0.819)	0.188 (0.182, 0.195)
phecode_333-2	Insomnia	0.544 (0.541, 0.547)	0.738 (0.735, 0.74)	0.194 (0.191, 0.198)
phecode_333-3	Hypersomnia	0.531 (0.507, 0.554)	0.78 (0.763, 0.799)	0.248 (0.22, 0.279)
phecode_333-4	Circadian rhythm sleep disorder	0.528 (0.514, 0.541)	0.805 (0.796, 0.815)	0.276 (0.261, 0.293)
phecode_333-43	Circadian rhythm sleep disorder, jet lag or shift work type	0.551 (0.509, 0.598)	0.748 (0.716, 0.775)	0.196 (0.147, 0.24)
phecode_333-5	Parasomnia and sleep arousal Disorders	0.548 (0.526, 0.568)	0.758 (0.745, 0.775)	0.21 (0.189, 0.236)
phecode_334	Disorders of other cranial nerves	0.552 (0.546, 0.558)	0.663 (0.657, 0.668)	0.111 (0.104, 0.118)

Supplementary Tables

Table 17 continued from previous page

Endpoint	PheCode string	C-Index Age+Sex (IQR)	C-Index Age+Sex+MedicalHistory (IQR)	Difference in C-Index (IQR)
phecode_334-1	Trigeminal nerve disorders [CN5]	0.598 (0.591, 0.606)	0.721 (0.715, 0.728)	0.122 (0.114, 0.131)
phecode_334-11	Trigeminal neuralgia	0.604 (0.595, 0.612)	0.726 (0.718, 0.733)	0.122 (0.114, 0.131)
phecode_334-12	Atypical face pain	0.588 (0.57, 0.609)	0.741 (0.726, 0.757)	0.153 (0.131, 0.175)
phecode_334-2	Facial nerve disorders and weakness	0.523 (0.513, 0.532)	0.634 (0.625, 0.642)	0.111 (0.101, 0.123)
phecode_334-21	Bell's palsy	0.525 (0.515, 0.534)	0.629 (0.621, 0.64)	0.105 (0.093, 0.117)
phecode_334-23	Facial weakness	0.529 (0.504, 0.553)	0.726 (0.705, 0.745)	0.196 (0.17, 0.225)
phecode_334-24	Clonic hemifacial spasm*	0.564 (0.533, 0.602)	0.686 (0.656, 0.717)	0.121 (0.076, 0.163)
phecode_334-4	Disorders of oculomotor nerves	0.625 (0.609, 0.641)	0.668 (0.652, 0.687)	0.042 (0.024, 0.063)
phecode_334-41	Third [oculomotor] nerve palsy	0.642 (0.614, 0.673)	0.679 (0.65, 0.711)	0.037 (0.004, 0.068)
phecode_334-42	Fourth [trochlear] nerve palsy	0.608 (0.576, 0.639)	0.652 (0.619, 0.686)	0.044 (0.009, 0.08)
phecode_334-44	Sixth [abducent] nerve palsy	0.617 (0.594, 0.639)	0.684 (0.659, 0.708)	0.067 (0.041, 0.095)
phecode_335	Nerve root and plexus disorders	0.533 (0.527, 0.538)	0.719 (0.715, 0.724)	0.186 (0.18, 0.193)
phecode_335-1	Nerve plexus lesions	0.589 (0.561, 0.62)	0.703 (0.676, 0.731)	0.114 (0.078, 0.15)
phecode_335-11	Brachial plexus lesions	0.6 (0.568, 0.631)	0.707 (0.676, 0.739)	0.107 (0.073, 0.141)
phecode_335-2	Nerve root lesions	0.504 (0.47, 0.536)	0.764 (0.741, 0.786)	0.26 (0.227, 0.292)
phecode_335-4	Phantom limb (syndrome)	0.621 (0.572, 0.667)	0.865 (0.831, 0.901)	0.247 (0.199, 0.299)
phecode_336	Mononeuropathies	0.549 (0.546, 0.551)	0.676 (0.673, 0.679)	0.127 (0.124, 0.131)
phecode_336-1	Carpal tunnel syndrome	0.553 (0.549, 0.556)	0.674 (0.671, 0.678)	0.122 (0.117, 0.126)
phecode_336-2	Lesion of median, ulnar, radial nerve	0.547 (0.539, 0.555)	0.706 (0.699, 0.712)	0.159 (0.149, 0.168)
phecode_336-4	Mononeuritis of upper limb	0.494 (0.47, 0.524)	0.721 (0.694, 0.746)	0.227 (0.19, 0.265)
phecode_336-5	Mononeuritis of lower limb	0.578 (0.574, 0.584)	0.704 (0.7, 0.709)	0.126 (0.12, 0.13)
phecode_336-51	Lesion of sciatic nerve	0.585 (0.567, 0.603)	0.716 (0.699, 0.729)	0.13 (0.112, 0.146)
phecode_336-52	Meralgia paresthetica	0.501 (0.488, 0.512)	0.694 (0.686, 0.703)	0.193 (0.179, 0.205)
phecode_336-54	Tarsal tunnel syndrome	0.555 (0.516, 0.588)	0.695 (0.664, 0.729)	0.138 (0.101, 0.185)
phecode_336-55	Lesion of plantar nerve	0.624 (0.618, 0.63)	0.724 (0.718, 0.73)	0.1 (0.095, 0.106)
phecode_337	Polyneuropathies	0.611 (0.606, 0.616)	0.736 (0.731, 0.74)	0.125 (0.119, 0.13)
phecode_337-1	Hereditary and idiopathic neuropathy	0.663 (0.647, 0.681)	0.74 (0.721, 0.757)	0.077 (0.06, 0.093)
phecode_337-11	Hereditary motor and sensory neuropathy	0.672 (0.65, 0.691)	0.74 (0.72, 0.761)	0.069 (0.047, 0.089)
phecode_337-2	Inflammatory polyneuropathy	0.585 (0.567, 0.607)	0.634 (0.615, 0.654)	0.049 (0.026, 0.069)
phecode_337-21	Guillain-Barre syndrome [Acute infective polyneuritis]	0.588 (0.563, 0.617)	0.591 (0.563, 0.621)	0.003 (-0.024, 0.029)
phecode_337-3	Toxic neuropathy	0.551 (0.527, 0.575)	0.702 (0.678, 0.728)	0.151 (0.121, 0.181)
phecode_337-31	Drug-induced polyneuropathy	0.572 (0.542, 0.598)	0.666 (0.635, 0.7)	0.097 (0.06, 0.134)
phecode_337-8	Polyneuropathy in diseases classified elsewhere	0.657 (0.644, 0.669)	0.929 (0.922, 0.936)	0.272 (0.26, 0.286)
phecode_338	Myasthenia gravis and other myoneural disorders	0.619 (0.594, 0.641)	0.671 (0.649, 0.693)	0.052 (0.026, 0.078)
phecode_338-1	Myasthenia gravis	0.632 (0.606, 0.658)	0.682 (0.657, 0.705)	0.049 (0.022, 0.076)
phecode_339	Primary disorders of muscles	0.495 (0.467, 0.526)	0.688 (0.659, 0.72)	0.194 (0.157, 0.233)
phecode_339-1	Muscular dystrophy	0.468 (0.429, 0.508)	0.657 (0.609, 0.705)	0.189 (0.129, 0.253)
phecode_340	Myopathies	0.604 (0.583, 0.624)	0.74 (0.719, 0.761)	0.136 (0.109, 0.162)
phecode_341	Cerebral palsy and other paralytic syndromes	0.635 (0.627, 0.642)	0.746 (0.739, 0.752)	0.111 (0.103, 0.118)
phecode_341-1	Cerebral palsy	0.516 (0.459, 0.568)	0.731 (0.688, 0.771)	0.214 (0.164, 0.268)
phecode_341-2	Hemiplegia and hemiparesis	0.652 (0.644, 0.66)	0.758 (0.752, 0.766)	0.107 (0.098, 0.115)
phecode_341-6	Cauda equina syndrome	0.537 (0.513, 0.557)	0.717 (0.698, 0.736)	0.181 (0.154, 0.21)
phecode_342	Plegia and unspecified paralysis	0.577 (0.567, 0.586)	0.734 (0.725, 0.743)	0.157 (0.146, 0.168)
phecode_342-1	Paraplegia/Diplegia	0.574 (0.549, 0.598)	0.773 (0.748, 0.795)	0.197 (0.169, 0.228)
phecode_342-2	Quadriplegia	0.55 (0.515, 0.582)	0.683 (0.647, 0.721)	0.134 (0.085, 0.182)
phecode_342-4	Monoplegia	0.59 (0.578, 0.6)	0.763 (0.753, 0.772)	0.173 (0.159, 0.186)
phecode_342-5	Paralysis NOS	0.522 (0.483, 0.563)	0.754 (0.719, 0.793)	0.232 (0.186, 0.275)
phecode_343	Disorders of autonomic nervous system	0.533 (0.522, 0.544)	0.743 (0.734, 0.753)	0.211 (0.196, 0.225)

4 Medical history predicts future health trajectories over the human phenome

Table 17 continued from previous page

Endpoint	PheCode string	C-Index Age+Sex (IQR)	C-Index Age+Sex+MedicalHistory (IQR)	Difference in C-Index (IQR)
phecode_343-1	Autonomic neuropathy	0.629 (0.607, 0.653)	0.865 (0.848, 0.885)	0.237 (0.206, 0.266)
phecode_343-3	Complex regional pain syndrome	0.536 (0.522, 0.551)	0.773 (0.761, 0.784)	0.237 (0.218, 0.255)
phecode_343-5	Horner's syndrome*	0.493 (0.459, 0.526)	0.617 (0.587, 0.651)	0.125 (0.083, 0.164)
phecode_343-6	Multi-system degeneration of the autonomic nervous system*	0.674 (0.643, 0.707)	0.752 (0.712, 0.792)	0.079 (0.046, 0.111)
phecode_344	Disorders of the circulation of the cerebrospinal fluid	0.572 (0.559, 0.586)	0.656 (0.644, 0.669)	0.084 (0.069, 0.1)
phecode_344-1	Hydrocephalus	0.604 (0.589, 0.62)	0.683 (0.668, 0.698)	0.079 (0.062, 0.096)
phecode_344-12	Obstructive hydrocephalus	0.586 (0.549, 0.622)	0.71 (0.674, 0.743)	0.123 (0.081, 0.167)
phecode_344-13	(Idiopathic) normal pressure hydrocephalus	0.727 (0.701, 0.754)	0.745 (0.718, 0.769)	0.018 (-0.009, 0.048)
phecode_344-2	Benign intracranial hypertension	0.518 (0.479, 0.558)	0.63 (0.59, 0.672)	0.114 (0.062, 0.162)
phecode_344-3	Cerebrospinal fluid leak	0.476 (0.444, 0.507)	0.592 (0.56, 0.625)	0.118 (0.074, 0.159)
phecode_345	Encephalopathy	0.619 (0.597, 0.641)	0.789 (0.769, 0.81)	0.169 (0.145, 0.196)
phecode_346	Brain damage and brain death	0.568 (0.555, 0.579)	0.673 (0.662, 0.686)	0.106 (0.092, 0.121)
phecode_346-1	Postconcussion syndrome	0.545 (0.519, 0.575)	0.762 (0.741, 0.782)	0.216 (0.186, 0.246)
phecode_346-3	Anoxic brain damage	0.668 (0.646, 0.69)	0.807 (0.789, 0.825)	0.139 (0.118, 0.16)
phecode_346-5	Compression of brain	0.552 (0.514, 0.589)	0.663 (0.627, 0.701)	0.112 (0.073, 0.153)
phecode_346-6	Cerebral edema	0.601 (0.578, 0.626)	0.735 (0.717, 0.755)	0.134 (0.11, 0.16)
phecode_347	Other disorders of the brain and CNS	0.585 (0.575, 0.595)	0.68 (0.67, 0.69)	0.095 (0.084, 0.106)
phecode_347-1	Cerebral cysts	0.465 (0.442, 0.487)	0.627 (0.6, 0.65)	0.161 (0.128, 0.194)
phecode_347-2	Disorders of meninges	0.52 (0.479, 0.554)	0.695 (0.659, 0.729)	0.176 (0.129, 0.225)
phecode_348	Other diseases of spinal cord	0.568 (0.56, 0.574)	0.7 (0.693, 0.706)	0.132 (0.124, 0.141)
phecode_348-2	Myelopathies	0.564 (0.556, 0.571)	0.703 (0.695, 0.711)	0.139 (0.131, 0.148)
phecode_348-21	Vascular myelopathies	0.682 (0.659, 0.706)	0.768 (0.749, 0.789)	0.087 (0.061, 0.111)
phecode_348-4	Spinal cord compression*	0.633 (0.617, 0.65)	0.718 (0.704, 0.732)	0.085 (0.069, 0.1)
phecode_349	Disorder of nervous system	0.61 (0.604, 0.616)	0.694 (0.689, 0.7)	0.084 (0.078, 0.09)
phecode_349-1	Abnormal findings on diagnostic test of central nervous system	0.627 (0.619, 0.634)	0.686 (0.678, 0.693)	0.059 (0.051, 0.067)
phecode_349-12	Abnormal electroencephalogram [EEG]	0.529 (0.492, 0.565)	0.798 (0.767, 0.827)	0.268 (0.232, 0.309)
phecode_349-13	Abnormal findings on diagnostic imaging of skull and head	0.642 (0.634, 0.651)	0.731 (0.724, 0.738)	0.089 (0.079, 0.096)
phecode_349-15	Intracranial space-occupying lesion found on diagnostic imaging of central nervous system*	0.565 (0.535, 0.594)	0.695 (0.667, 0.723)	0.13 (0.097, 0.162)
phecode_349-2	Abnormal results of function studies of peripheral nervous system	0.608 (0.597, 0.617)	0.744 (0.737, 0.753)	0.137 (0.127, 0.148)
phecode_349-3	Nonspecific abnormal electromyogram [EMG]	0.536 (0.507, 0.572)	0.794 (0.767, 0.819)	0.255 (0.22, 0.29)
phecode_350	Other symptoms involving nervous system	0.617 (0.614, 0.62)	0.706 (0.704, 0.709)	0.09 (0.087, 0.092)
phecode_350-3	Abnormal reflex	0.54 (0.517, 0.564)	0.752 (0.732, 0.772)	0.212 (0.184, 0.239)
phecode_350-5	Repeated falls*	0.711 (0.708, 0.715)	0.789 (0.786, 0.792)	0.078 (0.074, 0.081)
phecode_351	Disturbances of skin sensation	0.529 (0.526, 0.531)	0.735 (0.733, 0.737)	0.206 (0.204, 0.209)
phecode_351-1	Anesthesia of skin*	0.509 (0.504, 0.514)	0.731 (0.727, 0.735)	0.222 (0.216, 0.228)
phecode_351-2	Hypoesthesia of skin*	0.68 (0.672, 0.686)	0.905 (0.9, 0.909)	0.225 (0.218, 0.233)
phecode_351-3	Paresthesia of skin*	0.518 (0.515, 0.522)	0.708 (0.705, 0.712)	0.19 (0.186, 0.194)
phecode_351-4	Hyperesthesia*	0.519 (0.494, 0.544)	0.788 (0.769, 0.805)	0.269 (0.239, 0.299)
phecode_352	Disturbances of sensation of smell and taste	0.536 (0.529, 0.544)	0.703 (0.696, 0.709)	0.167 (0.159, 0.176)
phecode_352-1	Anosmia*	0.506 (0.495, 0.518)	0.684 (0.674, 0.694)	0.179 (0.164, 0.193)
phecode_352-2	Parosmia*	0.46 (0.43, 0.49)	0.691 (0.67, 0.714)	0.232 (0.198, 0.262)
phecode_352-3	Parageusia*	0.541 (0.527, 0.554)	0.719 (0.706, 0.729)	0.178 (0.163, 0.194)
phecode_353	Symptoms and signs involving general sensations and perceptions	0.585 (0.579, 0.59)	0.684 (0.679, 0.689)	0.1 (0.094, 0.105)
phecode_353-1	Hallucinations	0.63 (0.618, 0.64)	0.765 (0.756, 0.775)	0.136 (0.124, 0.148)
phecode_353-11	Auditory hallucinations*	0.525 (0.496, 0.551)	0.811 (0.786, 0.834)	0.286 (0.253, 0.319)
phecode_353-12	Visual hallucinations	0.656 (0.638, 0.673)	0.77 (0.756, 0.787)	0.114 (0.094, 0.133)
phecode_354	Dizziness and giddiness	0.576 (0.574, 0.578)	0.711 (0.71, 0.713)	0.135 (0.133, 0.137)
phecode_355	Coma and other alteration of consciousness	0.623 (0.616, 0.63)	0.731 (0.725, 0.738)	0.109 (0.101, 0.116)
phecode_355-1	Coma	0.601 (0.589, 0.615)	0.673 (0.661, 0.684)	0.072 (0.058, 0.085)
phecode_355-2	Alteration of consciousness	0.634 (0.625, 0.641)	0.758 (0.751, 0.765)	0.125 (0.117, 0.133)

Supplementary Tables

Table 17 continued from previous page

Endpoint	PheCode string	C-Index Age+Sex (IQR)	C-Index Age+Sex+MedicalHistory (IQR)	Difference in C-Index (IQR)
phecode_355-21	Transient alteration of awareness	0.502 (0.462, 0.538)	0.865 (0.849, 0.885)	0.365 (0.326, 0.405)
phecode_356	Speech disturbance	0.639 (0.632, 0.646)	0.712 (0.705, 0.718)	0.073 (0.066, 0.08)
phecode_356-1	Dysarthria	0.643 (0.627, 0.66)	0.729 (0.713, 0.744)	0.086 (0.068, 0.102)
phecode_356-2	Aphasia and dysphasia	0.667 (0.657, 0.677)	0.713 (0.703, 0.724)	0.046 (0.037, 0.056)
phecode_356-4	Slurred speech*	0.602 (0.58, 0.624)	0.774 (0.757, 0.791)	0.172 (0.148, 0.199)
phecode_360	Inflammation of eyelids	0.561 (0.559, 0.564)	0.699 (0.697, 0.701)	0.137 (0.134, 0.14)
phecode_360-1	Hordeolum	0.525 (0.521, 0.53)	0.692 (0.688, 0.696)	0.166 (0.161, 0.172)
phecode_360-11	Hordeolum externum	0.525 (0.521, 0.53)	0.692 (0.688, 0.695)	0.167 (0.162, 0.172)
phecode_360-12	Hordeolum internum	0.532 (0.522, 0.541)	0.709 (0.7, 0.716)	0.178 (0.166, 0.188)
phecode_360-13	Abscess of eyelid	0.508 (0.48, 0.536)	0.691 (0.669, 0.716)	0.184 (0.152, 0.216)
phecode_360-2	Chalazion	0.525 (0.52, 0.53)	0.669 (0.665, 0.673)	0.144 (0.138, 0.15)
phecode_360-4	Blepharitis	0.595 (0.592, 0.598)	0.725 (0.722, 0.728)	0.13 (0.127, 0.134)
phecode_360-5	Noninfectious dermatoses of eyelid	0.651 (0.641, 0.663)	0.766 (0.757, 0.776)	0.115 (0.103, 0.126)
phecode_360-51	Eczematous dermatitis of eyelid	0.659 (0.648, 0.672)	0.769 (0.76, 0.779)	0.11 (0.098, 0.121)
phecode_361	Disorders of eyelid function	0.651 (0.646, 0.656)	0.694 (0.688, 0.7)	0.043 (0.038, 0.048)
phecode_361-1	Entropion and trichiasis of eyelid	0.701 (0.691, 0.712)	0.727 (0.716, 0.738)	0.026 (0.018, 0.034)
phecode_361-15	Trichiasis of eyelid without entropion	0.636 (0.611, 0.661)	0.748 (0.724, 0.768)	0.11 (0.088, 0.135)
phecode_361-2	Lagophthalmos	0.518 (0.471, 0.562)	0.73 (0.688, 0.773)	0.21 (0.156, 0.27)
phecode_361-3	Ptosis of eyelid	0.631 (0.623, 0.639)	0.694 (0.686, 0.702)	0.063 (0.054, 0.071)
phecode_361-4	Blepharochalasis	0.654 (0.644, 0.668)	0.696 (0.682, 0.708)	0.041 (0.031, 0.052)
phecode_361-9	Ectropion of eyelid	0.724 (0.713, 0.736)	0.744 (0.732, 0.756)	0.02 (0.012, 0.028)
phecode_362	Other disorders of the eyelids	0.552 (0.547, 0.557)	0.66 (0.656, 0.665)	0.108 (0.102, 0.114)
phecode_362-1	Xanthelasma of eyelid	0.608 (0.594, 0.627)	0.691 (0.677, 0.705)	0.082 (0.065, 0.098)
phecode_362-5	Cysts of eyelid	0.542 (0.53, 0.554)	0.683 (0.674, 0.692)	0.141 (0.128, 0.155)
phecode_362-6	Dermatochalasis of eyelid	0.591 (0.553, 0.629)	0.865 (0.848, 0.882)	0.274 (0.236, 0.314)
phecode_363	Disorders of lacrimal system	0.639 (0.636, 0.642)	0.752 (0.749, 0.754)	0.113 (0.111, 0.115)
phecode_363-2	Dry eye syndrome [Tear film insufficiency]	0.644 (0.641, 0.646)	0.772 (0.77, 0.774)	0.128 (0.126, 0.131)
phecode_363-5	Epiphora	0.624 (0.618, 0.63)	0.72 (0.714, 0.725)	0.096 (0.09, 0.101)
phecode_363-51	Epiphora due to excess lacrimation	0.603 (0.565, 0.639)	0.748 (0.721, 0.776)	0.146 (0.108, 0.181)
phecode_363-6	Inflammation of lacrimal passages	0.61 (0.589, 0.632)	0.683 (0.662, 0.704)	0.074 (0.051, 0.094)
phecode_363-61	Dacryocystitis	0.604 (0.58, 0.628)	0.719 (0.699, 0.739)	0.115 (0.089, 0.14)
phecode_363-7	Stenosis and insufficiency of lacrimal passages	0.644 (0.634, 0.654)	0.698 (0.689, 0.708)	0.055 (0.045, 0.064)
phecode_365	Noninflammatory disorders of the orbit	0.553 (0.535, 0.57)	0.666 (0.649, 0.681)	0.113 (0.096, 0.131)
phecode_365-2	Orbital edema or congestion	0.587 (0.566, 0.608)	0.71 (0.692, 0.729)	0.122 (0.104, 0.144)
phecode_365-3	Exophthalmos [Proptosis]	0.506 (0.476, 0.537)	0.669 (0.643, 0.698)	0.163 (0.127, 0.202)
phecode_366	Noninflammatory disorders of conjunctiva	0.549 (0.541, 0.557)	0.674 (0.668, 0.681)	0.125 (0.117, 0.134)
phecode_366-1	Pterygium of eye	0.531 (0.515, 0.545)	0.664 (0.651, 0.677)	0.133 (0.115, 0.15)
phecode_366-2	Conjunctival degenerations and deposits	0.555 (0.538, 0.573)	0.685 (0.67, 0.7)	0.13 (0.111, 0.15)
phecode_366-21	Pinguecula	0.552 (0.533, 0.571)	0.679 (0.664, 0.695)	0.128 (0.104, 0.149)
phecode_366-4	Vascular abnormalities of conjunctiva	0.55 (0.539, 0.561)	0.691 (0.682, 0.7)	0.141 (0.129, 0.154)
phecode_366-42	Conjunctival hyperemia	0.545 (0.53, 0.559)	0.717 (0.707, 0.728)	0.172 (0.158, 0.189)
phecode_366-5	Conjunctival edema	0.421 (0.375, 0.464)	0.635 (0.596, 0.671)	0.214 (0.153, 0.27)
phecode_366-6	Conjunctival cysts	0.547 (0.523, 0.573)	0.694 (0.673, 0.714)	0.147 (0.118, 0.175)
phecode_367	Inflammation of the eye	0.561 (0.559, 0.564)	0.696 (0.694, 0.698)	0.135 (0.132, 0.137)
phecode_367-1	Conjunctivitis	0.567 (0.564, 0.57)	0.703 (0.701, 0.705)	0.136 (0.134, 0.139)
phecode_367-12	Allergic [atopic] conjunctivitis	0.563 (0.555, 0.57)	0.743 (0.737, 0.748)	0.18 (0.171, 0.187)
phecode_367-13	Blepharoconjunctivitis	0.591 (0.58, 0.602)	0.745 (0.737, 0.753)	0.154 (0.143, 0.164)
phecode_367-2	Keratitis	0.528 (0.519, 0.536)	0.675 (0.667, 0.682)	0.146 (0.136, 0.157)
phecode_367-21	Corneal ulcer	0.531 (0.516, 0.547)	0.685 (0.671, 0.699)	0.154 (0.135, 0.171)
phecode_367-22	Punctate keratitis	0.564 (0.525, 0.604)	0.717 (0.687, 0.751)	0.153 (0.103, 0.2)

4 Medical history predicts future health trajectories over the human phenome

Table 17 continued from previous page

Endpoint	PheCode string	C-Index Age+Sex (IQR)	C-Index Age+Sex+MedicalHistory (IQR)	Difference in C-Index (IQR)
phecode_367-3	Keratoconjunctivitis	0.542 (0.52, 0.565)	0.709 (0.689, 0.73)	0.167 (0.14, 0.194)
phecode_367-4	Inflammation of orbit	0.535 (0.513, 0.555)	0.675 (0.659, 0.691)	0.141 (0.118, 0.168)
phecode_367-41	Cellulitis of orbit	0.546 (0.523, 0.57)	0.704 (0.684, 0.722)	0.159 (0.129, 0.185)
phecode_367-5	Uveitis	0.531 (0.523, 0.54)	0.672 (0.664, 0.68)	0.14 (0.129, 0.151)
phecode_367-52	Iridocyclitis	0.531 (0.523, 0.54)	0.671 (0.663, 0.68)	0.139 (0.129, 0.15)
phecode_367-6	Episcleritis	0.564 (0.553, 0.576)	0.695 (0.686, 0.704)	0.131 (0.117, 0.143)
phecode_367-7	Scleritis	0.501 (0.475, 0.523)	0.68 (0.661, 0.702)	0.18 (0.153, 0.212)
phecode_367-9	Chorioretinal inflammation	0.471 (0.438, 0.504)	0.653 (0.621, 0.687)	0.182 (0.147, 0.22)
phecode_369	Noninflammatory disorders of the cornea	0.621 (0.614, 0.629)	0.687 (0.68, 0.693)	0.065 (0.059, 0.072)
phecode_369-1	Corneal scars and opacities	0.606 (0.589, 0.623)	0.698 (0.683, 0.713)	0.091 (0.073, 0.109)
phecode_369-2	Corneal edema	0.675 (0.651, 0.696)	0.774 (0.749, 0.797)	0.099 (0.072, 0.128)
phecode_369-4	Corneal degenerations	0.659 (0.646, 0.67)	0.713 (0.701, 0.724)	0.054 (0.046, 0.061)
phecode_369-42	Recurrent erosion of cornea	0.573 (0.545, 0.598)	0.69 (0.665, 0.713)	0.117 (0.088, 0.145)
phecode_369-44	Senile corneal changes including arcus senilis	0.527 (0.475, 0.577)	0.685 (0.645, 0.719)	0.156 (0.105, 0.215)
phecode_369-5	Hereditary corneal dystrophies	0.638 (0.625, 0.651)	0.678 (0.664, 0.691)	0.04 (0.027, 0.052)
phecode_369-51	Fuchs' dystrophy	0.654 (0.637, 0.674)	0.718 (0.701, 0.735)	0.064 (0.046, 0.077)
phecode_369-6	Corneal deformities	0.592 (0.553, 0.628)	0.676 (0.64, 0.717)	0.084 (0.049, 0.125)
phecode_369-62	Keratoconus	0.61 (0.574, 0.65)	0.682 (0.64, 0.727)	0.071 (0.031, 0.107)
phecode_370	Disorders of iris and ciliary body	0.623 (0.612, 0.633)	0.709 (0.699, 0.721)	0.086 (0.076, 0.098)
phecode_370-1	Degeneration of iris and ciliary body	0.576 (0.55, 0.602)	0.732 (0.706, 0.756)	0.155 (0.123, 0.187)
phecode_370-3	Vascular disorders of iris and ciliary body	0.643 (0.626, 0.662)	0.708 (0.693, 0.726)	0.065 (0.047, 0.081)
phecode_370-4	Adhesions of iris	0.64 (0.611, 0.674)	0.764 (0.735, 0.798)	0.123 (0.086, 0.157)
phecode_371	Cataract	0.734 (0.732, 0.736)	0.752 (0.75, 0.753)	0.018 (0.017, 0.019)
phecode_371-3	Nuclear cataract	0.744 (0.741, 0.747)	0.773 (0.771, 0.776)	0.029 (0.027, 0.031)
phecode_371-31	Age-related nuclear cataract	0.744 (0.742, 0.747)	0.773 (0.771, 0.776)	0.029 (0.027, 0.031)
phecode_372	Disorders of lens (excluding cataracts)	0.663 (0.644, 0.684)	0.74 (0.721, 0.758)	0.077 (0.055, 0.098)
phecode_372-1	Aphakia	0.64 (0.612, 0.673)	0.781 (0.75, 0.818)	0.141 (0.101, 0.177)
phecode_372-2	Dislocation of lens	0.628 (0.587, 0.674)	0.731 (0.686, 0.777)	0.101 (0.049, 0.161)
phecode_373	Noninflammatory disorders of choroid	0.624 (0.612, 0.638)	0.728 (0.717, 0.74)	0.104 (0.092, 0.117)
phecode_373-1	Chorioretinal scars	0.603 (0.582, 0.621)	0.743 (0.723, 0.763)	0.141 (0.118, 0.161)
phecode_373-2	Choroidal degenerations	0.6 (0.553, 0.651)	0.687 (0.637, 0.745)	0.086 (0.036, 0.141)
phecode_374	Disorders of the retina	0.631 (0.628, 0.633)	0.755 (0.752, 0.757)	0.124 (0.122, 0.126)
phecode_374-1	Retinal detachments and breaks	0.576 (0.569, 0.582)	0.616 (0.611, 0.623)	0.041 (0.034, 0.048)
phecode_374-11	Serous retinal detachment	0.582 (0.574, 0.591)	0.628 (0.619, 0.638)	0.046 (0.037, 0.056)
phecode_374-12	Traction detachment of retina	0.567 (0.531, 0.6)	0.684 (0.645, 0.724)	0.119 (0.072, 0.166)
phecode_374-13	Horseshoe tear of retina without detachment	0.364 (0.316, 0.416)	0.598 (0.555, 0.638)	0.234 (0.177, 0.292)
phecode_374-14	Round hole of retina without detachment	0.449 (0.402, 0.492)	0.695 (0.659, 0.733)	0.244 (0.192, 0.298)
phecode_374-2	Retinoschisis and retinal cysts	0.572 (0.555, 0.591)	0.65 (0.634, 0.666)	0.077 (0.059, 0.097)
phecode_374-21	Retinoschisis	0.569 (0.549, 0.591)	0.662 (0.644, 0.681)	0.092 (0.071, 0.115)
phecode_374-3	Retinal vascular changes and occlusions	0.649 (0.644, 0.654)	0.728 (0.723, 0.733)	0.079 (0.073, 0.083)
phecode_374-32	Retinal microaneurysms	0.601 (0.576, 0.627)	0.785 (0.765, 0.806)	0.183 (0.157, 0.213)
phecode_374-33	Retinal neovascularization	0.657 (0.637, 0.68)	0.768 (0.75, 0.786)	0.111 (0.09, 0.131)
phecode_374-37	Retinal arterial occlusions	0.668 (0.651, 0.684)	0.739 (0.724, 0.754)	0.071 (0.053, 0.087)
phecode_374-38	Retinal vein occlusions	0.661 (0.653, 0.669)	0.728 (0.721, 0.736)	0.067 (0.061, 0.073)
phecode_374-39	Transient retinal arterial occlusion [Amaurosis fugax]	0.674 (0.661, 0.686)	0.733 (0.722, 0.744)	0.06 (0.048, 0.071)
phecode_374-4	Retinal disorders in diseases classified elsewhere	0.623 (0.62, 0.627)	0.884 (0.882, 0.887)	0.261 (0.258, 0.265)

Supplementary Tables

Table 17 continued from previous page

Endpoint	PheCode string	C-Index Age+Sex (IQR)	C-Index Age+Sex+MedicalHistory (IQR)	Difference in C-Index (IQR)
phecode_374-41	Hypertensive retinopathy	0.515 (0.479, 0.557)	0.682 (0.646, 0.717)	0.167 (0.116, 0.218)
phecode_374-42	Diabetic retinopathy	0.621 (0.617, 0.624)	0.895 (0.893, 0.898)	0.274 (0.271, 0.278)
phecode_374-5	Macular degeneration	0.723 (0.72, 0.727)	0.755 (0.752, 0.758)	0.032 (0.029, 0.034)
phecode_374-51	Age-related macular degeneration	0.747 (0.742, 0.752)	0.803 (0.799, 0.807)	0.055 (0.052, 0.058)
phecode_374-511	Nonexudative (dry) age-related macular degeneration	0.756 (0.747, 0.764)	0.825 (0.818, 0.833)	0.069 (0.063, 0.076)
phecode_374-512	Exudative (wet) age-related macular degeneration	0.777 (0.765, 0.787)	0.848 (0.838, 0.856)	0.071 (0.062, 0.08)
phecode_374-52	Macular cyst, hole, or pseudohole	0.695 (0.686, 0.704)	0.749 (0.741, 0.757)	0.054 (0.047, 0.062)
phecode_374-54	Drusen (degenerative) of macula	0.669 (0.651, 0.686)	0.796 (0.781, 0.812)	0.128 (0.11, 0.144)
phecode_374-55	Puckering of macula	0.667 (0.659, 0.674)	0.752 (0.745, 0.759)	0.085 (0.078, 0.092)
phecode_374-6	Peripheral retinal degeneration	0.478 (0.457, 0.499)	0.631 (0.613, 0.652)	0.153 (0.124, 0.182)
phecode_374-61	Lattice degeneration of retina	0.522 (0.484, 0.558)	0.642 (0.612, 0.673)	0.119 (0.072, 0.172)
phecode_374-7	Hereditary retinal dystrophy	0.618 (0.594, 0.643)	0.673 (0.652, 0.698)	0.057 (0.038, 0.077)
phecode_374-8	Retinal edema	0.644 (0.634, 0.653)	0.804 (0.796, 0.812)	0.16 (0.149, 0.17)
phecode_374-9	Central serous chorioretinopathy	0.614 (0.586, 0.639)	0.677 (0.657, 0.697)	0.064 (0.041, 0.086)
phecode_375	Abnormal intraocular pressure	0.632 (0.629, 0.635)	0.676 (0.674, 0.679)	0.044 (0.042, 0.047)
phecode_375-1	Glaucoma	0.658 (0.654, 0.661)	0.699 (0.696, 0.703)	0.042 (0.038, 0.045)
phecode_375-11	Open angle glaucoma	0.678 (0.673, 0.684)	0.756 (0.75, 0.762)	0.078 (0.072, 0.083)
phecode_375-112	Pigmentary glaucoma	0.548 (0.511, 0.588)	0.794 (0.763, 0.823)	0.245 (0.204, 0.283)
phecode_375-113	Primary open angle glaucoma	0.672 (0.664, 0.681)	0.786 (0.778, 0.793)	0.113 (0.105, 0.121)
phecode_375-12	Angle-Closure Glaucoma	0.661 (0.651, 0.671)	0.713 (0.704, 0.724)	0.053 (0.044, 0.062)
phecode_375-14	Low-tension glaucoma (Normal-tension glaucoma)	0.663 (0.65, 0.677)	0.749 (0.737, 0.76)	0.086 (0.074, 0.098)
phecode_375-6	Ocular hypertension	0.596 (0.59, 0.602)	0.713 (0.708, 0.718)	0.117 (0.111, 0.121)
phecode_375-7	Hypotony of eye	0.6 (0.557, 0.649)	0.761 (0.718, 0.814)	0.16 (0.11, 0.211)
phecode_376	Disorders of vitreous body	0.617 (0.614, 0.619)	0.71 (0.707, 0.712)	0.093 (0.091, 0.095)
phecode_376-1	Vitreous degeneration	0.615 (0.612, 0.618)	0.716 (0.713, 0.719)	0.101 (0.098, 0.104)
phecode_376-2	Vitreous opacities	0.617 (0.614, 0.619)	0.71 (0.707, 0.713)	0.093 (0.091, 0.095)
phecode_376-21	Crystalline deposits in vitreous body	0.607 (0.602, 0.613)	0.711 (0.707, 0.716)	0.104 (0.099, 0.109)
phecode_376-4	Vitreomacular adhesion	0.672 (0.652, 0.692)	0.759 (0.739, 0.778)	0.087 (0.069, 0.106)
phecode_377	Hemorrhage of the eye	0.596 (0.593, 0.6)	0.691 (0.688, 0.694)	0.095 (0.092, 0.098)
phecode_377-2	Conjunctival hemorrhage	0.594 (0.59, 0.598)	0.696 (0.693, 0.7)	0.103 (0.099, 0.106)
phecode_377-4	Retinal hemorrhage	0.634 (0.623, 0.646)	0.709 (0.7, 0.72)	0.075 (0.066, 0.084)
phecode_377-5	Vitreous hemorrhage	0.6 (0.588, 0.612)	0.69 (0.678, 0.704)	0.091 (0.076, 0.105)
phecode_377-8	HypHEMA	0.569 (0.532, 0.607)	0.676 (0.636, 0.715)	0.106 (0.066, 0.149)
phecode_379	Infection of the eye	0.561 (0.552, 0.571)	0.669 (0.66, 0.677)	0.107 (0.097, 0.118)
phecode_379-2	Eye infection, viral	0.561 (0.552, 0.57)	0.664 (0.656, 0.673)	0.103 (0.092, 0.114)
phecode_379-21	Infection of the eye, herpes	0.591 (0.578, 0.603)	0.665 (0.653, 0.677)	0.074 (0.06, 0.089)
phecode_380	Disorders of optic nerve and visual pathways	0.573 (0.563, 0.581)	0.675 (0.666, 0.684)	0.103 (0.092, 0.112)
phecode_380-1	Optic neuropathy	0.559 (0.543, 0.578)	0.662 (0.642, 0.679)	0.101 (0.079, 0.126)
phecode_380-11	Optic neuritis	0.553 (0.517, 0.584)	0.681 (0.651, 0.717)	0.13 (0.085, 0.173)
phecode_380-12	Ischemic optic neuropathy	0.639 (0.615, 0.666)	0.724 (0.701, 0.745)	0.084 (0.057, 0.109)
phecode_380-2	Disorders of optic disc	0.569 (0.557, 0.579)	0.686 (0.676, 0.696)	0.117 (0.106, 0.129)
phecode_380-21	Papilledema	0.56 (0.517, 0.605)	0.603 (0.556, 0.655)	0.042 (-0.006, 0.094)
phecode_380-22	Optic disc drusen	0.437 (0.402, 0.473)	0.601 (0.572, 0.631)	0.164 (0.116, 0.211)
phecode_380-3	Optic atrophy	0.612 (0.586, 0.64)	0.746 (0.725, 0.771)	0.134 (0.103, 0.164)
phecode_381	Strabismus	0.604 (0.595, 0.613)	0.66 (0.65, 0.669)	0.057 (0.045, 0.067)

4 Medical history predicts future health trajectories over the human phenome

Table 17 continued from previous page

Endpoint	PheCode string	C-Index Age+Sex (IQR)	C-Index Age+Sex+MedicalHistory (IQR)	Difference in C-Index (IQR)
phecode_381-1	Paralytic strabismus [Neurogenic strabismus]	0.644 (0.629, 0.659)	0.671 (0.657, 0.687)	0.027 (0.011, 0.042)
phecode_381-11	Ophthalmoplegia	0.713 (0.688, 0.74)	0.73 (0.707, 0.757)	0.017 (-0.01, 0.043)
phecode_381-3	Esotropia	0.58 (0.556, 0.602)	0.66 (0.637, 0.687)	0.08 (0.054, 0.112)
phecode_381-4	Exotropia	0.554 (0.533, 0.577)	0.691 (0.67, 0.71)	0.136 (0.111, 0.163)
phecode_381-6	Duane's syndrome [Duane anomaly]	0.585 (0.564, 0.605)	0.69 (0.671, 0.708)	0.104 (0.083, 0.127)
phecode_381-8	Heterophoria	0.555 (0.53, 0.582)	0.722 (0.701, 0.742)	0.167 (0.138, 0.195)
phecode_381-81	Esophoria	0.57 (0.524, 0.614)	0.709 (0.668, 0.748)	0.138 (0.088, 0.181)
phecode_381-82	Exophoria	0.574 (0.54, 0.612)	0.761 (0.733, 0.792)	0.187 (0.15, 0.223)
phecode_382	Other disorders of binocular movement	0.515 (0.492, 0.537)	0.716 (0.695, 0.737)	0.201 (0.174, 0.229)
phecode_383	Irregular eye movements	0.55 (0.536, 0.565)	0.68 (0.668, 0.694)	0.13 (0.112, 0.148)
phecode_383-1	Nystagmus	0.57 (0.552, 0.586)	0.723 (0.709, 0.738)	0.154 (0.133, 0.172)
phecode_384	Anomalies of pupillary function	0.552 (0.534, 0.568)	0.675 (0.658, 0.69)	0.123 (0.105, 0.141)
phecode_384-1	Anisocoria	0.523 (0.492, 0.557)	0.649 (0.618, 0.674)	0.125 (0.085, 0.162)
phecode_384-3	Mydriasis	0.528 (0.498, 0.561)	0.77 (0.743, 0.796)	0.241 (0.204, 0.278)
phecode_384-4	Tonic pupil	0.528 (0.491, 0.575)	0.661 (0.621, 0.704)	0.132 (0.088, 0.173)
phecode_385	Abnormal results of function studies of eye	0.635 (0.623, 0.648)	0.788 (0.779, 0.798)	0.153 (0.141, 0.166)
phecode_386	Visual disturbances	0.594 (0.59, 0.597)	0.69 (0.687, 0.693)	0.096 (0.093, 0.1)
phecode_386-1	Amblyopia	0.631 (0.622, 0.639)	0.712 (0.704, 0.721)	0.081 (0.071, 0.09)
phecode_386-2	Diplopia	0.611 (0.604, 0.62)	0.692 (0.683, 0.7)	0.081 (0.071, 0.089)
phecode_386-3	Visual discomfort	0.603 (0.563, 0.639)	0.76 (0.73, 0.793)	0.158 (0.117, 0.195)
phecode_386-4	Visual field defects	0.6 (0.591, 0.609)	0.673 (0.665, 0.682)	0.073 (0.064, 0.082)
phecode_386-41	Scotoma	0.574 (0.554, 0.595)	0.74 (0.726, 0.755)	0.165 (0.146, 0.187)
phecode_386-42	Hemianopia	0.644 (0.625, 0.663)	0.731 (0.716, 0.747)	0.087 (0.068, 0.105)
phecode_386-8	Sudden or transient visual loss	0.603 (0.569, 0.635)	0.727 (0.702, 0.757)	0.125 (0.088, 0.163)
phecode_386-9	Visual distortions and subjective visual disturbances	0.632 (0.622, 0.644)	0.752 (0.742, 0.761)	0.119 (0.108, 0.13)
phecode_387	Disorders of refraction and accommodation	0.635 (0.63, 0.64)	0.679 (0.674, 0.684)	0.044 (0.04, 0.048)
phecode_387-1	Hypermetropia	0.676 (0.661, 0.689)	0.714 (0.701, 0.728)	0.038 (0.029, 0.049)
phecode_387-2	Myopia	0.593 (0.587, 0.6)	0.665 (0.659, 0.672)	0.072 (0.065, 0.079)
phecode_387-21	Progressive high (degenerative) myopia	0.616 (0.596, 0.637)	0.708 (0.684, 0.731)	0.092 (0.063, 0.121)
phecode_387-3	Astigmatism	0.7 (0.692, 0.709)	0.732 (0.724, 0.74)	0.032 (0.027, 0.036)
phecode_387-4	Presbyopia	0.512 (0.479, 0.544)	0.67 (0.64, 0.701)	0.159 (0.124, 0.198)
phecode_387-5	Anisometropia	0.613 (0.58, 0.649)	0.694 (0.658, 0.728)	0.08 (0.044, 0.115)
phecode_388	Blindness and low vision	0.672 (0.666, 0.677)	0.778 (0.773, 0.783)	0.106 (0.101, 0.112)
phecode_388-1	Blindness of both eyes	0.608 (0.587, 0.628)	0.869 (0.855, 0.887)	0.261 (0.239, 0.285)
phecode_389	Other disorders of eye	0.561 (0.559, 0.564)	0.72 (0.718, 0.722)	0.159 (0.157, 0.161)
phecode_389-1	Ocular pain	0.559 (0.551, 0.568)	0.731 (0.725, 0.738)	0.172 (0.163, 0.181)
phecode_390	Disorders of external ear	0.554 (0.553, 0.556)	0.686 (0.685, 0.688)	0.132 (0.13, 0.133)
phecode_390-1	Otitis externa	0.503 (0.5, 0.505)	0.698 (0.696, 0.7)	0.196 (0.192, 0.199)
phecode_390-4	Impacted cerumen	0.587 (0.585, 0.588)	0.697 (0.695, 0.699)	0.11 (0.109, 0.112)
phecode_390-5	Stenosis of external ear canal	0.615 (0.568, 0.66)	0.721 (0.671, 0.769)	0.107 (0.04, 0.168)
phecode_390-6	Perichondritis and chondritis of pinna	0.654 (0.644, 0.662)	0.727 (0.718, 0.735)	0.073 (0.064, 0.081)
phecode_391	Disorders of the middle ear	0.526 (0.523, 0.529)	0.684 (0.682, 0.686)	0.158 (0.155, 0.161)
phecode_391-1	Otitis media	0.542 (0.538, 0.545)	0.704 (0.702, 0.707)	0.162 (0.158, 0.166)
phecode_391-11	Acute otitis media	0.529 (0.517, 0.54)	0.739 (0.731, 0.747)	0.21 (0.197, 0.224)
phecode_391-12	Chronic otitis media	0.503 (0.495, 0.511)	0.716 (0.709, 0.724)	0.214 (0.203, 0.224)
phecode_391-2	Eustachian tube disorders	0.558 (0.554, 0.561)	0.699 (0.695, 0.702)	0.141 (0.137, 0.145)
phecode_391-21	Eustachian salpingitis	0.558 (0.547, 0.568)	0.714 (0.705, 0.723)	0.156 (0.144, 0.169)
phecode_391-22	Obstruction of Eustachian tube	0.521 (0.502, 0.539)	0.714 (0.697, 0.73)	0.193 (0.173, 0.214)
phecode_391-4	Tympanosclerosis	0.557 (0.527, 0.592)	0.745 (0.715, 0.778)	0.187 (0.144, 0.23)
phecode_391-6	Cholesteatoma of middle ear	0.545 (0.524, 0.565)	0.73 (0.711, 0.749)	0.186 (0.158, 0.216)
phecode_391-7	Perforation of tympanic membrane	0.533 (0.524, 0.542)	0.703 (0.695, 0.712)	0.17 (0.16, 0.181)
phecode_391-8	Otosclerosis	0.579 (0.557, 0.603)	0.656 (0.635, 0.677)	0.076 (0.053, 0.102)
phecode_391-9	Otorrhea	0.567 (0.558, 0.577)	0.768 (0.761, 0.776)	0.201 (0.189, 0.212)
phecode_392	Otalgia and effusion of ear	0.563 (0.56, 0.566)	0.734 (0.732, 0.736)	0.17 (0.167, 0.174)

Supplementary Tables

Table 17 continued from previous page

Endpoint	PheCode string	C-Index Age+Sex (IQR)	C-Index Age+Sex+MedicalHistory (IQR)	Difference in C-Index (IQR)
phecode_393	Mastoiditis and related conditions	0.519 (0.495, 0.542)	0.724 (0.701, 0.748)	0.206 (0.176, 0.238)
phecode_394	Disorders of vestibular function	0.6 (0.596, 0.603)	0.726 (0.723, 0.729)	0.127 (0.123, 0.13)
phecode_394-1	Meniere disease	0.575 (0.564, 0.586)	0.702 (0.692, 0.713)	0.127 (0.114, 0.14)
phecode_394-2	Vertigo	0.607 (0.603, 0.611)	0.738 (0.735, 0.742)	0.131 (0.127, 0.135)
phecode_394-21	Paroxysmal vertigo	0.613 (0.609, 0.617)	0.757 (0.753, 0.76)	0.144 (0.14, 0.147)
phecode_394-22	Vestibular neuritis	0.58 (0.57, 0.59)	0.708 (0.7, 0.716)	0.128 (0.118, 0.138)
phecode_394-4	Abnormal vestibular function study	0.6 (0.586, 0.613)	0.738 (0.727, 0.749)	0.138 (0.123, 0.153)
phecode_395	Other diseases of inner ear	0.581 (0.578, 0.585)	0.705 (0.702, 0.707)	0.124 (0.12, 0.127)
phecode_395-1	Labyrinthitis	0.583 (0.58, 0.587)	0.706 (0.702, 0.708)	0.122 (0.119, 0.126)
phecode_395-3	Noise effects on inner ear	0.619 (0.588, 0.656)	0.778 (0.752, 0.802)	0.158 (0.129, 0.186)
phecode_396	Hearing impairment	0.628 (0.626, 0.63)	0.714 (0.713, 0.716)	0.086 (0.085, 0.088)
phecode_396-1	Conductive hearing loss	0.536 (0.525, 0.548)	0.726 (0.715, 0.739)	0.19 (0.175, 0.205)
phecode_396-11	Conductive hearing loss, bilateral	0.516 (0.477, 0.554)	0.787 (0.755, 0.82)	0.272 (0.22, 0.316)
phecode_396-2	Sensorineural hearing loss	0.632 (0.628, 0.636)	0.726 (0.722, 0.729)	0.094 (0.09, 0.097)
phecode_396-21	Sensorineural hearing loss, bilateral	0.635 (0.628, 0.642)	0.746 (0.74, 0.752)	0.111 (0.105, 0.117)
phecode_396-22	Presbycusis	0.698 (0.692, 0.705)	0.759 (0.753, 0.765)	0.06 (0.056, 0.064)
phecode_396-3	Mixed conductive and sensorineural hearing loss	0.595 (0.577, 0.613)	0.765 (0.748, 0.781)	0.168 (0.147, 0.189)
phecode_396-5	Sudden idiopathic hearing loss	0.599 (0.557, 0.641)	0.701 (0.664, 0.733)	0.1 (0.052, 0.145)
phecode_397	Other hearing abnormality	0.505 (0.503, 0.509)	0.683 (0.68, 0.686)	0.178 (0.174, 0.182)
phecode_397-1	Tinnitus	0.506 (0.503, 0.509)	0.683 (0.68, 0.686)	0.177 (0.173, 0.181)
phecode_397-3	Hyperacusis	0.506 (0.469, 0.548)	0.725 (0.692, 0.76)	0.22 (0.165, 0.266)
phecode_398	Other disorders of ear	0.567 (0.561, 0.572)	0.707 (0.703, 0.712)	0.141 (0.136, 0.147)
phecode_398-1	Abnormal auditory function study	0.597 (0.586, 0.606)	0.733 (0.724, 0.741)	0.136 (0.127, 0.148)
phecode_400	Rheumatic fever and chronic rheumatic heart diseases	0.712 (0.707, 0.717)	0.769 (0.765, 0.774)	0.057 (0.052, 0.062)
phecode_400-2	Chronic rheumatic heart diseases	0.712 (0.707, 0.717)	0.77 (0.765, 0.774)	0.057 (0.053, 0.062)
phecode_401	Hypertension	0.632 (0.631, 0.634)	0.67 (0.668, 0.671)	0.037 (0.036, 0.038)
phecode_401-1	Essential hypertension	0.632 (0.631, 0.634)	0.67 (0.668, 0.671)	0.037 (0.036, 0.039)
phecode_401-2	Hypertensive heart disease	0.707 (0.693, 0.722)	0.838 (0.828, 0.848)	0.131 (0.118, 0.144)
phecode_401-3	Hypertensive chronic kidney disease	0.71 (0.7, 0.72)	0.899 (0.893, 0.905)	0.189 (0.178, 0.198)
phecode_401-6	Secondary hypertension	0.602 (0.594, 0.61)	0.692 (0.684, 0.698)	0.09 (0.083, 0.096)
phecode_402	Elevated blood pressure reading without diagnosis of hypertension	0.534 (0.531, 0.537)	0.658 (0.656, 0.66)	0.124 (0.121, 0.127)
phecode_403	Angina pectoris	0.676 (0.673, 0.679)	0.745 (0.743, 0.748)	0.07 (0.067, 0.072)
phecode_403-1	Coronary artery spasm [Coronary vasospasm]	0.6 (0.574, 0.628)	0.743 (0.719, 0.766)	0.142 (0.111, 0.176)
phecode_404	Ischemic heart disease	0.695 (0.693, 0.697)	0.742 (0.74, 0.744)	0.047 (0.045, 0.049)
phecode_404-1	Myocardial infarction [Heart attack]	0.707 (0.704, 0.71)	0.763 (0.76, 0.766)	0.056 (0.053, 0.059)
phecode_404-11	Acute myocardial infarction	0.7 (0.697, 0.704)	0.734 (0.731, 0.737)	0.034 (0.031, 0.037)
phecode_404-2	Coronary atherosclerosis [Atherosclerotic heart disease]	0.707 (0.705, 0.71)	0.746 (0.744, 0.748)	0.039 (0.037, 0.041)
phecode_406	Chronic pulmonary heart disease	0.711 (0.703, 0.719)	0.822 (0.815, 0.831)	0.112 (0.102, 0.12)
phecode_406-1	Pulmonary hypertension	0.713 (0.704, 0.721)	0.813 (0.806, 0.821)	0.101 (0.091, 0.109)
phecode_406-11	Primary pulmonary hypertension	0.7 (0.683, 0.717)	0.831 (0.816, 0.847)	0.131 (0.111, 0.15)
phecode_406-13	Secondary pulmonary arterial hypertension*	0.716 (0.679, 0.752)	0.864 (0.839, 0.889)	0.149 (0.11, 0.188)
phecode_408	Diseases of pulmonary vessels	0.583 (0.545, 0.623)	0.685 (0.636, 0.737)	0.104 (0.045, 0.161)
phecode_410	Inflammation of the heart [Carditis]	0.627 (0.62, 0.634)	0.702 (0.695, 0.708)	0.075 (0.068, 0.082)
phecode_410-1	Pericarditis	0.591 (0.574, 0.607)	0.68 (0.664, 0.695)	0.088 (0.071, 0.106)
phecode_410-2	Endocarditis	0.66 (0.652, 0.668)	0.731 (0.724, 0.74)	0.071 (0.063, 0.079)
phecode_410-3	Myocarditis	0.552 (0.53, 0.578)	0.627 (0.605, 0.653)	0.075 (0.044, 0.102)
phecode_411	Other diseases of pericardium	0.625 (0.616, 0.635)	0.704 (0.696, 0.713)	0.078 (0.069, 0.087)
phecode_411-1	Hemopericardium NEC	0.687 (0.659, 0.72)	0.818 (0.796, 0.842)	0.131 (0.105, 0.157)
phecode_411-2	Pericardial effusion (noninflammatory)*	0.635 (0.624, 0.647)	0.721 (0.71, 0.731)	0.085 (0.075, 0.096)
phecode_413	Heart valve disorders	0.701 (0.698, 0.704)	0.737 (0.734, 0.74)	0.036 (0.033, 0.038)
phecode_413-1	Mitral valve disorders	0.694 (0.69, 0.698)	0.736 (0.732, 0.74)	0.042 (0.038, 0.045)
phecode_413-11	Mitral valve insufficiency	0.683 (0.678, 0.688)	0.732 (0.727, 0.737)	0.049 (0.044, 0.054)
phecode_413-12	Mitral valve prolapse*	0.646 (0.633, 0.661)	0.688 (0.673, 0.702)	0.04 (0.028, 0.054)

4 Medical history predicts future health trajectories over the human phenome

Table 17 continued from previous page

Endpoint	PheCode string	C-Index Age+Sex (IQR)	C-Index Age+Sex+MedicalHistory (IQR)	Difference in C-Index (IQR)
phecode_413-13	Mitral valve stenosis	0.736 (0.712, 0.758)	0.824 (0.802, 0.847)	0.089 (0.064, 0.113)
phecode_413-2	Aortic valve disorders	0.723 (0.719, 0.727)	0.755 (0.751, 0.759)	0.032 (0.029, 0.035)
phecode_413-21	Aortic stenosis	0.747 (0.742, 0.751)	0.784 (0.779, 0.789)	0.037 (0.034, 0.041)
phecode_413-22	Aortic insufficiency	0.68 (0.673, 0.687)	0.728 (0.721, 0.734)	0.048 (0.042, 0.054)
phecode_413-3	Tricuspid valve disorders	0.71 (0.704, 0.716)	0.771 (0.765, 0.776)	0.061 (0.055, 0.066)
phecode_413-32	Tricuspid valve insufficiency*	0.686 (0.675, 0.698)	0.777 (0.767, 0.788)	0.091 (0.081, 0.102)
phecode_413-4	Pulmonary valve disorders	0.665 (0.648, 0.683)	0.732 (0.716, 0.749)	0.068 (0.049, 0.086)
phecode_413-42	Pulmonary valve insufficiency*	0.672 (0.655, 0.692)	0.738 (0.719, 0.755)	0.065 (0.047, 0.086)
phecode_413-6	Heart valve replaced	0.725 (0.717, 0.732)	0.829 (0.822, 0.836)	0.104 (0.097, 0.113)
phecode_414	Cardiomyopathy	0.652 (0.643, 0.658)	0.721 (0.713, 0.728)	0.069 (0.061, 0.077)
phecode_414-1	Hypertrophic cardiomyopathy	0.612 (0.594, 0.629)	0.691 (0.671, 0.711)	0.079 (0.061, 0.098)
phecode_414-11	Obstructive hypertrophic cardiomyopathy	0.617 (0.596, 0.64)	0.717 (0.69, 0.747)	0.101 (0.067, 0.129)
phecode_414-2	Dilated cardiomyopathy*	0.668 (0.654, 0.68)	0.733 (0.722, 0.746)	0.065 (0.053, 0.079)
phecode_414-5	Ischemic cardiomyopathy*	0.802 (0.789, 0.814)	0.908 (0.897, 0.918)	0.106 (0.094, 0.118)
phecode_414-9	Takotsubo syndrome [Stress cardiomyopathy]	0.767 (0.745, 0.792)	0.806 (0.781, 0.832)	0.039 (0.017, 0.06)
phecode_416	Cardiac arrhythmia and conduction disorders	0.693 (0.691, 0.695)	0.717 (0.715, 0.719)	0.024 (0.023, 0.025)
phecode_416-1	Paroxysmal tachycardia	0.614 (0.608, 0.62)	0.68 (0.674, 0.687)	0.066 (0.06, 0.072)
phecode_416-11	Supraventricular tachycardia	0.602 (0.595, 0.609)	0.675 (0.668, 0.681)	0.073 (0.065, 0.08)
phecode_416-12	Ventricular tachycardia	0.704 (0.694, 0.715)	0.763 (0.753, 0.774)	0.059 (0.05, 0.067)
phecode_416-2	Atrial fibrillation and flutter	0.73 (0.728, 0.732)	0.753 (0.751, 0.755)	0.024 (0.022, 0.025)
phecode_416-21	Atrial fibrillation	0.724 (0.721, 0.727)	0.777 (0.774, 0.779)	0.053 (0.05, 0.055)
phecode_416-211	Paroxysmal atrial fibrillation*	0.681 (0.676, 0.685)	0.747 (0.742, 0.751)	0.066 (0.062, 0.07)
phecode_416-212	Persistent atrial fibrillation*	0.708 (0.697, 0.72)	0.786 (0.774, 0.798)	0.078 (0.065, 0.09)
phecode_416-213	Chronic atrial fibrillation*	0.774 (0.759, 0.789)	0.859 (0.843, 0.875)	0.085 (0.067, 0.102)
phecode_416-214	Permanent atrial fibrillation*	0.774 (0.742, 0.811)	0.861 (0.832, 0.887)	0.086 (0.049, 0.122)
phecode_416-22	Atrial flutter	0.731 (0.725, 0.737)	0.8 (0.795, 0.807)	0.069 (0.065, 0.074)
phecode_416-221	Typical atrial flutter*	0.719 (0.69, 0.754)	0.771 (0.739, 0.804)	0.05 (0.026, 0.079)
phecode_416-222	Atypical atrial flutter*	0.712 (0.67, 0.748)	0.777 (0.721, 0.831)	0.065 (0, 0.127)
phecode_416-3	Ventricular fibrillation and flutter	0.723 (0.709, 0.738)	0.757 (0.743, 0.772)	0.034 (0.022, 0.046)
phecode_416-31	Ventricular fibrillation	0.708 (0.677, 0.744)	0.785 (0.757, 0.818)	0.079 (0.047, 0.109)
phecode_416-4	Heart block	0.711 (0.708, 0.715)	0.746 (0.743, 0.75)	0.035 (0.032, 0.038)
phecode_416-41	Atrioventricular block	0.743 (0.738, 0.748)	0.779 (0.774, 0.784)	0.036 (0.033, 0.04)
phecode_416-42	Left bundle branch block	0.7 (0.694, 0.705)	0.745 (0.739, 0.75)	0.045 (0.04, 0.05)
phecode_416-43	Right bundle branch block	0.707 (0.701, 0.712)	0.738 (0.732, 0.743)	0.031 (0.027, 0.036)
phecode_416-5	Premature depolarization [Premature beats]	0.603 (0.598, 0.609)	0.7 (0.695, 0.705)	0.097 (0.091, 0.102)
phecode_416-51	Atrial premature depolarization [Supraventricular premature beats]	0.632 (0.62, 0.646)	0.708 (0.698, 0.72)	0.075 (0.063, 0.089)
phecode_416-52	Ventricular premature depolarization*	0.603 (0.595, 0.609)	0.695 (0.687, 0.702)	0.092 (0.085, 0.1)
phecode_416-6	Long QT syndrome	0.557 (0.511, 0.607)	0.652 (0.609, 0.692)	0.095 (0.041, 0.141)
phecode_416-7	Sinoatrial node dysfunction	0.717 (0.705, 0.729)	0.78 (0.768, 0.793)	0.063 (0.052, 0.075)
phecode_416-71	Sick sinus syndrome*	0.716 (0.703, 0.728)	0.78 (0.767, 0.792)	0.063 (0.051, 0.075)
phecode_416-8	Pre-excitation syndrome [Anomalous atrioventricular excitation]	0.535 (0.499, 0.569)	0.588 (0.548, 0.628)	0.053 (0.001, 0.104)
phecode_417	Abnormalities of heart beat	0.585 (0.583, 0.587)	0.68 (0.678, 0.682)	0.095 (0.093, 0.097)

Supplementary Tables

Table 17 continued from previous page

Endpoint	PheCode string	C-Index Age+Sex (IQR)	C-Index Age+Sex+MedicalHistory (IQR)	Difference in C-Index (IQR)
phecode_417-1	Palpitations	0.577 (0.574, 0.579)	0.707 (0.704, 0.709)	0.13 (0.127, 0.133)
phecode_417-2	Tachycardia	0.588 (0.582, 0.593)	0.689 (0.683, 0.694)	0.101 (0.095, 0.107)
phecode_417-3	Bradycardia*	0.666 (0.661, 0.67)	0.716 (0.712, 0.72)	0.05 (0.047, 0.054)
phecode_418	Abnormal results of cardiovascular function studies	0.661 (0.657, 0.664)	0.749 (0.746, 0.752)	0.088 (0.085, 0.091)
phecode_418-1	Abnormal electrocardiogram [ECG] [EKG]	0.669 (0.664, 0.672)	0.76 (0.756, 0.763)	0.091 (0.088, 0.094)
phecode_419	Presence of cardiac device	0.745 (0.74, 0.75)	0.823 (0.819, 0.827)	0.078 (0.073, 0.082)
phecode_419-2	Presence of cardiac defibrillator	0.724 (0.691, 0.762)	0.886 (0.86, 0.912)	0.161 (0.133, 0.189)
phecode_420	Cardiac arrest	0.709 (0.7, 0.717)	0.801 (0.794, 0.809)	0.092 (0.085, 0.1)
phecode_423	Abnormal cardiac sounds	0.631 (0.625, 0.636)	0.711 (0.706, 0.716)	0.08 (0.075, 0.086)
phecode_423-1	Cardiac murmurs	0.639 (0.633, 0.646)	0.707 (0.701, 0.713)	0.067 (0.062, 0.074)
phecode_424	Heart failure	0.724 (0.72, 0.727)	0.797 (0.794, 0.8)	0.073 (0.07, 0.076)
phecode_424-1	Left heart failure	0.723 (0.719, 0.727)	0.793 (0.789, 0.797)	0.07 (0.066, 0.074)
phecode_424-2	Systolic heart failure	0.722 (0.714, 0.73)	0.825 (0.819, 0.832)	0.103 (0.095, 0.11)
phecode_424-3	Diastolic heart failure	0.699 (0.688, 0.711)	0.813 (0.803, 0.822)	0.114 (0.105, 0.125)
phecode_424-5	Right heart failure*	0.736 (0.689, 0.789)	0.864 (0.831, 0.901)	0.128 (0.071, 0.18)
phecode_424-6	Hypertensive heart disease with heart failure	0.682 (0.65, 0.713)	0.845 (0.822, 0.87)	0.163 (0.132, 0.195)
phecode_425	Cardiomegaly	0.68 (0.676, 0.684)	0.761 (0.758, 0.766)	0.081 (0.077, 0.086)
phecode_426	Other heart disorders in diseases NEC	0.708 (0.703, 0.712)	0.782 (0.778, 0.786)	0.074 (0.07, 0.079)
phecode_430	Nontraumatic Intracranial hemorrhage	0.651 (0.644, 0.659)	0.693 (0.686, 0.701)	0.043 (0.036, 0.049)
phecode_430-1	Nontraumatic subarachnoid hemorrhage	0.582 (0.572, 0.593)	0.645 (0.634, 0.656)	0.063 (0.052, 0.073)
phecode_430-2	Nontraumatic intracerebral hemorrhage	0.69 (0.68, 0.7)	0.737 (0.727, 0.748)	0.048 (0.039, 0.058)
phecode_430-3	Nontraumatic subdural hemorrhage	0.712 (0.698, 0.726)	0.746 (0.733, 0.76)	0.034 (0.022, 0.045)
phecode_431	Stroke and transient cerebral ischemic attacks	0.681 (0.678, 0.684)	0.708 (0.705, 0.711)	0.027 (0.025, 0.029)
phecode_431-1	Stroke	0.682 (0.678, 0.686)	0.713 (0.709, 0.717)	0.031 (0.027, 0.034)
phecode_431-11	Cerebral infarction [Ischemic stroke]	0.696 (0.692, 0.701)	0.73 (0.726, 0.734)	0.034 (0.03, 0.037)
phecode_431-12	Hemorrhagic stroke	0.634 (0.626, 0.642)	0.679 (0.671, 0.688)	0.045 (0.037, 0.053)
phecode_431-14	Cerebellar stroke syndrome*	0.69 (0.671, 0.714)	0.745 (0.726, 0.765)	0.054 (0.035, 0.073)
phecode_431-15	Lacunar syndrome*	0.688 (0.671, 0.706)	0.819 (0.807, 0.834)	0.132 (0.115, 0.149)
phecode_431-2	Transient cerebral ischemic attacks	0.682 (0.678, 0.686)	0.738 (0.734, 0.742)	0.056 (0.053, 0.06)
phecode_433	Other cerebrovascular disease	0.718 (0.715, 0.722)	0.763 (0.759, 0.767)	0.045 (0.041, 0.048)
phecode_433-1	Occlusion and stenosis of cerebral arteries	0.701 (0.692, 0.709)	0.738 (0.729, 0.747)	0.038 (0.03, 0.046)
phecode_433-2	Occlusion and stenosis of precerebral arteries	0.735 (0.727, 0.742)	0.781 (0.773, 0.789)	0.046 (0.038, 0.054)
phecode_433-21	Carotid artery stenosis	0.738 (0.73, 0.746)	0.785 (0.777, 0.793)	0.047 (0.039, 0.054)
phecode_433-3	Cerebrovascular insufficiency	0.709 (0.686, 0.731)	0.773 (0.755, 0.795)	0.064 (0.039, 0.088)
phecode_433-32	Vertebro-basilar artery syndrome	0.717 (0.687, 0.752)	0.775 (0.745, 0.807)	0.058 (0.029, 0.088)
phecode_433-34	Cerebral artery syndrome (middle, anterior, posterior)*	0.694 (0.66, 0.731)	0.778 (0.749, 0.81)	0.082 (0.044, 0.122)
phecode_436	Atherosclerosis [ASCVD]	0.706 (0.703, 0.709)	0.743 (0.74, 0.745)	0.037 (0.035, 0.039)
phecode_436-1	Atherosclerosis of renal artery	0.664 (0.62, 0.715)	0.801 (0.758, 0.844)	0.137 (0.084, 0.193)
phecode_436-2	Atherosclerosis of the extremities	0.742 (0.723, 0.759)	0.874 (0.86, 0.888)	0.133 (0.114, 0.152)
phecode_436-3	Atherosclerosis of aorta	0.742 (0.716, 0.766)	0.849 (0.825, 0.871)	0.108 (0.081, 0.131)
phecode_436-5	Cerebral atherosclerosis	0.729 (0.714, 0.745)	0.768 (0.753, 0.783)	0.039 (0.023, 0.054)
phecode_437	Vascular insufficiency of intestine	0.667 (0.656, 0.678)	0.756 (0.746, 0.766)	0.089 (0.078, 0.101)
phecode_438	Aneurysm or ectasia	0.724 (0.719, 0.728)	0.751 (0.745, 0.755)	0.027 (0.023, 0.03)
phecode_438-1	Aortic aneurysm and ectasia	0.783 (0.778, 0.788)	0.806 (0.802, 0.812)	0.024 (0.02, 0.028)
phecode_438-11	Abdominal aortic aneurysm	0.818 (0.813, 0.822)	0.848 (0.844, 0.853)	0.031 (0.027, 0.035)
phecode_438-12	Thoracic aneurysm	0.699 (0.688, 0.71)	0.751 (0.74, 0.763)	0.053 (0.041, 0.064)
phecode_438-14	Ruptured aortic aneurysm	0.848 (0.834, 0.864)	0.875 (0.861, 0.892)	0.027 (0.016, 0.04)
phecode_438-2	Aneurysm of iliac or artery of lower extremity	0.783 (0.769, 0.798)	0.814 (0.8, 0.83)	0.031 (0.02, 0.042)
phecode_438-3	Aneurysm of carotid artery	0.477 (0.444, 0.513)	0.615 (0.581, 0.651)	0.139 (0.097, 0.182)
phecode_438-4	Arterial dissection	0.639 (0.623, 0.657)	0.695 (0.676, 0.714)	0.055 (0.037, 0.074)
phecode_438-41	Aortic dissection	0.683 (0.661, 0.707)	0.762 (0.742, 0.783)	0.078 (0.056, 0.102)
phecode_438-43	Coronary artery dissection	0.619 (0.585, 0.655)	0.68 (0.649, 0.716)	0.061 (0.024, 0.1)

4 Medical history predicts future health trajectories over the human phenome

Table 17 continued from previous page

Endpoint	PheCode string	C-Index Age+Sex (IQR)	C-Index Age+Sex+MedicalHistory (IQR)	Difference in C-Index (IQR)
phecode_438-6	Cerebral aneurysm	0.603 (0.584, 0.619)	0.688 (0.669, 0.706)	0.085 (0.065, 0.105)
phecode_438-7	Aneurysm of heart	0.669 (0.639, 0.699)	0.737 (0.699, 0.772)	0.067 (0.032, 0.103)
phecode_439	Hemorrhoids	0.502 (0.499, 0.504)	0.678 (0.676, 0.68)	0.177 (0.174, 0.179)
phecode_440	Embolism and thrombosis	0.603 (0.6, 0.606)	0.648 (0.646, 0.651)	0.045 (0.043, 0.048)
phecode_440-1	Venous thromboembolism	0.574 (0.57, 0.578)	0.648 (0.644, 0.651)	0.074 (0.07, 0.077)
phecode_440-11	Deep vein thrombosis [DVT]	0.605 (0.599, 0.611)	0.721 (0.716, 0.726)	0.116 (0.11, 0.121)
phecode_440-12	Portal vein thrombosis	0.639 (0.616, 0.663)	0.748 (0.727, 0.77)	0.108 (0.084, 0.134)
phecode_440-13	Phlebitis and thrombophlebitis	0.597 (0.593, 0.6)	0.663 (0.659, 0.667)	0.066 (0.063, 0.07)
phecode_440-2	Arterial embolism and thrombosis	0.697 (0.691, 0.704)	0.776 (0.77, 0.782)	0.079 (0.073, 0.085)
phecode_440-21	Embolism of cerebral or precerebral arteries	0.726 (0.706, 0.749)	0.756 (0.735, 0.777)	0.03 (0.01, 0.049)
phecode_440-22	Thrombosis of cerebral or precerebral arteries	0.694 (0.677, 0.711)	0.709 (0.693, 0.727)	0.016 (0, 0.031)
phecode_440-3	Pulmonary embolism	0.636 (0.631, 0.641)	0.698 (0.693, 0.703)	0.062 (0.057, 0.067)
phecode_441	Arteriovenous malformation	0.537 (0.512, 0.562)	0.615 (0.588, 0.639)	0.077 (0.045, 0.108)
phecode_441-1	Arteriovenous fistula	0.608 (0.573, 0.649)	0.307 (0.262, 0.343)	-0.302 (-0.367, -0.242)
phecode_442	Disease of capillaries	0.601 (0.577, 0.625)	0.674 (0.651, 0.698)	0.074 (0.051, 0.095)
phecode_442-1	Hereditary hemorrhagic telangiectasia	0.595 (0.55, 0.637)	0.688 (0.643, 0.733)	0.091 (0.053, 0.143)
phecode_443	Other specified disorders of arteries and arterioles	0.696 (0.691, 0.701)	0.831 (0.827, 0.835)	0.135 (0.13, 0.14)
phecode_443-1	Stricture of artery [Arterial stenosis]	0.717 (0.708, 0.727)	0.845 (0.836, 0.853)	0.127 (0.117, 0.137)
phecode_444	Venous insufficiency	0.586 (0.583, 0.589)	0.662 (0.659, 0.665)	0.076 (0.072, 0.079)
phecode_444-1	Varicose veins	0.584 (0.581, 0.587)	0.66 (0.657, 0.663)	0.076 (0.073, 0.08)
phecode_444-11	Varicose veins of lower extremities	0.593 (0.59, 0.596)	0.671 (0.668, 0.674)	0.078 (0.075, 0.081)
phecode_444-12	Gastric varices*	0.592 (0.561, 0.624)	0.79 (0.755, 0.826)	0.198 (0.152, 0.246)
phecode_444-13	Esophageal varix	0.638 (0.624, 0.652)	0.805 (0.791, 0.818)	0.167 (0.151, 0.183)
phecode_444-15	Scrotal varices [Varicocele]	0.797 (0.791, 0.803)	0.857 (0.852, 0.863)	0.061 (0.054, 0.068)
phecode_444-5	Venous insufficiency (chronic) (peripheral)	0.668 (0.664, 0.673)	0.772 (0.768, 0.777)	0.104 (0.1, 0.108)
phecode_446	Hypotension	0.669 (0.665, 0.672)	0.736 (0.733, 0.739)	0.067 (0.064, 0.07)
phecode_446-2	Orthostatic hypotension	0.696 (0.691, 0.7)	0.764 (0.76, 0.768)	0.069 (0.064, 0.072)
phecode_447	Nonspecific low blood-pressure reading	0.634 (0.627, 0.642)	0.703 (0.696, 0.71)	0.069 (0.062, 0.076)
phecode_448	Peripheral vascular disease	0.619 (0.615, 0.623)	0.71 (0.706, 0.715)	0.091 (0.087, 0.096)
phecode_448-1	Raynaud's syndrome	0.559 (0.553, 0.566)	0.669 (0.663, 0.675)	0.109 (0.103, 0.117)
phecode_448-9	Peripheral vascular disease NOS [includes PAD]	0.717 (0.712, 0.722)	0.824 (0.82, 0.829)	0.108 (0.103, 0.113)
phecode_449	Other disorders of the circulatory system	0.622 (0.611, 0.635)	0.723 (0.711, 0.734)	0.1 (0.089, 0.112)
phecode_452	Hemorrhage, NOS	0.601 (0.59, 0.611)	0.743 (0.733, 0.75)	0.141 (0.13, 0.154)
phecode_460	Acute respiratory infection	0.53 (0.529, 0.532)	0.702 (0.701, 0.703)	0.171 (0.17, 0.173)
phecode_460-1	Acute upper respiratory infection	0.537 (0.535, 0.539)	0.712 (0.711, 0.713)	0.175 (0.173, 0.177)
phecode_460-2	Acute lower respiratory infection	0.563 (0.562, 0.564)	0.698 (0.697, 0.7)	0.135 (0.134, 0.137)
phecode_462	Sinusitis	0.554 (0.552, 0.557)	0.712 (0.71, 0.714)	0.157 (0.155, 0.16)
phecode_462-1	Acute sinusitis	0.581 (0.578, 0.584)	0.741 (0.738, 0.743)	0.16 (0.157, 0.163)
phecode_462-2	Chronic sinusitis	0.548 (0.546, 0.551)	0.717 (0.715, 0.72)	0.169 (0.166, 0.172)
phecode_463	Rhinitis and nasal congestion	0.525 (0.522, 0.527)	0.712 (0.71, 0.714)	0.187 (0.185, 0.19)
phecode_463-1	Chronic rhinitis	0.555 (0.551, 0.559)	0.722 (0.719, 0.725)	0.167 (0.163, 0.171)
phecode_463-2	Allergic rhinitis	0.511 (0.508, 0.514)	0.717 (0.714, 0.72)	0.206 (0.202, 0.209)
phecode_463-21	Seasonal allergic rhinitis	0.526 (0.522, 0.53)	0.723 (0.72, 0.727)	0.197 (0.192, 0.202)
phecode_463-23	Allergic rhinitis, due to animal hair and dander	0.641 (0.624, 0.658)	0.799 (0.785, 0.811)	0.157 (0.139, 0.173)
phecode_463-3	Vasomotor rhinitis*	0.586 (0.566, 0.602)	0.711 (0.697, 0.727)	0.126 (0.106, 0.145)
phecode_463-4	Nasal congestion*	0.523 (0.519, 0.527)	0.73 (0.727, 0.733)	0.207 (0.201, 0.211)
phecode_463-5	Postnasal drip	0.585 (0.579, 0.589)	0.75 (0.745, 0.753)	0.165 (0.16, 0.17)
phecode_464	Nasopharyngitis	0.543 (0.538, 0.547)	0.719 (0.715, 0.722)	0.176 (0.171, 0.181)
phecode_464-1	Acute nasopharyngitis	0.543 (0.538, 0.547)	0.719 (0.715, 0.722)	0.176 (0.171, 0.181)
phecode_465	Pharyngitis	0.559 (0.557, 0.562)	0.717 (0.715, 0.719)	0.157 (0.155, 0.16)
phecode_465-1	Acute Pharyngitis	0.56 (0.558, 0.562)	0.718 (0.716, 0.719)	0.157 (0.155, 0.16)
phecode_465-11	Streptococcal tonsillitis/pharyngitis	0.577 (0.55, 0.603)	0.713 (0.691, 0.736)	0.136 (0.11, 0.162)

Supplementary Tables

Table 17 continued from previous page

Endpoint	PheCode string	C-Index Age+Sex (IQR)	C-Index Age+Sex+MedicalHistory (IQR)	Difference in C-Index (IQR)
phecode_465-2	Chronic Pharyngitis	0.562 (0.559, 0.565)	0.73 (0.728, 0.733)	0.168 (0.165, 0.172)
phecode_466	Tonsillitis and adenoiditis	0.646 (0.642, 0.651)	0.718 (0.714, 0.721)	0.071 (0.068, 0.074)
phecode_466-1	Acute tonsillitis and adenoiditis	0.66 (0.655, 0.665)	0.732 (0.728, 0.736)	0.072 (0.069, 0.076)
phecode_466-3	Peritonsillar abscess	0.628 (0.61, 0.649)	0.649 (0.63, 0.668)	0.02 (0.002, 0.039)
phecode_466-4	Hypertrophy of tonsils and adenoids	0.629 (0.613, 0.646)	0.723 (0.711, 0.738)	0.095 (0.079, 0.11)
phecode_467	Laryngitis and tracheitis	0.618 (0.613, 0.623)	0.739 (0.735, 0.743)	0.121 (0.117, 0.125)
phecode_467-1	Acute laryngitis and tracheitis	0.625 (0.62, 0.629)	0.744 (0.741, 0.748)	0.12 (0.116, 0.124)
phecode_467-2	Chronic laryngitis and tracheitis	0.516 (0.487, 0.541)	0.729 (0.708, 0.752)	0.213 (0.185, 0.245)
phecode_467-4	Acute epiglottitis	0.619 (0.574, 0.663)	0.59 (0.541, 0.642)	-0.027 (-0.076, 0.015)
phecode_468	Pneumonia	0.662 (0.659, 0.665)	0.743 (0.741, 0.746)	0.081 (0.079, 0.084)
phecode_468-1	Viral pneumonia	0.592 (0.579, 0.603)	0.696 (0.683, 0.708)	0.104 (0.091, 0.116)
phecode_468-14	Influenza with pneumonia	0.608 (0.594, 0.623)	0.707 (0.691, 0.719)	0.099 (0.083, 0.112)
phecode_468-2	Bacterial pneumonia	0.61 (0.601, 0.619)	0.706 (0.698, 0.716)	0.096 (0.085, 0.106)
phecode_468-21	Pneumococcal pneumonia	0.588 (0.57, 0.604)	0.694 (0.675, 0.709)	0.107 (0.086, 0.125)
phecode_468-22	Pneumonia due to Klebsiella pneumoniae	0.732 (0.699, 0.765)	0.842 (0.822, 0.864)	0.112 (0.077, 0.143)
phecode_468-23	Pneumonia due to Pseudomonas	0.681 (0.656, 0.708)	0.864 (0.845, 0.884)	0.183 (0.152, 0.212)
phecode_468-24	Pneumonia due to Hemophilus influenzae	0.649 (0.62, 0.677)	0.785 (0.756, 0.816)	0.137 (0.106, 0.17)
phecode_468-25	Pneumonia due to Streptococcus	0.593 (0.555, 0.632)	0.667 (0.626, 0.712)	0.073 (0.027, 0.128)
phecode_468-26	Pneumonia due to Staphylococcus	0.644 (0.602, 0.688)	0.781 (0.748, 0.822)	0.136 (0.091, 0.182)
phecode_468-29	Tuberculosis pneumonia	0.541 (0.506, 0.582)	0.72 (0.68, 0.76)	0.178 (0.124, 0.228)
phecode_468-5	Fungal pneumonia	0.591 (0.563, 0.618)	0.861 (0.839, 0.884)	0.269 (0.235, 0.306)
phecode_468-51	Pulmonary aspergillosis	0.592 (0.562, 0.618)	0.861 (0.839, 0.885)	0.268 (0.238, 0.306)
phecode_468-8	Bronchopneumonia	0.711 (0.701, 0.721)	0.837 (0.83, 0.846)	0.127 (0.119, 0.135)
phecode_468-9	Lobar pneumonia*	0.668 (0.665, 0.672)	0.756 (0.752, 0.759)	0.088 (0.084, 0.091)
phecode_469	Bronchitis	0.544 (0.54, 0.547)	0.789 (0.787, 0.792)	0.246 (0.242, 0.25)
phecode_469-1	Acute bronchitis	0.564 (0.555, 0.573)	0.732 (0.725, 0.741)	0.168 (0.159, 0.178)
phecode_469-2	Chronic bronchitis	0.644 (0.637, 0.652)	0.867 (0.86, 0.874)	0.222 (0.213, 0.23)
phecode_470	Acute bronchiolitis	0.535 (0.507, 0.562)	0.682 (0.648, 0.714)	0.147 (0.105, 0.19)
phecode_471	Other disorders of nose and nasal sinuses	0.506 (0.504, 0.509)	0.696 (0.694, 0.699)	0.19 (0.187, 0.193)
phecode_471-1	Cyst and mucocele of nose and nasal sinus*	0.512 (0.47, 0.55)	0.653 (0.617, 0.688)	0.143 (0.091, 0.194)
phecode_471-2	Deviated nasal septum	0.591 (0.584, 0.599)	0.691 (0.684, 0.699)	0.099 (0.091, 0.108)
phecode_471-3	Hypertrophy of nasal turbinates	0.603 (0.591, 0.617)	0.711 (0.697, 0.724)	0.108 (0.093, 0.123)
phecode_471-5	Nasal polyps	0.578 (0.572, 0.586)	0.695 (0.688, 0.702)	0.116 (0.108, 0.125)
phecode_472	Diseases of vocal cords and larynx, not elsewhere classified	0.56 (0.55, 0.57)	0.672 (0.663, 0.681)	0.111 (0.099, 0.122)
phecode_472-1	Paralysis of vocal cords and larynx	0.599 (0.579, 0.615)	0.673 (0.654, 0.692)	0.075 (0.052, 0.1)
phecode_472-2	Polyp of vocal cord and larynx	0.563 (0.535, 0.591)	0.678 (0.653, 0.705)	0.116 (0.077, 0.151)
phecode_472-3	Edema of larynx	0.506 (0.481, 0.535)	0.749 (0.718, 0.773)	0.242 (0.209, 0.275)
phecode_472-4	Laryngeal spasm	0.597 (0.565, 0.629)	0.654 (0.624, 0.686)	0.059 (0.023, 0.094)
phecode_473	Other diseases of upper respiratory tract	0.547 (0.542, 0.552)	0.708 (0.704, 0.712)	0.161 (0.154, 0.166)
phecode_474	Chronic obstructive pulmonary disease [COPD]	0.662 (0.66, 0.666)	0.818 (0.815, 0.82)	0.155 (0.152, 0.158)
phecode_474-1	Emphysema	0.686 (0.68, 0.692)	0.847 (0.842, 0.852)	0.161 (0.154, 0.167)
phecode_474-13	Centrilobular emphysema*	0.648 (0.625, 0.672)	0.877 (0.861, 0.893)	0.23 (0.197, 0.256)
phecode_475	Asthma	0.531 (0.528, 0.534)	0.667 (0.664, 0.67)	0.136 (0.132, 0.14)
phecode_475-5	Exercise induced bronchospasm	0.588 (0.573, 0.601)	0.86 (0.85, 0.87)	0.272 (0.256, 0.291)
phecode_476	Bronchiectasis	0.673 (0.669, 0.678)	0.773 (0.768, 0.778)	0.1 (0.095, 0.105)
phecode_477	Inhalation lung injury	0.745 (0.732, 0.758)	0.796 (0.785, 0.808)	0.051 (0.042, 0.06)
phecode_477-1	Pneumoconiosis	0.857 (0.846, 0.868)	0.883 (0.871, 0.892)	0.025 (0.018, 0.033)
phecode_477-2	Hypersensitivity pneumonitis	0.605 (0.576, 0.632)	0.716 (0.692, 0.743)	0.112 (0.081, 0.148)
phecode_478	Aspiration pneumonia	0.724 (0.716, 0.732)	0.834 (0.828, 0.84)	0.11 (0.102, 0.117)
phecode_479	Pulmonary insufficiency and acute respiratory distress syndrome	0.658 (0.653, 0.663)	0.755 (0.751, 0.76)	0.098 (0.093, 0.102)

4 Medical history predicts future health trajectories over the human phenome

Table 17 continued from previous page

Endpoint	PheCode string	C-Index Age+Sex (IQR)	C-Index Age+Sex+MedicalHistory (IQR)	Difference in C-Index (IQR)
phecode_479-1	Acute respiratory distress syndrome [ARDS]*	0.652 (0.627, 0.677)	0.786 (0.767, 0.807)	0.134 (0.106, 0.163)
phecode_479-3	Respiratory failure	0.682 (0.675, 0.689)	0.833 (0.827, 0.838)	0.151 (0.143, 0.159)
phecode_479-31	Acute respiratory failure	0.686 (0.669, 0.703)	0.83 (0.816, 0.845)	0.143 (0.125, 0.162)
phecode_479-6	Pulmonary collapse [Atelectasis]	0.644 (0.638, 0.649)	0.718 (0.713, 0.723)	0.074 (0.069, 0.079)
phecode_480	Pulmonary edema	0.689 (0.677, 0.699)	0.806 (0.797, 0.817)	0.118 (0.106, 0.131)
phecode_480-1	Acute pulmonary edema	0.758 (0.716, 0.798)	0.867 (0.833, 0.906)	0.108 (0.071, 0.153)
phecode_480-2	Chronic pulmonary edema	0.693 (0.671, 0.717)	0.846 (0.83, 0.861)	0.152 (0.129, 0.176)
phecode_481	Interstitial pulmonary diseases	0.721 (0.714, 0.727)	0.781 (0.775, 0.788)	0.061 (0.055, 0.067)
phecode_481-1	Pulmonary eosinophilia	0.413 (0.374, 0.453)	0.724 (0.683, 0.767)	0.311 (0.252, 0.367)
phecode_481-4	Pulmonary fibrosis	0.736 (0.726, 0.747)	0.826 (0.817, 0.835)	0.089 (0.08, 0.099)
phecode_481-42	Idiopathic pulmonary fibrosis	0.77 (0.746, 0.797)	0.83 (0.809, 0.849)	0.059 (0.04, 0.076)
phecode_482	Suppurative and necrotic conditions of the lower respiratory tract	0.636 (0.62, 0.653)	0.729 (0.712, 0.747)	0.094 (0.075, 0.111)
phecode_482-1	Abscess of lung and mediastinum	0.627 (0.591, 0.663)	0.761 (0.726, 0.796)	0.133 (0.089, 0.173)
phecode_482-11	Abscess of lung	0.64 (0.609, 0.675)	0.78 (0.748, 0.818)	0.14 (0.092, 0.18)
phecode_482-2	Pyothorax (empyema)	0.63 (0.614, 0.647)	0.719 (0.701, 0.736)	0.089 (0.072, 0.108)
phecode_483	Pleural effusion	0.676 (0.672, 0.68)	0.745 (0.742, 0.749)	0.069 (0.066, 0.072)
phecode_484	Pneumothorax and air leak	0.648 (0.637, 0.66)	0.694 (0.683, 0.704)	0.046 (0.036, 0.055)
phecode_484-1	Spontaneous pneumothorax	0.61 (0.581, 0.636)	0.735 (0.708, 0.761)	0.124 (0.095, 0.154)
phecode_484-4	Interstitial emphysema [Pneumomediastinum]	0.691 (0.639, 0.746)	0.73 (0.686, 0.771)	0.039 (-0.011, 0.085)
phecode_485	Chronic respiratory failure	0.65 (0.629, 0.67)	0.882 (0.869, 0.895)	0.231 (0.209, 0.256)
phecode_486	Other respiratory disorders	0.54 (0.538, 0.542)	0.76 (0.758, 0.761)	0.22 (0.217, 0.222)
phecode_486-2	Other diseases of bronchus	0.525 (0.514, 0.537)	0.765 (0.754, 0.775)	0.24 (0.226, 0.256)
phecode_486-21	Bronchospasm	0.573 (0.56, 0.586)	0.821 (0.811, 0.83)	0.248 (0.233, 0.263)
phecode_486-3	Other diseases of mediastinum	0.583 (0.547, 0.618)	0.665 (0.621, 0.702)	0.081 (0.041, 0.125)
phecode_486-4	Disorders of diaphragm	0.743 (0.73, 0.756)	0.77 (0.758, 0.786)	0.028 (0.016, 0.039)
phecode_486-41	Hiccough	0.781 (0.768, 0.796)	0.793 (0.779, 0.808)	0.012 (0.002, 0.022)
phecode_486-5	Abnormal sputum	0.554 (0.551, 0.557)	0.759 (0.756, 0.762)	0.205 (0.202, 0.208)
phecode_487	Hemorrhage from respiratory passages	0.595 (0.592, 0.599)	0.689 (0.686, 0.693)	0.094 (0.091, 0.098)
phecode_487-1	Epistaxis	0.598 (0.594, 0.602)	0.686 (0.682, 0.689)	0.087 (0.083, 0.091)
phecode_487-3	Hemoptysis	0.594 (0.588, 0.6)	0.727 (0.722, 0.733)	0.134 (0.127, 0.14)
phecode_488	Abnormalities of breathing	0.567 (0.566, 0.569)	0.731 (0.729, 0.732)	0.163 (0.161, 0.165)
phecode_488-1	Dyspnea [Shortness of breath]	0.594 (0.592, 0.596)	0.75 (0.748, 0.751)	0.156 (0.153, 0.157)
phecode_488-11	Orthopnea	0.683 (0.664, 0.702)	0.868 (0.856, 0.881)	0.186 (0.167, 0.205)
phecode_488-2	Hyperventilation	0.541 (0.523, 0.564)	0.721 (0.702, 0.74)	0.18 (0.154, 0.203)
phecode_488-5	Stridor	0.577 (0.553, 0.604)	0.678 (0.651, 0.705)	0.1 (0.071, 0.13)
phecode_488-6	Wheezing	0.526 (0.522, 0.53)	0.791 (0.788, 0.794)	0.265 (0.26, 0.27)
phecode_488-8	Mouth breathing*	0.598 (0.593, 0.603)	0.709 (0.705, 0.714)	0.111 (0.106, 0.116)
phecode_488-9	Obesity hypoventilation syndrome [OHS]	0.511 (0.48, 0.543)	0.916 (0.903, 0.931)	0.405 (0.372, 0.44)
phecode_491	Pleurisy	0.665 (0.658, 0.672)	0.715 (0.709, 0.722)	0.05 (0.045, 0.055)
phecode_492	Tracheoesophageal fistula	0.652 (0.606, 0.693)	0.808 (0.762, 0.85)	0.157 (0.095, 0.209)
phecode_494	Voice disturbance	0.589 (0.584, 0.593)	0.731 (0.727, 0.735)	0.142 (0.138, 0.147)
phecode_495	Abnormal findings on diagnostic imaging of lung	0.642 (0.638, 0.646)	0.71 (0.706, 0.715)	0.068 (0.065, 0.073)
phecode_495-1	Solitary pulmonary nodule	0.624 (0.612, 0.636)	0.764 (0.755, 0.775)	0.141 (0.127, 0.151)
phecode_496	Abnormal results of pulmonary function studies	0.564 (0.559, 0.568)	0.796 (0.793, 0.8)	0.232 (0.228, 0.237)
phecode_497	Respiratory arrest	0.663 (0.64, 0.691)	0.806 (0.785, 0.83)	0.144 (0.116, 0.17)
phecode_498	Asphyxia and hypoxemia	0.622 (0.608, 0.639)	0.749 (0.735, 0.762)	0.127 (0.109, 0.143)
phecode_498-2	Cyanosis	0.576 (0.546, 0.604)	0.705 (0.681, 0.732)	0.129 (0.099, 0.167)
phecode_499	Other symptoms and disorders of the respiratory system	0.584 (0.562, 0.606)	0.831 (0.81, 0.85)	0.247 (0.219, 0.275)
phecode_500	Disorders of tooth development	0.579 (0.567, 0.594)	0.602 (0.59, 0.615)	0.023 (0.01, 0.036)
phecode_500-4	Disturbances in tooth eruption	0.582 (0.568, 0.598)	0.603 (0.589, 0.617)	0.021 (0.008, 0.034)
phecode_500-41	Impacted teeth*	0.594 (0.576, 0.607)	0.614 (0.6, 0.627)	0.02 (0.007, 0.033)
phecode_501	Dental caries	0.527 (0.519, 0.533)	0.668 (0.661, 0.674)	0.141 (0.132, 0.15)
phecode_502	Other diseases of teeth and supporting structures	0.515 (0.509, 0.52)	0.693 (0.689, 0.697)	0.179 (0.172, 0.185)

Supplementary Tables

Table 17 continued from previous page

Endpoint	PheCode string	C-Index Age+Sex (IQR)	C-Index Age+Sex+MedicalHistory (IQR)	Difference in C-Index (IQR)
phecode_502-6	Loss of teeth and edentulism	0.551 (0.515, 0.585)	0.666 (0.634, 0.7)	0.116 (0.079, 0.155)
phecode_503	Diseases of pulp and periapical tissues	0.523 (0.518, 0.529)	0.69 (0.686, 0.695)	0.167 (0.16, 0.174)
phecode_503-5	Periapical abscess	0.523 (0.518, 0.529)	0.693 (0.688, 0.698)	0.169 (0.162, 0.176)
phecode_504	Periodontal diseases	0.507 (0.499, 0.514)	0.66 (0.653, 0.668)	0.153 (0.143, 0.163)
phecode_504-1	Gingivitis	0.528 (0.515, 0.54)	0.709 (0.698, 0.721)	0.181 (0.165, 0.196)
phecode_504-3	Periodontitis	0.573 (0.56, 0.586)	0.635 (0.622, 0.648)	0.063 (0.046, 0.078)
phecode_504-31	Aggressive periodontitis	0.524 (0.484, 0.565)	0.612 (0.573, 0.651)	0.088 (0.034, 0.141)
phecode_504-32	Chronic periodontitis	0.581 (0.567, 0.595)	0.636 (0.621, 0.648)	0.055 (0.039, 0.071)
phecode_504-4	Gingival enlargement*	0.556 (0.521, 0.597)	0.735 (0.708, 0.764)	0.179 (0.127, 0.223)
phecode_505	Oral cysts	0.5 (0.485, 0.517)	0.61 (0.595, 0.626)	0.109 (0.09, 0.132)
phecode_506	Diseases of salivary glands	0.585 (0.58, 0.59)	0.733 (0.729, 0.738)	0.148 (0.143, 0.153)
phecode_506-3	Sialoadenitis	0.549 (0.537, 0.559)	0.705 (0.695, 0.715)	0.157 (0.144, 0.169)
phecode_506-4	Sialolithiasis	0.532 (0.516, 0.548)	0.66 (0.646, 0.676)	0.128 (0.11, 0.149)
phecode_506-5	Disturbances of salivary secretion	0.622 (0.616, 0.629)	0.786 (0.78, 0.791)	0.163 (0.157, 0.17)
phecode_506-6	Mucocele of salivary gland	0.459 (0.426, 0.493)	0.633 (0.605, 0.66)	0.175 (0.135, 0.214)
phecode_507	Lesions of mouth	0.567 (0.563, 0.571)	0.704 (0.7, 0.708)	0.137 (0.133, 0.141)
phecode_507-1	Stomatitis	0.565 (0.56, 0.57)	0.722 (0.718, 0.727)	0.158 (0.152, 0.163)
phecode_507-11	Recurrent oral aphthae [Recurrent aphthous stomatitis]	0.555 (0.544, 0.565)	0.748 (0.739, 0.756)	0.193 (0.18, 0.204)
phecode_507-2	Cellulitis and abscess of mouth	0.56 (0.538, 0.582)	0.638 (0.617, 0.66)	0.077 (0.053, 0.103)
phecode_507-3	Oral leukoplakia	0.54 (0.523, 0.557)	0.675 (0.659, 0.691)	0.134 (0.115, 0.155)
phecode_508	Diseases of lips	0.584 (0.578, 0.59)	0.715 (0.71, 0.721)	0.132 (0.125, 0.138)
phecode_509	Diseases of tongue	0.575 (0.57, 0.579)	0.719 (0.715, 0.723)	0.145 (0.139, 0.15)
phecode_509-1	Glossitis	0.608 (0.6, 0.616)	0.744 (0.737, 0.751)	0.136 (0.127, 0.144)
phecode_509-11	Glossodynia	0.63 (0.618, 0.641)	0.785 (0.776, 0.793)	0.155 (0.145, 0.165)
phecode_509-2	Geographic tongue	0.542 (0.522, 0.565)	0.705 (0.691, 0.72)	0.163 (0.14, 0.183)
phecode_509-3	Hypertrophy of tongue papillae	0.568 (0.552, 0.583)	0.76 (0.747, 0.773)	0.193 (0.172, 0.208)
phecode_510	Diseases of esophagus	0.565 (0.562, 0.567)	0.669 (0.667, 0.672)	0.105 (0.102, 0.107)
phecode_510-1	Achalasia of cardia	0.564 (0.537, 0.595)	0.679 (0.647, 0.709)	0.115 (0.078, 0.152)
phecode_510-2	Esophagitis	0.558 (0.556, 0.561)	0.665 (0.663, 0.668)	0.107 (0.104, 0.11)
phecode_510-5	Dyskinesia of esophagus	0.592 (0.581, 0.604)	0.745 (0.735, 0.755)	0.153 (0.139, 0.165)
phecode_510-6	Diverticulum of esophagus	0.68 (0.661, 0.699)	0.728 (0.708, 0.748)	0.048 (0.026, 0.067)
phecode_510-7	Gastroesophageal laceration-hemorrhage syndrome [Mallory-Weiss tear]	0.575 (0.554, 0.597)	0.679 (0.659, 0.704)	0.106 (0.08, 0.129)
phecode_510-8	Barrett's esophagus	0.644 (0.638, 0.649)	0.712 (0.707, 0.717)	0.068 (0.063, 0.073)
phecode_511	Gastro-esophageal reflux disease	0.556 (0.554, 0.558)	0.687 (0.685, 0.689)	0.131 (0.129, 0.134)
phecode_512	Dysphagia	0.571 (0.568, 0.575)	0.687 (0.684, 0.691)	0.116 (0.112, 0.12)
phecode_513	Peptic ulcer	0.611 (0.607, 0.615)	0.692 (0.688, 0.696)	0.081 (0.077, 0.085)
phecode_513-1	Esophageal ulcer	0.63 (0.623, 0.636)	0.698 (0.692, 0.705)	0.068 (0.062, 0.075)
phecode_513-2	Gastric ulcer	0.6 (0.595, 0.606)	0.705 (0.7, 0.711)	0.105 (0.099, 0.111)
phecode_513-3	Duodenal ulcer	0.647 (0.64, 0.655)	0.709 (0.703, 0.716)	0.062 (0.056, 0.069)
phecode_513-4	Gastrojejunal ulcer	0.549 (0.504, 0.584)	0.781 (0.745, 0.817)	0.234 (0.183, 0.284)
phecode_514	Gastrointestinal obstruction	0.624 (0.62, 0.629)	0.683 (0.678, 0.687)	0.058 (0.054, 0.063)
phecode_514-1	Esophageal obstruction (Stricture and stenosis of esophagus)	0.625 (0.617, 0.633)	0.67 (0.663, 0.68)	0.046 (0.038, 0.052)
phecode_514-2	Intestinal obstruction	0.628 (0.622, 0.634)	0.708 (0.701, 0.714)	0.08 (0.074, 0.086)
phecode_514-21	Impaction of intestine	0.599 (0.581, 0.615)	0.705 (0.69, 0.724)	0.107 (0.089, 0.126)
phecode_514-3	Ileus	0.663 (0.651, 0.675)	0.698 (0.687, 0.71)	0.036 (0.023, 0.046)
phecode_514-31	Paralytic ileus	0.653 (0.623, 0.681)	0.718 (0.692, 0.749)	0.065 (0.034, 0.094)
phecode_515	Heartburn and epigastric pain	0.555 (0.552, 0.559)	0.726 (0.724, 0.729)	0.17 (0.167, 0.174)
phecode_516	Other diseases of stomach and duodenum	0.593 (0.587, 0.599)	0.717 (0.711, 0.723)	0.124 (0.116, 0.13)
phecode_516-1	Pyloric stenosis	0.64 (0.617, 0.66)	0.744 (0.724, 0.765)	0.105 (0.085, 0.125)
phecode_517	Gastrointestinal angiodysplasia	0.676 (0.664, 0.687)	0.729 (0.719, 0.741)	0.053 (0.042, 0.065)
phecode_518	Appendicitis	0.532 (0.525, 0.542)	0.56 (0.551, 0.569)	0.027 (0.018, 0.038)
phecode_520	Hernia	0.618 (0.616, 0.62)	0.667 (0.665, 0.669)	0.049 (0.047, 0.051)
phecode_520-1	Hernia of the abdominal wall	0.713 (0.711, 0.716)	0.728 (0.726, 0.73)	0.015 (0.014, 0.016)
phecode_520-11	Inguinal hernia	0.782 (0.78, 0.784)	0.79 (0.788, 0.792)	0.008 (0.007, 0.009)
phecode_520-12	Femoral hernia	0.628 (0.615, 0.641)	0.653 (0.637, 0.666)	0.024 (0.012, 0.037)

4 Medical history predicts future health trajectories over the human phenome

Table 17 continued from previous page

Endpoint	PheCode string	C-Index Age+Sex (IQR)	C-Index Age+Sex+MedicalHistory (IQR)	Difference in C-Index (IQR)
phecode_520-13	Umbilical hernia	0.633 (0.628, 0.638)	0.707 (0.702, 0.713)	0.075 (0.069, 0.08)
phecode_520-14	Ventral hernia	0.593 (0.587, 0.599)	0.705 (0.699, 0.711)	0.112 (0.104, 0.119)
phecode_520-15	Incisional hernia	0.609 (0.6, 0.617)	0.759 (0.751, 0.766)	0.15 (0.141, 0.16)
phecode_520-2	Diaphragmatic hernia [Hiatal hernia]	0.595 (0.593, 0.597)	0.682 (0.68, 0.685)	0.087 (0.085, 0.09)
phecode_522	Gastrointestinal inflammation	0.559 (0.557, 0.561)	0.668 (0.666, 0.67)	0.109 (0.107, 0.111)
phecode_522-1	Inflammatory bowel disease	0.544 (0.536, 0.551)	0.662 (0.655, 0.669)	0.118 (0.11, 0.127)
phecode_522-11	Crohn's disease	0.496 (0.484, 0.508)	0.707 (0.695, 0.72)	0.211 (0.195, 0.228)
phecode_522-12	Ulcerative colitis	0.523 (0.513, 0.533)	0.685 (0.675, 0.696)	0.162 (0.149, 0.176)
phecode_522-13	Inflammatory polyps of colon [Pseudopolyposis of colon]	0.577 (0.553, 0.601)	0.816 (0.795, 0.837)	0.239 (0.211, 0.268)
phecode_522-14	Microscopic colitis*	0.659 (0.646, 0.671)	0.755 (0.745, 0.765)	0.096 (0.085, 0.106)
phecode_522-4	Toxic gastroenteritis and colitis	0.623 (0.605, 0.64)	0.705 (0.688, 0.721)	0.082 (0.065, 0.101)
phecode_522-6	Eosinophilic gastroenteritis and colitis	0.653 (0.61, 0.699)	0.743 (0.708, 0.779)	0.091 (0.048, 0.133)
phecode_522-61	Eosinophilic esophagitis	0.667 (0.627, 0.709)	0.746 (0.71, 0.782)	0.079 (0.037, 0.114)
phecode_522-7	Ulceration of the lower GI tract	0.558 (0.547, 0.569)	0.702 (0.691, 0.713)	0.144 (0.131, 0.157)
phecode_522-8	Duodenitis	0.595 (0.59, 0.6)	0.682 (0.678, 0.686)	0.087 (0.083, 0.092)
phecode_522-9	Gastritis	0.567 (0.565, 0.569)	0.686 (0.684, 0.688)	0.119 (0.116, 0.121)
phecode_522-91	Atrophic gastritis	0.643 (0.626, 0.66)	0.737 (0.723, 0.751)	0.094 (0.076, 0.11)
phecode_523	Diverticular disease [Bowel diverticulosis]	0.619 (0.617, 0.62)	0.668 (0.666, 0.67)	0.05 (0.048, 0.051)
phecode_523-1	Diverticula of small intestine	0.676 (0.661, 0.689)	0.73 (0.717, 0.744)	0.054 (0.041, 0.07)
phecode_523-2	Diverticula of colon	0.626 (0.624, 0.627)	0.676 (0.675, 0.679)	0.051 (0.049, 0.053)
phecode_523-3	Gastric diverticulum	0.58 (0.552, 0.609)	0.658 (0.631, 0.687)	0.079 (0.05, 0.109)
phecode_523-4	Diverticulitis	0.613 (0.608, 0.618)	0.744 (0.74, 0.749)	0.131 (0.126, 0.136)
phecode_524	Functional intestinal disorder	0.606 (0.601, 0.609)	0.772 (0.769, 0.776)	0.167 (0.162, 0.171)
phecode_524-1	Irritable bowel syndrome	0.611 (0.606, 0.616)	0.788 (0.784, 0.792)	0.177 (0.172, 0.182)
phecode_524-3	Megacolon (other than Hirschsprung's)	0.64 (0.617, 0.662)	0.703 (0.682, 0.722)	0.063 (0.038, 0.085)
phecode_525	Intestinal malabsorption	0.545 (0.538, 0.552)	0.667 (0.659, 0.675)	0.122 (0.113, 0.131)
phecode_525-1	Celiac disease	0.526 (0.516, 0.536)	0.635 (0.625, 0.646)	0.109 (0.097, 0.123)
phecode_525-3	Disorders of intestinal carbohydrate absorption	0.553 (0.536, 0.569)	0.712 (0.698, 0.724)	0.158 (0.14, 0.176)
phecode_526	Intestinal infection	0.557 (0.555, 0.56)	0.664 (0.661, 0.666)	0.106 (0.103, 0.109)
phecode_526-1	Bacterial enteritis	0.532 (0.527, 0.537)	0.656 (0.652, 0.661)	0.124 (0.118, 0.131)
phecode_526-11	Intestinal e.coli	0.525 (0.515, 0.536)	0.699 (0.69, 0.708)	0.174 (0.161, 0.185)
phecode_526-12	Intestinal infection due to C. difficile	0.648 (0.638, 0.659)	0.738 (0.728, 0.749)	0.09 (0.078, 0.101)
phecode_526-2	Viral enteritis	0.589 (0.578, 0.6)	0.74 (0.73, 0.749)	0.151 (0.137, 0.163)
phecode_526-3	Intestinal infection due to protozoa	0.512 (0.489, 0.531)	0.675 (0.657, 0.695)	0.164 (0.139, 0.19)
phecode_527	Abdominal pain	0.542 (0.541, 0.544)	0.684 (0.683, 0.685)	0.142 (0.14, 0.143)
phecode_528	Nausea and vomiting	0.584 (0.582, 0.587)	0.695 (0.693, 0.697)	0.11 (0.108, 0.113)
phecode_528-1	Nausea	0.584 (0.582, 0.587)	0.695 (0.693, 0.697)	0.111 (0.108, 0.113)
phecode_528-2	Vomiting	0.584 (0.582, 0.587)	0.695 (0.693, 0.697)	0.11 (0.108, 0.113)
phecode_529	Symptoms involving digestive system	0.56 (0.559, 0.562)	0.687 (0.685, 0.688)	0.127 (0.125, 0.128)
phecode_529-1	Diarrhea	0.545 (0.543, 0.548)	0.728 (0.726, 0.73)	0.183 (0.18, 0.185)
phecode_529-2	Abdominal distension and flatulence	0.575 (0.571, 0.578)	0.726 (0.723, 0.728)	0.151 (0.147, 0.155)
phecode_529-3	Fecal incontinence	0.648 (0.642, 0.654)	0.746 (0.74, 0.75)	0.097 (0.092, 0.103)
phecode_529-4	Abnormal bowel sounds and visible peristalsis	0.532 (0.508, 0.557)	0.728 (0.707, 0.748)	0.197 (0.166, 0.223)
phecode_529-5	Constipation	0.605 (0.602, 0.607)	0.7 (0.698, 0.702)	0.096 (0.093, 0.098)
phecode_529-6	Halitosis*	0.489 (0.474, 0.504)	0.689 (0.677, 0.702)	0.2 (0.182, 0.218)
phecode_530	Disease of anus and rectum	0.518 (0.514, 0.521)	0.67 (0.666, 0.673)	0.152 (0.148, 0.156)
phecode_530-1	Anal fissure	0.564 (0.558, 0.569)	0.696 (0.691, 0.701)	0.132 (0.125, 0.14)
phecode_530-2	Anorectal abscess	0.627 (0.616, 0.638)	0.706 (0.695, 0.716)	0.08 (0.069, 0.09)
phecode_530-3	Rectal prolapse	0.676 (0.667, 0.687)	0.732 (0.722, 0.742)	0.056 (0.047, 0.064)
phecode_530-4	Stenosis of anus and rectum	0.635 (0.611, 0.658)	0.741 (0.719, 0.764)	0.106 (0.081, 0.133)
phecode_532	Other disorders of the intestines	0.567 (0.562, 0.572)	0.672 (0.667, 0.676)	0.105 (0.099, 0.111)
phecode_532-1	Intestinal fistula	0.612 (0.601, 0.625)	0.714 (0.704, 0.726)	0.101 (0.089, 0.113)
phecode_532-2	Intestinal perforation	0.652 (0.638, 0.666)	0.727 (0.712, 0.742)	0.075 (0.061, 0.09)
phecode_532-3	Intussusception	0.681 (0.662, 0.702)	0.773 (0.751, 0.794)	0.092 (0.068, 0.115)

Supplementary Tables

Table 17 continued from previous page

Endpoint	PheCode string	C-Index Age+Sex (IQR)	C-Index Age+Sex+MedicalHistory (IQR)	Difference in C-Index (IQR)
phecode_532-4	Volvulus	0.617 (0.605, 0.629)	0.694 (0.683, 0.707)	0.078 (0.064, 0.089)
phecode_535	Intestinal or peritoneal adhesions	0.573 (0.567, 0.577)	0.67 (0.665, 0.676)	0.098 (0.091, 0.104)
phecode_537	Abnormality of the peritoneum	0.573 (0.565, 0.583)	0.632 (0.623, 0.642)	0.059 (0.048, 0.069)
phecode_537-1	Peritonitis	0.57 (0.56, 0.58)	0.629 (0.62, 0.639)	0.06 (0.048, 0.07)
phecode_537-2	Hemoperitoneum	0.626 (0.583, 0.667)	0.718 (0.681, 0.758)	0.093 (0.054, 0.135)
phecode_540	Hepatitis	0.527 (0.514, 0.54)	0.541 (0.528, 0.556)	0.014 (-0.003, 0.034)
phecode_540-1	Chronic hepatitis	0.497 (0.48, 0.512)	0.419 (0.402, 0.435)	-0.078 (-0.103, -0.055)
phecode_540-11	Autoimmune hepatitis	0.671 (0.649, 0.693)	0.656 (0.626, 0.687)	-0.015 (-0.049, 0.02)
phecode_540-2	Acute hepatitis	0.545 (0.524, 0.567)	0.424 (0.402, 0.447)	-0.121 (-0.154, -0.087)
phecode_540-3	Viral hepatitis	0.577 (0.559, 0.594)	0.488 (0.47, 0.508)	-0.089 (-0.111, -0.063)
phecode_540-4	Alcoholic hepatitis	0.69 (0.662, 0.716)	0.867 (0.848, 0.889)	0.178 (0.148, 0.206)
phecode_542	Chronic liver disease and sequelae	0.528 (0.524, 0.532)	0.722 (0.719, 0.726)	0.194 (0.189, 0.199)
phecode_542-1	Fibrosis and cirrhosis of liver	0.64 (0.63, 0.65)	0.822 (0.815, 0.831)	0.183 (0.172, 0.193)
phecode_542-2	Fatty liver disease (FLD)	0.522 (0.517, 0.527)	0.741 (0.737, 0.744)	0.219 (0.214, 0.224)
phecode_542-3	Hepatic failure	0.64 (0.628, 0.654)	0.821 (0.812, 0.831)	0.182 (0.167, 0.193)
phecode_542-4	Portal hypertension	0.643 (0.629, 0.655)	0.831 (0.82, 0.843)	0.188 (0.174, 0.203)
phecode_542-5	Hepatorenal syndrome	0.691 (0.656, 0.73)	0.887 (0.867, 0.911)	0.196 (0.161, 0.232)
phecode_544	Biliary cirrhosis	0.632 (0.611, 0.652)	0.724 (0.702, 0.747)	0.092 (0.071, 0.115)
phecode_544-1	Primary biliary cirrhosis*	0.651 (0.63, 0.672)	0.739 (0.717, 0.761)	0.088 (0.067, 0.111)
phecode_545	Nonspecific abnormal results of function study of liver	0.54 (0.536, 0.544)	0.655 (0.651, 0.659)	0.115 (0.11, 0.12)
phecode_546	Other disorders of liver	0.586 (0.581, 0.591)	0.679 (0.674, 0.684)	0.093 (0.087, 0.099)
phecode_546-2	Abscess of liver	0.636 (0.612, 0.659)	0.701 (0.677, 0.723)	0.065 (0.043, 0.089)
phecode_546-3	Hepatomegaly	0.59 (0.574, 0.605)	0.731 (0.716, 0.746)	0.141 (0.121, 0.161)
phecode_546-31	Hepatosplenomegaly*	0.609 (0.572, 0.644)	0.773 (0.738, 0.809)	0.164 (0.122, 0.207)
phecode_550	Disorders of the gallbladder	0.578 (0.575, 0.581)	0.649 (0.646, 0.651)	0.07 (0.067, 0.074)
phecode_550-1	Gallstones [Cholelithiasis]	0.583 (0.58, 0.586)	0.653 (0.65, 0.656)	0.07 (0.066, 0.073)
phecode_550-2	Cholecystitis	0.561 (0.555, 0.567)	0.656 (0.65, 0.662)	0.095 (0.088, 0.103)
phecode_550-3	Gallbladder perforation	0.686 (0.655, 0.719)	0.732 (0.702, 0.763)	0.045 (0.023, 0.066)
phecode_550-4	Cholesterosis of gallbladder	0.579 (0.568, 0.592)	0.677 (0.666, 0.689)	0.097 (0.086, 0.11)
phecode_552	Other diseases of biliary tract	0.637 (0.629, 0.644)	0.709 (0.703, 0.715)	0.073 (0.066, 0.08)
phecode_552-1	Cholangitis	0.661 (0.649, 0.673)	0.737 (0.726, 0.749)	0.076 (0.064, 0.087)
phecode_552-2	Obstruction of bile duct	0.651 (0.64, 0.661)	0.709 (0.698, 0.718)	0.057 (0.046, 0.067)
phecode_554	Diseases of the pancreas	0.605 (0.598, 0.612)	0.688 (0.682, 0.694)	0.083 (0.076, 0.09)
phecode_554-1	Pancreatitis	0.585 (0.578, 0.592)	0.676 (0.668, 0.683)	0.092 (0.082, 0.1)
phecode_554-11	Acute pancreatitis	0.575 (0.566, 0.584)	0.667 (0.659, 0.678)	0.092 (0.082, 0.103)
phecode_554-12	Chronic pancreatitis	0.63 (0.613, 0.648)	0.775 (0.762, 0.79)	0.145 (0.127, 0.163)
phecode_554-2	Cyst and pseudocyst of pancreas	0.627 (0.614, 0.641)	0.695 (0.681, 0.706)	0.067 (0.052, 0.083)
phecode_555	Ascites	0.618 (0.611, 0.626)	0.772 (0.765, 0.778)	0.153 (0.145, 0.162)
phecode_556	Other symptoms involving the digestive system and abdomen	0.558 (0.555, 0.561)	0.696 (0.693, 0.698)	0.138 (0.135, 0.141)
phecode_556-3	Abdominal or pelvic swelling, mass, or lump	0.542 (0.537, 0.548)	0.642 (0.636, 0.647)	0.099 (0.093, 0.106)
phecode_556-4	Abdominal rigidity	0.535 (0.497, 0.576)	0.77 (0.74, 0.805)	0.236 (0.196, 0.281)
phecode_556-8	Nonspecific abnormal findings in stool contents	0.564 (0.561, 0.567)	0.716 (0.713, 0.719)	0.152 (0.148, 0.155)
phecode_557	Gastrointestinal hemorrhage	0.533 (0.531, 0.536)	0.638 (0.636, 0.64)	0.104 (0.102, 0.107)
phecode_557-1	Hematemesis	0.605 (0.597, 0.614)	0.723 (0.715, 0.73)	0.118 (0.108, 0.127)
phecode_557-2	Blood in stool	0.607 (0.601, 0.613)	0.674 (0.669, 0.68)	0.067 (0.061, 0.073)
phecode_557-8	Hemorrhage of rectum and anus	0.513 (0.511, 0.516)	0.663 (0.66, 0.665)	0.149 (0.146, 0.153)
phecode_558	Abnormal findings on diagnostic imaging of the digestive tract	0.592 (0.588, 0.596)	0.692 (0.688, 0.696)	0.1 (0.095, 0.105)
phecode_559	Other disease of digestive system, NOS	0.536 (0.526, 0.545)	0.793 (0.785, 0.8)	0.257 (0.245, 0.269)
phecode_580	Glomerular diseases	0.635 (0.627, 0.643)	0.768 (0.76, 0.775)	0.133 (0.125, 0.141)
phecode_580-2	Chronic nephritic syndrome	0.714 (0.704, 0.725)	0.883 (0.874, 0.891)	0.169 (0.156, 0.18)
phecode_580-3	Nephrotic syndrome	0.567 (0.555, 0.58)	0.662 (0.65, 0.673)	0.094 (0.08, 0.109)
phecode_580-4	Nephritis and nephropathy, not specified as acute or chronic	0.642 (0.626, 0.658)	0.892 (0.881, 0.902)	0.25 (0.233, 0.267)
phecode_581	Renal tubulo-interstitial diseases	0.566 (0.561, 0.571)	0.651 (0.646, 0.656)	0.086 (0.08, 0.092)
phecode_581-1	Pyelonephritis	0.585 (0.576, 0.593)	0.722 (0.713, 0.731)	0.137 (0.126, 0.147)
phecode_581-11	Acute pyelonephritis	0.604 (0.595, 0.615)	0.744 (0.737, 0.753)	0.14 (0.13, 0.15)
phecode_581-3	Obstructive and reflux uropathy	0.632 (0.626, 0.638)	0.677 (0.671, 0.682)	0.045 (0.039, 0.05)

4 Medical history predicts future health trajectories over the human phenome

Table 17 continued from previous page

Endpoint	PheCode string	C-Index Age+Sex (IQR)	C-Index Age+Sex+MedicalHistory (IQR)	Difference in C-Index (IQR)
phecode_581-31	Hydronephrosis	0.618 (0.612, 0.625)	0.665 (0.658, 0.671)	0.046 (0.04, 0.052)
phecode_581-32	Hydroureter	0.618 (0.597, 0.638)	0.68 (0.66, 0.7)	0.063 (0.042, 0.083)
phecode_581-33	Stricture or kinking of ureter	0.619 (0.608, 0.632)	0.718 (0.707, 0.73)	0.098 (0.085, 0.113)
phecode_582	Acute kidney failure	0.702 (0.699, 0.706)	0.799 (0.796, 0.801)	0.096 (0.093, 0.1)
phecode_583	Chronic kidney disease	0.689 (0.686, 0.691)	0.751 (0.748, 0.753)	0.062 (0.06, 0.064)
phecode_583-1	End stage renal disease [CDK, stage 5]	0.689 (0.68, 0.699)	0.887 (0.881, 0.893)	0.198 (0.188, 0.208)
phecode_584	Renal failure	0.668 (0.66, 0.674)	0.774 (0.767, 0.779)	0.106 (0.1, 0.113)
phecode_584-1	Renal dialysis	0.664 (0.651, 0.678)	0.873 (0.863, 0.885)	0.209 (0.192, 0.224)
phecode_585	Kidney stone disease	0.592 (0.587, 0.596)	0.658 (0.653, 0.662)	0.066 (0.062, 0.071)
phecode_585-1	Renal colic	0.588 (0.58, 0.596)	0.701 (0.695, 0.708)	0.113 (0.106, 0.122)
phecode_586	Other disorders of the kidney and ureters	0.645 (0.64, 0.649)	0.72 (0.716, 0.724)	0.075 (0.071, 0.08)
phecode_586-1	Small kidney	0.597 (0.557, 0.64)	0.74 (0.699, 0.779)	0.143 (0.093, 0.192)
phecode_586-2	Cyst of kidney	0.655 (0.649, 0.661)	0.719 (0.713, 0.724)	0.064 (0.059, 0.069)
phecode_586-3	Vascular disorders of kidney	0.681 (0.649, 0.714)	0.77 (0.74, 0.805)	0.089 (0.055, 0.122)
phecode_586-5	Renal sclerosis, unspecified	0.632 (0.609, 0.658)	0.752 (0.726, 0.777)	0.12 (0.096, 0.148)
phecode_588	Disorders and findings from impaired renal function	0.651 (0.639, 0.663)	0.762 (0.751, 0.774)	0.111 (0.098, 0.124)
phecode_588-1	Disorders resulting from impaired renal function	0.615 (0.582, 0.643)	0.667 (0.632, 0.712)	0.053 (0.003, 0.11)
phecode_588-2	Abnormal results of function study of kidney	0.657 (0.645, 0.67)	0.759 (0.747, 0.772)	0.102 (0.089, 0.116)
phecode_589	Pathological lesions of kidney	0.618 (0.594, 0.643)	0.37 (0.34, 0.403)	-0.25 (-0.294, -0.201)
phecode_589-1	Membranous glomerulonephritis	0.636 (0.586, 0.681)	0.68 (0.625, 0.741)	0.044 (-0.018, 0.11)
phecode_591	Urinary tract infection	0.627 (0.625, 0.629)	0.7 (0.698, 0.702)	0.073 (0.071, 0.075)
phecode_592	Cystitis and urethritis	0.673 (0.67, 0.676)	0.762 (0.759, 0.765)	0.089 (0.086, 0.092)
phecode_592-1	Cystitis	0.682 (0.679, 0.685)	0.769 (0.766, 0.772)	0.087 (0.085, 0.09)
phecode_592-11	Acute cystitis	0.716 (0.711, 0.723)	0.806 (0.8, 0.812)	0.09 (0.084, 0.095)
phecode_592-12	Chronic cystitis	0.63 (0.618, 0.642)	0.737 (0.726, 0.748)	0.107 (0.094, 0.12)
phecode_592-13	Chronic interstitial cystitis	0.668 (0.652, 0.687)	0.803 (0.787, 0.821)	0.134 (0.116, 0.154)
phecode_592-2	Urethritis and urethral syndrome	0.539 (0.519, 0.561)	0.692 (0.676, 0.708)	0.152 (0.126, 0.178)
phecode_593	Hematuria	0.587 (0.586, 0.589)	0.691 (0.689, 0.692)	0.104 (0.102, 0.105)
phecode_593-1	Gross hematuria	0.679 (0.672, 0.688)	0.769 (0.763, 0.776)	0.09 (0.084, 0.095)
phecode_593-2	Microscopic hematuria	0.576 (0.57, 0.581)	0.712 (0.707, 0.716)	0.136 (0.131, 0.141)
phecode_593-3	Recurrent and persistent hematuria*	0.569 (0.556, 0.582)	0.704 (0.692, 0.716)	0.136 (0.12, 0.15)
phecode_594	Abnormality of urination	0.581 (0.58, 0.583)	0.683 (0.682, 0.685)	0.102 (0.1, 0.104)
phecode_594-1	Retention of urine	0.744 (0.741, 0.747)	0.764 (0.761, 0.766)	0.02 (0.018, 0.022)
phecode_594-11	Urinary hesitancy	0.765 (0.755, 0.776)	0.843 (0.834, 0.852)	0.077 (0.068, 0.086)
phecode_594-2	Dysuria	0.606 (0.602, 0.609)	0.743 (0.74, 0.745)	0.137 (0.134, 0.14)
phecode_594-3	Urinary incontinence	0.63 (0.627, 0.632)	0.722 (0.719, 0.725)	0.092 (0.089, 0.095)
phecode_594-31	Urge incontinence	0.661 (0.656, 0.666)	0.785 (0.781, 0.788)	0.124 (0.119, 0.127)
phecode_594-32	Stress incontinence	0.724 (0.72, 0.728)	0.801 (0.798, 0.805)	0.077 (0.074, 0.081)
phecode_594-33	Nocturnal enuresis	0.583 (0.556, 0.613)	0.763 (0.742, 0.785)	0.18 (0.149, 0.208)
phecode_594-4	Frequency of urination and polyuria	0.587 (0.585, 0.59)	0.726 (0.724, 0.728)	0.139 (0.136, 0.141)
phecode_594-41	Nocturia	0.71 (0.707, 0.715)	0.794 (0.791, 0.798)	0.084 (0.081, 0.086)
phecode_594-5	Oliguria and anuria	0.671 (0.647, 0.697)	0.761 (0.738, 0.783)	0.091 (0.064, 0.114)
phecode_594-6	Urinary urgency	0.593 (0.588, 0.597)	0.744 (0.741, 0.748)	0.152 (0.147, 0.156)
phecode_596	Other disorders of bladder	0.622 (0.618, 0.625)	0.712 (0.709, 0.715)	0.09 (0.087, 0.093)
phecode_596-1	Bladder neck obstruction	0.823 (0.819, 0.827)	0.857 (0.853, 0.862)	0.034 (0.031, 0.038)
phecode_596-2	Overactive bladder	0.597 (0.592, 0.601)	0.757 (0.753, 0.76)	0.16 (0.155, 0.165)
phecode_596-3	Diverticulum of bladder	0.782 (0.774, 0.792)	0.804 (0.795, 0.812)	0.021 (0.015, 0.027)
phecode_596-4	Vesical fistula	0.67 (0.645, 0.697)	0.769 (0.747, 0.792)	0.097 (0.071, 0.123)
phecode_596-5	Neuromuscular dysfunction of bladder	0.558 (0.546, 0.571)	0.79 (0.779, 0.801)	0.232 (0.216, 0.247)
phecode_597	Other disorders of urethra and urinary tract	0.609 (0.603, 0.614)	0.694 (0.689, 0.7)	0.085 (0.08, 0.092)
phecode_597-1	Urethral stricture	0.642 (0.635, 0.65)	0.723 (0.716, 0.729)	0.081 (0.074, 0.088)
phecode_597-3	Urethral fistula	0.755 (0.726, 0.787)	0.818 (0.788, 0.849)	0.063 (0.035, 0.091)

Supplementary Tables

Table 17 continued from previous page

Endpoint	PheCode string	C-Index Age+Sex (IQR)	C-Index Age+Sex+MedicalHistory (IQR)	Difference in C-Index (IQR)
phecode_597-5	Urethral caruncle	0.776 (0.766, 0.784)	0.827 (0.817, 0.837)	0.052 (0.044, 0.06)
phecode_598	Abnormal findings on radiological and other examination of genitourinary organs	0.666 (0.648, 0.683)	0.712 (0.694, 0.728)	0.046 (0.031, 0.062)
phecode_599	Other symptoms/disorders or the urinary system	0.575 (0.573, 0.577)	0.714 (0.712, 0.716)	0.139 (0.137, 0.142)
phecode_599-7	Urethral discharge	0.703 (0.681, 0.723)	0.811 (0.791, 0.83)	0.11 (0.09, 0.127)
phecode_600	Benign prostatic hyperplasia	0.658 (0.655, 0.66)	0.694 (0.692, 0.697)	0.037 (0.035, 0.038)
phecode_601	Inflammatory diseases of prostate	0.556 (0.549, 0.563)	0.669 (0.662, 0.677)	0.113 (0.104, 0.123)
phecode_601-1	Prostatitis	0.556 (0.548, 0.563)	0.669 (0.662, 0.676)	0.114 (0.104, 0.123)
phecode_601-11	Acute prostatitis	0.526 (0.51, 0.541)	0.7 (0.685, 0.715)	0.176 (0.155, 0.196)
phecode_601-12	Chronic prostatitis	0.587 (0.576, 0.598)	0.659 (0.647, 0.67)	0.072 (0.058, 0.087)
phecode_602	Other disorders of prostate	0.637 (0.633, 0.64)	0.721 (0.718, 0.725)	0.085 (0.081, 0.089)
phecode_602-3	Dysplasia of prostate	0.62 (0.585, 0.654)	0.688 (0.649, 0.729)	0.071 (0.025, 0.114)
phecode_602-4	Elevated prostate specific antigen [PSA]	0.628 (0.623, 0.632)	0.733 (0.729, 0.736)	0.105 (0.101, 0.109)
phecode_603	Disorders and symptoms of testis	0.525 (0.521, 0.529)	0.666 (0.662, 0.67)	0.141 (0.136, 0.147)
phecode_603-1	Hydrocele	0.573 (0.564, 0.581)	0.625 (0.616, 0.634)	0.052 (0.043, 0.061)
phecode_603-2	Spermatocele	0.525 (0.517, 0.533)	0.683 (0.677, 0.69)	0.159 (0.149, 0.168)
phecode_603-4	Atrophy of testis	0.505 (0.459, 0.548)	0.652 (0.603, 0.699)	0.146 (0.084, 0.216)
phecode_603-5	Orchitis and epididymitis	0.526 (0.519, 0.533)	0.685 (0.678, 0.69)	0.158 (0.15, 0.168)
phecode_603-6	Scrotal pain*	0.81 (0.807, 0.814)	0.883 (0.88, 0.885)	0.072 (0.069, 0.076)
phecode_604	Disorders of penis	0.512 (0.507, 0.517)	0.668 (0.663, 0.672)	0.156 (0.149, 0.162)
phecode_604-1	Redundant prepuce and phimosis	0.495 (0.483, 0.507)	0.681 (0.671, 0.691)	0.186 (0.17, 0.201)
phecode_604-3	Peyronie's disease	0.517 (0.507, 0.527)	0.659 (0.649, 0.67)	0.143 (0.128, 0.155)
phecode_604-5	Balanoposthitis	0.532 (0.524, 0.541)	0.702 (0.695, 0.709)	0.17 (0.161, 0.18)
phecode_605	Male sexual dysfunction	0.551 (0.549, 0.554)	0.718 (0.716, 0.72)	0.167 (0.164, 0.17)
phecode_605-1	Male erectile dysfunction	0.551 (0.549, 0.554)	0.718 (0.716, 0.72)	0.167 (0.164, 0.17)
phecode_605-2	Ejaculatory dysfunction	0.511 (0.471, 0.549)	0.734 (0.702, 0.767)	0.224 (0.181, 0.27)
phecode_608	Other disorders of male genital organs	0.527 (0.523, 0.53)	0.672 (0.669, 0.674)	0.145 (0.141, 0.148)
phecode_608-1	Abnormal findings in semen	0.821 (0.819, 0.823)	0.871 (0.87, 0.873)	0.05 (0.049, 0.052)
phecode_609	Male infertility	0.868 (0.856, 0.883)	0.876 (0.864, 0.89)	0.008 (0.001, 0.015)
phecode_610	Benign mammary dysplasias	0.649 (0.642, 0.657)	0.695 (0.688, 0.703)	0.046 (0.04, 0.051)
phecode_610-1	Cystic mastopathy	0.683 (0.675, 0.693)	0.735 (0.727, 0.742)	0.051 (0.046, 0.057)
phecode_610-2	Fibroadenosis of breast	0.651 (0.624, 0.68)	0.686 (0.66, 0.715)	0.035 (0.01, 0.062)
phecode_610-3	Fibrosclerosis of breast	0.581 (0.541, 0.62)	0.663 (0.631, 0.699)	0.082 (0.039, 0.129)
phecode_610-5	Mammary duct ectasia	0.46 (0.435, 0.483)	0.627 (0.6, 0.65)	0.165 (0.132, 0.2)
phecode_612	Breast conditions, congenital or relating to hormones	0.63 (0.607, 0.657)	0.649 (0.625, 0.671)	0.019 (0.001, 0.036)
phecode_612-2	Hypertrophy of breast (Gynecomastia)	0.58 (0.551, 0.607)	0.596 (0.565, 0.62)	0.016 (-0.01, 0.041)
phecode_613	Other nonmalignant breast conditions	0.542 (0.539, 0.546)	0.703 (0.701, 0.706)	0.161 (0.157, 0.165)
phecode_613-1	Inflammatory disease of breast	0.571 (0.56, 0.581)	0.691 (0.682, 0.7)	0.12 (0.108, 0.131)
phecode_613-5	Mastodynia	0.55 (0.546, 0.554)	0.728 (0.725, 0.731)	0.177 (0.173, 0.183)
phecode_613-7	Other signs and symptoms in breast	0.501 (0.493, 0.508)	0.679 (0.673, 0.685)	0.178 (0.169, 0.187)
phecode_613-8	Other specified disorders of breast	0.544 (0.525, 0.563)	0.659 (0.64, 0.675)	0.114 (0.09, 0.137)
phecode_614	Inflammatory diseases of female pelvic organs	0.56 (0.556, 0.564)	0.701 (0.698, 0.705)	0.141 (0.136, 0.146)
phecode_614-1	Pelvic peritoneal adhesions, female (postoperative) (postinfection)	0.563 (0.552, 0.574)	0.639 (0.627, 0.648)	0.075 (0.064, 0.089)
phecode_614-3	Pelvic inflammatory disease (PID)	0.664 (0.646, 0.682)	0.674 (0.657, 0.692)	0.01 (-0.006, 0.028)
phecode_614-32	Chronic inflammatory pelvic disease	0.684 (0.659, 0.711)	0.702 (0.676, 0.725)	0.017 (-0.008, 0.043)
phecode_614-33	Pelvic inflammatory disease, NOS	0.65 (0.623, 0.678)	0.686 (0.66, 0.711)	0.034 (0.009, 0.06)
phecode_614-4	Inflammatory diseases of uterus, except cervix	0.669 (0.645, 0.698)	0.695 (0.672, 0.724)	0.027 (-0.002, 0.056)
phecode_614-5	Inflammatory disease of cervix, vagina, and vulva	0.558 (0.554, 0.562)	0.719 (0.715, 0.722)	0.161 (0.156, 0.166)
phecode_614-51	Cervicitis and endocervicitis	0.651 (0.633, 0.669)	0.672 (0.654, 0.689)	0.021 (0.004, 0.039)
phecode_614-52	Vaginitis and vulvovaginitis	0.554 (0.549, 0.558)	0.733 (0.73, 0.736)	0.179 (0.174, 0.185)
phecode_614-53	Cyst or abscess of Bartholin's gland	0.627 (0.614, 0.639)	0.688 (0.677, 0.699)	0.061 (0.05, 0.073)
phecode_614-54	Abscess or ulceration of vulva	0.51 (0.495, 0.524)	0.703 (0.691, 0.715)	0.194 (0.175, 0.212)

4 Medical history predicts future health trajectories over the human phenome

Table 17 continued from previous page

Endpoint	PheCode string	C-Index Age+Sex (IQR)	C-Index Age+Sex+MedicalHistory (IQR)	Difference in C-Index (IQR)
phecode_614-55	Candidiasis of vulva and vagina	0.763 (0.761, 0.766)	0.867 (0.865, 0.869)	0.103 (0.1, 0.106)
phecode_615	Endometriosis	0.708 (0.698, 0.719)	0.724 (0.714, 0.734)	0.016 (0.01, 0.023)
phecode_617	Fistula involving female genital tract	0.481 (0.457, 0.504)	0.635 (0.605, 0.664)	0.154 (0.116, 0.189)
phecode_618	Genital prolapse	0.569 (0.565, 0.572)	0.637 (0.634, 0.64)	0.068 (0.065, 0.072)
phecode_618-1	Prolapse of vaginal walls	0.571 (0.567, 0.575)	0.661 (0.657, 0.664)	0.09 (0.086, 0.094)
phecode_618-11	Cystocele	0.581 (0.577, 0.586)	0.664 (0.66, 0.668)	0.082 (0.078, 0.087)
phecode_618-12	Rectocele	0.553 (0.548, 0.559)	0.682 (0.677, 0.687)	0.128 (0.122, 0.135)
phecode_618-2	Uterine/Uterovaginal prolapse	0.587 (0.582, 0.591)	0.64 (0.635, 0.645)	0.053 (0.048, 0.058)
phecode_618-5	Prolapse of vaginal vault after hysterectomy	0.633 (0.619, 0.648)	0.797 (0.784, 0.813)	0.164 (0.145, 0.185)
phecode_618-6	Vaginal enterocele, congenital or acquired	0.591 (0.576, 0.607)	0.732 (0.717, 0.746)	0.14 (0.123, 0.159)
phecode_619	Noninflammatory female genital disorders	0.562 (0.56, 0.565)	0.686 (0.684, 0.688)	0.123 (0.121, 0.126)
phecode_619-1	Noninflammatory disorders of ovary, fallopian tube, and broad ligament	0.556 (0.538, 0.574)	0.589 (0.57, 0.605)	0.033 (0.013, 0.054)
phecode_619-2	Disorders of uterus, NEC	0.582 (0.573, 0.59)	0.631 (0.622, 0.638)	0.049 (0.04, 0.058)
phecode_619-3	Noninflammatory disorders of cervix	0.597 (0.67, 0.685)	0.714 (0.706, 0.72)	0.037 (0.031, 0.042)
phecode_619-32	Erosion and ectropion of cervix	0.799 (0.788, 0.811)	0.809 (0.797, 0.82)	0.01 (0.002, 0.018)
phecode_619-4	Noninflammatory disorders of vagina	0.569 (0.566, 0.572)	0.709 (0.706, 0.711)	0.14 (0.136, 0.143)
phecode_619-5	Noninflammatory disorders of vulva and perineum	0.525 (0.52, 0.53)	0.718 (0.714, 0.722)	0.193 (0.187, 0.199)
phecode_620	Dysplasia of female genital organs	0.717 (0.707, 0.727)	0.754 (0.743, 0.765)	0.037 (0.028, 0.046)
phecode_621	Endometrial hyperplasia	0.565 (0.55, 0.581)	0.676 (0.661, 0.69)	0.111 (0.094, 0.128)
phecode_622	Polyp of female genital organs	0.593 (0.589, 0.598)	0.625 (0.621, 0.63)	0.033 (0.027, 0.038)
phecode_622-1	Polyp of corpus uteri	0.554 (0.548, 0.56)	0.611 (0.604, 0.618)	0.057 (0.05, 0.064)
phecode_622-2	Mucous polyp of cervix	0.66 (0.652, 0.668)	0.683 (0.675, 0.69)	0.023 (0.016, 0.029)
phecode_623	Hypertrophy of female genital organs	0.509 (0.504, 0.515)	0.724 (0.72, 0.728)	0.214 (0.207, 0.221)
phecode_624	Symptoms involving female genital tract	0.58 (0.571, 0.587)	0.72 (0.712, 0.727)	0.14 (0.13, 0.151)
phecode_624-1	Dystrophy of female genital tract	0.579 (0.569, 0.59)	0.733 (0.722, 0.742)	0.154 (0.142, 0.166)
phecode_624-2	Atrophy of female genital tract	0.581 (0.57, 0.592)	0.711 (0.699, 0.721)	0.13 (0.117, 0.145)
phecode_625	Pain and other symptoms associated with female genital organs	0.619 (0.615, 0.623)	0.729 (0.725, 0.733)	0.11 (0.106, 0.114)
phecode_625-1	Dyspareunia	0.65 (0.643, 0.656)	0.751 (0.745, 0.757)	0.101 (0.095, 0.107)
phecode_625-2	Postcoital bleeding	0.743 (0.735, 0.751)	0.787 (0.778, 0.794)	0.044 (0.039, 0.049)
phecode_625-3	Vulvodynia	0.474 (0.45, 0.495)	0.768 (0.75, 0.786)	0.296 (0.268, 0.324)
phecode_626	Disorders of menstruation and other abnormal bleeding from female genital tract	0.794 (0.791, 0.798)	0.818 (0.814, 0.821)	0.024 (0.022, 0.025)
phecode_626-1	Irregular menstrual cycle/bleeding	0.872 (0.869, 0.874)	0.887 (0.884, 0.889)	0.015 (0.014, 0.016)
phecode_626-11	Absent or infrequent menstruation	0.848 (0.844, 0.854)	0.873 (0.868, 0.878)	0.025 (0.022, 0.027)
phecode_626-13	Irregular menstrual cycle	0.884 (0.881, 0.886)	0.901 (0.899, 0.903)	0.017 (0.016, 0.019)
phecode_626-14	Irregular menstrual bleeding	0.869 (0.866, 0.873)	0.883 (0.88, 0.887)	0.014 (0.011, 0.016)
phecode_626-2	Dysmenorrhea	0.896 (0.892, 0.901)	0.909 (0.904, 0.914)	0.013 (0.01, 0.016)
phecode_627	Menopausal and postmenopausal disorders	0.679 (0.677, 0.681)	0.739 (0.737, 0.741)	0.06 (0.058, 0.061)
phecode_627-1	Postmenopausal bleeding	0.559 (0.556, 0.563)	0.652 (0.648, 0.656)	0.093 (0.089, 0.097)
phecode_627-2	Symptomatic menopause	0.762 (0.76, 0.765)	0.816 (0.814, 0.818)	0.054 (0.052, 0.055)
phecode_627-3	Postmenopausal atrophic vaginitis	0.523 (0.519, 0.526)	0.709 (0.705, 0.712)	0.186 (0.181, 0.19)
phecode_627-4	Menorrhagia/Excessive and frequent menstruation	0.88 (0.878, 0.882)	0.891 (0.889, 0.893)	0.011 (0.01, 0.012)
phecode_628	Ovarian cyst	0.589 (0.583, 0.596)	0.646 (0.641, 0.652)	0.057 (0.051, 0.064)
phecode_628-1	Follicular cyst of ovary	0.804 (0.781, 0.829)	0.804 (0.777, 0.83)	-0.001 (-0.015, 0.015)
phecode_628-2	Corpus luteum cyst or hematoma	0.582 (0.575, 0.589)	0.646 (0.64, 0.653)	0.064 (0.058, 0.07)
phecode_628-3	polycystic ovary syndrome	0.761 (0.729, 0.792)	0.786 (0.752, 0.821)	0.026 (-0.004, 0.055)
phecode_629	Female infertility	0.898 (0.885, 0.912)	0.901 (0.889, 0.914)	0.003 (-0.005, 0.011)
phecode_660	Infection of the skin	0.522 (0.521, 0.524)	0.686 (0.685, 0.687)	0.163 (0.162, 0.165)
phecode_660-1	Fungal infection of the skin	0.541 (0.538, 0.543)	0.7 (0.698, 0.702)	0.159 (0.157, 0.161)
phecode_660-11	Candidiasis of skin and nails	0.541 (0.534, 0.547)	0.757 (0.752, 0.762)	0.217 (0.208, 0.225)
phecode_660-12	Dermatophytosis	0.547 (0.545, 0.55)	0.7 (0.698, 0.701)	0.152 (0.15, 0.155)
phecode_660-13	Pityriasis versicolor	0.636 (0.625, 0.647)	0.711 (0.702, 0.721)	0.075 (0.067, 0.083)

Supplementary Tables

Table 17 continued from previous page

Endpoint	PheCode string	C-Index Age+Sex (IQR)	C-Index Age+Sex+MedicalHistory (IQR)	Difference in C-Index (IQR)
phecode_660-2	Bacterial infection of the skin	0.532 (0.527, 0.54)	0.68 (0.675, 0.686)	0.148 (0.14, 0.155)
phecode_660-21	Impetigo	0.546 (0.539, 0.553)	0.687 (0.682, 0.693)	0.141 (0.132, 0.149)
phecode_660-23	Cutaneous erysipeloid	0.549 (0.531, 0.567)	0.652 (0.633, 0.67)	0.102 (0.081, 0.124)
phecode_660-4	Carbuncle and furuncle	0.54 (0.534, 0.546)	0.723 (0.719, 0.728)	0.183 (0.176, 0.189)
phecode_660-6	Cellulitis and abscess	0.537 (0.535, 0.54)	0.673 (0.671, 0.675)	0.136 (0.133, 0.138)
phecode_660-7	Pilonidal cyst	0.646 (0.629, 0.664)	0.719 (0.701, 0.737)	0.073 (0.057, 0.09)
phecode_661	Viral exanthemata NOS	0.59 (0.568, 0.616)	0.683 (0.665, 0.703)	0.093 (0.069, 0.117)
phecode_662	Rosacea	0.539 (0.535, 0.543)	0.674 (0.671, 0.678)	0.136 (0.131, 0.14)
phecode_663	Bullous disorders	0.627 (0.609, 0.644)	0.7 (0.682, 0.716)	0.073 (0.053, 0.09)
phecode_663-1	Pemphigus	0.457 (0.411, 0.502)	0.639 (0.583, 0.691)	0.18 (0.12, 0.244)
phecode_663-2	Pemphigoid	0.68 (0.658, 0.703)	0.734 (0.71, 0.757)	0.054 (0.033, 0.076)
phecode_663-6	Acantholytic Disorders*	0.746 (0.704, 0.792)	0.793 (0.75, 0.839)	0.048 (0.011, 0.084)
phecode_664	Papulosquamous disorders	0.558 (0.551, 0.564)	0.679 (0.674, 0.685)	0.121 (0.116, 0.128)
phecode_664-1	Lichen planus, nitidus, or striatus	0.6 (0.592, 0.608)	0.704 (0.698, 0.711)	0.104 (0.098, 0.11)
phecode_664-2	Pityriasis	0.588 (0.576, 0.598)	0.682 (0.674, 0.692)	0.095 (0.084, 0.107)
phecode_664-21	Pityriasis rosea	0.598 (0.586, 0.61)	0.685 (0.675, 0.697)	0.087 (0.076, 0.099)
phecode_665	Psoriasis	0.512 (0.507, 0.518)	0.666 (0.661, 0.67)	0.153 (0.147, 0.16)
phecode_665-1	Psoriasis vulgaris*	0.532 (0.515, 0.551)	0.829 (0.816, 0.845)	0.297 (0.272, 0.324)
phecode_665-2	Psoriatic arthropathy	0.534 (0.522, 0.547)	0.755 (0.74, 0.768)	0.22 (0.2, 0.239)
phecode_665-3	Other psoriasis	0.537 (0.521, 0.552)	0.726 (0.713, 0.738)	0.189 (0.172, 0.206)
phecode_666	Urticaria	0.562 (0.557, 0.566)	0.695 (0.691, 0.699)	0.133 (0.128, 0.138)
phecode_666-1	Allergic urticaria	0.578 (0.567, 0.591)	0.702 (0.692, 0.711)	0.123 (0.113, 0.135)
phecode_666-2	Idiopathic urticaria	0.571 (0.554, 0.588)	0.707 (0.692, 0.72)	0.135 (0.116, 0.153)
phecode_666-3	Urticaria due to cold and heat	0.466 (0.432, 0.5)	0.706 (0.677, 0.738)	0.241 (0.202, 0.279)
phecode_666-4	Dermatographic urticaria	0.529 (0.489, 0.573)	0.707 (0.673, 0.741)	0.176 (0.124, 0.223)
phecode_667	Erythematous conditions	0.559 (0.551, 0.568)	0.662 (0.654, 0.67)	0.103 (0.093, 0.112)
phecode_667-1	Erythema multiforme	0.514 (0.486, 0.539)	0.637 (0.613, 0.662)	0.124 (0.093, 0.154)
phecode_667-3	Erythema nodosum	0.648 (0.625, 0.676)	0.706 (0.683, 0.729)	0.058 (0.031, 0.086)
phecode_668	Dermatitis [eczema]	0.531 (0.529, 0.534)	0.695 (0.694, 0.697)	0.164 (0.161, 0.167)
phecode_668-1	Atopic dermatitis	0.516 (0.511, 0.52)	0.727 (0.723, 0.73)	0.211 (0.205, 0.216)
phecode_668-2	Seborrheic dermatitis	0.529 (0.524, 0.533)	0.698 (0.694, 0.701)	0.169 (0.164, 0.174)
phecode_668-3	Contact dermatitis	0.526 (0.521, 0.53)	0.684 (0.681, 0.688)	0.159 (0.153, 0.164)
phecode_668-4	Dermatitis due to substances taken internally	0.568 (0.554, 0.582)	0.648 (0.634, 0.662)	0.08 (0.064, 0.096)
phecode_668-5	Lichen simplex chronicus	0.539 (0.527, 0.551)	0.716 (0.706, 0.725)	0.176 (0.163, 0.19)
phecode_668-6	Prurigo	0.549 (0.536, 0.564)	0.74 (0.729, 0.752)	0.19 (0.174, 0.207)
phecode_669	Neutrophilic and Eosinophilic dermatoses	0.552 (0.508, 0.596)	0.696 (0.655, 0.74)	0.144 (0.079, 0.204)
phecode_670	Seborrheic keratosis	0.627 (0.621, 0.634)	0.678 (0.671, 0.684)	0.051 (0.044, 0.056)
phecode_671	Sunburn	0.5 (0.481, 0.522)	0.728 (0.712, 0.743)	0.228 (0.202, 0.25)
phecode_672	Other acute skin changes due to ultraviolet radiation	0.596 (0.583, 0.61)	0.731 (0.719, 0.742)	0.135 (0.123, 0.149)
phecode_672-1	Acute dermatitis due to solar radiation	0.645 (0.62, 0.666)	0.751 (0.732, 0.773)	0.107 (0.088, 0.128)
phecode_672-4	Disseminated superficial actinic porokeratosis (DSAP)	0.721 (0.69, 0.753)	0.786 (0.758, 0.816)	0.066 (0.044, 0.087)
phecode_673	Skin changes due to chronic exposure to non-ionizing radiation	0.593 (0.592, 0.595)	0.718 (0.717, 0.719)	0.125 (0.123, 0.126)
phecode_673-1	Actinic keratosis	0.594 (0.593, 0.596)	0.718 (0.717, 0.72)	0.124 (0.123, 0.125)
phecode_674	Disorders of pigmentation	0.55 (0.545, 0.554)	0.689 (0.685, 0.692)	0.139 (0.134, 0.144)
phecode_674-1	Hypopigmentation	0.542 (0.528, 0.557)	0.673 (0.663, 0.686)	0.131 (0.114, 0.146)
phecode_674-11	Vitiligo	0.526 (0.511, 0.541)	0.664 (0.65, 0.676)	0.139 (0.12, 0.157)
phecode_674-2	Hypopigmentation	0.588 (0.581, 0.595)	0.71 (0.704, 0.717)	0.122 (0.115, 0.13)
phecode_674-22	Freckles*	0.607 (0.584, 0.629)	0.742 (0.723, 0.76)	0.134 (0.112, 0.16)
phecode_674-23	Acanthosis nigricans	0.558 (0.519, 0.59)	0.637 (0.609, 0.668)	0.081 (0.041, 0.122)
phecode_675	Atrophic conditions of skin	0.716 (0.711, 0.721)	0.789 (0.784, 0.793)	0.073 (0.069, 0.077)
phecode_675-1	Circumscribed scleroderma	0.718 (0.713, 0.722)	0.791 (0.787, 0.796)	0.073 (0.07, 0.078)
phecode_676	Hypertrophic conditions of skin	0.508 (0.504, 0.511)	0.68 (0.677, 0.682)	0.172 (0.168, 0.176)
phecode_676-1	Hypertrophic scar [Keloid scar]	0.533 (0.517, 0.547)	0.71 (0.698, 0.723)	0.177 (0.161, 0.194)
phecode_676-2	Scar conditions and fibrosis of skin	0.548 (0.541, 0.556)	0.658 (0.651, 0.666)	0.11 (0.1, 0.119)
phecode_678	Other skin and connective tissue disorders	0.549 (0.547, 0.55)	0.701 (0.7, 0.703)	0.153 (0.151, 0.154)

4 Medical history predicts future health trajectories over the human phenome

Table 17 continued from previous page

Endpoint	PheCode string	C-Index Age+Sex (IQR)	C-Index Age+Sex+MedicalHistory (IQR)	Difference in C-Index (IQR)
phecode_679	Skin symptoms	0.537 (0.535, 0.539)	0.714 (0.712, 0.715)	0.176 (0.175, 0.178)
phecode_679-1	Rash and other nonspecific skin eruption	0.531 (0.529, 0.533)	0.714 (0.712, 0.716)	0.183 (0.181, 0.185)
phecode_679-2	Pallor and flushing	0.595 (0.587, 0.603)	0.759 (0.751, 0.766)	0.164 (0.156, 0.172)
phecode_679-21	Pallor	0.571 (0.554, 0.587)	0.746 (0.733, 0.758)	0.175 (0.157, 0.192)
phecode_679-22	Flushing	0.649 (0.639, 0.659)	0.786 (0.777, 0.794)	0.137 (0.127, 0.146)
phecode_679-3	Changes in skin texture	0.578 (0.572, 0.584)	0.752 (0.747, 0.757)	0.175 (0.168, 0.181)
phecode_679-4	Pruritus	0.555 (0.552, 0.558)	0.731 (0.729, 0.733)	0.176 (0.174, 0.179)
phecode_679-5	Pyoderma	0.516 (0.494, 0.536)	0.703 (0.687, 0.721)	0.188 (0.164, 0.212)
phecode_679-7	Abnormal granulation tissue, NOS	0.613 (0.601, 0.623)	0.761 (0.753, 0.769)	0.148 (0.138, 0.159)
phecode_680	Epidermal thickening	0.59 (0.587, 0.594)	0.721 (0.718, 0.724)	0.131 (0.127, 0.134)
phecode_680-1	Corns and callosities	0.567 (0.562, 0.572)	0.724 (0.72, 0.729)	0.158 (0.152, 0.163)
phecode_680-2	Keratoderma	0.546 (0.515, 0.579)	0.683 (0.654, 0.713)	0.137 (0.097, 0.177)
phecode_680-3	Xerosis cutis*	0.586 (0.58, 0.591)	0.76 (0.756, 0.765)	0.175 (0.169, 0.18)
phecode_681	Localized swelling, mass and lump of skin and subcutaneous tissue	0.555 (0.553, 0.557)	0.711 (0.709, 0.713)	0.156 (0.154, 0.159)
phecode_682	Other follicular disorders	0.513 (0.51, 0.515)	0.656 (0.654, 0.658)	0.144 (0.14, 0.147)
phecode_682-1	Cutaneous cyst	0.501 (0.498, 0.503)	0.644 (0.641, 0.647)	0.144 (0.14, 0.147)
phecode_682-11	Sebaceous cyst [Epidermal cyst]	0.498 (0.495, 0.502)	0.658 (0.655, 0.661)	0.16 (0.155, 0.164)
phecode_682-12	Pilar and trichodermal cyst	0.497 (0.495, 0.501)	0.65 (0.647, 0.652)	0.152 (0.148, 0.156)
phecode_682-3	Hidradenitis suppurativa	0.692 (0.672, 0.711)	0.763 (0.746, 0.779)	0.07 (0.053, 0.089)
phecode_682-4	Acne	0.624 (0.615, 0.634)	0.711 (0.703, 0.718)	0.086 (0.077, 0.095)
phecode_683	Nail disorders	0.539 (0.536, 0.542)	0.704 (0.702, 0.707)	0.165 (0.162, 0.17)
phecode_683-1	Ingrowing nail	0.548 (0.543, 0.553)	0.72 (0.716, 0.724)	0.173 (0.167, 0.178)
phecode_683-2	Nail dystrophy*	0.565 (0.553, 0.575)	0.725 (0.716, 0.734)	0.16 (0.149, 0.174)
phecode_683-4	Onycholysis*	0.574 (0.554, 0.592)	0.708 (0.694, 0.723)	0.134 (0.114, 0.154)
phecode_684	Diseases of hair and hair follicles	0.692 (0.688, 0.697)	0.79 (0.787, 0.794)	0.098 (0.094, 0.102)
phecode_684-1	Alopecia	0.687 (0.683, 0.691)	0.789 (0.785, 0.792)	0.102 (0.097, 0.106)
phecode_684-11	Alopecia Areata	0.682 (0.675, 0.688)	0.796 (0.79, 0.802)	0.115 (0.108, 0.121)
phecode_684-12	Telogen effluvium	0.68 (0.657, 0.702)	0.849 (0.833, 0.866)	0.17 (0.147, 0.195)
phecode_684-13	Scarring hair loss*	0.598 (0.576, 0.622)	0.732 (0.71, 0.752)	0.133 (0.108, 0.159)
phecode_684-14	Androgenic alopecia*	0.651 (0.631, 0.671)	0.776 (0.756, 0.795)	0.125 (0.104, 0.147)
phecode_684-2	Hypertrichosis and hirsutism	0.77 (0.756, 0.784)	0.841 (0.83, 0.855)	0.071 (0.058, 0.085)
phecode_685	Disorders of sweat glands	0.547 (0.543, 0.551)	0.725 (0.721, 0.728)	0.177 (0.172, 0.182)
phecode_685-1	Dyshidrosis	0.568 (0.559, 0.576)	0.713 (0.705, 0.719)	0.145 (0.136, 0.154)
phecode_685-4	Prickly heat and miliaria	0.542 (0.527, 0.556)	0.723 (0.712, 0.736)	0.182 (0.163, 0.198)
phecode_685-8	Hyperhidrosis	0.542 (0.537, 0.547)	0.734 (0.731, 0.739)	0.193 (0.186, 0.198)
phecode_685-82	Generalized hyperhidrosis	0.542 (0.537, 0.546)	0.735 (0.73, 0.739)	0.193 (0.187, 0.198)
phecode_686	Chronic ulcer of skin	0.678 (0.674, 0.683)	0.795 (0.791, 0.798)	0.116 (0.112, 0.121)
phecode_686-1	Pressure ulcer	0.709 (0.702, 0.716)	0.834 (0.829, 0.84)	0.125 (0.119, 0.132)
phecode_686-2	Non-pressure chronic ulcer	0.663 (0.657, 0.668)	0.802 (0.797, 0.807)	0.139 (0.134, 0.145)
phecode_687	Vascular disorders of the skin	0.579 (0.555, 0.606)	0.7 (0.674, 0.728)	0.12 (0.089, 0.155)
phecode_688	Granulomatous disorder of the skin	0.521 (0.513, 0.528)	0.624 (0.617, 0.631)	0.104 (0.094, 0.112)
phecode_688-1	Sarcoidosis	0.524 (0.507, 0.541)	0.598 (0.579, 0.613)	0.073 (0.053, 0.094)
phecode_688-2	Foreign body granuloma	0.377 (0.348, 0.409)	0.566 (0.533, 0.603)	0.19 (0.145, 0.24)
phecode_688-3	Pyogenic granuloma of skin and subcutaneous tissue	0.522 (0.509, 0.536)	0.645 (0.633, 0.654)	0.123 (0.109, 0.138)
phecode_688-6	Rheumatoid nodule*	0.641 (0.608, 0.675)	0.913 (0.889, 0.938)	0.274 (0.235, 0.308)
phecode_700	Diffuse diseases of connective tissue	0.63 (0.621, 0.638)	0.727 (0.718, 0.735)	0.097 (0.088, 0.107)
phecode_700-1	Lupus	0.64 (0.623, 0.659)	0.731 (0.713, 0.752)	0.091 (0.07, 0.112)
phecode_700-11	Systemic lupus erythematosus [SLE]	0.656 (0.636, 0.68)	0.706 (0.677, 0.732)	0.048 (0.013, 0.083)
phecode_700-12	Cutaneous lupus erythematosus	0.641 (0.619, 0.664)	0.545 (0.513, 0.574)	-0.098 (-0.139, -0.057)
phecode_700-2	Sicca syndrome [Sjogren syndrome]	0.693 (0.682, 0.702)	0.769 (0.757, 0.78)	0.076 (0.064, 0.089)
phecode_700-3	Systemic sclerosis	0.646 (0.624, 0.669)	0.621 (0.592, 0.649)	-0.025 (-0.054, 0.006)
phecode_700-4	Inflammatory myopathy	0.599 (0.579, 0.617)	0.679 (0.661, 0.697)	0.081 (0.06, 0.101)

Supplementary Tables

Table 17 continued from previous page

Endpoint	PheCode string	C-Index Age+Sex (IQR)	C-Index Age+Sex+MedicalHistory (IQR)	Difference in C-Index (IQR)
phecode_700-41	Myositis	0.614 (0.595, 0.635)	0.714 (0.694, 0.734)	0.099 (0.077, 0.125)
phecode_700-43	Polymyositis	0.571 (0.529, 0.608)	0.623 (0.58, 0.677)	0.056 (0.003, 0.112)
phecode_700-9	Other diffuse diseases of connective tissue	0.61 (0.591, 0.628)	0.785 (0.764, 0.802)	0.175 (0.153, 0.198)
phecode_701	Osteomyelitis, periostitis, and other infections involving bone	0.616 (0.603, 0.627)	0.766 (0.753, 0.777)	0.15 (0.137, 0.165)
phecode_701-1	Osteomyelitis	0.612 (0.599, 0.625)	0.767 (0.756, 0.78)	0.156 (0.141, 0.17)
phecode_701-11	Acute/subacute osteomyelitis	0.651 (0.624, 0.685)	0.83 (0.808, 0.855)	0.178 (0.146, 0.212)
phecode_701-12	Chronic osteomyelitis	0.603 (0.576, 0.637)	0.777 (0.746, 0.808)	0.174 (0.134, 0.215)
phecode_702	Infective and reactive arthropathies	0.543 (0.532, 0.554)	0.762 (0.752, 0.772)	0.219 (0.207, 0.232)
phecode_702-1	Pyogenic arthritis [Septic arthritis]	0.611 (0.595, 0.629)	0.731 (0.715, 0.748)	0.121 (0.1, 0.138)
phecode_702-2	Reiter's disease [Reactive arthritis]	0.395 (0.364, 0.43)	0.724 (0.696, 0.751)	0.329 (0.283, 0.369)
phecode_702-3	Enteropathic arthropathies	0.491 (0.475, 0.505)	0.855 (0.844, 0.868)	0.364 (0.349, 0.382)
phecode_703	Chrysal arthropathies	0.692 (0.689, 0.695)	0.759 (0.756, 0.762)	0.067 (0.064, 0.069)
phecode_703-1	Hyperuricemia	0.697 (0.694, 0.7)	0.764 (0.761, 0.767)	0.067 (0.064, 0.07)
phecode_703-11	Gout	0.7 (0.698, 0.704)	0.767 (0.764, 0.769)	0.066 (0.063, 0.069)
phecode_703-2	Chondrocalcinosis	0.664 (0.651, 0.676)	0.762 (0.751, 0.773)	0.098 (0.086, 0.109)
phecode_704	Systemic vasculitis	0.67 (0.661, 0.678)	0.708 (0.7, 0.717)	0.039 (0.032, 0.046)
phecode_704-1	Polyarteritis nodosa and related conditions	0.616 (0.571, 0.661)	0.734 (0.682, 0.783)	0.119 (0.069, 0.165)
phecode_704-2	Wegener's granulomatosis	0.571 (0.539, 0.605)	0.659 (0.613, 0.7)	0.087 (0.042, 0.128)
phecode_704-5	Giant cell arteritis	0.733 (0.723, 0.743)	0.751 (0.74, 0.761)	0.018 (0.01, 0.025)
phecode_705	Rheumatoid arthritis and other inflammatory polyarthropathies	0.637 (0.633, 0.641)	0.704 (0.701, 0.708)	0.068 (0.065, 0.071)
phecode_705-1	Rheumatoid arthritis	0.614 (0.61, 0.619)	0.73 (0.725, 0.735)	0.115 (0.109, 0.122)
phecode_705-3	Polymyalgia rheumatica	0.722 (0.717, 0.726)	0.744 (0.739, 0.749)	0.022 (0.019, 0.025)
phecode_705-4	Palindromic rheumatism	0.58 (0.547, 0.611)	0.712 (0.682, 0.742)	0.132 (0.099, 0.171)
phecode_705-5	Rheumatism, unspecified	0.532 (0.521, 0.544)	0.827 (0.819, 0.834)	0.295 (0.28, 0.308)
phecode_706	Other inflammatory spondylopathies	0.593 (0.587, 0.599)	0.745 (0.739, 0.751)	0.152 (0.145, 0.159)
phecode_706-1	Sacroiliitis NEC	0.605 (0.593, 0.618)	0.739 (0.729, 0.75)	0.134 (0.122, 0.148)
phecode_706-2	Ankylosing spondylitis	0.535 (0.513, 0.557)	0.719 (0.698, 0.739)	0.184 (0.156, 0.214)
phecode_706-6	Discitis*	0.666 (0.642, 0.688)	0.736 (0.715, 0.759)	0.07 (0.046, 0.097)
phecode_707	Other arthropathies	0.632 (0.63, 0.635)	0.738 (0.736, 0.741)	0.106 (0.104, 0.109)
phecode_707-1	Villonodular synovitis	0.481 (0.439, 0.526)	0.597 (0.554, 0.642)	0.117 (0.058, 0.167)
phecode_707-4	Arthropathy associated with neurological disorders	0.616 (0.581, 0.648)	0.91 (0.887, 0.93)	0.295 (0.256, 0.333)
phecode_707-8	Polyarthritis	0.599 (0.587, 0.611)	0.749 (0.738, 0.759)	0.15 (0.137, 0.164)
phecode_707-9	Monoarthritis	0.498 (0.471, 0.527)	0.647 (0.623, 0.673)	0.151 (0.117, 0.184)
phecode_708	Osteoarthritis	0.61 (0.609, 0.612)	0.691 (0.69, 0.693)	0.081 (0.08, 0.083)
phecode_708-1	Primary osteoarthritis	0.609 (0.606, 0.612)	0.698 (0.695, 0.7)	0.089 (0.086, 0.091)
phecode_708-11	Primary osteoarthritis of hip, pelvic region and thigh	0.666 (0.659, 0.672)	0.712 (0.705, 0.718)	0.046 (0.041, 0.051)
phecode_708-12	Primary osteoarthritis of knee, lower leg	0.61 (0.605, 0.614)	0.707 (0.702, 0.711)	0.097 (0.092, 0.102)
phecode_708-13	Primary osteoarthritis of the hand	0.643 (0.638, 0.647)	0.747 (0.743, 0.751)	0.105 (0.101, 0.109)
phecode_708-14	Primary osteoarthritis of the shoulder, upper arm	0.617 (0.611, 0.624)	0.756 (0.75, 0.761)	0.138 (0.132, 0.145)
phecode_708-15	Primary osteoarthritis of the wrist, forearm	0.638 (0.625, 0.65)	0.769 (0.759, 0.78)	0.131 (0.12, 0.144)
phecode_708-16	Primary osteoarthritis ankle and foot	0.58 (0.575, 0.586)	0.735 (0.73, 0.74)	0.155 (0.149, 0.161)
phecode_708-7	Generalized osteoarthritis	0.661 (0.658, 0.664)	0.793 (0.79, 0.795)	0.132 (0.127, 0.135)
phecode_708-8	Secondary osteoarthritis	0.577 (0.569, 0.586)	0.724 (0.717, 0.731)	0.147 (0.138, 0.156)
phecode_708-9	Heberden's or Bouchard's nodes*	0.637 (0.629, 0.643)	0.732 (0.725, 0.738)	0.095 (0.089, 0.101)
phecode_709	Acquired deformities of fingers and toes	0.638 (0.635, 0.64)	0.697 (0.694, 0.7)	0.06 (0.057, 0.062)
phecode_709-1	Acquired deformities of fingers	0.496 (0.487, 0.506)	0.657 (0.648, 0.666)	0.161 (0.148, 0.174)
phecode_709-11	Mallet finger	0.513 (0.499, 0.526)	0.665 (0.654, 0.676)	0.151 (0.137, 0.167)
phecode_709-2	Acquired deformities of toe	0.651 (0.648, 0.654)	0.707 (0.704, 0.71)	0.056 (0.054, 0.059)
phecode_709-21	Hallux valgus (Bunion)	0.667 (0.664, 0.67)	0.722 (0.719, 0.725)	0.055 (0.052, 0.057)

4 Medical history predicts future health trajectories over the human phenome

Table 17 continued from previous page

Endpoint	PheCode string	C-Index Age+Sex (IQR)	C-Index Age+Sex+MedicalHistory (IQR)	Difference in C-Index (IQR)
phecode_709-22	Hallux rigidus	0.591 (0.583, 0.6)	0.69 (0.682, 0.699)	0.099 (0.09, 0.108)
phecode_709-23	Hallux varus (acquired)	0.626 (0.605, 0.646)	0.698 (0.682, 0.718)	0.072 (0.049, 0.098)
phecode_709-24	Hammer toe	0.69 (0.684, 0.696)	0.748 (0.742, 0.755)	0.058 (0.053, 0.064)
phecode_710	Acquired deformities of limbs	0.575 (0.569, 0.58)	0.72 (0.715, 0.724)	0.144 (0.138, 0.151)
phecode_710-1	Acquired deformities of the forearm and hand	0.495 (0.456, 0.536)	0.689 (0.652, 0.729)	0.195 (0.14, 0.251)
phecode_710-2	Acquired deformities of the hip	0.636 (0.598, 0.67)	0.703 (0.665, 0.739)	0.068 (0.024, 0.107)
phecode_710-3	Acquired deformities of the knee	0.639 (0.628, 0.65)	0.755 (0.746, 0.766)	0.117 (0.105, 0.128)
phecode_710-31	Genu valgum (acquired)	0.662 (0.646, 0.678)	0.75 (0.735, 0.765)	0.088 (0.072, 0.104)
phecode_710-32	Genu varum (acquired)	0.655 (0.639, 0.67)	0.753 (0.738, 0.77)	0.099 (0.081, 0.117)
phecode_710-4	Acquired deformities of the ankle and foot	0.55 (0.543, 0.558)	0.729 (0.723, 0.735)	0.179 (0.17, 0.186)
phecode_710-41	Flat foot [pes planus]	0.552 (0.542, 0.563)	0.717 (0.709, 0.726)	0.164 (0.152, 0.177)
phecode_710-43	Cavus deformity of foot	0.481 (0.447, 0.518)	0.708 (0.679, 0.738)	0.226 (0.179, 0.264)
phecode_710-44	Foot drop (acquired)*	0.599 (0.586, 0.615)	0.746 (0.733, 0.759)	0.146 (0.128, 0.163)
phecode_710-6	Unequal limb length (acquired)	0.57 (0.538, 0.6)	0.738 (0.715, 0.763)	0.169 (0.137, 0.202)
phecode_711	Disorder of patella	0.537 (0.533, 0.54)	0.669 (0.665, 0.672)	0.132 (0.127, 0.137)
phecode_711-1	Derangement of meniscus	0.538 (0.534, 0.542)	0.674 (0.671, 0.678)	0.136 (0.132, 0.142)
phecode_712	Other specific joint derangements	0.546 (0.54, 0.552)	0.704 (0.699, 0.709)	0.157 (0.15, 0.165)
phecode_712-1	Loose body in joint	0.517 (0.506, 0.529)	0.688 (0.676, 0.7)	0.171 (0.155, 0.187)
phecode_712-3	Articular cartilage disorder	0.62 (0.604, 0.636)	0.748 (0.735, 0.762)	0.128 (0.109, 0.147)
phecode_712-4	Ankylosis of joint	0.49 (0.456, 0.52)	0.724 (0.689, 0.755)	0.236 (0.191, 0.281)
phecode_712-5	Disorder of ligament	0.639 (0.623, 0.654)	0.741 (0.728, 0.755)	0.103 (0.086, 0.119)
phecode_712-51	Hypermobility syndrome	0.798 (0.779, 0.815)	0.858 (0.842, 0.874)	0.061 (0.043, 0.079)
phecode_712-6	Instability of joint	0.527 (0.515, 0.538)	0.716 (0.705, 0.728)	0.19 (0.174, 0.204)
phecode_712-61	Recurrent dislocation of joint	0.513 (0.493, 0.534)	0.721 (0.699, 0.739)	0.208 (0.18, 0.235)
phecode_713	Symptoms related to joints	0.523 (0.522, 0.525)	0.724 (0.723, 0.725)	0.201 (0.199, 0.202)
phecode_713-1	Hemarthrosis	0.653 (0.629, 0.675)	0.737 (0.709, 0.761)	0.084 (0.058, 0.108)
phecode_713-2	Effusion of joint	0.589 (0.585, 0.592)	0.733 (0.73, 0.736)	0.144 (0.141, 0.147)
phecode_713-3	Pain in joint	0.523 (0.522, 0.524)	0.725 (0.724, 0.726)	0.202 (0.2, 0.204)
phecode_713-4	Stiffness of joint	0.552 (0.545, 0.559)	0.702 (0.695, 0.709)	0.15 (0.142, 0.158)
phecode_714	Deforming dorsopathies	0.664 (0.657, 0.671)	0.745 (0.739, 0.751)	0.081 (0.074, 0.087)
phecode_714-1	Kyphosis	0.665 (0.648, 0.683)	0.752 (0.736, 0.768)	0.088 (0.073, 0.103)
phecode_714-2	Lordosis	0.54 (0.501, 0.584)	0.68 (0.647, 0.718)	0.14 (0.092, 0.19)
phecode_714-3	Scoliosis	0.687 (0.679, 0.694)	0.76 (0.753, 0.767)	0.073 (0.066, 0.08)
phecode_714-31	Idiopathic scoliosis	0.654 (0.627, 0.683)	0.798 (0.776, 0.818)	0.143 (0.117, 0.169)
phecode_714-32	Thoracogenic scoliosis	0.664 (0.632, 0.694)	0.777 (0.752, 0.799)	0.114 (0.082, 0.147)
phecode_714-7	Fusion of spine*	0.604 (0.561, 0.647)	0.803 (0.765, 0.848)	0.201 (0.149, 0.249)
phecode_715	Non-inflammatory spondylopathy	0.607 (0.605, 0.61)	0.724 (0.722, 0.727)	0.118 (0.115, 0.12)
phecode_715-1	Spondylolysis	0.601 (0.598, 0.603)	0.734 (0.731, 0.737)	0.133 (0.13, 0.136)
phecode_715-2	Spondylolysis	0.647 (0.628, 0.665)	0.789 (0.775, 0.805)	0.142 (0.121, 0.163)
phecode_715-3	Spondylolisthesis	0.658 (0.651, 0.666)	0.756 (0.749, 0.763)	0.098 (0.092, 0.105)
phecode_715-4	Spinal stenosis	0.637 (0.633, 0.641)	0.75 (0.746, 0.754)	0.114 (0.109, 0.118)
phecode_715-5	Ankylosing hyperostosis [Forestier]	0.73 (0.703, 0.761)	0.814 (0.787, 0.841)	0.083 (0.051, 0.114)
phecode_716	Intervertebral disc disorder	0.527 (0.523, 0.53)	0.709 (0.706, 0.712)	0.183 (0.178, 0.187)
phecode_716-1	Schmorl's nodes	0.513 (0.462, 0.566)	0.721 (0.674, 0.77)	0.21 (0.146, 0.268)
phecode_716-2	Degenerative disc disease	0.541 (0.534, 0.548)	0.732 (0.727, 0.738)	0.191 (0.183, 0.198)
phecode_716-3	Spinal disc displacement (herniation)	0.528 (0.524, 0.532)	0.719 (0.716, 0.723)	0.191 (0.186, 0.197)
phecode_717	Other and unspecified dorsopathies	0.575 (0.569, 0.58)	0.73 (0.726, 0.735)	0.155 (0.15, 0.162)
phecode_717-2	Sacrococcygeal disorders	0.606 (0.599, 0.613)	0.738 (0.732, 0.743)	0.132 (0.124, 0.139)
phecode_718	Back pain	0.519 (0.517, 0.521)	0.72 (0.719, 0.722)	0.201 (0.2, 0.203)
phecode_718-1	Radiculopathy	0.56 (0.556, 0.565)	0.751 (0.747, 0.755)	0.191 (0.186, 0.197)
phecode_718-2	Cervicalgia	0.541 (0.539, 0.543)	0.732 (0.73, 0.734)	0.191 (0.189, 0.194)

Supplementary Tables

Table 17 continued from previous page

Endpoint	PheCode string	C-Index Age+Sex (IQR)	C-Index Age+Sex+MedicalHistory (IQR)	Difference in C-Index (IQR)
phecode_718-3	Mid back pain	0.584 (0.578, 0.59)	0.743 (0.738, 0.748)	0.159 (0.153, 0.165)
phecode_718-4	Low back pain	0.511 (0.509, 0.513)	0.721 (0.72, 0.723)	0.21 (0.208, 0.213)
phecode_718-5	Sciatica	0.542 (0.539, 0.544)	0.698 (0.696, 0.7)	0.156 (0.153, 0.159)
phecode_719	Disorders of muscle	0.574 (0.571, 0.577)	0.733 (0.73, 0.736)	0.159 (0.156, 0.163)
phecode_719-1	Cramp and spasm	0.583 (0.579, 0.586)	0.753 (0.75, 0.756)	0.17 (0.166, 0.174)
phecode_719-11	Spasm of muscle	0.513 (0.5, 0.528)	0.709 (0.698, 0.724)	0.195 (0.178, 0.216)
phecode_719-3	Separation of muscle (nontraumatic)	0.712 (0.701, 0.722)	0.826 (0.817, 0.835)	0.114 (0.105, 0.122)
phecode_719-4	Rupture of muscle (nontraumatic)	0.673 (0.649, 0.695)	0.788 (0.77, 0.805)	0.115 (0.093, 0.141)
phecode_719-6	Muscle wasting and atrophy	0.613 (0.594, 0.632)	0.723 (0.706, 0.741)	0.111 (0.091, 0.131)
phecode_719-7	Muscle weakness (generalized)	0.569 (0.559, 0.58)	0.759 (0.751, 0.767)	0.19 (0.179, 0.201)
phecode_719-8	Rhabdomyolysis	0.578 (0.527, 0.628)	0.696 (0.658, 0.738)	0.118 (0.068, 0.17)
phecode_719-9	Contractures	0.531 (0.514, 0.548)	0.673 (0.659, 0.689)	0.142 (0.121, 0.164)
phecode_719-92	Contracture of joint	0.518 (0.499, 0.537)	0.673 (0.655, 0.692)	0.155 (0.13, 0.183)
phecode_719-93	Contracture of muscle*	0.563 (0.523, 0.596)	0.684 (0.642, 0.724)	0.122 (0.071, 0.17)
phecode_720	Spontaneous rupture of synovium and tendon	0.602 (0.595, 0.61)	0.704 (0.698, 0.711)	0.101 (0.095, 0.108)
phecode_721	Synoviopathy and bursopathy	0.532 (0.53, 0.534)	0.687 (0.686, 0.689)	0.155 (0.153, 0.157)
phecode_721-1	Synovitis and tenosynovitis	0.534 (0.532, 0.537)	0.694 (0.692, 0.697)	0.16 (0.157, 0.163)
phecode_721-11	Trigger finger	0.57 (0.566, 0.574)	0.7 (0.697, 0.704)	0.13 (0.126, 0.135)
phecode_721-12	Radial styloid tenosynovitis [de Quervain]	0.62 (0.612, 0.627)	0.734 (0.727, 0.741)	0.114 (0.107, 0.121)
phecode_721-15	Infective (teno)synovitis*	0.549 (0.511, 0.593)	0.657 (0.615, 0.697)	0.106 (0.054, 0.154)
phecode_721-2	Ganglion cyst	0.556 (0.552, 0.56)	0.678 (0.674, 0.681)	0.122 (0.118, 0.126)
phecode_721-4	Calcium deposits in tendon and bursa	0.579 (0.566, 0.591)	0.685 (0.674, 0.699)	0.107 (0.091, 0.122)
phecode_721-5	Bursitis	0.531 (0.528, 0.534)	0.697 (0.694, 0.699)	0.166 (0.163, 0.169)
phecode_721-6	Baker's cyst [popliteal cyst]	0.582 (0.576, 0.587)	0.693 (0.688, 0.698)	0.112 (0.106, 0.117)
phecode_721-7	Synovial hypertrophy*	0.521 (0.474, 0.561)	0.7 (0.662, 0.739)	0.18 (0.129, 0.234)
phecode_722	Fasciopathy	0.526 (0.523, 0.528)	0.675 (0.673, 0.677)	0.15 (0.147, 0.153)
phecode_722-1	Plantar fascial fibromatosis [Plantar fasciitis]	0.529 (0.526, 0.532)	0.697 (0.694, 0.699)	0.167 (0.164, 0.17)
phecode_722-4	Palmar fascial fibromatosis [Dupuytren]	0.633 (0.629, 0.637)	0.688 (0.684, 0.692)	0.055 (0.051, 0.058)
phecode_723	Enthesopathy/Enthesitis/Tendinopathy	0.54 (0.538, 0.542)	0.692 (0.691, 0.694)	0.152 (0.15, 0.154)
phecode_723-1	Adhesive capsulitis of shoulder	0.544 (0.54, 0.548)	0.69 (0.686, 0.694)	0.146 (0.141, 0.151)
phecode_723-2	Rotator cuff tear or rupture	0.544 (0.541, 0.547)	0.697 (0.694, 0.7)	0.153 (0.15, 0.157)
phecode_723-3	Medial epicondylitis (Golfer's elbow)	0.555 (0.549, 0.562)	0.722 (0.716, 0.727)	0.167 (0.16, 0.173)
phecode_723-4	Lateral epicondylitis (Tennis elbow)	0.639 (0.636, 0.641)	0.735 (0.733, 0.737)	0.097 (0.094, 0.099)
phecode_723-5	Tendinitis	0.543 (0.54, 0.547)	0.692 (0.689, 0.695)	0.148 (0.144, 0.153)
phecode_723-51	Achilles tendinitis	0.551 (0.547, 0.555)	0.692 (0.688, 0.695)	0.141 (0.137, 0.146)
phecode_723-52	Posterior tibial tendinitis	0.585 (0.56, 0.609)	0.725 (0.704, 0.745)	0.139 (0.116, 0.165)
phecode_723-6	Impingement syndrome of shoulder*	0.502 (0.499, 0.506)	0.701 (0.698, 0.704)	0.198 (0.193, 0.203)
phecode_723-7	Iliotibial band syndrome*	0.561 (0.544, 0.579)	0.722 (0.707, 0.736)	0.161 (0.142, 0.178)
phecode_724	Other symptoms and disorders of the soft tissue	0.548 (0.546, 0.551)	0.693 (0.692, 0.695)	0.145 (0.143, 0.147)
phecode_724-1	Myalgia	0.551 (0.548, 0.554)	0.729 (0.727, 0.732)	0.179 (0.175, 0.182)
phecode_724-3	Nontraumatic hematoma of soft tissue	0.583 (0.573, 0.592)	0.759 (0.75, 0.768)	0.177 (0.167, 0.187)
phecode_724-4	Panniculitis	0.497 (0.46, 0.54)	0.754 (0.726, 0.785)	0.258 (0.209, 0.305)
phecode_724-5	Exostosis	0.558 (0.552, 0.564)	0.703 (0.698, 0.709)	0.146 (0.139, 0.153)
phecode_724-51	Calcaneal spur	0.518 (0.504, 0.531)	0.718 (0.707, 0.729)	0.2 (0.184, 0.216)
phecode_724-52	Osteophyte*	0.576 (0.57, 0.583)	0.713 (0.707, 0.719)	0.136 (0.129, 0.144)
phecode_725	Developmental disorders of bone	0.548 (0.521, 0.573)	0.704 (0.681, 0.726)	0.156 (0.128, 0.186)
phecode_725-2	Juvenile osteochondrosis	0.551 (0.522, 0.579)	0.703 (0.682, 0.725)	0.153 (0.122, 0.185)
phecode_726	Osteoporosis and low bone density	0.745 (0.742, 0.747)	0.779 (0.777, 0.782)	0.035 (0.033, 0.037)
phecode_726-1	Osteoporosis	0.751 (0.749, 0.753)	0.786 (0.784, 0.788)	0.035 (0.033, 0.037)
phecode_726-2	Pathologic fracture	0.729 (0.721, 0.737)	0.783 (0.775, 0.791)	0.054 (0.047, 0.061)
phecode_726-4	Stress fracture	0.609 (0.581, 0.638)	0.708 (0.679, 0.743)	0.1 (0.069, 0.132)
phecode_726-5	Osteomalacia	0.52 (0.477, 0.562)	0.765 (0.73, 0.803)	0.245 (0.189, 0.302)
phecode_726-6	Osteolysis*	0.651 (0.616, 0.689)	0.809 (0.774, 0.842)	0.157 (0.112, 0.194)
phecode_727	Other disorders of bone	0.703 (0.701, 0.705)	0.758 (0.756, 0.76)	0.055 (0.053, 0.057)
phecode_727-1	Osteonecrosis	0.58 (0.565, 0.595)	0.696 (0.683, 0.712)	0.116 (0.097, 0.134)

4 Medical history predicts future health trajectories over the human phenome

Table 17 continued from previous page

Endpoint	PheCode string	C-Index Age+Sex (IQR)	C-Index Age+Sex+MedicalHistory (IQR)	Difference in C-Index (IQR)
phecode_727-3	Osteitis deformans [Paget's disease of bone]	0.689 (0.663, 0.716)	0.701 (0.676, 0.729)	0.012 (-0.009, 0.037)
phecode_727-5	Cyst of bone	0.574 (0.551, 0.597)	0.677 (0.655, 0.7)	0.103 (0.08, 0.129)
phecode_727-6	Hypertrophy of bone*	0.483 (0.443, 0.523)	0.768 (0.731, 0.805)	0.284 (0.239, 0.33)
phecode_728	Chondropathies	0.6 (0.595, 0.605)	0.727 (0.722, 0.731)	0.127 (0.121, 0.132)
phecode_728-1	Chondromalacia	0.63 (0.617, 0.643)	0.7 (0.689, 0.712)	0.07 (0.057, 0.083)
phecode_728-3	Costochondritis [Tietze's disease]	0.611 (0.605, 0.617)	0.75 (0.746, 0.756)	0.14 (0.134, 0.146)
phecode_729	Other acquired musculoskeletal deformity	0.511 (0.481, 0.538)	0.641 (0.61, 0.666)	0.13 (0.087, 0.164)
phecode_729-1	Acquired deformity of nose	0.617 (0.57, 0.669)	0.646 (0.608, 0.691)	0.028 (-0.018, 0.076)
phecode_730	Other disorders and symptoms of the musculoskeletal system	0.525 (0.518, 0.532)	0.837 (0.833, 0.842)	0.312 (0.305, 0.32)
phecode_730-2	Ehlers-Danlos syndrome	0.771 (0.739, 0.807)	0.852 (0.823, 0.886)	0.081 (0.051, 0.109)
phecode_731	Symptoms involving musculoskeletal systems	0.524 (0.519, 0.528)	0.731 (0.727, 0.734)	0.207 (0.202, 0.212)
phecode_731-1	Loss of height	0.742 (0.715, 0.774)	0.782 (0.757, 0.81)	0.04 (0.017, 0.06)
phecode_732	Nonspecific abnormal findings on radiological and other examination of musculoskeletal system	0.62 (0.616, 0.624)	0.739 (0.736, 0.742)	0.119 (0.116, 0.123)
phecode_733	Dentofacial anomalies, including malocclusion	0.584 (0.579, 0.59)	0.743 (0.739, 0.748)	0.159 (0.153, 0.165)
phecode_733-6	Temporomandibular joint disorders	0.587 (0.582, 0.592)	0.745 (0.74, 0.748)	0.157 (0.152, 0.163)
phecode_733-62	Arthralgia of temporomandibular joint	0.598 (0.587, 0.608)	0.765 (0.756, 0.773)	0.167 (0.155, 0.178)
phecode_734	Diseases of the jaws	0.571 (0.563, 0.579)	0.74 (0.734, 0.747)	0.169 (0.16, 0.178)
phecode_734-5	Inflammatory conditions of jaw	0.502 (0.46, 0.542)	0.749 (0.713, 0.78)	0.247 (0.195, 0.292)
phecode_734-9	Jaw pain	0.574 (0.566, 0.582)	0.744 (0.738, 0.751)	0.17 (0.162, 0.178)
phecode_800	Chest pain	0.518 (0.517, 0.52)	0.671 (0.67, 0.673)	0.153 (0.151, 0.155)
phecode_800-1	Chest pain on breathing	0.549 (0.544, 0.553)	0.735 (0.731, 0.739)	0.187 (0.181, 0.191)
phecode_800-11	Pleurodynia*	0.549 (0.544, 0.553)	0.739 (0.735, 0.742)	0.19 (0.185, 0.195)
phecode_800-2	Precordial pain	0.545 (0.54, 0.55)	0.709 (0.704, 0.714)	0.164 (0.157, 0.17)
phecode_800-3	Intercostal pain*	0.52 (0.494, 0.548)	0.714 (0.692, 0.735)	0.193 (0.162, 0.224)
phecode_801	Cough	0.533 (0.531, 0.534)	0.721 (0.72, 0.722)	0.188 (0.186, 0.189)
phecode_802	Throat pain	0.557 (0.553, 0.561)	0.718 (0.716, 0.722)	0.162 (0.158, 0.167)
phecode_803	Snoring*	0.629 (0.623, 0.635)	0.739 (0.734, 0.744)	0.11 (0.105, 0.115)
phecode_804	Other symptoms and signs involving the circulatory and respiratory system	0.537 (0.535, 0.539)	0.742 (0.741, 0.744)	0.205 (0.203, 0.208)
phecode_805	Fever of unknown origin	0.535 (0.53, 0.54)	0.641 (0.636, 0.645)	0.105 (0.099, 0.111)
phecode_806	Chills (without fever)	0.591 (0.565, 0.615)	0.745 (0.725, 0.762)	0.154 (0.128, 0.179)
phecode_807	Malaise and fatigue	0.554 (0.552, 0.556)	0.715 (0.713, 0.717)	0.161 (0.159, 0.163)
phecode_807-1	Chronic fatigue syndrome	0.601 (0.597, 0.604)	0.749 (0.747, 0.752)	0.149 (0.146, 0.152)
phecode_807-11	Postviral fatigue syndrome*	0.656 (0.645, 0.667)	0.76 (0.75, 0.771)	0.104 (0.094, 0.117)
phecode_808	Syncope and collapse	0.595 (0.593, 0.598)	0.65 (0.648, 0.653)	0.055 (0.052, 0.058)
phecode_809	Pain	0.529 (0.528, 0.531)	0.713 (0.712, 0.714)	0.184 (0.182, 0.186)
phecode_809-1	Acute pain	0.51 (0.494, 0.528)	0.781 (0.77, 0.792)	0.27 (0.251, 0.29)
phecode_809-3	Pain in limb	0.528 (0.526, 0.53)	0.715 (0.713, 0.716)	0.186 (0.185, 0.188)
phecode_810	Shock	0.677 (0.664, 0.69)	0.786 (0.775, 0.798)	0.11 (0.097, 0.122)
phecode_810-1	Cardiogenic shock	0.733 (0.712, 0.752)	0.829 (0.812, 0.848)	0.096 (0.074, 0.119)
phecode_811	Hypothermia and chills, not associated with low environmental temperature	0.627 (0.601, 0.655)	0.761 (0.737, 0.785)	0.135 (0.105, 0.162)
phecode_812	Edema	0.61 (0.607, 0.613)	0.728 (0.726, 0.731)	0.118 (0.115, 0.121)
phecode_812-1	Generalized edema*	0.591 (0.556, 0.623)	0.673 (0.638, 0.706)	0.082 (0.047, 0.118)
phecode_812-2	Angioneurotic edema	0.517 (0.505, 0.529)	0.674 (0.662, 0.684)	0.156 (0.141, 0.171)
phecode_813	Clubbing of fingers	0.602 (0.565, 0.639)	0.797 (0.764, 0.836)	0.196 (0.153, 0.232)
phecode_814	Jaundice (not of newborn)	0.628 (0.618, 0.638)	0.675 (0.663, 0.687)	0.048 (0.037, 0.058)
phecode_815	Symptoms and signs concerning food and fluid intake	0.556 (0.55, 0.562)	0.748 (0.743, 0.753)	0.192 (0.186, 0.199)
phecode_817	Motion sickness	0.637 (0.623, 0.649)	0.751 (0.741, 0.762)	0.115 (0.103, 0.127)
phecode_819	General symptoms and other findings	0.581 (0.579, 0.582)	0.648 (0.647, 0.65)	0.068 (0.066, 0.069)
phecode_820	Elevated erythrocyte sedimentation rate and abnormality of plasma viscosity	0.641 (0.628, 0.655)	0.766 (0.755, 0.776)	0.125 (0.113, 0.136)
phecode_821	Abnormality of red blood cells	0.518 (0.509, 0.528)	0.753 (0.746, 0.76)	0.234 (0.223, 0.246)
phecode_821-1	Precipitous drop in hematocrit	0.661 (0.617, 0.709)	0.795 (0.766, 0.829)	0.133 (0.095, 0.173)
phecode_822	Other and nonspecific abnormal cytological, histological and immunological findings	0.84 (0.833, 0.848)	0.889 (0.882, 0.896)	0.049 (0.041, 0.057)
phecode_822-8	Abnormal tumor markers	0.84 (0.832, 0.848)	0.889 (0.882, 0.896)	0.049 (0.041, 0.057)

Supplementary Tables

Table 17 continued from previous page

Endpoint	PheCode string	C-Index Age+Sex (IQR)	C-Index Age+Sex+MedicalHistory (IQR)	Difference in C-Index (IQR)
phecode_823	Abnormal serum enzyme levels	0.547 (0.544, 0.55)	0.727 (0.725, 0.729)	0.18 (0.177, 0.184)
phecode_823-1	Nonspecific elevation of levels of transaminase and lactic acid dehydrogenase [LDH]	0.506 (0.479, 0.538)	0.615 (0.583, 0.645)	0.109 (0.069, 0.15)
phecode_823-2	Abnormal levels of other serum enzymes	0.547 (0.544, 0.55)	0.729 (0.726, 0.731)	0.182 (0.178, 0.185)
phecode_824	Other abnormalities of plasma proteins*	0.553 (0.546, 0.561)	0.725 (0.719, 0.729)	0.172 (0.162, 0.179)
phecode_825	Elevated C-reactive protein (CRP)	0.532 (0.489, 0.571)	0.752 (0.719, 0.789)	0.22 (0.171, 0.273)
phecode_826	Other abnormal immunological findings in serum	0.555 (0.546, 0.564)	0.715 (0.708, 0.722)	0.16 (0.15, 0.169)
phecode_826-3	Raised antibody titer*	0.561 (0.551, 0.572)	0.721 (0.713, 0.728)	0.159 (0.148, 0.172)
phecode_826-4	Other and unspecified nonspecific immunological findings	0.614 (0.598, 0.631)	0.712 (0.698, 0.725)	0.098 (0.083, 0.115)
phecode_827	Toxicology findings	0.606 (0.599, 0.612)	0.736 (0.73, 0.742)	0.13 (0.124, 0.138)
phecode_827-1	Finding of alcohol in blood	0.614 (0.608, 0.622)	0.738 (0.731, 0.744)	0.124 (0.117, 0.13)
phecode_829	Nonspecific findings on examination of blood	0.634 (0.632, 0.637)	0.704 (0.702, 0.706)	0.07 (0.067, 0.072)
phecode_829-2	Abnormal level of blood mineral*	0.554 (0.548, 0.559)	0.73 (0.725, 0.736)	0.176 (0.17, 0.184)
phecode_830	Proteinuria	0.585 (0.583, 0.587)	0.743 (0.741, 0.744)	0.158 (0.156, 0.16)
phecode_831	Glycosuria	0.596 (0.591, 0.6)	0.819 (0.816, 0.822)	0.223 (0.219, 0.228)
phecode_832	Other abnormal findings in urine	0.64 (0.639, 0.642)	0.751 (0.75, 0.753)	0.111 (0.11, 0.112)
phecode_832-4	Hemoglobinuria	0.581 (0.538, 0.624)	0.74 (0.706, 0.771)	0.158 (0.112, 0.203)
phecode_832-5	Acetonuria	0.552 (0.547, 0.557)	0.749 (0.744, 0.753)	0.197 (0.191, 0.202)
phecode_832-6	Pyuria*	0.639 (0.634, 0.644)	0.792 (0.789, 0.795)	0.152 (0.148, 0.157)
phecode_835	Cytology and pathology findings	0.718 (0.715, 0.722)	0.788 (0.785, 0.791)	0.07 (0.068, 0.072)
phecode_840	Allergy	0.547 (0.545, 0.55)	0.692 (0.69, 0.694)	0.144 (0.142, 0.147)
phecode_840-1	Food allergy	0.584 (0.576, 0.592)	0.718 (0.711, 0.726)	0.134 (0.125, 0.144)
phecode_840-11	Peanut allergy	0.638 (0.595, 0.677)	0.814 (0.784, 0.847)	0.176 (0.134, 0.222)
phecode_840-12	Seafood allergy	0.53 (0.51, 0.55)	0.678 (0.661, 0.695)	0.148 (0.126, 0.172)
phecode_840-18	Egg allergy	0.586 (0.559, 0.615)	0.73 (0.707, 0.753)	0.144 (0.117, 0.171)
phecode_840-2	Allergy to insects	0.566 (0.561, 0.57)	0.705 (0.702, 0.709)	0.139 (0.135, 0.144)
phecode_840-4	Latex allergy	0.66 (0.641, 0.677)	0.792 (0.778, 0.804)	0.132 (0.116, 0.15)
phecode_840-8	Allergies related to other diseases/symptoms	0.518 (0.515, 0.521)	0.711 (0.708, 0.713)	0.193 (0.189, 0.196)
phecode_840-9	Anaphylactic reaction	0.522 (0.507, 0.537)	0.675 (0.658, 0.689)	0.153 (0.134, 0.174)
phecode_841	Drug and medical agent allergy	0.597 (0.595, 0.598)	0.683 (0.681, 0.685)	0.086 (0.084, 0.088)
phecode_841-1	Allergy to other anti-infective agents	0.599 (0.597, 0.602)	0.674 (0.672, 0.677)	0.075 (0.073, 0.078)
phecode_841-11	Penicillin allergy	0.584 (0.582, 0.587)	0.661 (0.658, 0.664)	0.077 (0.074, 0.08)
phecode_841-12	Allergy to antibiotic agent (excluding penicillin)	0.66 (0.656, 0.665)	0.755 (0.751, 0.76)	0.095 (0.09, 0.1)
phecode_841-13	Allergy to sulfonamides	0.685 (0.674, 0.696)	0.765 (0.755, 0.776)	0.08 (0.07, 0.093)
phecode_841-2	Allergy to anesthetic agent	0.64 (0.62, 0.658)	0.744 (0.724, 0.763)	0.103 (0.083, 0.125)
phecode_841-3	Allergy to narcotic agent	0.662 (0.655, 0.669)	0.779 (0.772, 0.787)	0.117 (0.109, 0.125)
phecode_841-4	Allergy to analgesic agent	0.622 (0.618, 0.626)	0.742 (0.738, 0.746)	0.12 (0.115, 0.125)
phecode_841-5	Allergy to serum and vaccine	0.62 (0.603, 0.638)	0.712 (0.696, 0.728)	0.092 (0.072, 0.111)
phecode_848	Nonspecific abnormal findings of other body structures	0.617 (0.614, 0.62)	0.729 (0.727, 0.732)	0.113 (0.11, 0.115)
phecode_848-2	Nonspecific abnormal findings on radiological and other examination of other intrathoracic organs (echo)	0.692 (0.687, 0.696)	0.772 (0.767, 0.776)	0.08 (0.076, 0.084)
phecode_900-1	Spontaneous abortion	0.947 (0.939, 0.956)	0.949 (0.941, 0.958)	0.002 (-0.002, 0.006)
phecode_901	Abnormal products of conception	0.688 (0.666, 0.713)	0.704 (0.684, 0.725)	0.015 (-0.002, 0.035)
phecode_901-3	Missed abortion	0.68 (0.656, 0.703)	0.704 (0.684, 0.726)	0.024 (0.004, 0.045)
phecode_904	Bleeding in pregnancy	0.927 (0.909, 0.951)	0.921 (0.903, 0.946)	-0.006 (-0.016, 0.004)
phecode_908	Edema, proteinuria and hypertensive disorders in pregnancy, childbirth and the puerperium	0.814 (0.775, 0.849)	0.805 (0.77, 0.84)	-0.009 (-0.041, 0.024)
phecode_930	Maternal care for abnormality of pelvic organs	0.946 (0.937, 0.963)	0.95 (0.938, 0.969)	0.004 (-0.005, 0.013)
phecode_932	Known or suspected fetal abnormality affecting management of mother	0.937 (0.926, 0.952)	0.933 (0.921, 0.948)	-0.004 (-0.011, 0.002)
phecode_938	Other conditions or status of the mother complicating pregnancy, childbirth, or the puerperium	0.835 (0.811, 0.861)	0.84 (0.82, 0.863)	0.005 (-0.01, 0.019)
phecode_940	Malposition and malpresentation of fetus	0.712 (0.68, 0.742)	0.788 (0.772, 0.808)	0.077 (0.048, 0.107)

4 Medical history predicts future health trajectories over the human phenome

Table 17 continued from previous page

Endpoint	PheCode string	C-Index Age+Sex (IQR)	C-Index Age+Sex+MedicalHistory (IQR)	Difference in C-Index (IQR)
phecode_965	Psychotropic agents causing adverse effects in therapeutic use	0.495 (0.446, 0.544)	0.691 (0.641, 0.745)	0.194 (0.128, 0.268)
phecode_965-5	Psychostimulants causing adverse effects in therapeutic use	0.529 (0.478, 0.58)	0.686 (0.635, 0.744)	0.157 (0.095, 0.224)
phecode_969	Adverse effects of agents primarily affecting gastrointestinal system	0.59 (0.588, 0.593)	0.745 (0.742, 0.747)	0.154 (0.152, 0.157)
phecode_973	Adverse effect of other drug	0.652 (0.641, 0.666)	0.785 (0.775, 0.796)	0.133 (0.122, 0.144)
phecode_974	Adverse effect of vaccine	0.565 (0.545, 0.586)	0.692 (0.674, 0.71)	0.127 (0.105, 0.149)
phecode_975	Primarily systemic agents causing adverse effects in therapeutic use	0.54 (0.524, 0.555)	0.669 (0.653, 0.683)	0.129 (0.108, 0.151)
phecode_975-2	Antineoplastic and immunosuppressive drugs causing adverse effects in therapeutic use	0.539 (0.521, 0.555)	0.669 (0.654, 0.686)	0.131 (0.107, 0.151)
phecode_975-23	Mucositis (ulcerative) due to antineoplastic therapy	0.54 (0.522, 0.555)	0.669 (0.653, 0.686)	0.13 (0.109, 0.152)
phecode_976	Adverse effect of anesthesia	0.556 (0.52, 0.592)	0.659 (0.62, 0.699)	0.103 (0.056, 0.152)
phecode_977	Long term (current) drug therapy	0.503 (0.501, 0.506)	0.823 (0.822, 0.825)	0.32 (0.317, 0.323)
phecode_977-1	Long term (current) use of anticoagulants and antithrombotics/antiplatelets	0.713 (0.697, 0.729)	0.853 (0.84, 0.867)	0.14 (0.124, 0.156)
phecode_977-11	Long term (current) use of anticoagulants	0.713 (0.697, 0.73)	0.853 (0.839, 0.868)	0.14 (0.123, 0.157)
phecode_977-4	Long term (current) use of steroids	0.546 (0.539, 0.552)	0.924 (0.921, 0.927)	0.378 (0.372, 0.385)
phecode_977-41	Long term (current) use of inhaled steroids*	0.546 (0.54, 0.553)	0.924 (0.921, 0.927)	0.378 (0.371, 0.385)
phecode_977-5	Long term (current) use of agents affecting hormones	0.896 (0.893, 0.898)	0.917 (0.915, 0.919)	0.021 (0.02, 0.023)
phecode_977-51	Long term (current) use of hormonal contraceptives	0.917 (0.914, 0.921)	0.939 (0.936, 0.943)	0.022 (0.019, 0.025)
phecode_977-52	Hormone replacement therapy (postmenopausal)	0.747 (0.742, 0.753)	0.811 (0.807, 0.816)	0.064 (0.06, 0.067)
phecode_977-7	Long term (current) use of insulin or oral hypoglycemic drugs	0.599 (0.596, 0.602)	0.874 (0.872, 0.876)	0.275 (0.272, 0.279)
phecode_977-71	Long term (current) use of insulin	0.584 (0.577, 0.591)	0.949 (0.946, 0.953)	0.366 (0.358, 0.373)
phecode_977-72	Long term (current) use of oral hypoglycemic drugs	0.602 (0.598, 0.605)	0.871 (0.869, 0.873)	0.27 (0.266, 0.273)
phecode_978	Tissue transplant or graft	0.588 (0.564, 0.613)	0.355 (0.321, 0.383)	-0.234 (-0.275, -0.195)
phecode_978-2	Corneal transplant or graft	0.595 (0.568, 0.626)	0.844 (0.822, 0.871)	0.249 (0.212, 0.283)
phecode_979	Transplanted organ	0.569 (0.557, 0.581)	0.385 (0.372, 0.401)	-0.183 (-0.204, -0.161)
phecode_979-2	Kidney transplant	0.593 (0.57, 0.62)	0.897 (0.88, 0.916)	0.303 (0.276, 0.334)
phecode_979-7	Bone marrow transplant	0.531 (0.494, 0.569)	0.706 (0.669, 0.743)	0.173 (0.128, 0.227)

Table 18: Adjusted and unadjusted hazard ratios for all endpoints.

Code	Name	unadjusted HR (per 1 SD Risk State)	adj. HR for Age and Sex (per 1 SD Risk State)
OMOP_4306655	All-Cause Death	2.87 (2.72, 3.06)	2.73 (2.58, 2.88)
phecode_001	Salmonella	3.43 (2.87, 3.98)	3.39 (2.93, 4.07)
phecode_002	Staphylococcus	2.8 (2.66, 2.92)	2.8 (2.64, 2.89)
phecode_002-1	Staphylococcus aureus	2.78 (2.69, 2.96)	2.76 (2.68, 2.95)
phecode_003	Escherichia coli	2.9 (2.72, 3.11)	2.83 (2.66, 3)
phecode_004	Streptococcus	2.96 (2.69, 3.24)	2.94 (2.67, 3.23)
phecode_004-1	Streptococcus pneumoniae	3.05 (2.88, 3.45)	3.04 (2.9, 3.44)
phecode_004-2	Group A Streptococcus	3.27 (3, 3.72)	3.32 (3, 3.75)
phecode_004-3	Group B Streptococcus	2.95 (2.83, 3.19)	3.01 (2.83, 3.2)
phecode_005	Mycobacteria	3.17 (2.93, 3.4)	3.19 (2.93, 3.4)
phecode_005-1	Mycobacterium tuberculosis	3.18 (3, 3.48)	3.16 (2.99, 3.47)
phecode_005-2	Nontuberculous mycobacteria	3.12 (2.82, 3.43)	3.08 (2.8, 3.38)
phecode_007	Hemophilus infection	2.85 (2.65, 3.07)	2.83 (2.63, 3.04)
phecode_007-1	Hemophilus influenzae	2.85 (2.65, 3.07)	2.84 (2.63, 3.04)
phecode_008	Helicobacter [H. pylori]	3 (2.75, 3.42)	3.07 (2.76, 3.46)
phecode_009	Pseudomonas	2.81 (2.65, 2.97)	2.77 (2.61, 2.92)
phecode_010	Corynebacterium	2.89 (2.67, 3.6)	3.7 (3.39, 4.19)
phecode_011	Klebsiella	2.85 (2.72, 3.07)	2.81 (2.67, 3.02)
phecode_012	Proteus	2.9 (2.71, 3.07)	2.87 (2.69, 3.04)
phecode_015	Clostridium	3.06 (2.76, 3.28)	3 (2.72, 3.22)
phecode_015-2	Clostridium difficile	2.99 (2.74, 3.27)	2.96 (2.7, 3.21)
phecode_019	Treponema	3.07 (2.84, 3.37)	3.14 (2.9, 3.58)
phecode_020	Borrelia	4.24 (3.9, 4.98)	4.41 (4.05, 4.99)
phecode_020-1	Lyme disease	4.25 (3.93, 4.99)	4.4 (4.07, 5.01)
phecode_024	Pertussis	4.32 (3.44, 5.53)	4.57 (3.53, 5.48)
phecode_025	Enterococcus	3.06 (2.83, 3.22)	3 (2.78, 3.2)
phecode_030	Campylobacter	2.8 (2.5, 3.21)	2.88 (2.62, 3.22)
phecode_050	Enterovirus	3.56 (3.13, 6.05)	3.65 (3.19, 6.04)
phecode_050-4	Hand, foot, and mouth disease	4.96 (4.14, 7.08)	5.32 (4.19, 7.16)
phecode_052	Herpesvirus	2.64 (2.43, 2.78)	2.63 (2.41, 2.75)
phecode_052-1	Herpes simplex	2.51 (2.38, 2.62)	2.52 (2.39, 2.64)
phecode_052-3	Varicella zoster virus	2.72 (2.46, 2.92)	2.64 (2.4, 2.81)
phecode_052-31	Varicella [chickenpox]	3.64 (3.16, 4.69)	3.25 (2.79, 4.1)
phecode_052-32	Herpes zoster	2.72 (2.46, 2.92)	2.63 (2.38, 2.78)
phecode_052-4	Infectious mononucleosis	4.43 (3.88, 5.39)	4.49 (3.84, 5.55)
phecode_052-5	Cytomegalovirus [CMV]	2.97 (2.76, 3.26)	2.98 (2.76, 3.27)
phecode_054	Hepatovirus	0.03 (0, 571.66)	0.02 (0, 2.91)
phecode_054-2	Hepatitis B	0.62 (0.11, 2.63)	0.21 (0.01, 2.63)
phecode_054-3	Hepatitis C	2.42 (0, 2.52)	2.46 (0.01, 2.56)
phecode_054-31	Chronic hepatitis C	2.46 (0, 2.59)	2.48 (2.41, 2.6)
phecode_054-5	Hepatitis E	4.17 (3.64, 19.4)	4.24 (3.66, 30.57)
phecode_055	Poxvirus	4.17 (3.71, 4.87)	4.3 (3.77, 4.74)
phecode_055-1	Molluscum contagiosum	4.19 (3.86, 5.01)	4.41 (3.8, 4.86)
phecode_056	Human papillomavirus	2.68 (2.52, 2.87)	2.69 (2.52, 2.9)
phecode_056-1	Plantar wart	2.81 (2.49, 3.06)	2.78 (2.57, 3.03)
phecode_057	Retrovirus	2.85 (2.7, 3.1)	2.76 (2.61, 3.04)
phecode_057-1	Human immunodeficiency virus	2.88 (2.73, 3.13)	2.79 (2.63, 3.06)
phecode_058	Pneumoviridae	3.24 (2.95, 3.55)	3.23 (2.93, 3.53)
phecode_058-1	Respiratory syncytial virus	3.29 (2.95, 3.57)	3.23 (2.93, 3.54)
phecode_059	Coronavirus	3.05 (2.65, 3.56)	3.04 (2.65, 3.57)
phecode_059-1	COVID-19*	3.02 (2.67, 3.6)	3.06 (2.67, 3.62)
phecode_060	Adenovirus	4.73 (3.83, 6.22)	4.89 (3.95, 6.32)
phecode_061	Influenza virus	2.92 (2.68, 3.29)	2.88 (2.68, 3.26)
phecode_066	Orthorubulavirus [Mumps]	4.09 (3.62, 4.81)	3.64 (3.17, 4.5)
phecode_069	Other specified viral infections	8.57 (5.5, 11)	8.94 (5.88, 11.98)
phecode_070	Candidiasis	2.58 (2.47, 2.73)	2.61 (2.48, 2.76)
phecode_074	Aspergillosis	2.75 (2.64, 2.97)	2.75 (2.65, 2.97)
phecode_076	Pneumocystosis	3.13 (2.97, 3.52)	3.12 (2.94, 3.48)
phecode_084	Parasites	3.19 (2.9, 3.52)	3.23 (3.06, 3.59)
phecode_084-2	Malaria [Plasmodium]	4.39 (3.8, 5)	4.42 (3.91, 4.96)

4 Medical history predicts future health trajectories over the human phenome

Table 18 continued from previous page

Code	Name	unadjusted HR (per 1 SD Risk State)	adj. HR for Age and Sex (per 1 SD Risk State)
phecode_084-4	Trichomoniasis	3.82 (3.43, 4.17)	3.8 (3.34, 4.2)
phecode_084-6	Enterobiasis	3.16 (2.87, 3.87)	3.09 (2.78, 3.86)
phecode_084-7	Giardiasis	5.73 (4.8, 6.88)	5.92 (4.97, 7.23)
phecode_086	Pediculosis, acariasis and other infestations	3.02 (2.64, 3.46)	3 (2.6, 3.44)
phecode_088	Sexually transmitted disease	2.82 (2.6, 3.08)	2.76 (2.52, 3.04)
phecode_089	Infections	2.72 (2.55, 2.86)	2.68 (2.53, 2.84)
phecode_089-1	Bacterial infections	2.98 (2.76, 3.3)	2.95 (2.74, 3.23)
phecode_089-2	Viral infections	2.59 (2.49, 2.76)	2.61 (2.51, 2.77)
phecode_089-3	Fungal infections	2.6 (2.44, 2.75)	2.6 (2.44, 2.74)
phecode_091	Gangrene	3.05 (2.75, 3.28)	3.04 (2.71, 3.32)
phecode_092	Bacteremia, Sepsis, and SIRS	2.9 (2.77, 3.2)	2.81 (2.68, 3.08)
phecode_092-1	Systemic inflammatory response syndrome	3.03 (2.84, 3.38)	2.96 (2.77, 3.28)
phecode_092-2	Sepsis	2.89 (2.78, 3.19)	2.8 (2.7, 3.07)
phecode_092-8	Bacteremia	3.3 (3.02, 3.64)	3.32 (3, 3.63)
phecode_095	Sequela of infection	2.41 (2.24, 2.64)	2.44 (2.27, 2.65)
phecode_096	Contact or exposure to infectious agent	2.93 (2.73, 3.25)	2.95 (2.8, 3.25)
phecode_097	Drug resistant microorganisms	2.93 (2.58, 3.14)	2.96 (2.6, 3.14)
phecode_097-1	Methicillin resistant Staphylococcus aureus	2.93 (2.64, 3.11)	2.95 (2.66, 3.14)
phecode_098	Carrier or suspected carrier of infectious diseases	2.76 (2.53, 3.04)	2.75 (2.52, 3.03)
phecode_098-2	Carrier or suspected carrier of Staphylococcus aureus	3.22 (2.91, 3.73)	3.24 (2.91, 3.73)
phecode_099	Lab findings related to infections	2.6 (2.46, 2.75)	2.62 (2.41, 2.7)
phecode_100	Malignant neoplasm of the head and neck	2.8 (2.61, 3.11)	2.74 (2.55, 3.03)
phecode_100-1	Malignant neoplasm of the oral cavity	2.97 (2.78, 3.17)	2.99 (2.76, 3.15)
phecode_100-12	Malignant neoplasm of the tongue	3.08 (2.87, 3.36)	3.04 (0.13, 3.36)
phecode_100-2	Malignant neoplasm of the oropharynx	2.88 (2.66, 3.12)	2.84 (2.61, 3.11)
phecode_100-5	Malignant neoplasm of nasal cavities, middle ear, and accessory sinuses	4.51 (3.88, 6.94)	4.58 (3.84, 7.12)
phecode_100-6	Malignant neoplasm of the larynx	2.8 (2.59, 2.98)	2.66 (2.48, 2.85)
phecode_100-7	Malignant neoplasm of the pharynx	3.05 (2.9, 3.44)	3.03 (2.87, 3.4)
phecode_100-8	Malignant neoplasm of the lip	4.13 (3.63, 5.12)	4.02 (3.53, 4.95)
phecode_100-9	Malignant neoplasm of the salivary glands	4.62 (3.93, 5.81)	4.63 (3.92, 5.77)
phecode_101	Malignant neoplasm of the digestive organs	2.52 (2.39, 2.62)	2.3 (2.17, 2.38)
phecode_101-1	Malignant neoplasm of the esophagus	2.9 (2.69, 3.12)	2.68 (2.49, 2.85)
phecode_101-2	Malignant neoplasm of stomach	3.17 (2.86, 3.57)	2.9 (2.6, 3.22)
phecode_101-21	Malignant neoplasm of cardia	3.41 (3.11, 3.98)	3.1 (2.77, 3.51)
phecode_101-3	Malignant neoplasm of the small intestine	3.32 (2.85, 3.49)	3.23 (2.76, 3.35)
phecode_101-4	Malignant neoplasm of the colon and rectum	2.68 (2.48, 2.92)	2.18 (2.06, 2.35)
phecode_101-41	Malignant neoplasm of the colon	2.73 (2.5, 3.07)	2.25 (2.09, 2.46)
phecode_101-42	Malignant neoplasm of the rectum	2.67 (2.51, 2.89)	2.22 (2.09, 2.35)
phecode_101-5	Malignant neoplasm of the anus and anal canal	3.33 (2.96, 3.68)	3.3 (2.93, 3.64)
phecode_101-6	Malignant neoplasm of the liver and intrahepatic bile ducts	3.01 (2.82, 3.26)	2.97 (2.8, 3.23)
phecode_101-61	Malignant neoplasm of the liver	2.9 (2.7, 3.17)	2.88 (2.7, 3.15)
phecode_101-62	Malignant neoplasm of the intrahepatic bile ducts	3.57 (3.2, 3.86)	3.45 (3.11, 3.75)
phecode_101-7	Malignant neoplasm of the gallbladder and extrahepatic bile ducts	3.74 (3.38, 4.02)	3.57 (3.22, 3.81)
phecode_101-71	Malignant neoplasm of the gallbladder	4.34 (3.77, 5.11)	4.12 (3.54, 4.85)
phecode_101-8	Malignant neoplasm of the pancreas	3.12 (2.88, 3.56)	2.93 (2.71, 3.31)
phecode_102	Malignant neoplasm of the thoracic and respiratory organs	2.79 (2.61, 2.91)	2.66 (2.5, 2.77)
phecode_102-1	Malignant neoplasm of the of bronchus and lung	2.86 (2.63, 3.01)	2.71 (2.51, 2.84)
phecode_102-3	Malignant neoplasm of the trachea	3.31 (3.05, 3.95)	3.17 (2.93, 3.77)
phecode_102-5	Malignant neoplasm of the heart, mediastinum, thymus, and pleura	3.23 (2.96, 3.59)	3.14 (2.87, 3.47)
phecode_103	Malignant neoplasm of the skin	2.68 (2.51, 2.81)	2.11 (2.02, 2.22)
phecode_103-1	Melanomas of skin	2.98 (2.62, 3.17)	2.64 (2.33, 2.78)
phecode_103-2	Keratinocyte carcinoma	2.74 (2.57, 2.83)	2.41 (2.31, 2.51)
phecode_103-21	Basal cell carcinoma	2.7 (2.54, 2.79)	2.45 (2.33, 2.52)
phecode_103-22	Squamous cell carcinoma of the skin	3.16 (2.9, 3.6)	2.71 (2.5, 2.96)

Supplementary Tables

Table 18 continued from previous page

Code	Name	unadjusted HR (per 1 SD Risk State)	adj. HR for Age and Sex (per 1 SD Risk State)
phecode_103-3	Carcinoma in situ of skin	3.02 (2.74, 3.31)	2.57 (2.35, 2.78)
phecode_104	Malignant sarcoma-related cancers	3.18 (2.74, 3.78)	3.11 (2.68, 3.64)
phecode_104-1	Malignant neoplasm of the bone and/or cartilage	3.28 (2.95, 3.93)	3.24 (2.95, 3.88)
phecode_104-2	Malignant neoplasm of retroperitoneum and peritoneum	3.35 (3.12, 3.77)	3.12 (2.94, 3.53)
phecode_104-3	Malignant neoplasm of other connective and soft tissue	3.42 (2.83, 3.88)	3.35 (2.83, 3.81)
phecode_104-5	Gastrointestinal stromal tumor*	5.55 (4.13, 7.22)	5.97 (4.55, 8.35)
phecode_105	Malignant neoplasm of the breast	2.3 (2.19, 2.41)	1.46 (1.42, 1.57)
phecode_105-1	Malignant neoplasm of the breast, female	2.02 (1.94, 2.14)	2 (1.92, 2.11)
phecode_106	Gynecological malignant neoplasms	3.03 (2.44, 3.3)	2.95 (2.37, 3.2)
phecode_106-1	Malignant neoplasm of external female genital organs and cervix	3.47 (3.06, 3.88)	3.53 (3.05, 4)
phecode_106-11	Malignant neoplasm of the vulva	3.57 (2.93, 4.14)	3.53 (2.91, 4.08)
phecode_106-13	Malignant neoplasm of the cervix	3.37 (2.94, 3.79)	3.35 (2.87, 3.88)
phecode_106-2	Malignant neoplasm of the uterus	3.46 (2.82, 3.87)	3.35 (2.69, 3.71)
phecode_106-21	Malignant neoplasm of endometrium	3.49 (2.95, 3.94)	3.36 (2.81, 3.76)
phecode_106-3	Malignant neoplasm of the ovary	2.75 (2.42, 2.99)	2.67 (2.35, 2.91)
phecode_106-4	Malignant neoplasm of the fallopian tube and uterine adnexa	4.51 (3.96, 5.71)	4.44 (3.98, 5.56)
phecode_107	Malignant neoplasm of male genitalia	2.06 (1.95, 2.17)	1.45 (1.4, 1.49)
phecode_107-1	Malignant neoplasm of the penis	2.83 (2.66, 3.09)	2.67 (2.51, 2.89)
phecode_107-2	Malignant neoplasm of the prostate	2.05 (1.93, 2.17)	1.44 (1.38, 1.48)
phecode_107-3	Malignant neoplasm of the testis	3.1 (2.7, 3.5)	2.99 (2.62, 3.41)
phecode_108	Malignant neoplasm of the urinary tract	2.4 (2.27, 2.61)	1.99 (1.91, 2.1)
phecode_108-4	Malignant neoplasm of the kidney	2.81 (2.57, 3.08)	2.54 (2.39, 2.81)
phecode_108-41	Malignant neoplasm of kidney, except pelvis	2.83 (2.61, 3.05)	2.61 (2.44, 2.78)
phecode_108-42	Malignant neoplasm of renal pelvis	3.64 (3.25, 4.06)	3.43 (3.13, 4.07)
phecode_108-5	Malignant neoplasm of the bladder	2.49 (2.32, 2.82)	1.91 (1.81, 2.04)
phecode_108-7	Malignant neoplasm of ureter	3.18 (2.97, 3.34)	2.87 (2.78, 3.16)
phecode_109	Malignant neoplasm of the eye, brain and other parts of central nervous system	3.26 (2.9, 3.53)	3.18 (2.83, 3.44)
phecode_109-1	Malignant neoplasm of eye	3.9 (3.41, 4.83)	3.93 (3.38, 4.89)
phecode_109-16	Malignant neoplasm of choroid	5.7 (4.42, 8.17)	5.73 (4.52, 8.57)
phecode_109-3	Malignant neoplasm of brain	3.19 (2.9, 3.59)	3.11 (2.83, 3.5)
phecode_110	Malignant neoplasm of the endocrine glands	2.78 (2.68, 2.99)	2.82 (2.69, 3)
phecode_110-1	Malignant neoplasm of the thyroid	2.78 (2.61, 3.01)	2.79 (2.61, 3.03)
phecode_110-4	Malignant neoplasm of the pituitary gland and craniopharyngeal duct	2.74 (0.48, 2.85)	0.23 (0.02, 2.83)
phecode_112	Malignant neoplasm of other and ill-defined sites	2.85 (2.64, 3.03)	2.56 (2.39, 2.68)
phecode_112-1	Mesothelioma*	3.54 (3.18, 4.05)	2.95 (2.67, 3.32)
phecode_114	Neuroendocrine tumors	3.11 (2.91, 3.25)	3.13 (2.92, 3.29)
phecode_114-4	Carcinoid tumors	3.64 (3.34, 4.19)	3.66 (3.32, 4.18)
phecode_114-6	Pheochromocytoma (including adrenal gland neoplasms)	3.17 (2.93, 3.5)	3.21 (2.93, 3.51)
phecode_116	Secondary malignant neoplasm	2.87 (2.73, 3.06)	2.73 (2.6, 2.89)
phecode_116-1	Secondary malignancy of lymph nodes	2.99 (2.77, 3.15)	2.87 (2.66, 3.01)
phecode_116-2	Secondary malignancy of respiratory organs	2.83 (2.67, 3.07)	2.75 (2.6, 2.98)
phecode_116-3	Secondary malignant neoplasm of digestive systems	2.94 (2.75, 3.15)	2.86 (2.68, 3.04)
phecode_116-4	Secondary malignant neoplasm of liver	2.83 (2.65, 3.03)	2.75 (2.57, 2.93)
phecode_116-5	Secondary malignancy of brain/spine	2.88 (2.73, 3.06)	2.83 (2.68, 3.02)
phecode_116-6	Secondary malignancy of bone	2.77 (2.61, 3.02)	2.71 (2.53, 2.91)
phecode_116-7	Secondary malignant neoplasm of skin	3.03 (2.87, 3.39)	3 (2.84, 3.36)
phecode_120	Hemo onc - by cell of origin	3.06 (2.85, 3.29)	2.99 (2.78, 3.2)
phecode_120-1	Myeloid	2.86 (2.59, 3.02)	2.81 (2.53, 2.96)
phecode_120-11	Plasma cell	3.04 (2.68, 3.33)	3.02 (2.66, 3.31)
phecode_120-12	Monocyte	3.47 (3.09, 3.97)	3.32 (2.98, 3.76)
phecode_120-13	Erythroid	0.01 (0, 2.62)	2.58 (0, 2.73)
phecode_120-2	Lymphoid	2.85 (2.71, 9.18)	2.79 (2.64, 2.95)
phecode_120-21	Mature B-cell	2.81 (2.66, 9.42)	2.75 (2.6, 2.9)

4 Medical history predicts future health trajectories over the human phenome

Table 18 continued from previous page

Code	Name	unadjusted HR (per 1 SD Risk State)	adj. HR for Age and Sex (per 1 SD Risk State)
phecode_120-22	Mature T-Cell	3.1 (2.82, 3.39)	3 (2.78, 3.31)
phecode_121	Leukemia	2.61 (2.49, 6.59)	2.54 (2.42, 2.62)
phecode_121-1	Acute leukemia	2.83 (2.61, 3.13)	2.77 (2.56, 3.05)
phecode_121-11	Acute lymphoid leukemia	3.06 (2.82, 3.31)	3.04 (2.81, 3.29)
phecode_121-12	Acute myeloid leukemia	2.83 (2.64, 3.21)	2.77 (2.59, 3.13)
phecode_121-2	Chronic leukemia	2.57 (2.46, 2.73)	2.5 (2.4, 2.63)
phecode_121-21	Chronic lymphoid leukemia	2.6 (2.48, 2.76)	2.53 (2.41, 2.67)
phecode_121-22	Chronic myloid leukemia	2.56 (2.4, 2.7)	2.52 (2.36, 2.67)
phecode_121-23	Chronic myelomonocytic (monocytic) leukemia	3.55 (3.3, 4.21)	3.41 (3.16, 3.95)
phecode_122	Lymphoma	2.59 (2.39, 2.75)	2.53 (2.33, 2.65)
phecode_122-1	Hodgkin lymphoma	2.55 (2.36, 2.62)	2.53 (2.34, 2.62)
phecode_122-2	Non-Hodgkin lymphoma	2.6 (2.39, 2.8)	2.54 (2.34, 2.71)
phecode_122-21	Follicular lymphoma	2.38 (2.3, 2.56)	2.35 (2.27, 2.52)
phecode_122-22	Diffuse large B-cell lymphoma*	2.8 (2.62, 2.97)	2.75 (2.56, 2.89)
phecode_122-24	T-cell lymphoma	3.35 (3.06, 3.66)	3.31 (3.01, 3.61)
phecode_123	Multiple myeloma and malignant plasma cell neoplasms	2.65 (2.5, 2.79)	2.58 (2.41, 2.7)
phecode_123-1	Multiple myeloma	2.67 (2.5, 2.82)	2.61 (2.41, 2.74)
phecode_124	Myeloproliferative disorder	2.93 (2.73, 3.1)	2.9 (2.69, 3.07)
phecode_124-3	Polycythemia vera	0 (0, 2.64)	2.57 (0, 2.73)
phecode_124-5	Essential thrombocythemia	2.77 (2.61, 2.92)	2.75 (2.6, 2.9)
phecode_124-6	Myelodysplastic syndrome	2.85 (2.67, 3.05)	2.77 (2.58, 2.93)
phecode_124-7	Chronic myeloproliferative disease	2.58 (0.05, 2.73)	0.08 (0, 2.51)
phecode_125	Other malignant neoplasms of lymphoid, hematopoietic and related tissue	3.09 (2.87, 3.36)	3.1 (2.86, 3.35)
phecode_130	Cancer (solid tumor, excluding BCC)	2.82 (2.59, 2.91)	2.31 (2.14, 2.39)
phecode_132	Sequelae of cancer	2.63 (2.51, 2.8)	2.58 (2.47, 2.76)
phecode_135	Benign neoplasm of the head and neck	2.47 (2.28, 2.58)	2.45 (2.29, 2.58)
phecode_135-1	Benign neoplasm of the oral cavity	3.38 (2.74, 4.63)	3.46 (2.79, 4.66)
phecode_135-12	Benign neoplasm of the tongue	7.24 (4.52, 15.35)	7.65 (5.17, 18.07)
phecode_135-16	Benign neoplasm of the salivary glands	3.68 (2.92, 4.78)	3.66 (3.12, 4.74)
phecode_135-5	Benign neoplasm of the paranasal sinus and nasal cavity	2.47 (2.37, 2.56)	2.43 (2.34, 2.56)
phecode_135-6	Benign neoplasm of vocal cord or larynx	3.58 (3.38, 4.08)	3.76 (3.48, 4.29)
phecode_136	Benign neoplasm of the digestive organs	2.61 (2.42, 2.78)	2.49 (2.33, 2.64)
phecode_136-1	Benign neoplasm of the esophagus	3.59 (3.02, 4.02)	3.65 (3.08, 4.1)
phecode_136-2	Benign neoplasm of stomach	2.82 (2.59, 2.92)	2.79 (2.54, 2.88)
phecode_136-3	Benign neoplasm of the small intestine	3.55 (2.72, 3.97)	3.44 (2.67, 3.87)
phecode_136-4	Benign neoplasm of colon, rectum, anus and anal canal	2.58 (2.38, 2.8)	2.43 (2.27, 2.63)
phecode_136-41	Benign neoplasm of the colon	2.57 (2.38, 2.81)	2.39 (2.25, 2.59)
phecode_136-42	Benign neoplasm of rectum and anus	2.7 (2.47, 2.9)	2.57 (2.37, 2.83)
phecode_136-6	Benign neoplasm of the liver and intrahepatic bile ducts	3.81 (3.26, 4.47)	3.85 (3.27, 4.54)
phecode_136-8	Benign neoplasm of the pancreas	3.95 (3.28, 5.65)	3.87 (3.08, 5.7)
phecode_137	Benign neoplasm of the thoracic and respiratory organs	3.52 (3.06, 3.95)	3.49 (2.98, 3.97)
phecode_137-5	Benign neoplasm of the heart, mediastinum, thymus, and pleura	4.14 (3.39, 4.76)	4.34 (3.55, 5.19)
phecode_138	Benign neoplasm of the skin	2.65 (2.54, 2.81)	2.65 (2.54, 2.84)
phecode_138-1	Nevus, non-neoplastic	2.99 (2.75, 3.32)	3.09 (2.82, 3.37)
phecode_138-2	Melanocytic nevi*	2.71 (2.59, 2.85)	2.71 (2.56, 2.86)
phecode_139	Benign sarcoma-related cancers	2.6 (2.46, 2.79)	2.59 (2.47, 2.81)
phecode_139-1	Benign neoplasms of the bone and/or cartilage	3.82 (2.92, 4.62)	3.85 (3.11, 4.7)
phecode_139-3	Benign neoplasm of other connective and soft tissue	3.13 (2.71, 3.7)	3.29 (2.94, 4.15)
phecode_139-4	Benign neoplasm of peripheral nerves*	3.57 (2.91, 4.53)	3.7 (2.95, 4.89)
phecode_139-5	Lipoma	2.54 (2.34, 2.83)	2.55 (2.32, 2.82)

Supplementary Tables

Table 18 continued from previous page

Code	Name	unadjusted HR (per 1 SD Risk State)	adj. HR for Age and Sex (per 1 SD Risk State)
phecode_139-51	Lipomatosis*	4.17 (3.36, 5.23)	4.16 (3.35, 5.18)
phecode_139-52	Lipoma of intrathoracic organs	3.14 (2.74, 3.67)	3.21 (2.79, 3.72)
phecode_139-53	Lipoma of other skin subcutaneous tissue	2.88 (2.42, 3.6)	2.9 (2.41, 3.64)
phecode_139-54	Testicular limpoma	2.54 (2.28, 3.16)	2.22 (1.98, 2.75)
phecode_139-6	Hemangioma and lymphangioma	2.92 (2.52, 3.45)	2.92 (2.57, 3.4)
phecode_139-61	Hemangioma	2.88 (2.53, 3.43)	2.92 (2.58, 3.38)
phecode_139-62	Lymphangioma	4.57 (3.8, 5.83)	4.65 (3.75, 6.02)
phecode_140	Benign neoplasm of the breast	2.87 (2.61, 3.22)	1.85 (1.73, 2.06)
phecode_142	Lump or mass in breast or nonspecific abnormal breast exam	2.5 (2.37, 2.58)	2.1 (2.03, 2.15)
phecode_142-1	Lump or mass in breast	2.45 (2.33, 2.56)	2.23 (2.16, 2.35)
phecode_142-2	Abnormal mammogram	6.96 (5.69, 9.61)	9.01 (7.06, 10.64)
phecode_142-21	Mammographic microcalcification	6.91 (5.69, 9.61)	8.95 (7.06, 10.64)
phecode_144	Gynecological benign neoplasms	2.4 (2.21, 2.58)	2.04 (1.92, 2.19)
phecode_144-1	Benign neoplasms of external female genital organs and cervix	2.66 (2.36, 2.85)	2.35 (2.17, 2.52)
phecode_144-11	Benign neoplasms of the vulva	3.35 (2.7, 4.38)	3.41 (2.63, 4.53)
phecode_144-12	Benign neoplasms of the vagina	3.27 (2.67, 3.74)	3.26 (2.66, 3.73)
phecode_144-13	Benign neoplasms of the cervix	2.7 (2.41, 2.9)	2.39 (2.2, 2.59)
phecode_144-2	Benign neoplasms of the uterus	2.45 (2.26, 2.6)	2.08 (1.94, 2.23)
phecode_144-21	Leiomyoma of uterus	2.43 (2.31, 2.62)	1.95 (1.86, 2.12)
phecode_144-3	Benign neoplasms of the ovary	2.37 (2.04, 2.81)	2.4 (2.04, 2.86)
phecode_146	Benign neoplasm of the genitourinary system	3.07 (2.93, 3.33)	2.81 (2.66, 3.01)
phecode_146-2	Benign neoplasm of the prostate	3.07 (2.93, 3.34)	2.92 (2.78, 3.15)
phecode_146-4	Benign neoplasm of the kidney	3.89 (3.26, 4.58)	4.11 (3.26, 4.55)
phecode_146-5	Benign neoplasm of the bladder	3.37 (3.09, 3.81)	3.25 (2.91, 3.57)
phecode_148	Benign neoplasm of the eye, brain and other parts of central nervous system	2.82 (2.68, 3.03)	2.75 (2.61, 3)
phecode_148-1	Benign neoplasm of eye	3.35 (3.04, 3.96)	3.59 (3.19, 4.12)
phecode_148-16	Benign neoplasm of choroid	3.4 (3.02, 4.19)	3.79 (3.16, 4.37)
phecode_148-2	Benign neoplasm of meninges (Meningioma)	2.76 (2.54, 3.07)	2.69 (2.48, 3)
phecode_148-3	Benign neoplasm of brain	3.15 (2.93, 3.66)	3.16 (2.93, 3.68)
phecode_148-5	Benign neoplasm of cranial nerve	2.49 (2.34, 2.66)	2.48 (2.37, 2.65)
phecode_149	Benign neoplasm of the endocrine glands	2.86 (0.09, 3.16)	2.87 (2.74, 3.17)
phecode_149-1	Benign neoplasm of the thyroid	4.36 (3.72, 5.06)	4.82 (3.79, 5.74)
phecode_149-3	Benign neoplasm of the parathyroid gland	3.82 (3.38, 4.32)	3.97 (3.55, 4.6)
phecode_149-4	Benign neoplasm of the pituitary gland and craniopharyngeal duct	2.5 (0.35, 2.69)	0.31 (0.03, 2.58)
phecode_153	Benign neoplasm of other or unspecified sites	3.57 (3.1, 4.34)	3.77 (3.22, 4.63)
phecode_159	Genetic susceptibility to malignant neoplasm	3.49 (3.16, 3.85)	3.68 (3.26, 4.07)
phecode_159-1	Genetic susceptibility to malignant neoplasm of breast	3.44 (3.13, 3.83)	3.61 (3.25, 4.09)
phecode_160	Nutritional anemias	2.72 (2.63, 2.87)	2.68 (2.59, 2.81)
phecode_160-1	Iron deficiency anemia	2.72 (2.61, 2.84)	2.68 (2.57, 2.78)
phecode_160-2	Megaloblastic anemia	2.74 (2.56, 2.9)	2.72 (2.56, 2.88)
phecode_160-4	Other deficiency anemia	2.97 (2.8, 3.25)	2.93 (2.8, 3.24)
phecode_161	Hemolytic anemias	2.72 (2.53, 2.89)	2.69 (2.51, 2.87)
phecode_161-1	Intrinsic (hereditary) hemolytic anemias	2.85 (2.66, 3.15)	2.85 (2.65, 3.16)
phecode_161-2	Extrinsic (acquired) hemolytic anemias	2.86 (2.63, 3.14)	2.82 (2.61, 3.1)
phecode_161-21	Autoimmune hemolytic anemias [AIHA]	2.88 (2.67, 3.24)	2.83 (2.63, 3.17)
phecode_162	Aplastic anemia	2.96 (2.88, 3.13)	2.92 (2.84, 3.09)
phecode_162-8	Pancytopenia	3.2 (3.03, 3.53)	3.26 (3.03, 3.53)
phecode_164	Anemia	2.8 (2.64, 2.99)	2.69 (2.55, 2.86)
phecode_164-1	Microcytic anemia	2.72 (2.61, 2.84)	2.67 (2.57, 2.78)
phecode_164-2	Macrocytic anemia	2.76 (2.57, 2.9)	2.73 (2.56, 2.87)
phecode_164-3	Acute posthemorrhagic anemia	3.37 (2.92, 3.72)	3.4 (2.91, 3.75)

4 Medical history predicts future health trajectories over the human phenome

Table 18 continued from previous page

Code	Name	unadjusted HR (per 1 SD Risk State)	adj. HR for Age and Sex (per 1 SD Risk State)
phecode_164-6	Anemia secondary to chronic diseases and conditions	2.67 (2.55, 2.87)	2.65 (2.53, 2.84)
phecode_164-62	Anemia in neoplastic disease	3.05 (2.85, 3.33)	2.96 (2.77, 3.23)
phecode_165	Hemoglobinopathies	2.79 (2.66, 2.96)	2.82 (2.66, 3.02)
phecode_165-2	Thalassemia	2.69 (2.55, 2.87)	2.72 (2.55, 2.93)
phecode_165-25	Thalassemia minor	2.83 (2.75, 3.04)	2.9 (2.75, 3.11)
phecode_165-3	Hemoglobin C trait [Sickle-cell trait]	2.59 (2.52, 2.7)	2.62 (2.54, 2.72)
phecode_168	Coagulation defects, purpura and other hemorrhagic conditions	2.76 (2.62, 2.9)	2.74 (2.61, 2.88)
phecode_168-1	Hypo-coagulability	2.84 (2.63, 2.9)	2.77 (2.6, 2.86)
phecode_168-11	Hereditary hypo-coagulability	2.52 (2.34, 2.64)	2.55 (2.39, 2.65)
phecode_168-12	Hemorrhagic disorder due to intrinsic circulating anticoagulants	3.33 (3.06, 3.47)	3.32 (3.05, 3.47)
phecode_168-15	Acquired coagulation factor deficiency	3.66 (3.32, 4.14)	3.69 (3.33, 4.16)
phecode_168-18	Other nonthrombocytopenic purpura	2.96 (2.76, 3.28)	2.77 (2.59, 3.08)
phecode_168-19	Spontaneous ecchymoses	2.79 (2.67, 3.02)	2.82 (2.68, 3.09)
phecode_168-2	Hyper-coagulability	0.54 (0.03, 2.82)	0.21 (0, 2.7)
phecode_168-21	Primary hypercoagulable state [Primary thrombophilia]	0.21 (0, 2.66)	0.07 (0, 0.75)
phecode_168-211	Activated protein C resistance*	3.09 (2.89, 3.44)	3.07 (2.88, 3.42)
phecode_168-214	Antiphospholipid syndrome*	2.79 (2.69, 2.95)	2.82 (2.69, 2.94)
phecode_168-3	Hereditary deficiency of other clotting factors	2.63 (0.37, 2.91)	2.63 (0.03, 2.91)
phecode_168-4	Abnormal coagulation profile	2.77 (2.63, 2.87)	2.77 (2.64, 2.88)
phecode_169	Platelet defects	2.91 (2.71, 3.08)	2.84 (2.64, 3.01)
phecode_169-1	Thrombocytopenia	2.9 (2.72, 3.09)	2.84 (2.66, 3.03)
phecode_169-11	Immune thrombocytopenic purpura [ITP]	2.67 (2.53, 2.83)	2.65 (2.5, 2.81)
phecode_169-14	Secondary thrombocytopenia	3.2 (2.94, 3.44)	3.24 (2.93, 3.42)
phecode_170	Decreased white blood cell count	3.13 (2.99, 3.47)	3.1 (2.97, 3.43)
phecode_170-1	Neutropenia	3.12 (3.02, 3.46)	3.08 (2.98, 3.42)
phecode_170-13	Neutropenia due to infection	3.64 (3.18, 4.02)	3.64 (3.18, 4.01)
phecode_170-19	Neutropenia NOS	3.12 (2.88, 3.59)	3.42 (3.08, 3.86)
phecode_170-2	Lymphocytopenia	3.55 (3.13, 4.44)	3.58 (3.16, 4.47)
phecode_171	Increased white blood cell count	3.17 (2.95, 3.42)	3.18 (2.97, 3.42)
phecode_171-1	Lymphocytosis (symptomatic)	3.52 (3.24, 4.06)	3.56 (3.27, 4.1)
phecode_171-7	Eosinophilia	3.21 (2.87, 3.78)	3.19 (2.87, 3.82)
phecode_171-9	Elevated white blood cell count [Leukocytosis] NOS	3.56 (2.96, 3.86)	3.59 (3.01, 3.89)
phecode_172	Other disorders of white blood cells	3.13 (2.86, 3.56)	3.15 (2.85, 3.53)
phecode_172-2	Genetic anomalies of leukocytes	3.36 (3.1, 3.86)	3.45 (3.15, 4)
phecode_174	Diseases of spleen	2.87 (2.56, 3.03)	2.87 (2.55, 3.03)
phecode_174-1	Hyposplenism*	2.76 (0.18, 3)	0.42 (0, 2.82)
phecode_174-2	Splenomegaly	2.95 (2.68, 3.1)	2.96 (2.68, 3.14)
phecode_174-6	Cyst of spleen*	4.01 (3.1, 4.78)	4.16 (3.14, 4.94)
phecode_174-7	Infarction of spleen*	3.33 (3.12, 3.83)	3.35 (3.13, 3.82)
phecode_175	Polycythemias	2.76 (2.64, 3.05)	2.76 (0.45, 3.07)
phecode_175-2	Secondary polycythemia	2.77 (2.66, 3.1)	2.77 (0.58, 3.1)
phecode_176	Other diseases of blood and blood-forming organs	3 (2.77, 3.4)	2.97 (2.77, 3.36)
phecode_177	Abnormality of the lymph nodes	2.97 (2.67, 3.14)	3 (2.71, 3.17)
phecode_177-1	Lymphadenitis	3.48 (2.84, 3.78)	3.48 (2.95, 3.94)
phecode_177-13	Acute lymphadenitis	3.65 (2.79, 4.21)	3.7 (2.9, 4.71)
phecode_177-2	Enlargement of lymph nodes [Lymphadenopathy]	3 (2.82, 3.38)	3.08 (2.87, 3.41)
phecode_177-3	Lymphangitis	3.58 (3.24, 4.2)	3.65 (3.3, 4.22)
phecode_177-4	Lymphedema	2.78 (2.65, 2.94)	2.83 (2.65, 2.98)
phecode_179	Immunodeficiencies	2.8 (2.63, 3.06)	2.82 (2.65, 3.07)
phecode_179-1	Hypogammaglobulinemia NOS	2.84 (2.55, 3.05)	2.84 (0.48, 3.05)
phecode_179-9	Immunodeficiency NOS	2.86 (2.57, 3.11)	2.85 (2.59, 3.11)

Supplementary Tables

Table 18 continued from previous page

Code	Name	unadjusted HR (per 1 SD Risk State)	adj. HR for Age and Sex (per 1 SD Risk State)
phecode_180	Other disorders involving the immune mechanism	2.85 (2.65, 2.99)	2.82 (2.64, 2.96)
phecode_180-3	Paraproteinemias	3.01 (2.81, 3.25)	2.85 (2.67, 3.07)
phecode_180-31	Monoclonal gammopathy	3.05 (2.79, 3.34)	2.84 (2.62, 3.15)
phecode_180-33	Macroglobulinemia [Waldenstrom macroglobulinemia]	2.83 (2.66, 3.02)	2.76 (2.61, 2.96)
phecode_181	Autoimmune disease	2.2 (2.14, 2.31)	2.19 (2.12, 2.3)
phecode_200	Disorders of thyroid gland	2.53 (2.44, 2.7)	2.47 (2.37, 2.68)
phecode_200-1	Hypothyroidism	2.58 (2.51, 2.74)	2.58 (2.49, 2.76)
phecode_200-12	Hypothyroidism due to drugs or iatrogenic causes	3.57 (3.15, 4.17)	3.62 (3.19, 4.26)
phecode_200-13	Postprocedural hypothyroidism	2.7 (2.57, 2.93)	2.73 (2.58, 2.93)
phecode_200-14	Hypothyroidism, not specified as secondary	2.58 (2.48, 2.73)	2.54 (2.46, 2.74)
phecode_200-2	Goiter	2.62 (2.5, 2.76)	2.65 (2.42, 2.81)
phecode_200-21	Diffuse goiter	2.78 (2.66, 3.01)	2.8 (2.68, 3.05)
phecode_200-22	Uninodular goiter [single thyroid nodule]	2.8 (2.65, 3.07)	2.9 (2.71, 3.27)
phecode_200-23	Multinodular goiter	2.79 (2.59, 3.07)	2.76 (2.54, 3.16)
phecode_200-3	Thyrotoxicosis [hyperthyroidism]	2.58 (2.4, 2.74)	2.51 (2.41, 2.72)
phecode_200-31	Graves' disease [Toxic diffuse goiter]	2.8 (2.68, 3.04)	2.8 (0.39, 3.04)
phecode_200-4	Thyroiditis	2.77 (2.62, 2.98)	2.82 (2.67, 3.01)
phecode_200-41	Hashimoto thyroiditis [Chronic lymphocytic thyroiditis]	2.76 (2.65, 2.92)	2.8 (2.69, 2.97)
phecode_200-7	Iodine-deficiency related thyroid disorders*	2.83 (2.59, 3.05)	2.95 (2.56, 3.37)
phecode_200-9	Abnormal thyroid function studies	2.97 (2.69, 3.3)	3.19 (2.75, 3.54)
phecode_202	Diabetes mellitus	2.55 (2.51, 2.65)	2.53 (2.5, 2.63)
phecode_202-1	Type 1 diabetes	2.37 (2.32, 2.47)	2.38 (2.32, 2.47)
phecode_202-2	Type 2 diabetes	2.57 (2.51, 2.66)	2.56 (2.51, 2.64)
phecode_202-3	Secondary diabetes	3.7 (3.39, 4.1)	3.77 (3.41, 4.18)
phecode_202-32	Drug or chemical induced diabetes mellitus*	3.93 (3.47, 4.29)	3.98 (3.51, 4.31)
phecode_202-4	Other specified diabetes*	2.57 (2.49, 2.64)	2.57 (2.49, 2.64)
phecode_203	Metabolic syndrome [Dysmetabolic syndrome X]	3.37 (3.09, 3.88)	3.41 (3.09, 3.86)
phecode_204	Elevated blood glucose level	2.72 (2.6, 2.97)	2.68 (2.54, 2.9)
phecode_204-1	Impaired fasting glucose	2.77 (2.65, 3.05)	2.81 (2.65, 3.04)
phecode_204-2	Impaired glucose tolerance (oral)	2.82 (2.62, 2.96)	2.81 (2.62, 2.93)
phecode_204-4	Prediabetes*	4.29 (3.93, 4.71)	4.49 (4.06, 4.98)
phecode_205	Hypoglycemia	2.51 (2.38, 2.7)	2.5 (2.37, 2.69)
phecode_205-3	Drug-induced hypoglycemia*	2.94 (2.74, 3.02)	2.91 (2.75, 3.02)
phecode_206	Disorders of pancreatic internal secretion (excl. DM)	4.94 (4.39, 6.05)	5 (4.49, 6.49)
phecode_208	Disorders of parathyroid gland	2.76 (2.63, 2.87)	2.73 (2.6, 2.86)
phecode_208-1	Hypoparathyroidism	3.04 (2.87, 3.25)	3.08 (2.89, 3.33)
phecode_208-2	Hyperparathyroidism	2.77 (2.61, 2.93)	2.73 (2.58, 2.92)
phecode_208-21	Primary hyperparathyroidism	3 (2.86, 3.32)	3.04 (2.77, 3.39)
phecode_208-22	Secondary hyperparathyroidism	2.73 (2.62, 2.9)	2.75 (2.63, 2.9)
phecode_209	Disorders of the pituitary gland and its hypothalamic control	2.77 (0.09, 2.96)	2.79 (2.66, 2.99)
phecode_209-1	Pituitary hyperfunction	2.68 (2.57, 2.97)	2.67 (2.55, 2.97)
phecode_209-12	Syndrome of inappropriate secretion of antidiuretic hormone	3.49 (3.15, 3.84)	3.38 (3.07, 3.74)
phecode_209-13	Hyperprolactinemia*	2.68 (2.53, 2.92)	2.66 (2.53, 2.91)
phecode_209-2	Hypopituitarism	2.53 (2.41, 2.7)	0.17 (0.02, 2.59)
phecode_209-21	Diabetes insipidus	2.45 (2.28, 2.63)	2.42 (0.25, 2.61)
phecode_209-22	Hypopituitarism NOS	2.49 (2.4, 2.63)	0.06 (0, 2.57)

4 Medical history predicts future health trajectories over the human phenome

Table 18 continued from previous page

Code	Name	unadjusted HR (per 1 SD Risk State)	adj. HR for Age and Sex (per 1 SD Risk State)
phecode_209-23	Iatrogenic hypopituitarism*	2.54 (2.41, 2.69)	2.54 (2.41, 2.7)
phecode_210	Cushing's syndrome	2.4 (2.26, 2.55)	2.39 (2.26, 2.55)
phecode_211	Disorders of adrenal glands	2.68 (2.56, 2.8)	2.68 (2.57, 2.8)
phecode_211-1	Hyperaldosteronism	3.13 (2.85, 3.31)	3.12 (2.87, 3.34)
phecode_211-2	Adrenocortical insufficiency	2.58 (2.49, 2.74)	2.58 (2.5, 2.74)
phecode_211-21	Primary adrenocortical insufficiency	2.54 (0.47, 2.74)	2.49 (0.14, 2.67)
phecode_211-22	Drug-induced adrenocortical insufficiency*	3.31 (3.16, 3.59)	3.32 (3.16, 3.59)
phecode_214	Ovarian dysfunction	4.05 (3.28, 4.45)	4.01 (3.29, 4.4)
phecode_214-1	Primary ovarian failure	4.08 (3.31, 4.56)	4.13 (3.35, 4.49)
phecode_214-11	Premature menopause	4.25 (3.29, 4.82)	4.39 (3.45, 4.95)
phecode_215	Testicular dysfunction	2.82 (2.7, 3.16)	2.82 (2.7, 3.17)
phecode_215-1	Testicular hypofunction	2.82 (2.71, 3.18)	2.81 (2.71, 3.19)
phecode_229	Other endocrine disorders	3.33 (3.06, 3.77)	3.33 (3.06, 3.77)
phecode_230	Malnutrition and underweight	2.85 (2.68, 2.99)	2.7 (2.56, 2.85)
phecode_230-1	Protein-calorie malnutrition	3.04 (2.78, 3.21)	3.02 (2.78, 3.24)
phecode_230-2	Abnormal loss of weight and underweight	2.83 (2.69, 3.01)	2.68 (2.55, 2.84)
phecode_230-21	Abnormal weight loss	2.82 (2.68, 3.03)	2.67 (2.53, 2.84)
phecode_230-22	Underweight	3.5 (3.18, 3.94)	3.55 (3.27, 4.1)
phecode_230-3	Anorexia	2.82 (2.58, 3.06)	2.74 (2.52, 2.98)
phecode_230-4	Cachexia	3.16 (3.02, 3.42)	3.12 (2.98, 3.38)
phecode_230-5	Early satiety	3.97 (3.62, 4.64)	4.31 (3.98, 4.87)
phecode_232	Vitamin deficiencies	2.66 (2.54, 2.79)	2.67 (2.54, 2.8)
phecode_232-1	Vitamin A deficiency	4.98 (4.43, 6.31)	5.29 (4.72, 6.71)
phecode_232-2	Vitamin B group deficiency	2.6 (2.45, 2.73)	2.57 (2.43, 2.71)
phecode_232-27	Vitamin B12 deficiency	2.65 (2.51, 2.81)	2.65 (2.53, 2.8)
phecode_232-29	Folate deficiency [Vitamin B9]	2.97 (2.76, 3.2)	2.91 (2.73, 3.18)
phecode_232-4	Vitamin D deficiency	2.73 (2.61, 2.85)	2.73 (2.69, 2.89)
phecode_234	Other nutritional deficiencies	4.21 (3.5, 4.65)	4.25 (3.53, 4.8)
phecode_236	Overweight and obesity	2.61 (2.53, 2.77)	2.64 (2.57, 2.8)
phecode_236-1	Obesity	2.61 (2.53, 2.77)	2.64 (2.57, 2.81)
phecode_236-11	Morbid obesity	2.83 (2.63, 3.05)	2.81 (2.63, 3.04)
phecode_236-2	Localized adiposity	2.69 (2.48, 2.9)	2.7 (2.5, 2.9)
phecode_237	Abnormal weight gain	2.69 (2.51, 2.97)	2.71 (2.53, 3.02)
phecode_239	Hyperlipidemia	2.45 (2.32, 2.59)	2.26 (2.15, 2.37)
phecode_239-1	Hypercholesterolemia	2.49 (2.4, 2.67)	2.3 (2.21, 2.42)
phecode_239-11	Pure hypercholesterolemia	2.51 (2.43, 2.68)	2.32 (2.22, 2.43)
phecode_239-12	Familial hypercholesterolemia*	3.08 (2.85, 3.85)	3.24 (2.98, 3.97)
phecode_239-2	Hyperglyceridemia	2.65 (2.48, 2.89)	2.73 (2.57, 2.94)
phecode_239-21	Pure hyperglyceridemia	3.6 (3.28, 4.01)	3.83 (3.51, 4.46)
phecode_239-3	Mixed hyperlipidemia	2.37 (2.25, 2.57)	2.47 (2.31, 2.68)
phecode_240	Disorders of amino-acid transport and metabolism	3.11 (2.85, 3.83)	3.08 (2.86, 3.79)
phecode_241	Disorders of carbohydrate metabolism	3.32 (3.03, 3.71)	3.36 (3.02, 3.79)
phecode_242	Lipid storage disorders	3.64 (2.9, 4.25)	3.7 (3.03, 4.5)
phecode_244	Disorders of lipoprotein metabolism and other lipidemias	2.68 (2.53, 2.85)	2.7 (2.54, 2.91)
phecode_244-4	Lipodystrophy, not elsewhere classified	2.73 (2.55, 3)	2.74 (2.55, 2.99)
phecode_247	Disorders of mineral metabolism and mineral deficiencies	2.79 (2.68, 2.95)	2.73 (2.63, 2.88)
phecode_247-3	Disorder of phosphorus metabolism	3.17 (2.81, 3.41)	3.13 (2.77, 3.36)
phecode_247-4	Disorders of magnesium metabolism	2.92 (2.74, 3.19)	2.89 (2.71, 3.15)
phecode_247-42	Hypomagnesemia*	3.01 (2.77, 3.2)	3 (2.76, 3.19)
phecode_247-5	Disorders of calcium metabolism	2.97 (2.64, 3.34)	2.91 (2.59, 3.25)
phecode_247-51	Hypocalcemia	3.1 (2.84, 3.57)	3.15 (2.86, 3.59)
phecode_247-52	Hypercalcemia	2.99 (2.89, 3.48)	3.02 (2.86, 3.47)
phecode_247-7	Disorders of iron metabolism	2.73 (2.64, 2.87)	2.69 (2.6, 2.81)

Supplementary Tables

Table 18 continued from previous page

Code	Name	unadjusted HR (per 1 SD Risk State)	adj. HR for Age and Sex (per 1 SD Risk State)
phecode_247-71	Hemochromatosis	2.99 (2.66, 3.34)	3.06 (2.79, 3.36)
phecode_247-711	Hereditary hemochromatosis	4 (3.65, 4.44)	4.09 (3.73, 4.68)
phecode_247-72	Iron deficiency	2.71 (2.65, 2.83)	2.68 (2.61, 2.78)
phecode_248	Disorders of plasma-protein metabolism, NEC	3.04 (2.87, 3.32)	3 (2.82, 3.25)
phecode_248-1	Alpha-1-antitrypsin deficiency	3.63 (3.17, 5.27)	3.84 (3.41, 5.3)
phecode_249	Amyloidosis	3.45 (3.16, 3.78)	3.13 (2.89, 3.35)
phecode_249-1	Cerebral amyloid angiopathy*	4.54 (4.04, 5.18)	3.68 (3.25, 4.33)
phecode_251	Disorders of bilirubin excretion	2.58 (2.34, 2.92)	2.63 (2.36, 2.98)
phecode_251-1	Gilbert syndrome*	2.61 (2.38, 2.94)	2.63 (2.4, 3.02)
phecode_252	Other and unspecified disorders of metabolism	3.12 (2.81, 3.64)	3.07 (2.82, 3.62)
phecode_256	Disorders of fluid, electrolyte and acid-base balance	2.86 (2.75, 3.05)	2.78 (2.64, 2.91)
phecode_256-1	Hyperosmolality and/or hyponatremia	2.97 (2.65, 3.38)	2.92 (2.6, 3.26)
phecode_256-2	Hyposmolality and/or hyponatremia	2.91 (2.74, 3.27)	2.75 (2.59, 3.03)
phecode_256-3	Mixed disorder of acid-base balance	2.78 (2.69, 2.98)	2.75 (2.65, 2.94)
phecode_256-31	Acidosis	2.79 (2.69, 2.98)	2.75 (2.66, 2.94)
phecode_256-32	Alkalosis	3.18 (2.9, 3.46)	3.11 (2.87, 3.43)
phecode_256-4	Hyperkalemia [Hyperpotassemia]	2.67 (2.54, 2.88)	2.6 (2.49, 2.82)
phecode_256-5	Hypokalemia [Hypopotassemia]	2.96 (2.74, 3.17)	2.85 (2.67, 3.07)
phecode_256-6	Fluid overload	2.82 (2.61, 3.14)	2.78 (2.58, 3.08)
phecode_256-7	Volume depletion	2.85 (2.71, 2.98)	2.75 (2.62, 2.89)
phecode_256-71	Dehydration	2.95 (2.56, 3.33)	2.91 (2.54, 3.28)
phecode_257	Polydipsia	2.77 (2.53, 3.03)	2.81 (2.51, 3.06)
phecode_280	Substance related disorders	2.75 (2.56, 2.98)	2.73 (2.53, 2.95)
phecode_280-1	Alcohol use disorders	2.73 (2.55, 2.89)	2.72 (2.53, 2.86)
phecode_280-11	Alcohol abuse	2.76 (2.63, 2.92)	2.76 (2.58, 2.9)
phecode_280-12	Alcohol dependence	2.63 (2.52, 2.75)	2.62 (2.5, 2.71)
phecode_280-13	Alcoholic liver disease	2.62 (2.51, 2.8)	2.61 (2.48, 2.77)
phecode_280-14	Alcoholic pancreatitis	2.61 (2.49, 2.83)	2.57 (2.44, 2.77)
phecode_280-2	Opioid related disorders	2.77 (2.61, 2.92)	2.77 (2.6, 2.91)
phecode_280-22	Opioid dependence	2.87 (2.65, 3.02)	2.86 (2.63, 3.02)
phecode_280-3	Cannabis related disorders	2.81 (2.6, 3.02)	2.73 (2.52, 2.95)
phecode_280-31	Cannabis abuse	2.88 (2.7, 3.16)	2.81 (2.64, 3.08)
phecode_280-4	Sedative, hypnotic, or anxiolytic related disorders	2.94 (2.72, 3.17)	3.03 (2.8, 3.27)
phecode_280-42	Sedative, hypnotic or anxiolytic-related dependence	2.95 (2.76, 3.18)	3.06 (2.83, 3.28)
phecode_280-8	Other psychoactive substance related disorders	2.67 (2.59, 2.88)	2.66 (2.58, 2.88)
phecode_280-81	Other psychoactive substance abuse	2.75 (2.59, 2.9)	2.69 (2.54, 2.81)
phecode_280-82	Other psychoactive substance dependence	2.7 (2.62, 2.91)	2.69 (2.61, 2.92)
phecode_281	Substance abuse, dependence, and withdrawal	2.7 (2.57, 2.79)	2.61 (2.5, 2.71)
phecode_281-1	Substance abuse	2.77 (2.64, 2.89)	2.74 (2.59, 2.88)
phecode_281-2	Substance dependence	2.72 (2.59, 2.84)	2.67 (2.54, 2.8)
phecode_281-21	Substance withdrawal	2.71 (2.61, 2.9)	2.75 (2.63, 2.91)
phecode_282-1	Current tobacco use and nicotine dependence	2.48 (2.36, 2.62)	2.44 (2.32, 2.57)
phecode_283	Other behavioral problems	2.66 (2.6, 2.82)	2.67 (2.62, 2.83)
phecode_283-3	High risk sexual behavior	2.96 (2.85, 3.32)	3.04 (2.73, 3.31)
phecode_283-4	Patient's noncompliance with medical treatment and regimen	2.7 (2.55, 2.83)	2.7 (2.56, 2.84)
phecode_283-8	Other problems related to lifestyle	2.65 (2.56, 2.82)	2.67 (2.58, 2.84)
phecode_284	Suicide ideation and attempt or self harm	2.51 (2.42, 2.62)	2.49 (2.4, 2.59)
phecode_284-1	Suicidal ideations	2.45 (2.35, 2.58)	2.43 (2.33, 2.54)
phecode_284-2	Suicide and self-inflicted harm	2.62 (2.45, 2.69)	2.61 (2.46, 2.69)

4 Medical history predicts future health trajectories over the human phenome

Table 18 continued from previous page

Code	Name	unadjusted HR (per 1 SD Risk State)	adj. HR for Age and Sex (per 1 SD Risk State)
phecode_284-29	Intentional self-harm*	2.62 (2.46, 2.69)	2.61 (2.46, 2.69)
phecode_286	Mood [affective] disorders	2.56 (2.46, 2.61)	2.54 (2.44, 2.59)
phecode_286-1	Bipolar disorder	2.26 (2.15, 2.38)	2.26 (2.16, 2.38)
phecode_286-2	Major depressive disorder	2.55 (2.45, 2.6)	2.53 (2.45, 2.59)
phecode_286-21	Major depressive disorder, recurrent	2.46 (2.4, 2.62)	2.46 (2.39, 2.62)
phecode_286-3	Premenstrual dysphoric disorder	2.97 (2.87, 3.14)	2.09 (2.02, 2.16)
phecode_286-4	Dysthymic disorder	2.47 (2.35, 2.52)	2.46 (2.35, 2.51)
phecode_287	Psychotic disorder	2.37 (2.23, 2.45)	2.36 (2.23, 2.46)
phecode_287-1	Schizophrenia	2.37 (2.24, 2.47)	2.36 (2.24, 2.47)
phecode_287-2	Schizoaffective disorder	2.44 (2.26, 8.62)	2.43 (2.27, 2.52)
phecode_287-4	Delusional disorders	2.47 (2.33, 2.58)	2.46 (2.34, 2.6)
phecode_287-5	Drug-induced psychotic disorder	2.68 (2.5, 2.81)	2.66 (2.49, 2.79)
phecode_288	Anxiety disorders	2.45 (2.37, 2.53)	2.42 (2.34, 2.48)
phecode_288-2	Panic disorder [episodic paroxysmal anxiety]	2.6 (2.51, 2.76)	2.64 (2.51, 2.79)
phecode_288-3	Generalized anxiety disorder	2.6 (2.48, 2.78)	2.6 (2.49, 2.8)
phecode_288-4	Phobic disorders	2.58 (2.49, 2.8)	2.65 (2.52, 2.82)
phecode_288-41	Agoraphobia	2.78 (2.6, 2.99)	2.76 (2.56, 2.94)
phecode_289	Obsessive-compulsive disorder	2.44 (2.35, 2.57)	2.42 (2.32, 2.54)
phecode_290	Reaction to severe stress, and adjustment disorders	2.61 (2.47, 2.7)	2.38 (2.26, 2.47)
phecode_290-1	Posttraumatic stress disorder	2.64 (2.48, 2.84)	2.5 (2.37, 2.73)
phecode_291	Dissociative, conversion and factitious disorders	2.75 (2.63, 3.13)	2.76 (2.63, 3.12)
phecode_292	Somatoform disorders	2.82 (2.56, 3.21)	2.85 (2.55, 3.15)
phecode_293	Eating disorders	2.55 (2.38, 2.7)	2.63 (2.48, 2.86)
phecode_293-1	Anorexia nervosa	2.78 (2.6, 2.96)	2.86 (2.66, 3.1)
phecode_293-4	Polyphagia	3.11 (2.94, 3.5)	3.18 (3.02, 3.66)
phecode_294	Sexual dysfunction and disorders	2.54 (2.49, 2.63)	2.21 (2.16, 2.27)
phecode_296	Specific personality disorders	2.61 (2.44, 2.71)	2.6 (2.42, 2.68)
phecode_296-4	Borderline personality disorder	2.8 (2.67, 2.88)	2.75 (2.61, 2.84)
phecode_299	Mental disorder, not otherwise specified	2.57 (2.4, 2.66)	2.58 (2.41, 2.67)
phecode_308	Signs and symptoms involving emotional state	2.63 (2.51, 2.69)	2.63 (2.52, 2.68)
phecode_308-1	Irritability	2.87 (2.68, 3.19)	2.78 (2.61, 3.14)
phecode_308-3	Emotional lability	2.78 (2.64, 2.99)	2.8 (2.55, 2.97)
phecode_308-4	Demoralization and apathy	3.54 (3.17, 4.17)	3.54 (3.25, 4.18)
phecode_308-5	Nervousness	3.09 (2.65, 3.49)	3.11 (2.71, 3.56)
phecode_308-6	Excessive crying of child, adolescent, or adult	2.86 (2.66, 3.06)	2.86 (2.62, 3.11)
phecode_308-7	Restlessness and agitation*	2.87 (2.7, 3.22)	2.82 (2.65, 3.16)
phecode_320	Meningitis	3.43 (3.04, 3.91)	3.44 (3.04, 3.92)
phecode_320-1	Infective meningitis	3.77 (3.41, 4.55)	3.81 (3.44, 4.56)
phecode_320-11	Bacterial meningitis	4 (3.43, 4.84)	4.01 (3.43, 4.91)
phecode_320-12	Viral meningitis	3.87 (3.57, 4.51)	4.04 (3.7, 4.83)
phecode_320-3	Meningitis NOS	3.46 (3.14, 3.9)	3.47 (3.14, 3.9)
phecode_321	Encephalitis, myelitis and encephalomyelitis	2.96 (2.8, 3.29)	3.01 (2.8, 3.29)
phecode_321-1	Encephalitis	3.34 (3.06, 3.68)	3.33 (3.06, 3.67)
phecode_321-12	Viral encephalitis	3.96 (3.47, 4.37)	3.99 (3.45, 4.38)
phecode_321-2	Myelitis	2.82 (2.71, 3.08)	2.83 (2.69, 3.1)
phecode_321-21	Acute (transverse) myelitis	2.83 (2.7, 3.04)	2.83 (2.68, 3.03)
phecode_322	Other CNS infection	3.28 (2.9, 3.51)	3.27 (2.9, 3.51)
phecode_322-4	Intracranial and intraspinal abscess	3.75 (3.2, 4.29)	3.73 (3.18, 4.24)
phecode_323	Systemic atrophies primarily affecting the central nervous system	2.95 (2.74, 3.19)	2.9 (2.69, 3.11)
phecode_323-1	Hereditary ataxia	2.77 (2.52, 3.08)	2.78 (2.52, 3.07)
phecode_323-3	Motor neuron disease	3.35 (2.99, 3.86)	3.18 (2.85, 3.63)
phecode_323-31	Amyotrophic lateral sclerosis [ALS]	5.16 (4.62, 5.86)	5.26 (4.72, 5.87)
phecode_324	Extrapyramidal and movement disorders	2.6 (2.48, 2.77)	2.5 (2.39, 2.64)
phecode_324-1	Parkinsonism	2.78 (2.64, 2.98)	2.32 (2.18, 2.44)
phecode_324-11	Parkinson's disease (Primary)	2.76 (2.66, 2.97)	2.27 (2.14, 2.39)
phecode_324-12	Secondary parkinsonism	2.95 (2.84, 3.35)	2.9 (2.8, 3.3)

Supplementary Tables

Table 18 continued from previous page

Code	Name	unadjusted HR (per 1 SD Risk State)	adj. HR for Age and Sex (per 1 SD Risk State)
phecode_324-2	Degenerative diseases of the basal ganglia (excluding parkinsons)	3.88 (3.47, 4.44)	3.77 (3.36, 4.19)
phecode_324-21	Progressive supranuclear ophthalmoplegia [Steele-Richardson-Olszewski]*	4.47 (3.92, 5.03)	4.21 (3.7, 4.74)
phecode_324-3	Dystonia	2.74 (2.49, 2.91)	2.75 (2.56, 2.93)
phecode_324-34	Torticollis	2.77 (2.53, 3.11)	2.85 (2.57, 3.12)
phecode_324-36	Blepharospasm	3.1 (2.89, 3.62)	3.19 (2.85, 3.68)
phecode_324-4	Tremor	2.66 (2.4, 2.84)	2.51 (2.29, 2.71)
phecode_324-41	Essential tremor*	2.71 (2.56, 2.89)	2.47 (2.35, 2.64)
phecode_324-5	Myoclonus	3.03 (2.76, 3.22)	3.03 (2.76, 3.22)
phecode_324-8	Restless legs syndrome	2.57 (2.38, 2.81)	2.55 (2.38, 2.83)
phecode_325	Symptoms and signs related to movement disorders	2.81 (2.7, 2.98)	2.66 (2.57, 2.81)
phecode_325-1	Abnormal involuntary movements	3.21 (2.94, 3.76)	3.24 (2.95, 3.79)
phecode_325-12	Fasciculation*	3.73 (3.32, 4.47)	3.82 (3.34, 4.52)
phecode_325-2	Abnormality of gait and mobility	2.83 (2.69, 2.94)	2.68 (2.56, 2.79)
phecode_325-21	Ataxic gait*	3.14 (3, 3.65)	3.1 (3.01, 3.62)
phecode_325-23	Unsteadiness on feet*	2.86 (2.66, 3.09)	2.72 (2.53, 2.87)
phecode_325-3	Lack of coordination	2.78 (2.55, 3.04)	2.69 (2.48, 2.9)
phecode_326	Demyelinating diseases of the central nervous system	2.32 (2.22, 2.43)	2.29 (2.19, 2.41)
phecode_326-1	Multiple sclerosis	2.33 (2.25, 2.43)	2.3 (0.86, 2.41)
phecode_327	Other degenerative diseases of nervous system	2.99 (2.77, 3.12)	2.85 (2.63, 3)
phecode_328	Dementias and cerebral degeneration	3.07 (2.92, 3.3)	2.41 (2.29, 2.56)
phecode_328-1	Alzheimer's disease	3.2 (2.92, 3.47)	1.97 (1.87, 2.17)
phecode_328-2	Frontotemporal dementia	3.69 (3.37, 4.75)	3.44 (3.12, 4.32)
phecode_328-4	Dementia with Lewy bodies	2.96 (2.79, 3.26)	2.85 (2.73, 3.19)
phecode_328-7	Vascular dementia	2.99 (2.71, 3.23)	2.58 (2.36, 2.77)
phecode_328-8	Dementia in conditions classified elsewhere	3.11 (2.95, 3.45)	2.27 (2.18, 2.54)
phecode_328-9	Dementia NOS	3.12 (2.92, 3.27)	2.6 (2.45, 2.71)
phecode_329	Symptoms and signs involving cognitive functions and awareness	2.89 (2.71, 3)	2.59 (2.46, 2.71)
phecode_329-1	Memory loss	2.7 (2.51, 2.79)	2.47 (2.3, 2.53)
phecode_329-4	Other specified cognitive deficit	2.89 (2.69, 3.14)	2.92 (2.72, 3.22)
phecode_329-41	Attention and concentration deficit	2.95 (2.67, 3.12)	2.91 (2.58, 3.1)
phecode_329-42	Cognitive communication deficit	2.94 (2.61, 3.25)	3.01 (2.64, 3.32)
phecode_329-5	Mild cognitive impairment, so stated	2.95 (2.75, 3.22)	2.4 (2.24, 2.6)
phecode_329-6	Transient global amnesia	2.91 (2.59, 3.13)	2.54 (2.22, 2.85)
phecode_329-8	Altered mental status, unspecified	3.79 (3.34, 4.2)	3.85 (3.3, 4.35)
phecode_329-9	Delirium	2.98 (2.83, 3.13)	2.75 (2.61, 2.88)
phecode_330	Epilepsy, recurrent seizures, convulsions	2.75 (2.57, 2.9)	2.74 (2.56, 2.86)
phecode_330-1	Epilepsy	2.7 (0, 2.82)	2.69 (2.56, 2.81)
phecode_330-11	Generalized epilepsy	2.41 (2.24, 2.56)	2.43 (2.25, 2.56)
phecode_330-12	Partial epilepsy	2.32 (2.19, 2.41)	2.33 (2.19, 2.41)
phecode_330-3	Convulsions	2.55 (2.4, 2.64)	2.53 (2.39, 2.62)
phecode_331	Headache	2.45 (2.37, 2.59)	2.45 (2.39, 2.55)
phecode_331-1	Tension headache	2.68 (2.51, 2.89)	2.66 (2.48, 2.86)
phecode_331-3	Headache syndromes, non migraine	3.05 (2.73, 3.36)	3.05 (2.78, 3.47)
phecode_331-4	Cluster headaches	3.14 (2.87, 3.32)	3.1 (2.8, 3.24)
phecode_331-6	Migraine	2.3 (2.21, 2.43)	2.21 (2.14, 2.37)
phecode_331-61	Migraine with aura	2.74 (2.55, 2.98)	2.78 (2.57, 2.99)
phecode_331-62	Hemiplegic migraine	3.36 (3.13, 3.66)	3.44 (3.08, 3.82)
phecode_331-7	Drug induced headache	3.15 (2.88, 3.44)	3.17 (2.93, 3.48)
phecode_331-8	Headache NOS	2.55 (2.46, 2.6)	2.55 (2.47, 2.59)
phecode_333	Sleep disorders	2.61 (2.52, 2.72)	2.63 (2.53, 2.74)
phecode_333-1	Sleep apnea	2.54 (2.43, 2.69)	2.56 (2.45, 2.73)
phecode_333-11	Obstructive sleep apnea	2.83 (2.63, 3.07)	2.84 (2.66, 3.08)

4 Medical history predicts future health trajectories over the human phenome

Table 18 continued from previous page

Code	Name	unadjusted HR (per 1 SD Risk State)	adj. HR for Age and Sex (per 1 SD Risk State)
phecode_333-2	Insomnia	2.58 (2.46, 2.7)	2.6 (2.45, 2.73)
phecode_333-3	Hypersomnia	3 (2.89, 3.21)	3 (2.88, 3.21)
phecode_333-4	Circadian rhythm sleep disorder	3.41 (3.25, 3.71)	3.42 (3.27, 3.74)
phecode_333-43	Circadian rhythm sleep disorder, jet lag or shift work type	6.25 (4.96, 7.49)	6.62 (5.12, 7.39)
phecode_333-5	Parasomnia and sleep arousal Disorders	2.87 (2.64, 3.22)	2.88 (2.64, 3.22)
phecode_334	Disorders of other cranial nerves	2.66 (2.41, 2.82)	2.7 (2.41, 2.85)
phecode_334-1	Trigeminal nerve disorders [CN5]	2.59 (2.38, 2.75)	2.59 (2.35, 2.79)
phecode_334-11	Trigeminal neuralgia	2.6 (2.37, 2.81)	2.56 (2.34, 2.89)
phecode_334-12	Atypical face pain	2.84 (2.64, 3.14)	2.92 (2.71, 3.36)
phecode_334-2	Facial nerve disorders and weakness	2.75 (2.43, 3.07)	2.77 (2.43, 3.07)
phecode_334-21	Bell's palsy	2.72 (2.45, 3.31)	2.75 (2.51, 3.33)
phecode_334-23	Facial weakness	3.38 (2.89, 3.76)	3.35 (2.9, 3.82)
phecode_334-24	Clonic hemifacial spasm*	3.26 (2.98, 3.86)	3.35 (3.09, 3.94)
phecode_334-4	Disorders of oculomotor nerves	3.67 (3.21, 4.45)	3.72 (3.28, 4.44)
phecode_334-41	Third [oculomotor] nerve palsy	3.92 (3.45, 4.37)	3.9 (3.48, 4.47)
phecode_334-42	Fourth [trochlear] nerve palsy	4.42 (3.91, 5.41)	4.48 (4.1, 5.6)
phecode_334-44	Sixth [abducent] nerve palsy	3.76 (3.24, 4.52)	3.95 (3.39, 4.78)
phecode_335	Nerve root and plexus disorders	2.81 (2.69, 2.94)	2.82 (2.7, 2.98)
phecode_335-1	Nerve plexus lesions	3.44 (3.15, 3.88)	3.56 (3.4, 4.01)
phecode_335-11	Brachial plexus lesions	3.53 (3.27, 3.95)	3.74 (3.5, 4.19)
phecode_335-2	Nerve root lesions	3.49 (3.24, 4.19)	3.59 (3.38, 4.41)
phecode_335-4	Phantom limb (syndrome)	3.33 (3.11, 3.62)	3.36 (3.14, 3.71)
phecode_336	Mononeuropathies	2.59 (2.41, 2.75)	2.6 (2.43, 2.77)
phecode_336-1	Carpal tunnel syndrome	2.54 (2.33, 2.75)	2.5 (2.33, 2.74)
phecode_336-2	Lesion of median, ulnar, radial nerve	2.78 (2.61, 3.09)	2.82 (2.6, 3.05)
phecode_336-4	Mononeuritis of upper limb	3.32 (2.95, 3.66)	3.34 (2.96, 3.68)
phecode_336-5	Mononeuritis of lower limb	2.67 (2.55, 2.82)	2.68 (2.56, 2.98)
phecode_336-51	Lesion of sciatic nerve	3.35 (2.92, 3.92)	3.61 (3.14, 4.53)
phecode_336-52	Meralgia paresthetica	2.95 (2.73, 3.33)	2.97 (2.75, 3.33)
phecode_336-54	Tarsal tunnel syndrome	3.8 (3.4, 4.71)	4.11 (3.58, 5.04)
phecode_336-55	Lesion of plantar nerve	2.65 (2.51, 2.83)	2.66 (2.49, 2.98)
phecode_337	Polyneuropathies	2.6 (2.49, 2.68)	2.57 (2.46, 2.65)
phecode_337-1	Hereditary and idiopathic neuropathy	2.85 (2.57, 3.12)	2.79 (2.5, 3.12)
phecode_337-11	Hereditary motor and sensory neuropathy	2.92 (2.63, 3.14)	2.88 (2.55, 3.17)
phecode_337-2	Inflammatory polyneuropathy	2.74 (2.4, 2.92)	2.73 (2.37, 2.91)
phecode_337-21	Guillain-Barre syndrome [Acute infective polyneuritis]	2.78 (2.43, 2.95)	2.79 (2.4, 2.93)
phecode_337-3	Toxic neuropathy	3.37 (2.95, 3.71)	3.38 (2.98, 3.73)
phecode_337-31	Drug-induced polyneuropathy	3.46 (3.14, 3.88)	3.47 (3.14, 3.86)
phecode_337-8	Polyneuropathy in diseases classified elsewhere	2.59 (2.47, 2.67)	2.59 (2.47, 2.66)
phecode_338	Myasthenia gravis and other myoneural disorders	2.62 (2.49, 2.77)	2.59 (2.47, 2.75)
phecode_338-1	Myasthenia gravis	2.6 (2.46, 2.8)	2.58 (2.42, 2.76)
phecode_339	Primary disorders of muscles	2.97 (2.74, 3.28)	2.97 (2.74, 3.28)
phecode_339-1	Muscular dystrophy	3.01 (2.74, 3.28)	2.99 (0.43, 3.19)
phecode_340	Myopathies	2.99 (2.75, 3.29)	2.99 (2.76, 3.28)
phecode_341	Cerebral palsy and other paralytic syndromes	2.72 (2.57, 2.87)	2.67 (2.51, 2.81)
phecode_341-1	Cerebral palsy	2.48 (2.38, 2.61)	2.48 (2.38, 2.62)
phecode_341-2	Hemiplegia and hemiparesis	2.75 (2.66, 2.94)	2.69 (2.6, 2.87)
phecode_341-6	Cauda equina syndrome	3 (2.55, 3.47)	3.02 (2.55, 3.48)
phecode_342	Plegia and unspecified paralysis	2.71 (2.55, 2.91)	2.71 (2.54, 2.9)
phecode_342-1	Paraplegia/Diplegia	2.53 (2.44, 2.66)	2.53 (2.45, 2.67)

Supplementary Tables

Table 18 continued from previous page

Code	Name	unadjusted HR (per 1 SD Risk State)	adj. HR for Age and Sex (per 1 SD Risk State)
phecode_342-2	Quadriplegia	2.67 (2.55, 2.8)	2.63 (0.43, 2.83)
phecode_342-4	Monoplegia	3 (2.79, 3.26)	2.99 (2.8, 3.26)
phecode_342-5	Paralysis NOS	3.2 (2.97, 3.4)	3.17 (2.94, 3.4)
phecode_343	Disorders of autonomic nervous system	2.79 (2.55, 3.07)	2.75 (2.56, 3.06)
phecode_343-1	Autonomic neuropathy	2.7 (2.51, 2.84)	2.69 (2.52, 2.84)
phecode_343-3	Complex regional pain syndrome	2.9 (2.75, 3.16)	2.9 (2.75, 3.18)
phecode_343-5	Horner's syndrome*	3.94 (3.04, 5.24)	4.08 (3.09, 5.23)
phecode_343-6	Multi-system degeneration of the autonomic nervous system*	3.71 (3.33, 4.31)	3.69 (3.29, 4.29)
phecode_344	Disorders of the circulation of the cerebrospinal fluid	3.17 (2.92, 3.53)	3.18 (2.91, 3.51)
phecode_344-1	Hydrocephalus	3.25 (2.94, 3.67)	3.23 (2.93, 3.6)
phecode_344-12	Obstructive hydrocephalus	4.45 (3.91, 5.47)	4.44 (3.92, 5.48)
phecode_344-13	(Idiopathic) normal pressure hydrocephalus	3.73 (3.25, 4.11)	3.53 (3.05, 3.91)
phecode_344-2	Benign intracranial hypertension	3.39 (3.02, 3.9)	3.52 (3.23, 4.15)
phecode_344-3	Cerebrospinal fluid leak	3.81 (3.49, 4.47)	3.81 (0.03, 4.48)
phecode_345	Encephalopathy	2.93 (2.76, 3.09)	2.93 (2.77, 3.08)
phecode_346	Brain damage and brain death	3.37 (3.02, 4.2)	3.37 (3.02, 4.17)
phecode_346-1	Postconcussion syndrome	3.34 (2.93, 3.65)	3.51 (3.03, 3.78)
phecode_346-3	Anoxic brain damage	3.17 (2.96, 3.37)	3.13 (2.93, 3.31)
phecode_346-5	Compression of brain	4.35 (3.8, 5.71)	4.45 (3.85, 5.65)
phecode_346-6	Cerebral edema	3.66 (3.4, 4.29)	3.62 (3.4, 4.19)
phecode_347	Other disorders of the brain and CNS	3.09 (2.91, 3.46)	3.09 (2.91, 3.43)
phecode_347-1	Cerebral cysts	3.18 (2.99, 3.6)	3.23 (3.04, 3.62)
phecode_347-2	Disorders of meninges	3.82 (3.07, 4.36)	3.88 (3.13, 4.48)
phecode_348	Other diseases of spinal cord	2.8 (2.48, 3.11)	2.78 (2.46, 3.07)
phecode_348-2	Myelopathies	2.8 (2.5, 3.19)	2.78 (2.49, 3.15)
phecode_348-21	Vascular myelopathies	3.21 (2.94, 3.88)	3.13 (2.94, 3.76)
phecode_348-4	Spinal cord compression*	3.24 (2.98, 3.71)	3.18 (2.92, 3.56)
phecode_349	Disorder of nervous system	2.98 (2.72, 3.24)	2.9 (2.68, 3.13)
phecode_349-1	Abnormal findings on diagnostic test of central nervous system	3.1 (2.83, 3.39)	3.04 (2.76, 3.36)
phecode_349-12	Abnormal electroencephalogram [EEG]	2.94 (2.71, 3.21)	2.91 (2.71, 3.23)
phecode_349-13	Abnormal findings on diagnostic imaging of skull and head	3.09 (2.81, 3.34)	3.06 (2.74, 3.24)
phecode_349-15	Intracranial space-occupying lesion found on diagnostic imaging of central nervous system*	3.81 (3.49, 4.61)	3.81 (3.48, 4.66)
phecode_349-2	Abnormal results of function studies of peripheral nervous system	3.71 (3.14, 4.65)	3.87 (3.2, 4.55)
phecode_349-3	Nonspecific abnormal electromyogram [EMG]	3.62 (3.36, 4.34)	3.61 (3.35, 4.34)
phecode_350	Other symptoms involving nervous system	2.87 (2.77, 2.98)	2.7 (2.62, 2.82)
phecode_350-3	Abnormal reflex	2.81 (2.63, 3.03)	2.85 (2.64, 3.07)
phecode_350-5	Repeated falls*	2.84 (2.68, 2.98)	2.62 (2.49, 2.73)
phecode_351	Disturbances of skin sensation	2.51 (2.41, 2.64)	2.52 (2.42, 2.65)
phecode_351-1	Anesthesia of skin*	2.85 (2.65, 3.07)	2.86 (2.66, 3.13)
phecode_351-2	Hypoesthesia of skin*	2.42 (2.33, 2.5)	2.41 (2.31, 2.48)
phecode_351-3	Paresthesia of skin*	2.66 (2.53, 2.78)	2.68 (2.54, 2.79)
phecode_351-4	Hyperesthesia*	2.84 (2.61, 3.17)	2.9 (2.67, 3.24)
phecode_352	Disturbances of sensation of smell and taste	2.81 (2.57, 3.12)	2.85 (2.62, 3.14)
phecode_352-1	Anosmia*	2.95 (2.74, 3.24)	3.05 (2.79, 3.37)
phecode_352-2	Parosmia*	3.73 (3.17, 4.5)	3.81 (3.27, 4.55)
phecode_352-3	Parageusia*	3.14 (2.85, 3.66)	3.26 (2.91, 3.73)
phecode_353	Symptoms and signs involving general sensations and perceptions	3.06 (2.67, 3.24)	2.91 (2.59, 3.13)
phecode_353-1	Hallucinations	2.84 (2.69, 3.11)	2.85 (2.67, 3.09)
phecode_353-11	Auditory hallucinations*	2.74 (2.59, 2.97)	2.75 (2.58, 2.97)
phecode_353-12	Visual hallucinations	2.97 (2.74, 3.23)	2.95 (2.71, 3.19)
phecode_354	Dizziness and giddiness	2.53 (2.42, 2.68)	2.46 (2.36, 2.6)
phecode_355	Coma and other alteration of consciousness	2.85 (2.62, 3.07)	2.75 (2.54, 2.98)
phecode_355-1	Coma	2.99 (2.71, 3.35)	2.92 (2.64, 3.22)
phecode_355-2	Alteration of consciousness	2.82 (2.62, 3.03)	2.69 (2.55, 2.94)
phecode_355-21	Transient alteration of awareness	5.73 (4.48, 7.73)	5.85 (4.56, 7.8)
phecode_356	Speech disturbance	2.9 (2.72, 3.37)	2.78 (2.64, 3.24)

4 Medical history predicts future health trajectories over the human phenome

Table 18 continued from previous page

Code	Name	unadjusted HR (per 1 SD Risk State)	adj. HR for Age and Sex (per 1 SD Risk State)
phecode_356-1	Dysarthria	3.05 (2.88, 3.39)	3.02 (2.82, 3.27)
phecode_356-2	Aphasia and dysphasia	3.07 (2.8, 3.49)	2.9 (2.66, 3.23)
phecode_356-4	Slurred speech*	3.03 (2.81, 3.62)	2.98 (2.81, 3.62)
phecode_360	Inflammation of eyelids	2.71 (2.59, 2.9)	2.67 (2.55, 2.85)
phecode_360-1	Hordeolum	2.87 (2.61, 3.18)	2.9 (2.68, 3.22)
phecode_360-11	Hordeolum externum	2.86 (2.61, 3.18)	2.9 (2.68, 3.22)
phecode_360-12	Hordeolum internum	3.17 (2.95, 3.76)	3.32 (3.05, 3.84)
phecode_360-13	Abscess of eyelid	4.04 (3.32, 4.88)	4.37 (3.42, 5.59)
phecode_360-2	Chalazion	2.85 (2.68, 3.19)	2.85 (2.66, 3.2)
phecode_360-4	Blepharitis	2.68 (2.55, 2.88)	2.62 (2.51, 2.81)
phecode_360-5	Noninfectious dermatoses of eyelid	3.21 (2.85, 3.55)	3.76 (3.3, 4.14)
phecode_360-51	Eczematous dermatitis of eyelid	3.21 (2.86, 3.57)	3.77 (3.34, 4.13)
phecode_361	Disorders of eyelid function	2.87 (2.61, 3.15)	2.51 (2.32, 2.78)
phecode_361-1	Entropion and trichiasis of eyelid	2.95 (2.71, 3.36)	2.38 (2.25, 2.67)
phecode_361-15	Trichiasis of eyelid without entropion	3.52 (3.28, 4.08)	3.65 (3.37, 4.28)
phecode_361-2	Lagophthalmos	4.07 (3.69, 4.5)	4.08 (3.71, 4.5)
phecode_361-3	Ptosis of eyelid	2.93 (2.66, 3.34)	2.78 (2.55, 3.14)
phecode_361-4	Blepharochalasis	3.17 (2.75, 3.8)	2.88 (2.55, 3.47)
phecode_361-9	Ectropion of eyelid	3.06 (2.69, 3.78)	2.47 (2.17, 2.94)
phecode_362	Other disorders of the eyelids	2.78 (2.49, 3.09)	2.75 (2.49, 3.02)
phecode_362-1	Xanthelasma of eyelid	3 (2.81, 3.83)	3.04 (2.87, 3.52)
phecode_362-5	Cysts of eyelid	3.41 (3.05, 4.02)	3.47 (3.17, 4.14)
phecode_362-6	Dermatochalasis of eyelid	5.01 (4.43, 6.48)	5.21 (4.7, 6.64)
phecode_363	Disorders of lacrimal system	2.61 (2.48, 2.66)	2.5 (2.39, 2.57)
phecode_363-2	Dry eye syndrome [Tear film insufficiency]	2.63 (2.48, 2.74)	2.57 (2.42, 2.64)
phecode_363-5	Epiphora	2.67 (2.55, 2.84)	2.53 (2.4, 2.75)
phecode_363-51	Epiphora due to excess lacrimation	3.79 (3.4, 4.71)	4.06 (3.6, 5.26)
phecode_363-6	Inflammation of lacrimal passages	3.25 (2.96, 3.68)	3.49 (3.06, 3.96)
phecode_363-61	Dacryocystitis	3.59 (3.18, 3.95)	4.18 (3.47, 4.75)
phecode_363-7	Stenosis and insufficiency of lacrimal passages	2.8 (2.52, 3.22)	2.73 (2.38, 3.2)
phecode_365	Noninflammatory disorders of the orbit	3.51 (3, 3.96)	3.59 (3.1, 4.16)
phecode_365-2	Orbital edema or congestion	3.41 (2.88, 4.04)	3.78 (3.11, 4.56)
phecode_365-3	Exophthalmos [Proptosis]	3.65 (3.33, 4.06)	3.68 (3.35, 4.07)
phecode_366	Noninflammatory disorders of conjunctiva	3.04 (2.66, 3.61)	3.12 (2.71, 3.63)
phecode_366-1	Pterygium of eye	3.31 (2.85, 3.67)	3.33 (2.88, 3.72)
phecode_366-2	Conjunctival degenerations and deposits	4.3 (3.3, 5.62)	4.8 (3.84, 5.94)
phecode_366-21	Pinguecula	4.6 (3.35, 5.94)	5.42 (3.94, 7.16)
phecode_366-4	Vascular abnormalities of conjunctiva	3.17 (2.91, 4.01)	3.27 (2.98, 4.1)
phecode_366-42	Conjunctival hyperemia	3.45 (3.07, 4.72)	3.62 (3.3, 4.86)
phecode_366-5	Conjunctival edema	4.95 (4.2, 6.12)	5.42 (4.78, 6.55)
phecode_366-6	Conjunctival cysts	4.11 (3.56, 4.74)	4.48 (4, 5.2)
phecode_367	Inflammation of the eye	2.61 (2.54, 2.74)	2.6 (2.49, 2.76)
phecode_367-1	Conjunctivitis	2.59 (2.49, 2.72)	2.58 (2.48, 2.73)
phecode_367-12	Allergic [atopic] conjunctivitis	2.82 (2.68, 2.98)	2.89 (2.67, 3.06)
phecode_367-13	Blepharoconjunctivitis	3.01 (2.73, 3.21)	2.97 (2.72, 3.25)
phecode_367-2	Keratitis	2.85 (2.58, 3.12)	2.88 (2.61, 3.12)
phecode_367-21	Corneal ulcer	3.21 (3.03, 3.67)	3.28 (3.07, 3.67)
phecode_367-22	Punctate keratitis	4.26 (3.32, 5.11)	4.32 (3.57, 5.46)
phecode_367-3	Keratoconjunctivitis	3.24 (2.87, 3.54)	3.36 (2.93, 3.6)
phecode_367-4	Inflammation of orbit	3.43 (2.79, 5.07)	3.66 (2.95, 5.23)
phecode_367-41	Cellulitis of orbit	3.61 (3.14, 5.47)	3.98 (3.28, 5.46)

Supplementary Tables

Table 18 continued from previous page

Code	Name	unadjusted HR (per 1 SD Risk State)	adj. HR for Age and Sex (per 1 SD Risk State)
phecode_367-5	Uveitis	2.56 (2.46, 2.75)	2.55 (2.45, 2.75)
phecode_367-52	Iridocyclitis	2.56 (2.45, 2.75)	2.55 (2.45, 2.75)
phecode_367-6	Episcleritis	3.09 (2.78, 3.52)	3.26 (2.97, 3.59)
phecode_367-7	Scleritis	3.39 (3.07, 3.86)	3.54 (3.13, 4.05)
phecode_367-9	Chorioretinal inflammation	3.2 (2.87, 3.47)	3.25 (2.92, 3.5)
phecode_369	Noninflammatory disorders of the cornea	2.72 (2.57, 2.97)	2.5 (2.36, 2.72)
phecode_369-1	Corneal scars and opacities	3.22 (2.94, 3.4)	3.17 (2.91, 3.37)
phecode_369-2	Corneal edema	3.01 (2.86, 3.19)	2.98 (2.84, 3.16)
phecode_369-4	Corneal degenerations	2.96 (2.61, 3.26)	2.35 (2.08, 2.56)
phecode_369-42	Recurrent erosion of cornea	4.12 (3.52, 4.84)	4.4 (3.81, 5.23)
phecode_369-44	Senile corneal changes including arcus senilis	6.53 (4.57, 8.71)	6.24 (4.74, 8.37)
phecode_369-5	Hereditary corneal dystrophies	2.64 (2.5, 2.91)	2.54 (2.42, 2.8)
phecode_369-51	Fuchs' dystrophy	3.13 (2.94, 3.37)	3.05 (2.9, 3.36)
phecode_369-6	Corneal deformities	2.54 (2.42, 2.67)	2.54 (2.42, 2.69)
phecode_369-62	Keratoconus	2.56 (2.38, 2.68)	2.57 (2.39, 2.7)
phecode_370	Disorders of iris and ciliary body	2.93 (2.73, 3.23)	2.91 (2.7, 3.18)
phecode_370-1	Degeneration of iris and ciliary body	3.15 (3, 3.43)	3.28 (3.1, 3.67)
phecode_370-3	Vascular disorders of iris and ciliary body	3.09 (2.89, 3.36)	3.06 (2.88, 3.38)
phecode_370-4	Adhesions of iris	3.22 (2.97, 3.42)	3.18 (2.96, 3.38)
phecode_371	Cataract	2.81 (2.66, 2.99)	1.93 (1.88, 2.04)
phecode_371-3	Nuclear cataract	2.81 (2.7, 3.05)	2.05 (1.98, 2.17)
phecode_371-31	Age-related nuclear cataract	2.81 (2.7, 3.05)	2.05 (1.98, 2.17)
phecode_372	Disorders of lens (excluding cataracts)	3 (2.72, 3.22)	2.89 (2.63, 3.1)
phecode_372-1	Aphakia	2.91 (2.73, 3.12)	2.89 (2.71, 3.1)
phecode_372-2	Dislocation of lens	3.66 (3.28, 3.95)	3.62 (3.25, 3.91)
phecode_373	Noninflammatory disorders of choroid	2.92 (2.77, 3.19)	2.89 (2.77, 3.13)
phecode_373-1	Chorioretinal scars	2.83 (2.65, 3.03)	2.83 (2.66, 3.03)
phecode_373-2	Choroidal degenerations	3.62 (3.16, 3.93)	3.61 (3.14, 3.92)
phecode_374	Disorders of the retina	2.46 (2.36, 2.51)	2.39 (2.3, 2.45)
phecode_374-1	Retinal detachments and breaks	2.64 (2.48, 2.98)	2.62 (2.47, 2.88)
phecode_374-11	Serous retinal detachment	2.67 (2.48, 3.07)	2.64 (2.52, 3)
phecode_374-12	Traction detachment of retina	3.2 (2.87, 3.8)	3.25 (2.96, 3.95)
phecode_374-13	Horseshoe tear of retina without detachment	11.83 (8.34, 17.9)	13.22 (8.78, 24.67)
phecode_374-14	Round hole of retina without detachment	6.15 (4.65, 8.79)	8.84 (6.66, 10.95)
phecode_374-2	Retinoschisis and retinal cysts	3.45 (2.85, 4.14)	3.69 (3.06, 4.36)
phecode_374-21	Retinoschisis	4.39 (3.32, 5.33)	5.59 (4.47, 6.45)
phecode_374-3	Retinal vascular changes and occlusions	2.59 (2.36, 2.78)	2.4 (2.24, 2.56)
phecode_374-32	Retinal microaneurysms	2.7 (2.43, 2.95)	2.69 (2.45, 2.95)
phecode_374-33	Retinal neovascularization	3.21 (2.99, 3.5)	3.19 (2.91, 3.43)
phecode_374-37	Retinal arterial occlusions	2.96 (2.5, 3.21)	2.84 (2.41, 3.01)
phecode_374-38	Retinal vein occlusions	2.63 (2.41, 2.98)	2.33 (2.18, 2.57)
phecode_374-39	Transient retinal arterial occlusion [Amaurosis fugax]	2.94 (2.58, 3.4)	2.59 (2.3, 3)
phecode_374-4	Retinal disorders in diseases classified elsewhere	2.4 (2.31, 2.49)	2.41 (2.31, 2.5)
phecode_374-41	Hypertensive retinopathy	3.94 (3.22, 4.75)	4.24 (3.31, 5.22)
phecode_374-42	Diabetic retinopathy	2.41 (2.32, 2.49)	2.42 (2.32, 2.51)
phecode_374-5	Macular degeneration	2.69 (2.55, 2.81)	2.07 (1.97, 2.13)

4 Medical history predicts future health trajectories over the human phenome

Table 18 continued from previous page

Code	Name	unadjusted HR (per 1 SD Risk State)	adj. HR for Age and Sex (per 1 SD Risk State)
phecode_374-51	Age-related macular degeneration	2.8 (2.61, 2.95)	2.17 (2.03, 2.25)
phecode_374-511	Nonexudative (dry) age-related macular degeneration	3.11 (2.85, 3.33)	2.54 (2.35, 2.73)
phecode_374-512	Exudative (wet) age-related macular degeneration	3.27 (2.87, 3.45)	2.75 (2.51, 2.9)
phecode_374-52	Macular cyst, hole, or pseudohole	3.29 (2.84, 3.86)	2.85 (2.41, 3.23)
phecode_374-54	Drusen (degenerative) of macula	2.71 (2.53, 2.92)	2.62 (2.44, 2.76)
phecode_374-55	Puckering of macula	2.99 (2.76, 3.27)	2.66 (2.5, 2.89)
phecode_374-6	Peripheral retinal degeneration	3.71 (3.44, 4.8)	3.91 (3.64, 4.95)
phecode_374-61	Lattice degeneration of retina	9.03 (6.73, 11.48)	9.91 (8.38, 11.94)
phecode_374-7	Hereditary retinal dystrophy	2.29 (2.17, 2.44)	2.16 (2.03, 2.29)
phecode_374-8	Retinal edema	2.54 (2.46, 2.76)	2.51 (2.44, 2.72)
phecode_374-9	Central serous chorioretinopathy	3.89 (3.34, 4.63)	3.92 (3.33, 4.39)
phecode_375	Abnormal intraocular pressure	2.35 (2.25, 2.45)	2.12 (2.02, 2.18)
phecode_375-1	Glaucoma	2.43 (2.36, 2.52)	2.22 (2.14, 2.27)
phecode_375-11	Open angle glaucoma	2.6 (2.52, 2.72)	2.49 (2.42, 2.59)
phecode_375-112	Pigmentary glaucoma	3 (2.77, 3.3)	3.12 (2.87, 3.5)
phecode_375-113	Primary open angle glaucoma	2.66 (2.51, 2.76)	2.58 (2.44, 2.67)
phecode_375-12	Angle-Closure Glaucoma	2.78 (2.57, 2.95)	2.64 (2.43, 2.82)
phecode_375-14	Low-tension glaucoma (Normal-tension glaucoma)	2.79 (2.69, 3)	2.63 (2.53, 2.78)
phecode_375-6	Ocular hypertension	2.55 (2.4, 2.81)	2.5 (2.4, 2.72)
phecode_375-7	Hypotony of eye	3.33 (3.14, 3.56)	3.34 (3.15, 3.55)
phecode_376	Disorders of vitreous body	2.81 (2.61, 2.99)	2.59 (2.43, 2.73)
phecode_376-1	Vitreous degeneration	2.95 (2.67, 3.24)	2.82 (2.62, 3.12)
phecode_376-2	Vitreous opacities	2.82 (2.61, 2.98)	2.6 (2.43, 2.73)
phecode_376-21	Crystalline deposits in vitreous body	2.81 (2.67, 3.09)	2.6 (2.45, 2.76)
phecode_376-4	Vitreomacular adhesion	3.54 (2.73, 3.98)	3.34 (2.56, 3.79)
phecode_377	Hemorrhage of the eye	2.74 (2.58, 2.98)	2.57 (2.46, 2.79)
phecode_377-2	Conjunctival hemorrhage	2.68 (2.58, 2.91)	2.52 (2.43, 2.72)
phecode_377-4	Retinal hemorrhage	3.01 (2.66, 3.59)	2.87 (2.5, 3.28)
phecode_377-5	Vitreous hemorrhage	2.73 (2.6, 3)	2.74 (2.58, 3.02)
phecode_377-8	HypHEMA	3.5 (3.29, 4.14)	3.6 (3.41, 4.28)
phecode_379	Infection of the eye	3.08 (2.75, 3.5)	3.07 (2.77, 3.45)
phecode_379-2	Eye infection, viral	3.07 (2.66, 3.52)	3.11 (2.71, 3.49)
phecode_379-21	Infection of the eye, herpes	2.97 (2.6, 3.17)	2.93 (2.54, 3.09)
phecode_380	Disorders of optic nerve and visual pathways	2.93 (2.71, 3.24)	2.92 (2.71, 3.17)
phecode_380-1	Optic neuropathy	3.02 (2.73, 3.31)	3.02 (2.73, 3.31)
phecode_380-11	Optic neuritis	2.83 (2.59, 3.02)	2.84 (2.59, 3.03)
phecode_380-12	Ischemic optic neuropathy	3.77 (3.23, 4.55)	3.99 (3.42, 5.16)
phecode_380-2	Disorders of optic disc	3.05 (2.8, 3.27)	3.04 (2.85, 3.3)
phecode_380-21	Papilledema	4.2 (3.59, 5.73)	4.17 (3.49, 5.73)
phecode_380-22	Optic disc drusen	5.6 (4.34, 8.78)	6.37 (5.24, 10.68)
phecode_380-3	Optic atrophy	3.08 (2.88, 3.36)	3.07 (2.87, 3.35)
phecode_381	Strabismus	3.28 (2.96, 3.43)	3.26 (2.94, 3.43)
phecode_381-1	Paralytic strabismus [Neurogenic strabismus]	3.58 (3.24, 4.2)	3.61 (3.25, 4.2)
phecode_381-11	Ophthalmoplegia	3.86 (3.66, 4.51)	3.75 (3.57, 4.3)
phecode_381-3	Esotropia	3.48 (3.18, 3.68)	3.49 (3.18, 3.74)
phecode_381-4	Exotropia	3.29 (2.83, 3.51)	3.32 (2.83, 3.52)
phecode_381-6	Duane's syndrome [Duane anomaly]	3.31 (3.09, 3.94)	3.31 (3.08, 3.91)
phecode_381-8	Heterophoria	3.58 (3.22, 4.13)	3.62 (3.27, 4.15)

Supplementary Tables

Table 18 continued from previous page

Code	Name	unadjusted HR (per 1 SD Risk State)	adj. HR for Age and Sex (per 1 SD Risk State)
phecode_381-81	Esophoria	4.9 (3.85, 5.5)	5.15 (4.05, 5.93)
phecode_381-82	Exophoria	3.85 (3.39, 4.33)	4.01 (3.55, 4.67)
phecode_382	Other disorders of binocular movement	3.64 (2.98, 4.17)	3.83 (2.99, 4.26)
phecode_383	Irregular eye movements	3.08 (2.72, 3.45)	3.1 (2.71, 3.45)
phecode_383-1	Nystagmus	3.28 (2.81, 3.61)	3.37 (2.84, 3.6)
phecode_384	Anomalies of pupillary function	3.05 (2.76, 3.24)	3.01 (2.76, 3.22)
phecode_384-1	Anisocoria	6.2 (4.88, 9.43)	6.3 (5.12, 9.73)
phecode_384-3	Mydriasis	2.75 (2.51, 3.14)	2.84 (2.6, 3.24)
phecode_384-4	Tonic pupil	4.36 (3.52, 5.5)	6.27 (4.9, 8.99)
phecode_385	Abnormal results of function studies of eye	2.87 (2.66, 3)	2.83 (2.63, 2.99)
phecode_386	Visual disturbances	2.68 (2.49, 2.95)	2.59 (2.39, 2.82)
phecode_386-1	Amblyopia	2.74 (2.45, 2.99)	2.61 (2.36, 2.77)
phecode_386-2	Diplopia	2.81 (2.59, 3.34)	2.79 (2.53, 3.26)
phecode_386-3	Visual discomfort	3.44 (3.14, 3.72)	3.55 (3.24, 3.92)
phecode_386-4	Visual field defects	3 (2.69, 3.46)	2.94 (2.68, 3.45)
phecode_386-41	Scotoma	3.95 (3.26, 4.32)	4.01 (3.56, 4.35)
phecode_386-42	Hemianopia	3.11 (2.87, 3.47)	3.15 (2.93, 3.66)
phecode_386-8	Sudden or transient visual loss	3.65 (3.13, 4.36)	3.82 (3.24, 4.64)
phecode_386-9	Visual distortions and subjective visual disturbances	3.12 (2.75, 3.53)	3.17 (2.88, 3.57)
phecode_387	Disorders of refraction and accommodation	2.82 (2.59, 3.03)	2.29 (2.12, 2.45)
phecode_387-1	Hypermetropia	3.07 (2.82, 3.61)	2.57 (2.36, 2.88)
phecode_387-2	Myopia	2.77 (2.65, 2.99)	2.69 (2.58, 2.92)
phecode_387-21	Progressive high (degenerative) myopia	3.04 (2.87, 3.18)	3.06 (2.9, 3.24)
phecode_387-3	Astigmatism	2.92 (2.53, 3.32)	2.05 (1.86, 2.24)
phecode_387-4	Presbyopia	4.7 (3.97, 5.9)	6.09 (4.93, 7.11)
phecode_387-5	Anisometropia	4.1 (3.49, 5.43)	4.38 (3.61, 5.63)
phecode_388	Blindness and low vision	2.68 (2.55, 2.81)	2.56 (2.43, 2.67)
phecode_388-1	Blindness of both eyes	2.87 (2.72, 2.98)	2.87 (2.72, 2.99)
phecode_389	Other disorders of eye	2.65 (2.54, 2.81)	2.64 (2.54, 2.78)
phecode_389-1	Ocular pain	2.78 (2.64, 3.08)	2.78 (2.69, 3.08)
phecode_390	Disorders of external ear	2.45 (2.33, 2.54)	2.39 (2.29, 2.47)
phecode_390-1	Otitis externa	2.58 (2.43, 2.67)	2.6 (2.46, 2.72)
phecode_390-4	Impacted cerumen	2.48 (2.39, 2.62)	2.32 (2.24, 2.44)
phecode_390-5	Stenosis of external ear canal	5.31 (4.33, 6.33)	5.2 (4.28, 6.28)
phecode_390-6	Perichondritis and chondritis of pinna	2.83 (2.64, 3.26)	2.61 (2.45, 2.94)
phecode_391	Disorders of the middle ear	2.68 (2.58, 2.79)	2.69 (2.57, 2.82)
phecode_391-1	Otitis media	2.78 (2.63, 2.92)	2.77 (2.63, 2.9)
phecode_391-11	Acute otitis media	3.12 (2.79, 3.35)	3.05 (2.83, 3.38)
phecode_391-12	Chronic otitis media	2.96 (2.8, 3.4)	2.97 (2.83, 3.44)
phecode_391-2	Eustachian tube disorders	2.7 (2.5, 2.82)	2.72 (2.55, 2.85)
phecode_391-21	Eustachian salpingitis	2.99 (2.74, 3.58)	3.11 (2.92, 3.83)
phecode_391-22	Obstruction of Eustachian tube	3.44 (3, 3.84)	3.83 (3.42, 4.03)
phecode_391-4	Tympanosclerosis	4.09 (3.64, 4.55)	4.1 (3.66, 4.61)
phecode_391-6	Cholesteatoma of middle ear	3.21 (3.02, 3.45)	3.22 (3.01, 3.45)
phecode_391-7	Perforation of tympanic membrane	2.85 (2.66, 3.07)	2.88 (2.66, 3.07)
phecode_391-8	Otosclerosis	2.72 (2.56, 3.05)	2.73 (2.55, 3.02)
phecode_391-9	Otorrhea	3.01 (2.83, 3.13)	3.06 (2.86, 3.16)
phecode_392	Otalgia and effusion of ear	2.65 (2.53, 2.8)	2.66 (2.55, 2.81)
phecode_393	Mastoiditis and related conditions	3.77 (3.52, 4.68)	3.82 (3.57, 4.71)
phecode_394	Disorders of vestibular function	2.58 (2.41, 2.83)	2.55 (2.41, 2.75)
phecode_394-1	Meniere disease	2.5 (2.2, 2.69)	2.53 (2.28, 2.71)
phecode_394-2	Vertigo	2.61 (2.5, 2.76)	2.59 (2.47, 2.72)
phecode_394-21	Paroxysmal vertigo	2.67 (2.52, 2.86)	2.63 (2.53, 2.78)
phecode_394-22	Vestibular neuronitis	2.94 (2.57, 3.25)	2.95 (2.63, 3.25)

4 Medical history predicts future health trajectories over the human phenome

Table 18 continued from previous page

Code	Name	unadjusted HR (per 1 SD Risk State)	adj. HR for Age and Sex (per 1 SD Risk State)
phecode_394-4	Abnormal vestibular function study	3.98 (3.37, 4.86)	4.2 (3.6, 4.85)
phecode_395	Other diseases of inner ear	2.57 (2.45, 2.74)	2.57 (2.41, 2.76)
phecode_395-1	Labyrinthitis	2.59 (2.43, 2.75)	2.6 (2.38, 2.76)
phecode_395-3	Noise effects on inner ear	3.58 (3.1, 4.23)	4.81 (3.92, 6.09)
phecode_396	Hearing impairment	2.57 (2.49, 2.69)	2.25 (2.18, 2.34)
phecode_396-1	Conductive hearing loss	3.25 (3.02, 3.49)	3.28 (3.05, 3.5)
phecode_396-11	Conductive hearing loss, bilateral	3.86 (3.43, 4.35)	3.86 (3.43, 4.36)
phecode_396-2	Sensorineural hearing loss	2.76 (2.56, 2.99)	2.5 (2.35, 2.67)
phecode_396-21	Sensorineural hearing loss, bilateral	2.91 (2.7, 3.07)	2.74 (2.55, 2.87)
phecode_396-22	Presbycusis	2.84 (2.6, 3.21)	2.18 (2.04, 2.34)
phecode_396-3	Mixed conductive and sensorineural hearing loss	3.4 (3.13, 3.61)	3.42 (3.15, 3.6)
phecode_396-5	Sudden idiopathic hearing loss	4.66 (3.62, 5.76)	4.79 (3.64, 6.11)
phecode_397	Other hearing abnormality	2.6 (2.45, 2.77)	2.62 (2.48, 2.78)
phecode_397-1	Tinnitus	2.59 (2.46, 2.78)	2.63 (2.49, 2.78)
phecode_397-3	Hyperacusis	3.92 (3.32, 4.76)	4.02 (3.39, 4.85)
phecode_398	Other disorders of ear	3.05 (2.76, 3.2)	3.02 (2.71, 3.16)
phecode_398-1	Abnormal auditory function study	3.74 (3.11, 4.7)	3.84 (3.24, 4.57)
phecode_400	Rheumatic fever and chronic rheumatic heart diseases	2.8 (2.6, 2.89)	2.64 (2.46, 2.73)
phecode_400-2	Chronic rheumatic heart diseases	2.79 (2.6, 2.89)	2.64 (2.46, 2.74)
phecode_401	Hypertension	2.19 (2.13, 2.29)	1.94 (1.9, 2.02)
phecode_401-1	Essential hypertension	2.18 (2.13, 2.28)	1.94 (1.9, 2.02)
phecode_401-2	Hypertensive heart disease	2.84 (2.66, 3.04)	2.81 (2.59, 2.95)
phecode_401-3	Hypertensive chronic kidney disease	2.6 (2.51, 2.73)	2.58 (2.49, 2.72)
phecode_401-6	Secondary hypertension	2.98 (2.78, 3.22)	2.93 (2.66, 3.14)
phecode_402	Elevated blood pressure reading without diagnosis of hypertension	2.56 (2.37, 2.71)	2.53 (2.35, 2.66)
phecode_403	Angina pectoris	2.36 (2.28, 2.43)	2.22 (2.15, 2.29)
phecode_403-1	Coronary artery spasm [Coronary vasospasm]	2.99 (2.75, 3.27)	3.09 (2.84, 3.4)
phecode_404	Ischemic heart disease	2.42 (2.31, 2.51)	2.2 (2.12, 2.29)
phecode_404-1	Myocardial infarction [Heart attack]	2.33 (2.24, 2.44)	2.16 (2.1, 2.24)
phecode_404-11	Acute myocardial infarction	2.6 (2.43, 2.77)	2.33 (2.2, 2.47)
phecode_404-2	Coronary atherosclerosis [Atherosclerotic heart disease]	2.42 (2.3, 2.5)	2.16 (2.06, 2.22)
phecode_406	Chronic pulmonary heart disease	2.83 (2.69, 3.04)	2.76 (2.63, 2.94)
phecode_406-1	Pulmonary hypertension	2.82 (2.72, 3.06)	2.74 (2.65, 2.96)
phecode_406-11	Primary pulmonary hypertension	2.85 (2.63, 3.12)	2.82 (2.6, 3.08)
phecode_406-13	Secondary pulmonary arterial hypertension*	3.43 (3.17, 3.72)	3.41 (3.15, 3.67)
phecode_408	Diseases of pulmonary vessels	3.46 (3.24, 3.82)	3.51 (3.29, 3.87)
phecode_410	Inflammation of the heart [Carditis]	2.87 (2.5, 3.02)	2.81 (2.47, 2.99)
phecode_410-1	Pericarditis	3.2 (3, 3.47)	3.3 (3.11, 3.72)
phecode_410-2	Endocarditis	2.86 (2.55, 3.13)	2.74 (2.46, 3.03)
phecode_410-3	Myocarditis	3.69 (3.17, 4.87)	3.79 (3.23, 5.05)
phecode_411	Other diseases of pericardium	3.08 (2.81, 3.32)	2.98 (2.78, 3.29)
phecode_411-1	Hemopericardium NEC	4.08 (3.45, 5.22)	3.76 (3.22, 4.79)
phecode_411-2	Pericardial effusion (noninflammatory)*	3.04 (2.79, 3.32)	3.01 (2.76, 3.27)
phecode_413	Heart valve disorders	2.6 (2.48, 2.73)	2.36 (2.25, 2.44)
phecode_413-1	Mitral valve disorders	2.73 (2.61, 2.82)	2.54 (2.44, 2.61)
phecode_413-11	Mitral valve insufficiency	2.7 (2.52, 2.83)	2.53 (2.37, 2.63)
phecode_413-12	Mitral valve prolapse*	2.57 (2.44, 2.73)	2.43 (2.3, 2.55)
phecode_413-13	Mitral valve stenosis	2.7 (2.6, 2.94)	2.65 (2.54, 2.87)
phecode_413-2	Aortic valve disorders	2.59 (2.5, 2.71)	2.28 (2.19, 2.37)
phecode_413-21	Aortic stenosis	2.63 (2.47, 2.75)	2.28 (2.13, 2.38)
phecode_413-22	Aortic insufficiency	2.68 (2.51, 2.82)	2.48 (2.36, 2.61)
phecode_413-3	Tricuspid valve disorders	2.86 (2.62, 3)	2.74 (2.5, 2.86)
phecode_413-32	Tricuspid valve insufficiency*	2.94 (2.72, 3.25)	2.86 (2.66, 3.17)
phecode_413-4	Pulmonary valve disorders	3 (2.82, 3.3)	2.95 (2.79, 3.3)
phecode_413-42	Pulmonary valve insufficiency*	3 (2.8, 3.2)	2.96 (2.73, 3.2)

Supplementary Tables

Table 18 continued from previous page

Code	Name	unadjusted HR (per 1 SD Risk State)	adj. HR for Age and Sex (per 1 SD Risk State)
phecode_413-6	Heart valve replaced	2.53 (2.45, 2.7)	2.49 (2.41, 2.65)
phecode_414	Cardiomyopathy	2.55 (2.41, 2.65)	2.51 (2.38, 2.64)
phecode_414-1	Hypertrophic cardiomyopathy	2.47 (2.3, 2.63)	2.51 (2.34, 2.67)
phecode_414-11	Obstructive hypertrophic cardiomyopathy	2.65 (2.55, 2.78)	2.69 (2.61, 2.83)
phecode_414-2	Dilated cardiomyopathy*	2.52 (2.41, 2.65)	2.52 (2.38, 2.62)
phecode_414-5	Ischemic cardiomyopathy*	2.82 (2.68, 2.99)	2.77 (2.62, 2.96)
phecode_414-9	Takotsubo syndrome [Stress cardiomyopathy]	3.73 (3.35, 4.19)	3.49 (3.08, 4.03)
phecode_416	Cardiac arrhythmia and conduction disorders	2.57 (2.46, 2.69)	2.2 (2.12, 2.29)
phecode_416-1	Paroxysmal tachycardia	2.69 (2.53, 2.86)	2.64 (2.48, 2.76)
phecode_416-11	Supraventricular tachycardia	2.65 (2.58, 2.86)	2.58 (2.53, 2.78)
phecode_416-12	Ventricular tachycardia	2.67 (2.48, 2.87)	2.6 (2.43, 2.79)
phecode_416-2	Atrial fibrillation and flutter	2.48 (2.34, 2.58)	2.13 (2.04, 2.22)
phecode_416-21	Atrial fibrillation	2.52 (2.41, 2.63)	2.22 (2.15, 2.36)
phecode_416-211	Paroxysmal atrial fibrillation*	2.53 (2.4, 2.71)	2.34 (2.21, 2.52)
phecode_416-212	Persistent atrial fibrillation*	2.83 (2.72, 2.98)	2.79 (2.64, 2.93)
phecode_416-213	Chronic atrial fibrillation*	2.87 (2.63, 3.03)	2.73 (2.51, 2.89)
phecode_416-214	Permanent atrial fibrillation*	3.15 (3.03, 3.4)	3.11 (2.99, 3.34)
phecode_416-22	Atrial flutter	2.59 (2.44, 2.71)	2.39 (2.27, 2.47)
phecode_416-221	Typical atrial flutter*	3.28 (2.98, 3.52)	3.18 (2.97, 3.45)
phecode_416-222	Atypical atrial flutter*	3.36 (3.13, 3.86)	3.43 (3.07, 3.87)
phecode_416-3	Ventricular fibrillation and flutter	2.84 (2.67, 3.06)	2.81 (2.62, 3.04)
phecode_416-31	Ventricular fibrillation	3.2 (2.95, 3.47)	3.34 (3.06, 3.62)
phecode_416-4	Heart block	2.66 (2.53, 2.73)	2.3 (2.21, 2.38)
phecode_416-41	Atrioventricular block	2.69 (2.58, 2.84)	2.31 (2.22, 2.41)
phecode_416-42	Left bundle branch block	2.61 (2.47, 2.72)	2.37 (2.25, 2.45)
phecode_416-43	Right bundle branch block	2.66 (2.55, 2.87)	2.33 (2.26, 2.53)
phecode_416-5	Premature depolarization [Premature beats]	2.85 (2.51, 3.06)	2.72 (2.44, 2.88)
phecode_416-51	Atrial premature depolarization [Supraventricular premature beats]	3.1 (2.8, 3.38)	2.93 (2.66, 3.26)
phecode_416-52	Ventricular premature depolarization*	2.9 (2.47, 3.32)	2.8 (2.42, 3.19)
phecode_416-6	Long QT syndrome	4.33 (3.48, 5.2)	4.27 (3.39, 5.24)
phecode_416-7	Sinoatrial node dysfunction	2.81 (2.67, 3.07)	2.67 (2.55, 2.93)
phecode_416-71	Sick sinus syndrome*	2.82 (2.67, 3.07)	2.68 (2.55, 2.93)
phecode_416-8	Pre-excitation syndrome [Anomalous atrioventricular excitation]	2.1 (2, 2.24)	2.12 (2.02, 2.25)
phecode_417	Abnormalities of heart beat	2.66 (2.52, 2.84)	2.54 (2.42, 2.71)
phecode_417-1	Palpitations	2.48 (2.33, 2.62)	2.48 (2.35, 2.66)
phecode_417-2	Tachycardia	2.89 (2.59, 3.18)	2.87 (2.58, 3.12)
phecode_417-3	Bradycardia*	2.69 (2.53, 2.81)	2.45 (2.29, 2.56)
phecode_418	Abnormal results of cardiovascular function studies	2.67 (2.51, 2.78)	2.52 (2.37, 2.6)
phecode_418-1	Abnormal electrocardiogram [ECG] [EKG]	2.73 (2.57, 2.89)	2.57 (2.43, 2.71)
phecode_419	Presence of cardiac device	2.53 (2.42, 2.64)	2.4 (2.29, 2.5)
phecode_419-2	Presence of cardiac defibrillator	3.13 (2.98, 3.3)	3.25 (3.06, 3.41)
phecode_420	Cardiac arrest	2.79 (2.66, 2.94)	2.68 (2.55, 2.81)
phecode_423	Abnormal cardiac sounds	2.79 (2.52, 3.01)	2.65 (2.35, 2.83)
phecode_423-1	Cardiac murmurs	2.82 (2.52, 3.08)	2.65 (2.32, 2.85)
phecode_424	Heart failure	2.65 (2.54, 2.76)	2.52 (2.4, 2.59)
phecode_424-1	Left heart failure	2.64 (2.46, 2.76)	2.51 (2.35, 2.63)
phecode_424-2	Systolic heart failure	2.68 (2.5, 2.86)	2.65 (2.47, 2.8)
phecode_424-3	Diastolic heart failure	2.85 (2.68, 3.18)	2.68 (2.53, 3.01)

4 Medical history predicts future health trajectories over the human phenome

Table 18 continued from previous page

Code	Name	unadjusted HR (per 1 SD Risk State)	adj. HR for Age and Sex (per 1 SD Risk State)
phecode_424-5	Right heart failure*	3.44 (3.06, 3.64)	3.42 (3.04, 3.6)
phecode_424-6	Hypertensive heart disease with heart failure	2.96 (2.79, 3.26)	2.96 (2.78, 3.24)
phecode_425	Cardiomegaly	2.77 (2.57, 2.97)	2.67 (2.48, 2.87)
phecode_426	Other heart disorders in diseases NEC	2.71 (2.61, 2.81)	2.59 (2.5, 2.67)
phecode_430	Nontraumatic Intracranial hemorrhage	3.01 (2.68, 3.1)	2.85 (2.57, 2.95)
phecode_430-1	Nontraumatic subarachnoid hemorrhage	3.01 (2.74, 3.24)	2.95 (2.69, 3.17)
phecode_430-2	Nontraumatic intracerebral hemorrhage	3.11 (2.87, 3.35)	2.92 (2.73, 3.16)
phecode_430-3	Nontraumatic subdural hemorrhage	3.12 (2.92, 3.36)	2.88 (2.72, 3.09)
phecode_431	Stroke and transient cerebral ischemic attacks	2.57 (2.45, 2.69)	2.26 (2.15, 2.36)
phecode_431-1	Stroke	2.79 (2.62, 2.94)	2.51 (2.38, 2.66)
phecode_431-11	Cerebral infarction [Ischemic stroke]	2.76 (2.6, 2.88)	2.46 (2.31, 2.58)
phecode_431-12	Hemorrhagic stroke	2.99 (2.71, 3.14)	2.9 (2.63, 3.03)
phecode_431-14	Cerebellar stroke syndrome*	3.15 (2.74, 3.47)	3.04 (2.66, 3.22)
phecode_431-15	Lacunar syndrome*	3.33 (2.97, 3.71)	3.12 (2.78, 3.46)
phecode_431-2	Transient cerebral ischemic attacks	2.54 (2.42, 2.68)	2.27 (2.17, 2.37)
phecode_433	Other cerebrovascular disease	2.81 (2.69, 2.94)	2.53 (2.43, 2.62)

Supplementary Tables

Table 19: C-index deltas for cardiovascular endpoints. Displayed are C-index deltas between a simplified risk score (Age+Sex+MedicalHistory) and established cardiovascular predictors for investigated cardiovascular endpoints.

Endpoint	Conventional Risk Score	Difference in C-Index	Significance
Ischemic stroke	QRISK3	0.002 (-0.001, 0.005)	n.s.
	ASCVD	0.01 (0.007, 0.014)	*
	SCORE2	0.017 (0.013, 0.02)	*
	Age+Sex	0.034 (0.03, 0.037)	*
Coronary heart disease	QRISK3	0.003 (0.001, 0.005)	*
	ASCVD	0.017 (0.015, 0.018)	*
	SCORE2	0.025 (0.023, 0.027)	*
	Age+Sex	0.047 (0.045, 0.049)	*
Myocardial infarction	QRISK3	0.007 (0.004, 0.009)	*
	ASCVD	0.017 (0.015, 0.02)	*
	SCORE2	0.026 (0.023, 0.028)	*
	Age+Sex	0.056 (0.053, 0.059)	*
Heart failure	QRISK3	0.02 (0.017, 0.022)	*
	ASCVD	0.042 (0.04, 0.045)	*
	SCORE2	0.05 (0.048, 0.053)	*
	Age+Sex	0.073 (0.07, 0.076)	*
Cardiac arrest	QRISK3	0.055 (0.049, 0.062)	*
	ASCVD	0.063 (0.057, 0.07)	*
	SCORE2	0.07 (0.063, 0.077)	*
	Age+Sex	0.092 (0.085, 0.1)	*
All-Cause Death	QRISK3	0.136 (0.134, 0.138)	*
	ASCVD	0.151 (0.149, 0.153)	*
	SCORE2	0.152 (0.15, 0.154)	*
	Age+Sex	0.177 (0.175, 0.179)	*

5 Phenome-wide prediction of disease onset from retinal fundus photographs

Supplementary Tables

Table 20: Significance ratios for all PheCode categories of the retinal risk model.

Phecode category	Number of PheCodes with significant C-index delta	Number of PheCodes in category	Significance Ratio
Metab	34	43	0.791
Eye	60	93	0.645
Rx	7	11	0.636
Derm	47	76	0.618
Signs/Symptoms	29	55	0.527
Mental	19	39	0.487
ID	19	40	0.475
Resp	32	68	0.471
Neuro	40	91	0.44
Blood	16	37	0.432
Musc/Skel	49	118	0.415
Endo	10	29	0.345
GI	37	123	0.301
Hearing	6	30	0.2
Neoplasms	18	94	0.191
Cardio	20	107	0.187
Genitourinary	12	115	0.104

5 Phenome-wide prediction of disease onset from retinal fundus photographs

Table 21: Incident event counts per deciles for all endpoints of the retinal risk model.

Endpoint	PheCode string	1	2	3	4	5	6	7	8	9	10	Total
OMOP_4306655	All-Cause Death	65	78	141	202	241	325	411	493	669	865	3490
phecode_002	Staphylococcus	34	24	39	39	60	73	58	86	96	149	658
phecode_002-1	Staphylococcus aureus	25	26	26	41	41	42	47	66	65	107	486
phecode_003	Escherichia coli	43	59	55	65	91	106	100	121	136	183	959
phecode_004	Streptococcus	29	24	33	33	45	39	62	57	84	88	494
phecode_005	Mycobacteria	6	6	7	10	9	9	5	11	8	42	113
phecode_007	Hemophilus infection	7	2	9	3	13	10	12	16	16	19	107
phecode_007-1	Hemophilus influenzae	6	4	6	5	11	10	12	13	18	18	103
phecode_008	Helicobacter [H. pylori]	24	36	27	28	27	34	41	26	35	86	364
phecode_009	Pseudomonas	11	7	10	18	16	22	26	32	34	55	231
phecode_011	Klebsiella	4	6	9	10	13	8	17	28	29	33	157
phecode_015	Clostridium	12	5	13	17	14	19	30	20	31	37	198
phecode_015-2	Clostridium difficile	4	5	10	8	18	15	22	22	27	35	166
phecode_030	Campylobacter	36	36	38	46	42	39	50	37	38	45	407
phecode_052	Herpesvirus	265	325	304	389	407	424	414	468	422	454	3872
phecode_052-1	Herpes simplex	59	57	64	77	72	78	78	91	91	130	797
phecode_052-3	Varicella zoster virus	197	275	283	313	345	382	400	376	374	397	3342
phecode_052-32	Herpes zoster	182	261	271	309	364	378	367	382	394	393	3301
phecode_054	Hepatitis virus	17	11	1	10	11	9	11	12	9	20	111
phecode_056	Human papillomavirus	203	253	242	222	245	238	250	239	255	250	2397
phecode_056-1	Plantar wart	67	71	75	90	93	71	100	89	99	85	840
phecode_059	Coronavirus	81	67	100	100	89	82	127	104	135	170	1055
phecode_059-1	COVID-19*	78	69	95	85	96	74	113	120	115	172	1017
phecode_061	Influenza virus	18	33	31	26	33	33	37	41	49	61	362
phecode_070	Candidiasis	208	218	263	314	306	351	347	368	383	461	3219
phecode_084	Parasites	20	16	23	28	22	28	24	30	24	52	267
phecode_086	Pediculosis, acariasis and other infestations	27	25	18	25	18	20	25	29	30	30	247
phecode_088	Sexually transmitted disease	13	16	13	23	25	18	25	29	45	41	248
phecode_089	Infections	1030	1129	1126	1246	1188	1228	1140	1226	1247	1302	11862
phecode_089-1	Bacterial infections	263	295	297	342	362	412	407	473	517	679	4047
phecode_089-2	Viral infections	742	784	761	833	870	863	836	920	891	966	8466
phecode_089-3	Fungal infections	596	645	673	723	722	704	781	763	757	859	7223
phecode_092	Bacteremia, Sepsis, and SIRS	47	65	79	92	117	180	200	211	293	368	1652
phecode_092-2	Sepsis	48	61	78	89	121	175	197	209	292	365	1635
phecode_095	Sequela of infection	21	25	19	29	21	15	24	24	36	27	241
phecode_096	Contact or exposure to infectious agent	42	45	42	33	39	34	49	50	60	73	467
phecode_097	Drug resistant microorganisms	10	9	7	12	12	22	15	15	18	17	137
phecode_097-1	Methicillin resistant Staphylococcus aureus	8	6	6	9	22	15	18	8	13	12	117
phecode_098	Carrier or suspected carrier of infectious diseases	12	16	25	16	28	22	35	37	39	59	289
phecode_099	Lab findings related to infections	39	56	64	66	63	90	98	87	104	109	776
phecode_100	Malignant neoplasm of the head and neck	10	12	20	15	33	19	29	25	40	47	250
phecode_101	Malignant neoplasm of the digestive organs	47	56	98	139	129	170	189	210	269	345	1652
phecode_101-1	Malignant neoplasm of the esophagus	1	2	6	17	13	26	18	22	26	45	176
phecode_101-2	Malignant neoplasm of stomach	1	5	11	15	11	13	20	14	24	24	138
phecode_101-4	Malignant neoplasm of the colon and rectum	25	31	48	69	72	104	106	121	128	178	882
phecode_101-41	Malignant neoplasm of the colon	17	24	36	47	60	77	84	90	99	136	670
phecode_101-42	Malignant neoplasm of the rectum	10	10	16	26	16	28	31	43	40	57	277
phecode_101-6	Malignant neoplasm of the liver and intrahepatic bile ducts	3	1	8	6	5	7	16	18	20	34	118
phecode_101-8	Malignant neoplasm of the pancreas	3	8	10	12	23	10	23	20	33	39	181
phecode_102	Malignant neoplasm of the thoracic and respiratory organs	7	22	26	36	54	72	89	98	114	168	686
phecode_102-1	Malignant neoplasm of the of bronchus and lung	3	16	25	32	46	59	73	87	95	158	594
phecode_103	Malignant neoplasm of the skin	57	176	215	241	343	410	476	499	566	637	3620
phecode_103-1	Melanomas of skin	11	34	48	53	52	75	67	82	71	81	574
phecode_103-2	Keratinocyte carcinoma	32	107	96	144	183	243	241	293	322	336	1997
phecode_103-21	Basal cell carcinoma	27	102	91	120	159	218	214	243	284	279	1737

Supplementary Tables

Table 21 continued from previous page

Endpoint	PheCode string	1	2	3	4	5	6	7	8	9	10	Total
phecode_103-22	Squamous cell carcinoma of the skin	6	11	13	20	26	45	46	54	62	85	368
phecode_103-3	Carcinoma in situ of skin	9	16	22	44	43	77	73	106	95	112	597
phecode_104	Malignant sarcoma-related cancers	15	13	17	24	17	28	31	28	25	26	224
phecode_105	Malignant neoplasm of the breast	10	30	60	114	137	154	183	227	242	263	1420
phecode_105-1	Malignant neoplasm of the breast, female	86	107	98	124	125	109	94	118	105	114	1080
phecode_106	Gynecological malignant neoplasms	31	44	31	34	51	43	49	62	57	40	442
phecode_106-2	Malignant neoplasm of the uterus	8	14	25	19	21	26	27	33	28	25	226
phecode_106-21	Malignant neoplasm of endometrium	8	13	23	17	22	24	25	35	25	23	215
phecode_106-3	Malignant neoplasm of the ovary	12	11	8	15	19	18	15	23	25	23	169
phecode_107	Malignant neoplasm of male genitalia	28	57	122	165	192	213	229	233	263	245	1747
phecode_107-2	Malignant neoplasm of the prostate	19	52	122	161	187	218	205	233	254	248	1699
phecode_108	Malignant neoplasm of the urinary tract	12	20	34	36	49	68	71	90	113	148	641
phecode_108-4	Malignant neoplasm of the kidney	7	11	18	6	20	26	27	28	42	42	227
phecode_108-41	Malignant neoplasm of kidney, except pelvis	4	9	16	10	17	24	24	25	34	38	201
phecode_108-5	Malignant neoplasm of the bladder	6	14	12	25	19	35	36	44	62	91	344
phecode_109	Malignant neoplasm of the eye, brain and other parts of central nervous system	15	9	8	9	17	23	27	28	22	30	188
phecode_109-3	Malignant neoplasm of brain	12	5	5	10	11	16	26	14	19	33	151
phecode_110	Malignant neoplasm of the endocrine glands	7	13	9	7	10	12	12	6	8	18	102
phecode_112	Malignant neoplasm of other and ill-defined sites	149	213	246	316	404	478	513	631	751	801	4502
phecode_114	Neuroendocrine tumors	4	4	6	10	12	16	13	18	14	27	124
phecode_116	Secondary malignant neoplasm	63	80	129	146	162	232	280	280	339	409	2120
phecode_116-1	Secondary malignancy of lymph nodes	32	46	64	72	110	112	99	114	117	143	909
phecode_116-2	Secondary malignancy of respiratory organs	18	29	41	45	45	62	74	96	96	115	621
phecode_116-3	Secondary malignant neoplasm of digestive systems	13	14	32	40	46	42	52	38	61	78	416
phecode_116-4	Secondary malignant neoplasm of liver	15	25	46	56	65	79	89	91	115	154	735
phecode_116-5	Secondary malignancy of brain/spine	6	18	13	13	17	36	30	34	47	36	250
phecode_116-6	Secondary malignancy of bone	9	12	34	49	63	68	86	102	109	149	681
phecode_120	Hemo onc - by cell of origin	32	33	46	55	82	114	109	144	142	140	897
phecode_120-1	Myeloid	13	21	13	27	36	42	62	62	76	77	429
phecode_120-2	Lymphoid	15	21	18	30	40	51	57	80	68	63	443
phecode_120-21	Mature B-cell	9	11	18	30	30	50	48	67	61	57	381
phecode_121	Leukemia	7	9	22	19	18	15	33	35	51	54	263
phecode_121-2	Chronic leukemia	1	9	12	8	12	14	17	24	29	39	165
phecode_121-21	Chronic lymphoid leukemia	0	6	12	6	11	12	15	23	28	29	142
phecode_122	Lymphoma	14	11	16	16	34	49	41	44	59	45	329
phecode_122-2	Non-Hodgkin lymphoma	7	10	15	18	30	41	40	41	53	41	296
phecode_122-22	Diffuse large B-cell lymphoma*	4	2	3	9	16	22	18	14	20	13	121
phecode_123	Multiple myeloma and malignant plasma cell neoplasms	15	3	4	6	9	17	14	20	29	30	147
phecode_123-1	Multiple myeloma	15	3	2	9	11	10	14	24	21	31	140
phecode_124	Myeloproliferative disorder	8	8	8	19	23	25	37	32	48	38	246
phecode_124-5	Essential thrombocythemia	5	2	5	11	7	11	16	14	16	16	103
phecode_130	Cancer (solid tumor, excluding BCC)	285	423	487	617	714	850	934	1024	1191	1276	7801
phecode_132	Sequelae of cancer	8	15	26	66	55	71	97	78	114	137	667
phecode_135	Benign neoplasm of the head and neck	65	57	48	69	66	66	72	60	76	98	677
phecode_135-1	Benign neoplasm of the oral cavity	12	9	7	12	17	16	12	15	19	26	145
phecode_135-5	Benign neoplasm of the paranasal sinus and nasal cavity	45	49	44	36	55	43	48	45	56	75	496
phecode_136	Benign neoplasm of the digestive organs	304	445	483	491	542	643	698	719	756	852	5933
phecode_136-2	Benign neoplasm of stomach	48	73	96	117	136	148	173	156	179	173	1299
phecode_136-4	Benign neoplasm of colon, rectum, anus and anal canal	263	384	381	392	453	517	590	583	688	753	5004
phecode_136-41	Benign neoplasm of the colon	214	273	295	350	388	457	496	544	580	634	4231
phecode_136-42	Benign neoplasm of rectum and anus	113	134	166	158	178	195	237	233	249	327	1990
phecode_138	Benign neoplasm of the skin	298	382	411	395	433	464	437	476	503	529	4328
phecode_138-1	Nevus, non-neoplastic	32	46	45	47	41	39	51	41	63	59	464
phecode_138-2	Melanocytic nevi*	239	277	282	301	350	373	391	366	416	471	3466

5 Phenome-wide prediction of disease onset from retinal fundus photographs

Table 21 continued from previous page

Endpoint	PheCode string	1	2	3	4	5	6	7	8	9	10	Total
phecode_139	Benign sarcoma-related cancers	204	206	210	240	207	216	204	210	245	232	2174
phecode_139-3	Benign neoplasm of other connective and soft tissue	10	6	6	9	17	8	12	10	10	15	103
phecode_139-5	Lipoma	154	177	152	149	155	153	132	152	156	205	1585
phecode_139-52	Lipoma of intrathoracic organs	13	8	7	10	14	9	11	16	12	17	117
phecode_139-53	Lipoma of other skin subcutaneous tissue	57	66	50	48	56	44	47	43	46	63	520
phecode_139-54	Testicular lipoma	12	7	15	12	17	21	8	15	10	12	129
phecode_139-6	Hemangioma and lymphangioma	42	43	51	47	56	44	47	53	62	49	494
phecode_139-61	Hemangioma	42	43	54	45	56	42	46	42	70	48	488
phecode_140	Benign neoplasm of the breast	3	4	4	7	14	19	20	23	39	37	170
phecode_142	Lump or mass in breast or nonspecific abnormal breast exam	36	77	96	181	245	275	359	422	497	625	2813
phecode_142-1	Lump or mass in breast	34	62	74	132	141	159	224	227	272	400	1725
phecode_144	Gynecological benign neoplasms	128	115	103	142	156	161	183	241	255	385	1869
phecode_144-1	Benign neoplasms of external female genital organs and cervix	26	19	22	28	38	39	44	46	54	57	373
phecode_144-13	Benign neoplasms of the cervix	19	18	22	23	32	39	36	49	52	47	337
phecode_144-2	Benign neoplasms of the uterus	109	86	92	129	110	149	144	186	210	360	1575
phecode_144-21	Leiomyoma of uterus	73	42	61	67	88	96	105	149	176	308	1165
phecode_144-3	Benign neoplasms of the ovary	13	19	16	10	11	18	22	23	16	10	158
phecode_146	Benign neoplasm of the genitourinary system	7	17	26	36	52	68	87	143	179	227	842
phecode_146-2	Benign neoplasm of the prostate	32	41	64	63	78	88	97	100	111	123	797
phecode_148	Benign neoplasm of the eye, brain and other parts of central nervous system	21	19	28	42	42	42	51	49	44	56	394
phecode_148-1	Benign neoplasm of eye	10	11	11	18	26	25	22	25	35	33	216
phecode_148-16	Benign neoplasm of choroid	9	8	13	15	31	20	19	24	33	34	206
phecode_148-2	Benign neoplasm of meninges (Meningioma)	8	4	12	5	15	15	12	17	14	19	121
phecode_149	Benign neoplasm of the endocrine glands	11	14	16	15	24	29	32	20	29	29	219
phecode_149-3	Benign neoplasm of the parathyroid gland	4	4	8	9	8	15	16	15	18	12	109
phecode_160	Nutritional anemias	123	155	195	220	245	279	299	379	395	586	2876
phecode_160-1	Iron deficiency anemia	112	148	169	195	209	277	277	324	356	533	2600
phecode_160-2	Megaloblastic anemia	19	18	15	28	33	42	51	64	59	88	417
phecode_162	Aplastic anemia	4	6	6	9	16	21	16	15	27	31	151
phecode_164	Anemia	201	283	301	352	424	497	539	598	738	954	4887
phecode_164-1	Microcytic anemia	110	148	161	216	209	268	281	323	354	533	2603
phecode_164-2	Macrocytic anemia	21	16	18	31	33	48	56	63	62	96	444
phecode_164-6	Anemia secondary to chronic diseases and conditions	2	5	7	9	8	13	23	14	23	50	154
phecode_165	Hemoglobinopathies	1	2	3	1	2	7	2	5	10	103	136
phecode_168	Coagulation defects, purpura and other hemorrhagic conditions	47	47	55	50	73	87	102	104	110	137	812
phecode_168-1	Hypo-coagulability	20	27	24	31	43	48	46	49	69	89	446
phecode_168-19	Spontaneous ecchymoses	16	20	25	17	24	30	26	27	33	40	258
phecode_168-2	Hyper-coagulability	6	14	10	12	15	16	7	13	16	14	123
phecode_168-4	Abnormal coagulation profile	7	2	7	5	10	12	28	19	29	28	147
phecode_169	Platelet defects	23	29	34	40	48	66	100	93	121	121	675
phecode_169-1	Thrombocytopenia	22	28	36	36	51	67	93	98	117	124	672
phecode_169-11	Immune thrombocytopenic purpura [ITP]	4	9	4	7	9	7	15	18	10	19	102
phecode_170	Decreased white blood cell count	44	48	79	79	88	97	106	110	107	148	906
phecode_170-1	Neutropenia	41	52	68	78	88	93	106	109	107	140	882
phecode_170-19	Neutropenia NOS	21	21	24	17	20	24	28	33	44	50	282
phecode_171	Increased white blood cell count	14	6	14	11	16	26	20	26	25	21	179
phecode_171-1	Lymphocytosis (symptomatic)	6	8	6	5	15	10	10	16	13	13	102
phecode_172	Other disorders of white blood cells	13	17	17	23	27	26	29	35	37	43	267
phecode_174	Diseases of spleen	12	13	18	18	24	27	21	43	41	37	254
phecode_174-2	Splenomegaly	7	10	6	7	12	15	17	22	25	23	144
phecode_175	Polycythemias	1	6	12	7	13	16	16	16	24	28	139
phecode_175-2	Secondary polycythemia	1	7	6	10	9	11	21	11	24	25	125

Supplementary Tables

Table 21 continued from previous page

Endpoint	PheCode string	1	2	3	4	5	6	7	8	9	10	Total
phecode_176	Other diseases of blood and blood-forming organs	21	36	42	38	37	47	50	57	61	57	446
phecode_177	Abnormality of the lymph nodes	117	146	185	168	203	207	235	222	218	283	1984
phecode_177-2	Enlargement of lymph nodes [Lymphadenopathy]	107	107	161	141	154	138	144	157	175	198	1482
phecode_177-4	Lymphedema	12	23	24	32	33	45	57	60	69	62	417
phecode_179	Immunodeficiencies	20	17	17	15	24	18	21	19	17	19	187
phecode_179-9	Immunodeficiency NOS	16	5	8	12	10	13	11	7	18	15	115
phecode_180	Other disorders involving the immune mechanism	25	22	24	26	39	35	49	54	55	61	390
phecode_180-3	Paraproteinemias	6	5	15	13	18	29	34	31	45	46	242
phecode_180-31	Monoclonal gammopathy	6	5	10	11	16	25	33	30	41	43	220
phecode_181	Autoimmune disease	114	122	141	133	135	165	168	190	196	214	1578
phecode_200	Disorders of thyroid gland	138	183	216	313	317	312	382	414	422	425	3122
phecode_200-1	Hypothyroidism	109	140	203	221	289	255	284	324	322	391	2538
phecode_200-13	Postprocedural hypothyroidism	10	9	15	23	28	29	23	32	27	43	239
phecode_200-14	Hypothyroidism, not specified as secondary	94	122	173	196	253	221	247	294	292	341	2233
phecode_200-2	Goiter	32	50	57	67	69	67	89	68	91	114	704
phecode_200-21	Diffuse goiter	11	8	13	17	20	20	17	22	14	15	157
phecode_200-22	Uninodular goiter [single thyroid nodule]	24	19	25	27	24	34	38	37	47	40	315
phecode_200-23	Multinodular goiter	6	14	21	25	24	27	30	22	26	48	243
phecode_200-3	Thyrotoxicosis [hyperthyroidism]	33	34	43	50	67	57	52	56	50	57	499
phecode_200-31	Graves' disease [Toxic diffuse goiter]	8	9	11	11	13	15	10	17	15	8	117
phecode_200-4	Thyroiditis	9	5	6	12	8	8	19	6	16	18	107
phecode_200-7	Iodine-deficiency related thyroid disorders*	27	20	38	26	39	41	51	47	39	39	367
phecode_200-9	Abnormal thyroid function studies	21	24	26	27	20	31	25	25	30	39	268
phecode_202	Diabetes mellitus	128	190	255	332	380	458	472	487	654	903	4259
phecode_202-1	Type 1 diabetes	9	21	30	24	42	44	42	48	57	133	450
phecode_202-2	Type 2 diabetes	126	187	259	320	375	441	454	472	657	872	4163
phecode_202-4	Other specified diabetes*	107	182	279	339	396	485	548	594	788	1288	5006
phecode_204	Elevated blood glucose level	183	250	364	417	470	518	544	581	656	840	4823
phecode_204-1	Impaired fasting glucose	18	40	53	59	83	75	82	95	95	120	720
phecode_204-2	Impaired glucose tolerance (oral)	57	96	113	140	165	179	209	221	223	295	1698
phecode_204-4	Prediabetes*	6	4	7	14	10	8	13	19	17	10	108
phecode_205	Hypoglycemia	12	10	15	14	20	25	30	47	80	163	416
phecode_208	Disorders of parathyroid gland	10	21	21	29	37	40	44	61	43	68	374
phecode_208-2	Hyperparathyroidism	6	22	22	22	43	39	40	53	54	61	362
phecode_208-21	Primary hyperparathyroidism	5	11	13	20	24	20	33	34	39	33	232
phecode_209	Disorders of the pituitary gland and its hypothalamic control	15	10	10	20	17	19	18	20	37	31	197
phecode_209-1	Pituitary hyperfunction	3	6	8	8	7	11	12	11	15	21	102
phecode_211	Disorders of adrenal glands	10	6	8	6	14	23	22	20	28	38	175
phecode_215	Testicular dysfunction	9	10	16	8	6	4	12	12	11	15	103
phecode_230	Malnutrition and underweight	168	178	202	244	266	312	378	381	404	532	3065
phecode_230-2	Abnormal loss of weight and underweight	125	124	182	196	209	250	312	288	339	427	2452
phecode_230-21	Abnormal weight loss	111	119	175	171	200	246	292	296	309	422	2341
phecode_230-22	Underweight	13	6	14	21	11	16	16	28	19	10	154
phecode_230-3	Anorexia	39	55	45	54	71	84	104	91	99	134	776
phecode_232	Vitamin deficiencies	297	344	349	404	432	499	471	597	652	1072	5117
phecode_232-2	Vitamin B group deficiency	70	78	91	120	128	184	197	188	221	308	1585
phecode_232-27	Vitamin B12 deficiency	14	21	10	18	27	28	33	37	31	53	272
phecode_232-29	Folate deficiency [Vitamin B9]	6	2	2	7	7	13	23	27	26	40	153
phecode_232-4	Vitamin D deficiency	237	218	281	319	331	366	379	412	495	873	3911
phecode_234	Other nutritional deficiencies	80	97	127	99	126	139	143	114	165	190	1280
phecode_236	Overweight and obesity	364	471	530	616	608	612	684	737	797	921	6340
phecode_236-1	Obesity	358	474	522	620	590	638	687	714	797	924	6324
phecode_236-11	Morbid obesity	8	20	14	19	19	36	40	28	49	50	283
phecode_237	Abnormal weight gain	20	27	32	41	35	44	37	39	43	51	369
phecode_239	Hyperlipidemia	452	589	774	841	944	1067	1168	1249	1396	1570	10050
phecode_239-1	Hypercholesterolemia	404	526	726	759	887	1013	1063	1203	1312	1514	9407
phecode_239-11	Pure hypercholesterolemia	402	527	708	767	889	1010	1107	1193	1325	1535	9463

5 Phenome-wide prediction of disease onset from retinal fundus photographs

Table 21 continued from previous page

Endpoint	PheCode string	1	2	3	4	5	6	7	8	9	10	Total
phecode_239-12	Familial hypercholesterolemia*	12	13	9	12	15	23	18	32	26	26	186
phecode_239-2	Hyperglyceridemia	10	20	28	19	28	21	28	32	36	36	258
phecode_239-3	Mixed hyperlipidemia	5	18	25	13	18	20	24	24	25	23	195
phecode_244	Disorders of lipoprotein metabolism and other lipidemias	17	24	29	25	49	47	41	59	78	62	431
phecode_247	Disorders of mineral metabolism and mineral deficiencies	183	227	264	315	329	393	421	478	540	727	3877
phecode_247-3	Disorder of phosphorus metabolism	10	8	12	15	17	23	27	26	33	47	218
phecode_247-4	Disorders of magnesium metabolism	10	7	15	22	31	31	37	45	69	91	358
phecode_247-5	Disorders of calcium metabolism	34	38	44	58	56	74	77	116	120	120	737
phecode_247-51	Hypocalcemia	12	10	9	9	8	12	19	21	15	17	132
phecode_247-52	Hypercalcemia	9	17	12	15	26	27	19	32	33	32	222
phecode_247-7	Disorders of iron metabolism	155	172	210	238	254	308	313	359	392	584	2985
phecode_247-72	Iron deficiency	134	161	195	231	225	296	318	341	383	568	2852
phecode_248	Disorders of plasma-protein metabolism, NEC	4	8	8	7	8	20	12	10	16	36	129
phecode_251	Disorders of bilirubin excretion	9	11	12	13	14	19	34	22	23	25	182
phecode_251-1	Gilbert syndrome*	6	12	10	14	12	19	31	25	14	30	173
phecode_256	Disorders of fluid, electrolyte and acid-base balance	261	274	360	383	433	473	519	619	662	833	4817
phecode_256-1	Hyperosmolality and/or hyponatremia	21	31	28	28	28	37	31	31	31	53	319
phecode_256-2	Hyposmolality and/or hyponatremia	160	190	211	204	229	224	237	255	287	318	2315
phecode_256-3	Mixed disorder of acid-base balance	12	17	21	28	34	42	51	70	85	132	492
phecode_256-31	Acidosis	10	13	24	17	30	39	43	71	77	113	437
phecode_256-4	Hyperkalemia [Hyperpotassemia]	9	16	21	31	50	50	83	96	96	164	616
phecode_256-5	Hypokalemia [Hypopotassemia]	35	31	57	62	89	77	134	128	157	190	960
phecode_256-6	Fluid overload	3	6	16	12	26	24	47	50	64	95	343
phecode_256-7	Volume depletion	30	40	56	76	90	118	141	147	202	269	1169
phecode_257	Polydipsia	14	23	21	21	30	15	18	17	17	14	190
phecode_280	Substance related disorders	131	157	154	214	238	241	273	285	352	368	2413
phecode_280-1	Alcohol use disorders	114	163	136	200	203	230	265	275	319	361	2266
phecode_280-11	Alcohol abuse	55	68	68	88	128	121	147	149	168	220	1212
phecode_280-12	Alcohol dependence	33	48	48	55	47	70	87	69	103	124	684
phecode_280-13	Alcoholic liver disease	3	6	14	11	18	9	23	19	34	49	186
phecode_280-8	Other psychoactive substance related disorders	13	12	7	10	8	17	9	19	19	31	145
phecode_280-82	Other psychoactive substance dependence	13	7	6	8	4	9	11	13	16	19	106
phecode_281	Substance abuse, dependence, and withdrawal	223	259	304	311	320	354	396	464	516	602	3749
phecode_281-1	Substance abuse	60	68	79	88	140	127	153	153	181	225	1274
phecode_281-2	Substance dependence	185	204	249	256	277	297	352	412	444	539	3215
phecode_282-1	Current tobacco use and nicotine dependence	135	158	166	183	204	238	266	311	315	439	2415
phecode_283	Other behavioral problems	759	835	823	867	908	926	952	985	1027	1061	9143
phecode_283-4	Patient's noncompliance with medical treatment and regimen	166	170	211	212	202	231	260	262	316	384	2414
phecode_283-8	Other problems related to lifestyle	680	711	757	750	800	814	849	805	881	924	7971
phecode_284	Suicide attempt	49	44	40	55	76	63	68	75	101	93	664
phecode_284-1	Suicidal ideations	34	25	41	46	53	63	50	69	86	69	536
phecode_284-2	Suicide and self-inflicted harm	13	20	17	20	17	18	18	22	19	42	206
phecode_284-29	Intentional self-harm*	15	17	18	20	20	13	23	19	17	41	203
phecode_286	Mood [affective] disorders	357	362	426	438	406	443	475	530	594	699	4730
phecode_286-1	Bipolar disorder	8	10	4	7	14	12	8	19	15	13	110
phecode_286-2	Major depressive disorder	362	362	437	415	432	466	488	517	617	685	4781
phecode_286-21	Major depressive disorder, recurrent	24	37	15	28	21	20	22	29	32	40	268
phecode_287	Psychotic disorder	15	16	16	13	13	11	8	21	23	33	169
phecode_288	Anxiety disorders	299	385	402	425	443	499	518	583	599	692	4845
phecode_288-2	Panic disorder [episodic paroxysmal anxiety]	43	48	41	51	42	67	55	69	47	83	546
phecode_288-3	Generalized anxiety disorder	78	83	99	109	104	97	119	133	130	168	1120
phecode_288-4	Phobic disorders	44	44	67	65	51	65	64	70	82	121	673
phecode_290	Reaction to severe stress, and adjustment disorders	120	135	155	159	168	204	253	295	338	406	2233
phecode_290-1	Posttraumatic stress disorder	7	5	4	12	7	12	8	14	14	27	110

Supplementary Tables

Table 21 continued from previous page

Endpoint	PheCode string	1	2	3	4	5	6	7	8	9	10	Total
phecode_292	Somatoform disorders	13	9	16	16	10	8	7	14	12	21	126
phecode_294	Sexual dysfunction and disorders	48	75	111	164	251	334	405	549	676	792	3405
phecode_299	Mental disorder, not otherwise specified	27	24	21	34	25	29	29	35	37	48	309
phecode_308	Signs and symptoms involving emotional state	319	358	363	383	395	398	427	443	546	637	4269
phecode_308-1	Irritability	7	11	8	12	14	17	17	19	10	20	135
phecode_308-3	Emotional lability	8	8	6	10	16	19	9	17	23	47	163
phecode_308-4	Demoralization and apathy	14	4	13	13	18	12	19	21	25	37	176
phecode_308-5	Nervousness	18	14	19	11	17	18	18	19	23	25	182
phecode_308-6	Excessive crying of child, adolescent, or adult	6	6	6	7	16	12	7	9	19	23	111
phecode_308-7	Restlessness and agitation*	14	13	20	10	21	18	22	32	24	35	209
phecode_323	Systemic atrophies primarily affecting the central nervous system	5	6	9	21	19	17	16	10	12	15	130
phecode_324	Extrapyramidal and movement disorders	129	151	193	215	239	274	310	320	320	372	2523
phecode_324-1	Parkinsonism	4	11	23	33	47	58	75	71	89	84	495
phecode_324-11	Parkinson's disease	4	7	24	31	44	58	70	73	90	80	481
phecode_324-3	Dystonia	57	75	74	82	98	87	94	104	122	120	913
phecode_324-34	Torticollis	44	55	54	56	61	57	57	84	66	80	614
phecode_324-4	Tremor	39	45	61	86	103	99	128	121	144	149	975
phecode_324-41	Essential tremor*	7	16	16	19	31	32	45	46	38	55	305
phecode_324-8	Restless legs syndrome	18	19	46	30	45	46	34	50	68	57	413
phecode_325	Symptoms and signs related to movement disorders	105	98	131	158	185	242	284	345	409	535	2492
phecode_325-1	Abnormal involuntary movements	20	14	8	21	17	13	16	12	9	17	147
phecode_325-2	Abnormality of gait and mobility	80	80	110	143	170	208	263	326	382	510	2272
phecode_325-23	Unsteadiness on feet*	13	12	13	19	25	27	38	54	55	44	300
phecode_325-3	Lack of coordination	29	25	22	25	30	41	67	57	63	69	428
phecode_327	Other degenerative diseases of nervous system	3	9	13	18	18	23	50	42	47	80	303
phecode_328	Dementias and cerebral degeneration	6	14	18	39	54	102	120	156	210	249	968
phecode_328-1	Alzheimer's disease	1	3	9	29	27	56	77	86	114	133	535
phecode_328-7	Vascular dementia	3	0	2	7	11	21	24	29	44	57	198
phecode_328-8	Dementia in conditions classified elsewhere	0	5	10	30	34	64	66	83	107	121	520
phecode_328-9	Dementia NOS	5	10	9	15	26	44	65	74	117	126	491
phecode_329	Symptoms and signs involving cognitive functions and awareness	151	187	219	275	322	392	468	529	644	776	3963
phecode_329-1	Memory loss	71	83	113	138	146	215	213	268	322	347	1916
phecode_329-4	Other specified cognitive deficit	23	43	31	38	40	27	49	39	44	31	365
phecode_329-41	Attention and concentration deficit	5	12	6	11	7	14	10	15	20	23	123
phecode_329-42	Cognitive communication deficit	18	20	25	25	26	31	24	27	30	20	246
phecode_329-5	Mild cognitive impairment, so stated	2	2	15	17	24	37	49	61	67	75	349
phecode_329-6	Transient global amnesia	5	16	8	19	18	14	33	22	16	20	171
phecode_329-9	Delirium	4	4	13	18	46	48	84	87	131	174	609
phecode_330	Epilepsy, recurrent seizures, convulsions	43	35	48	45	61	68	86	78	81	103	648
phecode_330-1	Epilepsy	26	29	32	33	39	37	54	50	49	68	417
phecode_330-11	Generalized epilepsy	9	11	11	8	15	12	13	15	20	23	137
phecode_330-12	Partial epilepsy	4	5	10	8	8	22	9	11	18	17	112
phecode_330-3	Convulsions	35	30	30	32	54	46	63	59	56	70	475
phecode_331	Headache	440	416	445	511	475	596	564	626	641	815	5529
phecode_331-1	Tension headache	50	53	54	75	66	64	71	69	83	135	720
phecode_331-3	Headache syndromes, non migraine	20	14	12	9	8	8	19	20	10	21	141
phecode_331-6	Migraine	87	110	118	123	126	148	147	184	241	306	1590
phecode_331-61	Migraine with aura	22	37	26	23	32	39	38	48	48	49	362
phecode_331-8	Headache NOS	359	371	403	450	465	486	527	499	542	778	4880
phecode_333	Sleep disorders	402	407	439	464	446	472	475	523	574	620	4822
phecode_333-1	Sleep apnea	56	58	94	111	107	116	136	154	179	215	1226
phecode_333-11	Obstructive sleep apnea	32	39	50	47	51	59	83	85	86	131	663
phecode_333-2	Insomnia	165	166	177	210	219	226	229	251	274	280	2197
phecode_333-4	Circadian rhythm sleep disorder	9	12	6	13	11	15	11	6	15	19	117
phecode_334	Disorders of other cranial nerves	52	68	77	71	78	78	93	91	84	102	794
phecode_334-1	Trigeminal nerve disorders [CNS]	22	28	25	39	36	39	47	40	52	40	368

5 Phenome-wide prediction of disease onset from retinal fundus photographs

Table 21 continued from previous page

Endpoint	PheCode string	1	2	3	4	5	6	7	8	9	10	Total
phecode_334-11	Trigeminal neuralgia	14	23	29	29	28	32	35	31	35	42	298
phecode_334-2	Facial nerve disorders and weakness	26	16	27	33	31	23	39	38	43	48	324
phecode_334-21	Bell's palsy	24	14	23	28	16	22	30	38	32	40	267
phecode_335	Nerve root and plexus disorders	86	84	76	86	100	99	102	126	114	147	1020
phecode_336	Mononeuropathies	209	277	277	314	310	343	326	338	381	360	3135
phecode_336-1	Carpal tunnel syndrome	121	157	177	211	198	228	217	243	229	259	2040
phecode_336-2	Lesion of median, ulnar, radial nerve	34	28	41	34	36	45	45	41	41	71	416
phecode_336-5	Mononeuritis of lower limb	69	82	97	75	107	92	93	93	116	115	939
phecode_336-52	Meralgia paresthetica	35	29	18	29	24	18	19	21	27	30	250
phecode_336-55	Lesion of plantar nerve	22	40	40	54	56	62	53	78	84	84	573
phecode_337	Polyneuropathies	41	58	91	74	79	94	100	146	158	205	1046
phecode_337-8	Polyneuropathy in diseases classified elsewhere	2	2	4	7	7	8	11	16	29	83	169
phecode_341	Cerebral palsy and other paralytic syndromes	17	21	25	28	43	49	40	64	83	112	482
phecode_341-2	Hemiplegia and hemiparesis	14	17	19	19	33	40	32	52	74	102	402
phecode_342	Plegia and unspecified paralysis	13	20	25	26	35	35	20	31	34	38	277
phecode_342-4	Monoplegia	12	11	15	21	26	19	29	18	30	23	204
phecode_343	Disorders of autonomic nervous system	15	8	19	16	17	22	22	28	27	38	212
phecode_344	Disorders of the circulation of the cerebrospinal fluid	8	8	10	12	13	16	17	16	23	25	148
phecode_344-1	Hydrocephalus	2	2	12	5	7	15	10	11	19	27	110
phecode_346	Brain damage and brain death	8	12	13	11	26	17	14	21	17	26	165
phecode_347	Other disorders of the brain and CNS	14	15	17	19	23	32	37	40	38	41	276
phecode_348	Other diseases of spinal cord	36	33	31	50	51	66	67	66	59	81	540
phecode_348-2	Myelopathies	25	24	29	39	35	49	63	43	48	52	407
phecode_349	Disorder of nervous system	29	44	35	42	70	67	86	87	77	98	635
phecode_349-1	Abnormal findings on diagnostic test of central nervous system	18	18	17	27	33	41	34	58	46	61	353
phecode_349-13	Abnormal findings on diagnostic imaging of skull and head	10	15	14	20	27	25	33	47	37	52	280
phecode_349-2	Abnormal results of function studies of peripheral nervous system	13	17	17	10	20	29	23	29	23	24	205
phecode_350	Other symptoms involving nervous system	182	218	237	256	327	391	398	463	519	691	3682
phecode_350-5	Repeated falls*	27	50	60	102	132	203	216	270	372	494	1926
phecode_351	Disturbances of skin sensation	326	343	415	431	432	450	489	525	610	676	4697
phecode_351-1	Anesthesia of skin*	101	98	98	104	118	112	107	125	89	139	1091
phecode_351-2	Hypoesthesia of skin*	2	14	8	24	25	36	47	66	68	140	430
phecode_351-3	Paresthesia of skin*	201	204	223	214	215	235	236	222	235	275	2260
phecode_352	Disturbances of sensation of smell and taste	45	40	55	48	46	54	55	57	58	63	521
phecode_352-1	Anosmia*	15	22	21	22	26	15	37	18	13	38	227
phecode_352-3	Parageusia*	9	19	15	23	17	15	16	21	19	15	169
phecode_353	Symptoms and signs involving general sensations and perceptions	60	88	70	85	115	108	116	130	120	144	1036
phecode_353-1	Hallucinations	8	11	9	20	23	28	28	31	35	37	230
phecode_354	Dizziness and giddiness	375	492	514	589	651	636	680	744	733	916	6330
phecode_355	Coma and other alteration of consciousness	28	49	41	51	46	67	82	81	91	134	670
phecode_355-1	Coma	20	17	17	20	15	18	24	20	27	33	211
phecode_355-2	Alteration of consciousness	19	21	31	24	32	57	61	54	70	101	470
phecode_356	Speech disturbance	17	22	31	41	63	72	63	70	92	97	568
phecode_356-1	Dysarthria	2	3	3	7	10	12	11	16	17	26	107
phecode_356-2	Aphasia and dysphasia	4	5	17	9	33	22	29	40	36	40	235
phecode_360	Inflammation of eyelids	296	309	323	327	356	371	389	415	478	446	3710
phecode_360-1	Hordeolum	123	133	114	158	164	136	133	155	138	154	1408
phecode_360-11	Hordeolum externum	126	128	129	151	153	139	127	161	138	156	1408
phecode_360-12	Hordeolum internum	24	12	16	15	34	27	26	23	32	30	239
phecode_360-2	Chalazion	82	85	110	82	96	91	104	110	90	102	952
phecode_360-4	Blepharitis	122	118	159	174	213	195	212	278	250	270	1991
phecode_360-5	Noninfectious dermatoses of eyelid	8	6	14	13	9	12	16	20	22	29	149
phecode_360-51	Eczematous dermatitis of eyelid	8	8	9	11	14	8	15	21	18	27	139
phecode_361	Disorders of eyelid function	30	32	39	38	59	64	87	96	118	147	710
phecode_361-1	Entropion and trichiasis of eyelid	3	5	3	8	14	23	22	21	23	36	158

Supplementary Tables

Table 21 continued from previous page

Endpoint	PheCode string	1	2	3	4	5	6	7	8	9	10	Total
phecode_361-3	Ptosis of eyelid	26	19	17	27	28	38	34	47	61	90	387
phecode_361-4	Blepharochalasis	4	4	3	9	9	15	18	19	23	20	124
phecode_361-9	Ectropion of eyelid	7	2	6	6	7	13	13	17	24	30	125
phecode_362	Other disorders of the eyelids	58	80	74	77	104	80	89	104	97	98	861
phecode_362-5	Cysts of eyelid	14	15	15	12	23	15	17	14	33	22	180
phecode_363	Disorders of lacrimal system	126	182	189	255	329	320	410	437	467	550	3265
phecode_363-2	Dry eye syndrome [Tear film insufficiency]	94	148	156	209	271	270	349	371	371	465	2704
phecode_363-5	Epiphora	32	32	34	55	49	60	67	67	87	68	551
phecode_363-7	Stenosis and insufficiency of lacrimal passages	9	10	16	9	28	17	22	32	23	27	193
phecode_366	Noninflammatory disorders of conjunctiva	30	30	34	35	40	55	57	52	52	62	447
phecode_366-1	Pterygium of eye	8	6	5	10	16	9	17	8	9	25	113
phecode_366-4	Vascular abnormalities of conjunctiva	16	12	20	17	27	31	22	20	30	24	219
phecode_366-42	Conjunctival hyperemia	13	5	10	12	16	15	20	9	16	16	132
phecode_367	Inflammation of the eye	269	335	338	348	384	392	414	495	526	538	4039
phecode_367-1	Conjunctivitis	221	277	290	323	336	348	354	407	411	431	3398
phecode_367-12	Allergic [atopic] conjunctivitis	22	37	51	48	40	47	52	47	57	80	481
phecode_367-13	Blepharoconjunctivitis	5	15	16	16	16	24	29	31	35	33	220
phecode_367-2	Keratitis	28	29	30	26	34	27	33	32	36	34	309
phecode_367-21	Corneal ulcer	10	9	5	10	15	5	14	12	12	9	101
phecode_367-5	Uveitis	30	25	24	21	25	31	36	38	54	79	363
phecode_367-52	Iridocyclitis	30	24	26	22	23	30	35	37	60	76	363
phecode_367-6	Episcleritis	17	9	15	10	13	14	18	23	23	26	168
phecode_369	Noninflammatory disorders of the cornea	15	18	33	34	47	51	68	88	97	196	647
phecode_369-1	Corneal scars and opacities	4	4	4	5	4	7	15	12	14	34	103
phecode_369-4	Corneal degenerations	3	7	17	15	22	22	33	45	55	71	290
phecode_369-5	Hereditary corneal dystrophies	2	7	5	3	16	6	14	18	26	76	173
phecode_370	Disorders of iris and ciliary body	5	8	9	8	13	16	22	22	33	91	227
phecode_371	Cataract	37	85	147	265	422	562	753	998	1361	2168	6798
phecode_371-31	Nuclear cataract	4	15	26	53	112	121	219	302	431	740	2023
phecode_371-31	Age-related nuclear cataract	4	15	27	51	111	122	223	294	432	744	2023
phecode_373	Noninflammatory disorders of choroid	7	5	7	7	5	11	13	15	25	94	189
phecode_374	Disorders of the retina	81	172	235	303	364	427	514	581	720	1075	4472
phecode_374-1	Retinal detachments and breaks	22	25	35	51	65	61	82	72	112	147	672
phecode_374-11	Serous retinal detachment	5	17	11	23	32	24	38	35	56	83	324
phecode_374-3	Retinal vascular changes and occlusions	11	18	31	49	50	93	80	78	116	210	736
phecode_374-38	Retinal vein occlusions	3	6	16	19	22	32	42	43	55	76	314
phecode_374-39	Transient retinal arterial occlusion [Amaurosis fugax]	3	3	9	8	12	17	16	16	23	24	131
phecode_374-4	Retinal disorders in diseases classified elsewhere	28	51	99	121	141	182	177	226	307	507	1839
phecode_374-42	Diabetic retinopathy	27	50	98	118	150	170	170	224	291	462	1760
phecode_374-5	Macular degeneration	13	28	48	91	111	137	173	228	320	667	1816
phecode_374-51	Age-related macular degeneration	3	9	9	20	35	38	61	80	99	350	704
phecode_374-511	Nonexudative (dry) age-related macular degeneration	1	3	6	7	11	19	20	29	35	115	246
phecode_374-512	Exudative (wet) age-related macular degeneration	1	1	2	3	7	7	11	12	20	86	150
phecode_374-52	Macular cyst, hole, or pseudohole	3	9	12	19	26	30	33	33	40	46	251
phecode_374-55	Puckering of macula	3	11	15	32	40	31	46	62	72	106	418
phecode_374-8	Retinal edema	3	6	12	14	14	19	22	29	43	105	267
phecode_375	Abnormal intraocular pressure	78	99	167	178	220	223	290	333	358	519	2465
phecode_375-1	Glaucoma	32	63	90	97	114	183	190	238	250	407	1664
phecode_375-11	Open angle glaucoma	6	22	36	42	58	72	75	101	108	232	752
phecode_375-113	Primary open angle glaucoma	2	10	23	14	28	29	43	39	47	92	327
phecode_375-12	Angle-Closure Glaucoma	7	2	11	14	22	31	22	34	33	46	222
phecode_375-14	Low-tension glaucoma (Normal-tension glaucoma)	0	4	5	6	11	16	14	15	19	30	120
phecode_375-6	Ocular hypertension	29	35	65	59	75	85	81	96	103	120	748
phecode_376	Disorders of vitreous body	39	96	193	255	256	280	308	342	352	380	2501
phecode_376-1	Vitreous degeneration	24	82	140	189	202	208	223	267	225	295	1855
phecode_376-2	Vitreous opacities	37	99	190	252	249	301	301	323	352	384	2488

5 Phenome-wide prediction of disease onset from retinal fundus photographs

Table 21 continued from previous page

Endpoint	PheCode string	1	2	3	4	5	6	7	8	9	10	Total
phecode_376-21	Crystalline deposits in vitreous body	18	39	48	59	64	71	75	84	85	103	646
phecode_377	Hemorrhage of the eye	83	97	130	152	168	188	168	196	207	232	1621
phecode_377-2	Conjunctival hemorrhage	78	93	106	136	141	150	155	138	165	159	1321
phecode_377-4	Retinal hemorrhage	5	8	6	8	18	16	24	22	19	45	171
phecode_377-5	Vitreous hemorrhage	4	4	4	12	11	11	14	17	12	55	144
phecode_379	Infection of the eye	26	26	22	14	36	33	36	44	40	34	311
phecode_379-2	Eye infection, viral	23	29	18	20	30	32	31	31	42	28	284
phecode_379-21	Infection of the eye, herpes	9	8	9	11	14	21	21	14	25	23	155
phecode_380	Disorders of optic nerve and visual pathways	20	10	18	21	23	21	26	43	46	77	305
phecode_380-2	Disorders of optic disc	10	14	18	9	15	13	15	30	35	46	205
phecode_381	Strabismus	14	13	15	22	30	30	35	42	49	52	302
phecode_381-1	Paralytic strabismus [Neurogenic strabismus]	8	2	6	15	17	11	12	20	22	19	132
phecode_383	Irregular eye movements	10	3	8	14	7	10	9	16	13	18	108
phecode_384	Anomalies of pupillary function	6	5	5	10	8	13	9	10	14	22	102
phecode_385	Abnormal results of function studies of eye	5	10	12	10	13	20	21	21	22	52	186
phecode_386	Visual disturbances	103	142	167	194	192	192	243	263	308	397	2201
phecode_386-1	Amblyopia	6	8	18	17	21	19	32	38	59	87	305
phecode_386-2	Diplopia	24	29	26	30	37	38	37	49	51	55	376
phecode_386-4	Visual field defects	12	13	18	23	24	40	24	44	41	68	307
phecode_386-9	Visual distortions and subjective visual disturbances	6	11	8	11	18	25	21	10	21	25	156
phecode_387	Disorders of refraction and accommodation	17	25	27	79	79	109	164	183	239	441	1363
phecode_387-1	Hypermetropia	2	2	2	1	6	8	14	17	29	47	128
phecode_387-2	Myopia	6	11	10	16	26	31	32	59	83	283	557
phecode_387-3	Astigmatism	10	6	20	29	55	54	78	113	136	184	685
phecode_388	Blindness and low vision	16	25	35	43	57	81	99	120	170	238	884
phecode_389	Other disorders of eye	298	344	359	422	432	425	479	514	550	649	4472
phecode_389-1	Ocular pain	38	30	35	43	35	41	33	42	48	34	379
phecode_390	Disorders of external ear	594	690	769	855	936	988	989	960	1087	1078	8946
phecode_390-1	Otitis externa	329	368	379	399	414	460	422	436	440	510	4157
phecode_390-4	Impacted cerumen	362	483	537	589	748	800	749	841	828	881	6818
phecode_390-6	Perichondritis and chondritis of pinna	3	12	6	15	35	24	32	28	42	47	244
phecode_391	Disorders of the middle ear	332	396	422	421	415	422	414	425	448	456	4151
phecode_391-1	Otitis media	157	224	201	219	244	191	230	241	251	249	2207
phecode_391-11	Acute otitis media	12	20	19	16	20	25	19	19	22	17	189
phecode_391-12	Chronic otitis media	33	27	38	29	32	35	41	27	33	26	321
phecode_391-2	Eustachian tube disorders	151	171	166	158	180	192	198	221	221	225	1883
phecode_391-21	Eustachian salpingitis	16	19	19	19	25	23	17	22	13	11	184
phecode_391-7	Perforation of tympanic membrane	20	25	31	25	36	35	34	38	34	29	307
phecode_391-9	Otorrhea	16	26	23	36	32	31	37	29	42	36	308
phecode_392	Otalgia and effusion of ear	228	245	229	235	266	290	326	315	308	372	2814
phecode_394	Disorders of vestibular function	102	146	162	180	189	188	197	214	244	294	1916
phecode_394-1	Meniere disease	12	13	23	21	18	22	28	23	28	23	211
phecode_394-2	Vertigo	87	129	136	158	137	150	138	210	178	259	1582
phecode_394-21	Paroxysmal vertigo	68	111	117	131	116	120	131	170	157	223	1344
phecode_394-22	Vestibular neuronitis	30	18	26	24	27	27	28	33	30	31	274
phecode_394-4	Abnormal vestibular function study	4	9	12	12	8	15	20	11	13	15	119
phecode_395	Other diseases of inner ear	121	126	147	166	167	177	158	173	178	201	1614
phecode_395-1	Labyrinthitis	111	125	142	150	137	186	142	160	172	186	1511
phecode_396	Hearing impairment	248	321	378	486	597	670	728	793	892	962	6075
phecode_396-1	Conductive hearing loss	12	13	11	11	13	15	17	19	19	23	153
phecode_396-2	Sensorineural hearing loss	57	51	70	104	118	138	144	152	173	197	1204
phecode_396-21	Sensorineural hearing loss, bilateral	15	19	29	30	52	46	43	51	60	71	416
phecode_396-22	Presbycusis	6	5	14	23	28	31	44	47	53	62	313
phecode_397	Other hearing abnormality	214	221	221	250	230	255	271	239	225	250	2376
phecode_397-1	Tinnitus	209	224	212	238	228	253	251	241	229	248	2333
phecode_398	Other disorders of ear	65	79	69	90	100	92	120	110	126	126	977
phecode_398-1	Abnormal auditory function study	11	19	18	13	27	31	37	26	29	27	238
phecode_400	Rheumatic heart disease	19	26	40	62	64	106	142	146	197	232	1034
phecode_400-2	Chronic rheumatic heart diseases	20	23	39	63	64	106	137	154	191	233	1030
phecode_401	Hypertension	421	605	717	914	981	1170	1339	1501	1622	1844	11114

Supplementary Tables

Table 21 continued from previous page

Endpoint	PheCode string	1	2	3	4	5	6	7	8	9	10	Total
phecode_401-1	Essential hypertension	420	605	719	915	981	1178	1330	1501	1626	1836	11111
phecode_401-3	Hypertensive chronic kidney disease	3	4	1	8	5	5	17	15	16	41	115
phecode_401-6	Secondary hypertension	22	11	32	38	46	53	70	62	60	56	450
phecode_402	Elevated blood pressure reading without diagnosis of hypertension	184	229	296	296	334	373	344	363	331	361	3111
phecode_403	Angina pectoris	40	113	139	180	176	250	312	304	365	506	2385
phecode_404	Coronary heart disease	67	165	228	275	328	378	490	576	661	887	4055
phecode_404-1	Myocardial infarction	25	54	98	117	130	168	192	277	306	433	1800
phecode_404-11	Acute myocardial infarction	18	44	75	101	107	126	130	191	201	272	1265
phecode_404-2	Coronary atherosclerosis [Atherosclerotic heart disease]	44	111	167	210	230	296	362	438	504	682	3044
phecode_406	Chronic pulmonary heart disease	1	5	17	12	15	31	31	60	54	90	316
phecode_406-1	Pulmonary hypertension	1	5	17	8	14	29	30	53	50	85	292
phecode_410	Inflammation of the heart [Carditis]	34	21	34	53	38	45	69	67	56	81	498
phecode_410-2	Endocarditis	17	14	26	40	15	42	55	49	58	60	376
phecode_411	Other diseases of pericardium	15	13	25	24	34	31	41	48	47	80	358
phecode_411-2	Pericardial effusion (noninflammatory)*	5	12	13	19	24	19	30	33	45	57	257
phecode_413	Heart valve disorders	57	68	97	125	185	239	285	336	422	561	2375
phecode_413-1	Mitral valve disorders	41	36	63	86	115	143	182	206	272	320	1464
phecode_413-11	Mitral insufficiency	28	30	32	56	76	88	103	140	150	196	899
phecode_413-12	Mitral valve prolapse*	3	3	8	10	11	19	12	11	12	19	108
phecode_413-2	Aortic valve disorders	23	29	43	64	80	130	143	195	230	312	1249
phecode_413-21	Aortic stenosis	8	14	15	25	25	56	82	97	131	172	625
phecode_413-22	Aortic insufficiency	11	21	23	31	42	45	60	72	68	96	469
phecode_413-3	Tricuspid valve disorders	11	13	28	47	52	75	98	132	161	186	803
phecode_413-32	Tricuspid valve insufficiency*	1	3	8	11	25	18	21	32	45	43	207
phecode_413-4	Pulmonary valve disorders	9	4	1	7	6	18	18	16	22	18	119
phecode_413-42	Pulmonary valve insufficiency*	9	3	2	6	8	18	15	16	22	15	114
phecode_413-6	Heart valve replaced	7	6	14	17	29	24	37	49	56	74	313
phecode_414	Cardiomyopathy	8	19	34	21	33	44	47	47	69	92	414
phecode_414-2	Dilated cardiomyopathy*	7	1	13	9	13	13	16	16	16	28	132
phecode_416	Cardiac arrhythmia and conduction disorders	153	212	288	368	466	627	767	819	996	1242	5938
phecode_416-1	Paroxysmal tachycardia	24	34	59	66	71	89	94	98	108	126	769
phecode_416-11	Supraventricular tachycardia	24	15	51	47	53	72	69	70	90	83	574
phecode_416-12	Ventricular tachycardia	6	11	16	10	22	17	26	30	41	51	230
phecode_416-2	Atrial fibrillation and flutter	48	88	145	192	261	377	483	534	694	870	3692
phecode_416-21	Atrial fibrillation	27	65	106	141	208	277	334	415	505	629	2707
phecode_416-211	Paroxysmal atrial fibrillation*	22	33	46	79	112	140	155	165	183	231	1166
phecode_416-212	Persistent atrial fibrillation*	3	8	6	10	13	16	29	19	48	42	194
phecode_416-22	Atrial flutter	1	13	18	28	36	56	61	53	99	95	460
phecode_416-4	Heart block	43	54	84	126	148	181	283	298	347	462	2026
phecode_416-41	Atrioventricular block	13	28	23	45	72	94	101	138	145	220	879
phecode_416-42	Left bundle branch block	14	10	23	37	33	72	86	89	119	139	622
phecode_416-43	Right bundle branch block	20	27	39	58	62	72	102	77	138	139	734
phecode_416-5	Premature depolarization [Premature beats]	37	55	48	68	82	102	114	103	107	141	857
phecode_416-51	Atrial premature depolarization [Supraventricular premature beats]	7	7	7	11	19	21	34	24	29	22	181
phecode_416-52	Ventricular premature depolarization*	23	31	34	41	41	65	70	54	83	89	531
phecode_416-7	Sinoatrial node dysfunction	7	0	3	4	11	17	13	16	18	27	116
phecode_416-71	Sick sinus syndrome*	7	0	3	4	11	19	12	17	16	27	116
phecode_417	Abnormalities of heart beat	440	489	521	581	663	701	803	817	887	1010	6912
phecode_417-1	Palpitations	216	240	274	306	322	371	337	398	402	441	3307
phecode_417-2	Tachycardia	50	52	77	62	81	119	117	112	121	174	965
phecode_417-3	Bradycardia*	81	71	104	109	141	145	155	181	222	244	1453
phecode_418	Abnormal results of cardiovascular function studies	83	116	142	155	185	216	266	278	332	454	2227
phecode_418-1	Abnormal electrocardiogram [ECG] [EKG]	71	75	104	114	142	185	190	198	251	319	1649
phecode_419	Presence of cardiac device	12	14	13	37	50	69	88	118	125	162	688
phecode_420	Cardiac arrest	4	11	10	23	15	34	39	48	63	94	341
phecode_423	Abnormal cardiac sounds	26	34	39	47	61	81	82	83	122	131	706
phecode_423-1	Cardiac murmurs	26	26	33	36	56	61	78	79	96	111	602

5 Phenome-wide prediction of disease onset from retinal fundus photographs

Table 21 continued from previous page

Endpoint	PheCode string	1	2	3	4	5	6	7	8	9	10	Total
phecode_424	Heart failure	24	48	55	99	122	187	243	273	351	594	1996
phecode_424-1	Left heart failure	11	22	35	46	56	80	109	133	136	256	884
phecode_424-2	Systolic heart failure	4	12	10	17	14	32	35	46	57	93	320
phecode_424-3	Diastolic heart failure	3	2	5	8	12	21	18	23	26	36	154
phecode_425	Cardiomegaly	23	38	71	80	97	126	169	192	236	286	1318
phecode_426	Other heart disorders in diseases NEC	22	30	52	63	86	111	141	180	190	288	1163
phecode_430	Nontraumatic Intracranial hemorrhage	10	20	24	39	42	34	35	81	67	91	443
phecode_430-1	Nontraumatic subarachnoid hemorrhage	5	7	8	18	17	18	16	27	18	25	159
phecode_430-2	Nontraumatic intracerebral hemorrhage	4	7	8	16	18	17	21	39	41	65	236
phecode_430-3	Nontraumatic subdural hemorrhage	4	9	5	11	10	8	12	24	18	17	118
phecode_431	Stroke and transient cerebral ischemic attacks	48	79	101	135	192	200	246	327	384	485	2197
phecode_431-1	Stroke	29	47	49	78	116	131	145	225	246	322	1388
phecode_431-11	Ischemic stroke	26	37	34	61	93	117	129	180	216	264	1157
phecode_431-12	Hemorrhagic stroke	8	16	13	30	28	26	27	57	55	73	333
phecode_431-2	Transient cerebral ischemic attacks	19	49	69	66	93	134	140	162	181	213	1126
phecode_433	Other cerebrovascular disease	26	35	54	68	101	159	192	231	293	375	1534
phecode_433-1	Occlusion and stenosis of cerebral arteries	4	10	9	14	19	30	32	47	47	71	283
phecode_433-2	Occlusion and stenosis of precerebral arteries	4	3	9	14	14	29	41	32	50	75	271
phecode_433-21	Carotid artery stenosis	4	3	8	11	19	25	39	29	47	70	255
phecode_436	Atherosclerosis [ASCVD]	42	116	161	201	235	297	376	419	491	680	3018
phecode_437	Vascular insufficiency of intestine	5	5	5	10	12	22	25	27	33	33	177
phecode_438	Aneurysm or ectasia	18	35	52	52	87	91	87	127	137	201	887
phecode_438-1	Aortic aneurysm and ectasia	3	13	21	33	61	69	68	72	107	162	609
phecode_438-11	Abdominal aortic aneurysm	2	4	9	22	34	42	39	44	76	118	390
phecode_438-12	Thoracic aneurysm	2	9	9	17	13	32	23	22	25	27	179
phecode_439	Hemorrhoids	480	502	464	487	501	510	508	522	515	495	4984
phecode_440	Embolism and thrombosis	152	184	223	232	293	321	375	404	396	494	3074
phecode_440-1	Venous thromboembolism	137	139	157	187	184	212	259	231	241	291	2038
phecode_440-11	Deep vein thrombosis [DVT]	30	36	51	54	73	79	68	80	87	101	659
phecode_440-13	Phlebitis and thrombophlebitis	83	103	121	136	151	196	207	190	209	239	1635
phecode_440-2	Arterial embolism and thrombosis	13	10	21	20	46	33	56	63	77	100	439
phecode_440-3	Pulmonary embolism	38	56	56	79	99	111	108	128	150	170	995
phecode_443	Other specified disorders of arteries and arterioles	15	32	35	52	56	87	99	107	154	161	798
phecode_443-1	Stricture of artery [Arterial stenosis]	1	2	7	4	6	15	16	24	18	48	141
phecode_444	Venous insufficiency	145	168	170	238	244	295	299	330	324	369	2582
phecode_444-1	Varicose veins	152	154	172	228	252	262	308	294	328	340	2490
phecode_444-11	Varicose veins of lower extremities	119	125	168	199	241	235	268	273	289	321	2238
phecode_444-15	Scrotal varices [Varicocele]	3	6	8	8	12	17	30	27	40	48	199
phecode_444-5	Venous insufficiency (chronic) (peripheral)	23	33	50	74	97	92	132	146	159	182	988
phecode_446	Hypotension	81	70	101	147	225	247	317	315	390	516	2409
phecode_446-2	Orthostatic hypotension	27	26	39	59	92	112	144	165	172	234	1070
phecode_447	Nonspecific low blood-pressure reading	13	21	23	31	37	38	54	50	71	94	432
phecode_448	Peripheral vascular disease	54	60	90	92	107	127	127	163	174	241	1235
phecode_448-1	Raynaud's syndrome	40	43	50	63	75	70	60	81	77	83	642
phecode_448-9	Peripheral vascular disease NOS [includes PAD]	5	15	15	31	42	53	69	88	117	186	621
phecode_449	Other disorders of the circulatory system	6	8	10	15	9	13	12	28	32	40	173
phecode_452	Hemorrhage, NOS	13	20	23	26	21	25	32	30	40	39	269
phecode_460	Acute respiratory infection	995	1100	1153	1227	1266	1198	1321	1322	1372	1418	12372
phecode_460-1	Acute upper respiratory infection	783	840	862	913	893	957	1027	1008	988	1192	9463
phecode_460-2	Acute lower respiratory infection	580	729	782	862	928	989	1065	1156	1269	1345	9705
phecode_462	Sinusitis	354	408	394	405	419	446	432	494	500	529	4381
phecode_462-1	Acute sinusitis	167	210	190	212	198	216	228	238	255	298	2212
phecode_462-2	Chronic sinusitis	291	289	315	325	360	347	386	388	413	422	3536
phecode_463	Rhinitis and nasal congestion	395	421	444	467	476	514	499	520	497	563	4796
phecode_463-1	Chronic rhinitis	126	121	119	151	132	170	173	173	169	177	1511
phecode_463-2	Allergic rhinitis	224	219	232	252	222	258	257	245	258	381	2548
phecode_463-21	Seasonal allergic rhinitis	140	156	143	162	182	195	156	176	167	292	1769

Supplementary Tables

Table 21 continued from previous page

Endpoint	PheCode string	1	2	3	4	5	6	7	8	9	10	Total
phecode_463-4	Nasal congestion*	139	143	147	159	158	144	161	166	162	169	1548
phecode_463-5	Postnasal drip	68	78	99	94	96	114	115	127	115	123	1029
phecode_464	Nasopharyngitis	82	83	111	105	101	118	105	96	130	124	1055
phecode_464-1	Acute nasopharyngitis	80	87	106	108	108	101	108	110	114	132	1054
phecode_465	Pharyngitis	326	354	388	403	413	471	424	491	551	614	4435
phecode_465-1	Acute Pharyngitis	315	367	386	376	411	465	453	478	530	628	4409
phecode_465-2	Chronic Pharyngitis	269	302	322	327	328	375	371	414	413	523	3644
phecode_466	Tonsillitis and adenoiditis	64	88	99	92	101	99	116	161	172	255	1247
phecode_466-1	Acute tonsillitis and adenoiditis	57	68	75	76	87	82	95	130	153	236	1059
phecode_466-4	Hypertrophy of tonsils and adenoids	8	14	4	14	7	8	17	8	23	18	121
phecode_467	Laryngitis and tracheitis	58	57	98	81	83	96	101	102	111	129	916
phecode_467-1	Acute laryngitis and tracheitis	53	62	82	78	80	105	88	102	111	129	890
phecode_468	Pneumonia	100	136	157	227	222	308	346	437	526	740	3199
phecode_468-1	Viral pneumonia	22	25	29	33	31	48	44	72	70	90	464
phecode_468-2	Bacterial pneumonia	6	11	17	20	24	33	33	38	38	66	286
phecode_468-8	Bronchopneumonia	9	3	1	11	11	15	9	23	38	42	162
phecode_468-9	Lobar pneumonia*	45	75	85	99	131	168	208	242	310	445	1808
phecode_469	Bronchitis	146	152	164	177	212	209	225	211	290	271	2057
phecode_469-1	Acute bronchitis	19	18	25	26	21	34	32	26	48	33	282
phecode_469-2	Chronic bronchitis	9	7	9	24	31	26	30	40	57	70	303
phecode_471	Other disorders of nose and nasal sinuses	310	342	346	346	350	340	380	351	352	396	3513
phecode_471-2	Deviated nasal septum	38	33	34	32	35	43	48	45	64	63	435
phecode_471-3	Hypertrophy of nasal turbinates	11	7	13	10	11	7	14	14	20	17	124
phecode_471-5	Nasal polyps	42	46	38	42	60	37	45	48	52	75	485
phecode_472	Diseases of vocal cords and larynx, not elsewhere classified	10	20	19	23	27	31	33	35	31	39	268
phecode_473	Other diseases of upper respiratory tract	99	123	105	114	119	113	99	119	114	135	1140
phecode_474	Chronic obstructive pulmonary disease	41	59	94	147	186	239	271	341	422	606	2406
phecode_474-1	Emphysema	10	14	19	20	38	47	53	83	112	165	561
phecode_475	Asthma	199	205	228	250	244	243	280	278	309	301	2537
phecode_475-5	Exercise induced bronchospasm	14	14	6	6	7	11	12	11	18	11	110
phecode_476	Bronchiectasis	22	32	38	47	77	89	98	109	135	145	792
phecode_477	Inhalation lung injury	8	5	2	5	3	6	16	11	23	44	123
phecode_478	Aspiration pneumonia	3	11	8	21	23	31	38	51	54	85	325
phecode_479	Pulmonary insufficiency and acute respiratory distress syndrome	43	45	75	74	114	124	146	166	196	294	1277
phecode_479-3	Respiratory failure	11	12	9	22	28	25	34	32	58	100	331
phecode_479-6	Pulmonary collapse [Atelectasis]	30	45	58	64	84	105	109	118	162	193	968
phecode_480	Pulmonary edema	5	8	3	6	13	12	14	25	20	47	153
phecode_481	Interstitial pulmonary diseases	8	16	14	28	36	47	39	69	96	129	482
phecode_481-4	Pulmonary fibrosis	2	1	9	9	10	21	17	21	34	42	166
phecode_483	Pleural effusion	38	62	86	93	131	170	186	226	301	393	1686
phecode_484	Pneumothorax and air leak	14	14	12	16	21	22	27	34	30	53	243
phecode_486	Other respiratory disorders	401	491	523	555	600	651	675	727	776	908	6307
phecode_486-2	Other diseases of bronchus	20	12	17	12	24	21	16	32	14	29	197
phecode_486-21	Bronchospasm	17	11	14	11	13	13	9	17	19	20	144
phecode_486-5	Abnormal sputum	160	186	232	197	267	249	288	297	362	370	2608
phecode_487	Hemorrhage from respiratory passages	113	119	160	163	180	193	234	243	265	297	1967
phecode_487-1	Epistaxis	89	88	117	115	141	162	182	170	180	221	1465
phecode_487-3	Hemoptysis	33	38	34	44	49	63	53	79	77	102	572
phecode_488	Abnormalities of breathing	553	592	674	715	871	875	970	981	1178	1331	8740
phecode_488-1	Dyspnea [Shortness of breath]	385	443	509	572	706	761	805	861	1038	1237	7317
phecode_488-6	Wheezing	120	161	138	151	156	171	204	187	232	258	1778
phecode_488-8	Mouth breathing*	52	76	64	77	66	73	83	74	95	121	781
phecode_491	Pleurisy	23	30	22	37	38	38	63	58	75	98	482
phecode_494	Voice disturbance	62	88	114	129	133	131	138	147	151	149	1242
phecode_495	Abnormal findings on diagnostic imaging of lung	34	49	59	72	112	136	141	177	186	236	1202
phecode_495-1	Solitary pulmonary nodule	5	7	7	12	24	27	28	33	29	24	196
phecode_496	Abnormal results of pulmonary function studies	78	72	106	118	109	150	133	158	175	232	1331
phecode_498	Asphyxia and hypoxemia	5	5	4	8	4	14	9	19	19	33	120
phecode_500	Disorders of tooth development	14	15	11	16	15	15	18	18	22	17	161
phecode_500-4	Disturbances in tooth eruption	17	9	7	18	13	13	19	14	22	20	152

5 Phenome-wide prediction of disease onset from retinal fundus photographs

Table 21 continued from previous page

Endpoint	PheCode string	1	2	3	4	5	6	7	8	9	10	Total
phecode_500-41	Impacted teeth*	7	8	8	14	5	16	14	19	17	15	123
phecode_501	Dental caries	39	55	55	45	61	58	60	66	66	74	579
phecode_502	Other diseases of teeth and supporting structures	76	82	97	86	98	88	88	89	113	118	935
phecode_503	Diseases of pulp and periapical tissues	65	68	83	89	82	78	74	52	91	111	793
phecode_503-5	Periapical abscess	59	73	87	83	79	79	65	48	93	107	773
phecode_504	Periodontal diseases	34	40	32	35	40	31	49	35	53	64	413
phecode_504-1	Gingivitis	11	12	14	10	14	13	14	13	12	16	129
phecode_504-3	Periodontitis	8	21	24	14	7	13	19	12	12	29	159
phecode_504-32	Chronic periodontitis	14	20	15	8	13	12	14	9	14	20	139
phecode_506	Diseases of salivary glands	53	73	84	87	84	117	117	102	132	138	987
phecode_506-3	Sialoadenitis	15	16	23	20	18	27	27	24	30	30	230
phecode_506-5	Disturbances of salivary secretion	24	36	38	49	58	56	71	72	76	88	568
phecode_507	Lesions of mouth	137	139	144	161	150	192	181	213	222	235	1774
phecode_507-1	Stomatitis	72	63	71	70	95	90	96	103	101	100	861
phecode_507-11	Recurrent oral aphthae [Recurrent aphthous stomatitis]	14	14	13	20	18	19	18	21	31	30	198
phecode_508	Diseases of lips	39	45	58	59	59	63	75	60	79	94	631
phecode_509	Diseases of tongue	93	102	73	85	105	100	127	103	123	114	1025
phecode_509-1	Glossitis	24	26	32	24	33	37	33	40	38	40	327
phecode_509-11	Glossodynia	10	15	12	20	20	15	19	26	17	23	177
phecode_509-3	Hypertrophy of tongue papillae	11	8	11	10	12	13	19	13	15	15	127
phecode_510	Diseases of esophagus	228	297	300	370	416	441	446	439	496	568	4001
phecode_510-2	Esophagitis	186	196	216	286	283	346	321	346	334	418	2932
phecode_510-5	Dyskinesia of esophagus	15	13	16	18	9	22	23	20	23	37	196
phecode_510-8	Barrett's esophagus	25	33	51	69	75	87	80	90	96	96	702
phecode_511	Gastro-esophageal reflux disease	506	576	619	700	744	807	755	841	895	966	7409
phecode_512	Dysphagia	126	149	179	178	226	234	262	272	310	319	2255
phecode_513	Peptic ulcer	60	79	108	99	142	135	171	167	171	221	1353
phecode_513-1	Esophageal ulcer	22	15	29	31	34	43	48	68	59	69	418
phecode_513-2	Gastric ulcer	33	29	73	56	72	84	90	84	100	123	744
phecode_513-3	Duodenal ulcer	22	25	22	34	37	38	39	48	56	70	391
phecode_514	Gastrointestinal obstruction	58	49	66	107	117	124	171	145	172	200	1209
phecode_514-1	Esophageal obstruction (Stricture and stenosis of esophagus)	26	16	27	40	50	60	55	57	67	68	466
phecode_514-2	Intestinal obstruction	31	28	40	46	60	69	89	82	87	118	650
phecode_514-21	Impaction of intestine	11	10	5	10	12	11	19	22	9	16	125
phecode_514-3	Ileus	7	6	5	10	10	17	15	23	28	32	153
phecode_515	Heartburn and epigastric pain	162	175	211	232	228	272	239	240	269	326	2354
phecode_516	Other diseases of stomach and duodenum	44	40	37	56	54	66	84	80	90	123	674
phecode_517	Gastrointestinal angiodysplasia	3	2	10	4	19	13	22	26	33	24	156
phecode_518	Appendicitis	38	40	31	33	38	35	32	47	46	51	391
phecode_520	Hernia	346	408	472	588	656	740	824	868	944	1038	6884
phecode_520-1	Hernia of the abdominal wall	132	151	186	238	260	331	395	510	616	685	3504
phecode_520-11	Inguinal hernia	34	71	85	120	166	204	303	322	479	549	2333
phecode_520-12	Femoral hernia	7	11	9	8	9	13	15	9	16	14	111
phecode_520-13	Umbilical hernia	35	52	57	58	72	72	98	101	108	134	787
phecode_520-14	Ventral hernia	52	37	53	40	51	68	73	83	79	108	644
phecode_520-15	Incisional hernia	28	18	20	19	28	32	38	51	43	58	335
phecode_520-2	Diaphragmatic hernia [Hiatal hernia]	237	245	318	333	465	435	517	535	566	654	4305
phecode_522	Gastrointestinal inflammation	446	516	636	630	670	717	722	831	832	965	6965
phecode_522-1	Inflammatory bowel disease	39	57	44	51	45	47	60	64	58	77	542
phecode_522-11	Crohn's disease	24	15	13	18	15	21	16	16	19	21	178
phecode_522-12	Ulcerative colitis	17	16	26	19	19	28	25	18	34	27	229
phecode_522-14	Microscopic colitis*	6	8	16	11	13	10	15	22	23	26	150
phecode_522-7	Ulceration of the lower GI tract	18	20	15	10	27	22	27	22	32	32	225
phecode_522-8	Duodenitis	62	66	95	119	100	124	142	134	144	190	1176
phecode_522-9	Gastritis	255	335	398	428	450	473	479	538	590	664	4610
phecode_523	Diverticular disease [Bowel diverticulosis]	245	394	431	507	585	729	762	878	865	1047	6443
phecode_523-1	Diverticula of small intestine	5	6	5	6	13	16	20	22	17	15	125
phecode_523-2	Diverticula of colon	221	313	385	452	539	656	712	780	839	920	5817
phecode_523-4	Diverticulitis	35	66	65	79	85	106	103	125	117	138	919
phecode_524	Functional intestinal disorder	71	106	112	126	122	136	136	148	171	185	1313

Supplementary Tables

Table 21 continued from previous page

Endpoint	PheCode string	1	2	3	4	5	6	7	8	9	10	Total
phecode_524-1	Irritable bowel syndrome	62	87	91	109	119	114	133	115	147	170	1147
phecode_525	Intestinal malabsorption	41	40	40	56	45	44	36	48	46	59	455
phecode_525-1	Celiac disease	14	17	19	27	28	23	20	23	23	31	225
phecode_525-3	Disorders of intestinal carbohydrate absorption	11	8	10	9	7	5	7	16	17	18	108
phecode_526	Intestinal infection	236	255	269	308	339	347	376	433	450	552	3565
phecode_526-1	Bacterial enteritis	99	85	101	93	103	115	118	107	119	123	1063
phecode_526-11	Intestinal e.coli	16	18	9	17	20	13	25	14	28	24	184
phecode_526-12	Intestinal infection due to C. difficile	4	6	10	9	16	11	28	18	30	34	166
phecode_526-2	Viral enteritis	16	10	12	12	11	15	24	19	23	40	182
phecode_527	Abdominal pain	792	839	862	927	942	1022	1004	1033	1128	1167	9716
phecode_528	Nausea and vomiting	281	339	363	376	440	402	519	548	587	686	4541
phecode_528-1	Nausea	282	332	370	374	433	407	527	541	579	697	4542
phecode_528-2	Vomiting	279	337	357	389	429	415	517	542	586	690	4541
phecode_529	Symptoms involving digestive system	805	895	956	1113	1105	1208	1248	1325	1405	1499	11559
phecode_529-1	Diarrhea	333	328	330	403	385	386	430	464	485	529	4073
phecode_529-2	Abdominal distension and flatulence	150	162	175	246	216	243	246	253	258	317	2266
phecode_529-3	Fecal incontinence	20	43	53	56	83	89	107	100	120	180	851
phecode_529-5	Constipation	233	293	413	417	502	540	641	782	770	893	5484
phecode_529-6	Halitosis*	7	11	15	10	11	12	12	7	10	16	111
phecode_530	Disease of anus and rectum	160	154	166	172	189	181	192	200	196	213	1823
phecode_530-1	Anal fissure	57	69	75	80	59	70	58	75	81	80	704
phecode_530-2	Anorectal abscess	11	10	10	15	20	5	13	20	28	31	163
phecode_530-3	Rectal prolapse	6	14	14	10	16	21	24	20	24	46	195
phecode_532	Other disorders of the intestines	83	65	77	75	100	100	104	132	114	137	987
phecode_532-1	Intestinal fistula	14	19	10	14	17	16	25	24	15	28	182
phecode_532-4	Volvulus	8	9	8	7	10	17	21	23	18	17	138
phecode_535	Intestinal or peritoneal adhesions	45	60	68	77	84	84	93	88	96	139	834
phecode_537	Abnormality of the peritoneum	33	28	26	33	35	29	33	43	49	60	369
phecode_537-1	Peritonitis	29	27	24	31	20	24	31	32	36	47	301
phecode_540	Hepatitis	15	11	17	17	14	21	16	26	24	21	182
phecode_540-1	Chronic hepatitis	6	7	9	7	16	14	8	15	24	16	122
phecode_540-3	Viral hepatitis	17	10	2	11	10	10	12	10	10	20	112
phecode_542	Chronic liver disease and sequelae	121	158	170	183	201	222	198	258	277	336	2124
phecode_542-1	Fibrosis and cirrhosis of liver	4	9	12	15	11	20	27	26	35	71	230
phecode_542-2	Fatty liver disease (FLD)	115	126	152	172	191	192	170	221	253	268	1860
phecode_542-3	Hepatic failure	4	7	6	10	10	11	14	19	27	45	153
phecode_542-4	Portal hypertension	7	2	2	12	6	6	19	7	32	29	122
phecode_545	Nonspecific abnormal results of function study of liver	123	141	162	147	156	196	196	208	210	296	1835
phecode_546	Other disorders of liver	56	59	64	73	95	98	130	129	146	146	996
phecode_546-3	Hepatomegaly	7	9	7	6	11	11	15	11	16	23	116
phecode_550	Disorders of the gallbladder	176	175	226	246	273	277	314	311	324	374	2696
phecode_550-1	Gallstones [Cholelithiasis]	142	154	176	221	237	262	259	267	292	337	2347
phecode_550-2	Cholecystitis	47	44	48	66	54	63	55	75	68	96	616
phecode_550-4	Cholesterosis of gallbladder	16	15	16	16	19	11	24	21	27	30	195
phecode_552	Other diseases of biliary tract	20	24	36	41	35	48	59	51	74	88	476
phecode_552-1	Cholangitis	8	7	9	13	26	22	13	21	20	36	175
phecode_552-2	Obstruction of bile duct	6	9	13	15	23	14	22	32	29	42	205
phecode_554	Diseases of the pancreas	22	27	40	48	61	54	71	67	101	87	578
phecode_554-1	Pancreatitis	13	19	26	43	26	35	37	37	48	54	338
phecode_554-11	Acute pancreatitis	12	18	23	39	25	33	34	33	40	43	300
phecode_554-2	Cyst and pseudocyst of pancreas	4	5	7	13	11	17	16	22	17	30	142
phecode_555	Ascites	13	15	20	27	24	39	53	51	65	79	386
phecode_556	Other symptoms involving the digestive system and abdomen	217	250	281	290	318	331	366	418	357	456	3284
phecode_556-3	Abdominal or pelvic swelling, mass, or lump	62	55	87	77	81	91	77	88	95	93	806
phecode_556-8	Nonspecific abnormal findings in stool contents	179	205	213	218	241	280	299	311	331	361	2638
phecode_557	Gastrointestinal hemorrhage	378	378	411	452	448	467	510	515	488	610	4657
phecode_557-1	Hematemesis	17	18	19	20	33	26	30	41	49	80	333
phecode_557-2	Blood in stool	28	39	53	51	54	59	81	94	100	127	686
phecode_557-8	Hemorrhage of rectum and anus	306	299	306	295	312	330	341	313	310	349	3161

5 Phenome-wide prediction of disease onset from retinal fundus photographs

Table 21 continued from previous page

Endpoint	PheCode string	1	2	3	4	5	6	7	8	9	10	Total
phecode_558	Abnormal findings on diagnostic imaging of the digestive tract	74	93	113	129	140	161	167	195	215	231	1518
phecode_559	Other disease of digestive system, NOS	12	21	26	33	30	24	32	31	36	28	273
phecode_580	Glomerular diseases	13	13	17	27	27	27	26	41	44	85	320
phecode_580-3	Nephrotic syndrome	14	13	16	20	17	12	22	25	21	22	182
phecode_581	Renal tubulo-interstitial diseases	79	104	98	99	111	116	141	129	141	162	1180
phecode_581-1	Pyelonephritis	20	26	30	24	32	32	27	28	38	38	295
phecode_581-11	Acute pyelonephritis	15	18	15	29	22	32	25	28	27	34	245
phecode_581-3	Obstructive and reflux uropathy	43	41	40	65	86	81	86	101	96	115	754
phecode_581-31	Hydronephrosis	41	36	36	63	77	71	63	83	83	103	656
phecode_581-33	Stricture or kinking of ureter	15	11	9	14	19	11	19	21	24	27	170
phecode_582	Acute kidney failure	34	58	106	103	149	210	258	304	401	636	2259
phecode_583	Chronic kidney disease	75	104	153	167	249	332	425	436	543	722	3206
phecode_583-1	End stage renal disease [CDK, stage 5]	7	5	4	7	8	18	18	24	25	71	187
phecode_584	Renal failure	17	14	22	21	31	44	54	63	78	109	453
phecode_585	Kidney stone disease	70	68	74	88	112	122	130	119	150	151	1084
phecode_585-1	Renal colic	23	17	36	33	40	33	48	46	47	80	403
phecode_586	Other disorders of the kidney and ureters	45	58	77	99	128	128	158	164	200	260	1317
phecode_586-2	Cyst of kidney	27	34	48	62	80	79	82	108	134	145	799
phecode_588	Disorders and findings from impaired renal function	2	7	6	9	9	14	18	23	25	31	144
phecode_588-2	Abnormal results of function study of kidney	1	4	5	6	11	8	16	20	23	31	125
phecode_591	Urinary tract infection	256	372	453	521	564	586	677	736	827	899	5891
phecode_592	Cystitis and urethritis	60	86	111	131	151	184	161	215	296	278	1673
phecode_592-1	Cystitis	51	78	97	129	143	171	168	215	295	270	1617
phecode_592-11	Acute cystitis	12	10	13	13	24	37	39	57	43	55	303
phecode_592-12	Chronic cystitis	6	7	17	11	13	16	16	12	30	26	154
phecode_593	Hematuria	433	575	725	741	733	880	964	982	985	1086	8104
phecode_593-1	Gross hematuria	18	18	25	21	34	33	43	48	71	62	373
phecode_593-2	Microscopic hematuria	40	57	52	77	83	66	81	84	83	106	729
phecode_593-3	Recurrent and persistent hematuria*	7	7	14	14	13	12	16	15	18	9	125
phecode_594	Abnormality of urination	456	617	648	696	809	914	965	983	1045	1166	8299
phecode_594-1	Retention of urine	45	81	120	145	192	214	270	363	431	524	2385
phecode_594-11	Urinary hesitancy	4	4	8	11	7	15	20	19	27	31	146
phecode_594-2	Dysuria	96	161	209	213	219	265	253	305	307	336	2364
phecode_594-3	Urinary incontinence	108	189	224	277	305	306	316	393	364	463	2945
phecode_594-31	Urge incontinence	33	69	65	79	93	101	116	104	116	134	910
phecode_594-32	Stress incontinence	19	26	57	51	88	91	124	132	142	153	883
phecode_594-4	Frequency of urination and polyuria	241	266	322	317	393	469	453	487	480	490	3918
phecode_594-41	Nocturia	42	63	75	104	107	133	160	200	213	269	1366
phecode_594-6	Urinary urgency	78	106	128	154	160	157	192	173	205	162	1515
phecode_596	Other disorders of bladder	118	153	171	190	201	251	260	292	375	345	2356
phecode_596-1	Bladder neck obstruction	8	6	6	14	19	27	37	47	58	71	293
phecode_596-2	Overactive bladder	74	96	115	110	111	126	144	153	146	154	1229
phecode_596-3	Diverticulum of bladder	4	0	10	10	11	14	29	26	25	35	164
phecode_596-5	Neuromuscular dysfunction of bladder	6	9	9	10	13	6	12	21	18	23	127
phecode_597	Other disorders of urethra and urinary tract	22	33	44	54	69	66	72	68	86	80	594
phecode_597-1	Urethral stricture	18	16	19	31	35	49	49	58	61	70	406
phecode_597-5	Urethral caruncle	2	4	5	9	12	6	16	18	21	25	118
phecode_599	Other symptoms/disorders or the urinary system	431	489	520	600	615	716	717	739	802	793	6422
phecode_600	Benign prostatic hyperplasia	119	169	249	326	387	408	427	469	464	491	3509
phecode_601	Inflammatory diseases of prostate	34	45	59	58	54	51	48	35	42	59	485
phecode_601-1	Prostatitis	35	43	56	60	56	52	47	35	41	60	485
phecode_601-11	Acute prostatitis	8	14	14	15	17	8	12	21	13	16	138
phecode_601-12	Chronic prostatitis	7	10	22	21	15	22	19	13	18	23	170
phecode_602	Other disorders of prostate	55	79	125	138	139	149	175	178	183	204	1425
phecode_602-4	Elevated prostate specific antigen [PSA]	39	62	104	109	114	122	126	130	147	163	1116
phecode_603	Disorders and symptoms of testis	130	133	135	109	139	135	130	159	151	139	1360
phecode_603-1	Hydrocele	31	32	31	32	24	23	45	40	29	30	317
phecode_603-2	Spermatocele	35	37	45	42	33	45	33	44	53	53	420
phecode_603-5	Orchitis and epididymitis	64	53	57	42	56	63	53	67	66	58	579
phecode_603-6	Scrotal pain*	3	4	10	22	28	52	60	74	103	124	480

Supplementary Tables

Table 21 continued from previous page

Endpoint	PheCode string	1	2	3	4	5	6	7	8	9	10	Total
phecode_604	Disorders of penis	129	93	105	88	99	100	93	94	111	108	1020
phecode_604-1	Redundant prepuce and phimosis	27	14	15	17	16	19	20	25	13	17	183
phecode_604-3	Peyronie's disease	25	19	17	19	21	13	18	21	17	21	191
phecode_604-5	Balanoposthitis	44	43	45	41	43	41	41	51	44	46	439
phecode_605	Male sexual dysfunction	212	256	300	304	341	319	351	399	370	397	3249
phecode_605-1	Male erectile dysfunction	210	256	294	312	327	326	344	396	379	395	3239
phecode_608	Other disorders of male genital organs	210	262	266	242	244	252	264	237	248	260	2485
phecode_608-1	Abnormal findings in semen	5	27	41	61	103	116	179	232	273	361	1398
phecode_610	Benign mammary dysplasias	34	31	19	21	34	30	38	40	56	74	377
phecode_610-1	Cystic mastopathy	18	13	18	11	18	20	29	35	47	68	277
phecode_613	Other nonmalignant breast conditions	215	216	205	204	203	213	237	236	206	278	2213
phecode_613-1	Inflammatory disease of breast	32	18	23	20	31	17	24	22	20	48	255
phecode_613-5	Mastodynia	158	139	139	157	172	153	164	169	181	214	1646
phecode_613-7	Other signs and symptoms in breast	48	68	61	67	56	67	56	61	53	46	583
phecode_614	Inflammatory diseases of female pelvic organs	144	154	144	151	155	149	148	172	206	242	1665
phecode_614-1	Pelvic peritoneal adhesions, female (postoperative) (postinfection)	22	16	12	20	17	16	15	21	17	34	190
phecode_614-5	Inflammatory disease of cervix, vagina, and vulva	129	142	128	140	131	134	139	161	183	214	1501
phecode_614-52	Vaginitis and vulvovaginitis	109	121	128	109	122	99	137	131	154	177	1287
phecode_614-53	Cyst or abscess of Bartholin's gland	9	10	12	8	12	12	18	15	23	21	140
phecode_614-54	Abscess or ulceration of vulva	14	8	15	7	14	9	7	10	10	10	104
phecode_614-55	Candidiasis of vulva and vagina	6	18	35	65	89	108	111	132	160	213	937
phecode_615	Endometriosis	10	8	12	14	9	26	26	30	58	66	259
phecode_618	Genital prolapse	127	160	167	164	189	211	198	202	224	209	1851
phecode_618-1	Prolapse of vaginal walls	107	153	136	129	164	174	178	164	194	172	1571
phecode_618-11	Cystocele	73	87	114	97	122	110	129	137	144	131	1144
phecode_618-12	Rectocele	52	71	66	68	75	85	68	74	90	80	729
phecode_618-2	Uterine/Uterovaginal prolapse	41	77	70	83	97	90	97	88	99	73	815
phecode_619	Noninflammatory female genital disorders	310	328	330	316	337	334	400	427	466	529	3777
phecode_619-2	Disorders of uterus, NEC	37	36	38	27	43	44	42	45	43	56	411
phecode_619-3	Noninflammatory disorders of cervix	18	31	26	34	26	42	55	69	81	110	492
phecode_619-4	Noninflammatory disorders of vagina	217	209	263	217	252	252	319	323	372	417	2841
phecode_619-5	Noninflammatory disorders of vulva and perineum	108	87	100	90	81	96	113	97	83	100	955
phecode_620	Dysplasia of female genital organs	6	6	7	5	11	11	18	25	32	39	160
phecode_621	Endometrial hyperplasia	8	9	15	4	12	5	9	13	16	18	109
phecode_622	Polyp of female genital organs	74	71	75	77	93	90	90	100	104	116	890
phecode_622-1	Polyp of corpus uteri	56	62	64	68	64	64	61	57	65	69	630
phecode_622-2	Mucous polyp of cervix	20	15	21	24	27	39	35	47	50	47	325
phecode_623	Hypertrophy of female genital organs	98	83	94	81	80	83	104	79	89	92	883
phecode_624	Symptoms involving female genital tract	32	38	44	44	43	48	46	44	49	57	445
phecode_624-1	Dystrophy of female genital tract	16	18	35	20	19	22	30	30	25	39	254
phecode_624-2	Atrophy of female genital tract	14	20	20	14	25	26	25	30	16	17	207
phecode_625	Pain and other symptoms associated with female genital organs	110	92	113	116	133	122	156	173	199	230	1444
phecode_625-1	Dyspareunia	19	25	27	37	30	47	53	55	73	77	443
phecode_625-2	Postcoital bleeding	9	7	7	11	9	22	43	50	57	65	280
phecode_626	Disorders of menstruation and other abnormal bleeding from female genital tract	49	40	48	42	47	69	115	198	322	459	1389
phecode_626-1	Irregular menstrual cycle/bleeding	4	5	13	12	24	34	61	148	235	365	901
phecode_626-11	Absent or infrequent menstruation	3	3	7	8	5	15	28	61	75	116	321
phecode_626-13	Irregular menstrual cycle	1	2	3	1	7	16	32	67	130	203	462
phecode_626-14	Irregular menstrual bleeding	0	1	3	6	17	20	26	56	84	159	372
phecode_626-2	Dysmenorrhea	0	0	1	1	0	5	6	21	26	57	117
phecode_627	Menopausal and postmenopausal disorders	240	219	253	262	324	369	477	650	761	873	4428
phecode_627-1	Postmenopausal bleeding	157	137	136	181	161	180	186	188	228	226	1780
phecode_627-2	Symptomatic menopause	58	62	72	104	117	191	316	434	564	670	2588
phecode_627-3	Postmenopausal atrophic vaginitis	109	138	151	147	166	161	174	190	191	178	1605

5 Phenome-wide prediction of disease onset from retinal fundus photographs

Table 21 continued from previous page

Endpoint	PheCode string	1	2	3	4	5	6	7	8	9	10	Total
phecode_627-4	Menorrhagia/Excessive and frequent menstruation	2	3	10	15	20	46	68	163	278	423	1028
phecode_628	Ovarian cyst	58	69	49	63	60	74	75	60	79	122	709
phecode_628-2	Corpus luteum cyst or hematoma	51	71	53	68	60	62	75	68	65	119	692
phecode_660	Infection of the skin	927	967	1000	1051	1092	1151	1196	1189	1223	1277	11073
phecode_660-1	Fungal infection of the skin	389	484	528	504	504	613	575	599	623	683	5502
phecode_660-11	Candidiasis of skin and nails	29	34	28	70	55	49	48	65	56	65	499
phecode_660-12	Dermatophytosis	317	356	374	432	415	428	460	500	541	543	4366
phecode_660-13	Pityriasis versicolor	9	8	18	15	14	12	28	29	23	48	204
phecode_660-2	Bacterial infection of the skin	45	41	64	48	42	56	58	51	51	52	508
phecode_660-21	Impetigo	39	41	47	31	47	45	56	43	44	41	434
phecode_660-4	Carbuncle and furuncle	69	73	79	66	80	77	85	98	79	125	831
phecode_660-6	Cellulitis and abscess	468	449	495	543	561	620	663	667	715	831	6012
phecode_662	Rosacea	83	130	129	155	128	171	163	178	178	215	1530
phecode_664	Papulosquamous disorders	66	64	71	70	79	63	59	48	72	89	681
phecode_664-1	Lichen planus, nitidus, or striatus	40	45	46	51	44	53	50	42	46	85	502
phecode_664-2	Pityriasis	21	20	16	17	15	21	9	25	26	19	189
phecode_664-21	Pityriasis rosea	19	9	12	15	14	13	13	18	25	20	158
phecode_665	Psoriasis	85	98	100	104	104	115	114	123	139	128	1110
phecode_665-2	Psoriatic arthropathy	7	14	18	13	12	9	9	9	16	14	121
phecode_665-3	Other psoriasis	10	11	14	13	13	16	15	9	21	11	133
phecode_666	Urticaria	99	98	101	106	142	141	131	126	148	147	1239
phecode_666-1	Allergic urticaria	11	15	13	17	18	18	21	16	27	38	194
phecode_666-2	Idiopathic urticaria	8	9	8	7	4	13	10	16	15	17	107
phecode_667	Erythematous conditions	37	38	30	34	43	57	44	50	48	54	435
phecode_668	Dermatitis [eczema]	445	466	534	552	553	575	663	630	652	684	5754
phecode_668-1	Atopic dermatitis	112	130	128	113	107	108	130	125	128	147	1228
phecode_668-2	Seborrheic dermatitis	102	107	113	134	131	153	154	173	159	172	1398
phecode_668-3	Contact dermatitis	99	99	118	124	137	130	134	128	143	153	1265
phecode_668-4	Dermatitis due to substances taken internally	13	3	11	14	17	10	19	13	21	21	142
phecode_668-5	Lichen simplex chronicus	16	19	16	19	23	25	23	31	28	44	244
phecode_668-6	Prurigo	14	9	10	16	10	15	22	19	17	17	149
phecode_670	Seborrheic keratosis	21	41	48	48	82	79	74	86	62	65	606
phecode_672	Other acute skin changes due to ultraviolet radiation	13	11	10	14	26	19	12	26	19	19	169
phecode_673	Skin changes due to chronic exposure to nonionizing radiation	320	647	789	899	1027	1135	1167	1303	1304	1303	9894
phecode_673-1	Actinic keratosis	320	642	749	917	1002	1115	1201	1265	1267	1300	9778
phecode_674	Disorders of pigmentation	108	97	124	122	122	137	139	114	148	172	1283
phecode_674-1	Hypopigmentation	11	14	11	10	13	13	10	8	12	41	143
phecode_674-11	Vitiligo	10	17	7	10	7	9	6	5	13	32	116
phecode_674-2	Hypopigmentation	39	41	39	39	42	46	55	50	65	71	487
phecode_675	Atrophic conditions of skin	24	32	30	54	65	76	78	86	122	138	705
phecode_675-1	Circumscribed scleroderma	24	32	30	48	69	71	79	86	119	136	694
phecode_676	Hypertrophic conditions of skin	209	201	200	198	208	210	230	226	231	249	2162
phecode_676-1	Hypertrophic scar [Keloid scar]	8	8	9	12	14	12	9	11	14	39	136
phecode_676-2	Scar conditions and fibrosis of skin	27	43	29	40	39	54	41	50	42	41	406
phecode_678	Other skin and connective tissue disorders	691	916	1012	1100	1127	1175	1199	1282	1267	1363	11132
phecode_679	Skin symptoms	793	908	972	1014	1035	1082	1106	1167	1169	1250	10496
phecode_679-1	Rash and other nonspecific skin eruption	632	635	707	687	729	752	801	773	804	849	7369
phecode_679-2	Pallor and flushing	17	17	24	31	23	35	41	35	25	30	278
phecode_679-21	Pallor	11	3	10	7	15	12	9	20	11	16	114
phecode_679-22	Flushing	11	14	16	22	16	18	21	23	19	9	169
phecode_679-3	Changes in skin texture	23	41	39	50	54	69	73	60	67	101	577
phecode_679-4	Pruritus	238	254	344	333	375	389	402	441	441	533	3750
phecode_679-7	Abnormal granulation tissue, NOS	7	13	9	21	23	24	33	41	41	39	251
phecode_680	Epidermal thickening	91	108	145	177	200	205	235	230	245	306	1942
phecode_680-1	Corns and callosities	44	54	77	63	91	82	100	102	101	125	839
phecode_680-3	Xerosis cutis*	25	46	56	63	72	68	101	71	122	134	758
phecode_681	Localized swelling, mass and lump of skin and subcutaneous tissue	405	470	513	534	592	619	638	669	767	824	6031
phecode_682	Other follicular disorders	448	406	380	424	388	426	399	390	453	442	4156

Supplementary Tables

Table 21 continued from previous page

Endpoint	PhCode string	1	2	3	4	5	6	7	8	9	10	Total
phecode_682-1	Cutaneous cyst	271	303	303	292	293	277	289	298	303	296	2925
phecode_682-11	Sebaceous cyst [Epidermal cyst]	219	205	223	206	236	208	212	237	237	229	2212
phecode_682-12	Pilar and trichodermal cyst	261	284	273	238	277	281	261	261	268	267	2671
phecode_682-4	Acne	29	17	17	19	20	29	32	33	43	64	303
phecode_683	Nail disorders	193	207	220	245	262	254	253	259	270	267	2430
phecode_683-1	Ingrowing nail	65	85	86	93	85	99	84	117	97	105	916
phecode_683-2	Nail dystrophy*	11	17	15	14	9	14	22	23	33	31	189
phecode_684	Diseases of hair and hair follicles	17	42	46	75	98	111	128	123	135	208	983
phecode_684-1	Alopecia	19	42	45	78	87	105	136	96	137	202	947
phecode_684-11	Alopecia Areata	12	20	21	28	32	26	39	48	48	54	328
phecode_685	Disorders of sweat glands	100	131	148	104	144	141	185	149	165	191	1458
phecode_685-1	Dyshidrosis	35	20	22	27	25	43	26	34	32	29	293
phecode_685-4	Prickly heat and miliaria	12	12	12	16	12	20	13	17	14	10	138
phecode_685-8	Hyperhidrosis	63	82	101	99	98	115	137	111	115	138	1059
phecode_685-82	Generalized hyperhidrosis	63	90	95	89	109	118	128	115	112	141	1060
phecode_686	Chronic ulcer of skin	22	26	54	60	91	103	144	185	236	339	1260
phecode_686-1	Pressure ulcer	7	10	18	25	44	56	76	94	121	183	634
phecode_686-2	Non-pressure chronic ulcer	17	22	29	39	45	61	95	105	132	205	750
phecode_688	Granulomatous disorder of the skin	37	46	43	34	49	47	36	41	47	51	431
phecode_688-1	Sarcoidosis	12	9	9	11	5	10	9	14	12	18	109
phecode_688-3	Pyogenic granuloma of skin and subcutaneous tissue	8	13	17	14	23	8	16	16	23	20	158
phecode_700	Diffuse diseases of connective tissue	17	22	23	30	27	38	51	40	50	59	357
phecode_700-2	Sicca syndrome [Sjogren syndrome]	5	5	14	11	11	16	16	14	21	23	136
phecode_701	Osteomyelitis, periostitis, and other infections involving bone	10	11	5	6	5	13	15	15	18	41	139
phecode_701-1	Osteomyelitis	11	10	4	7	6	10	16	9	20	37	130
phecode_702	Infective and reactive arthropathies	21	29	24	17	22	19	24	38	27	34	255
phecode_702-3	Enteropathic arthropathies	11	8	11	15	12	11	12	12	14	16	122
phecode_703	Chrysal arthropathies	49	82	129	159	186	219	222	300	338	458	2142
phecode_703-1	Hyperuricemia	47	76	100	160	174	184	219	279	321	447	2007
phecode_703-11	Gout	41	71	106	146	171	181	212	284	319	429	1960
phecode_703-2	Chondrocalcinosis	4	1	8	13	16	16	19	17	30	35	159
phecode_704	Systemic vasculitis	11	14	13	20	23	39	45	44	46	53	308
phecode_704-5	Giant cell arteritis	1	7	4	9	6	17	17	18	36	31	146
phecode_705	Rheumatoid arthritis and other inflammatory polyarthropathies	75	92	121	139	130	166	195	222	241	274	1655
phecode_705-1	Rheumatoid arthritis	44	48	63	73	75	67	110	115	101	124	820
phecode_705-3	Polymyalgia rheumatica	11	22	34	50	63	81	98	117	129	148	753
phecode_705-5	Rheumatism, unspecified	15	14	17	19	12	13	17	11	16	17	151
phecode_706	Other inflammatory spondylopathies	24	35	34	43	51	65	61	69	75	86	543
phecode_706-1	Sacroiliitis NEC	8	12	15	9	13	17	15	20	19	37	165
phecode_707	Other arthropathies	123	184	230	310	372	411	460	557	655	753	4055
phecode_707-8	Polyarthritis	14	19	18	19	20	44	20	36	34	37	261
phecode_708	Osteoarthritis	514	713	875	1047	1163	1219	1352	1472	1569	1668	11592
phecode_708-1	Primary osteoarthritis	163	254	278	324	360	417	457	475	576	565	3869
phecode_708-11	Primary osteoarthritis of hip, pelvic region and thigh	20	16	43	41	49	74	56	85	80	89	553
phecode_708-12	Primary osteoarthritis of knee, lower leg	43	58	79	91	105	129	160	148	146	152	1111
phecode_708-13	Primary osteoarthritis of the hand	25	56	82	77	75	121	130	140	146	139	991
phecode_708-14	Primary osteoarthritis of the shoulder, upper arm	22	42	52	33	65	63	65	82	97	86	607
phecode_708-15	Primary osteoarthritis of the wrist, forearm	4	7	11	12	13	11	21	27	22	22	150
phecode_708-16	Primary osteoarthritis ankle and foot	44	59	74	72	61	83	93	101	86	98	771
phecode_708-7	Generalized osteoarthritis	40	86	89	125	143	169	195	254	241	267	1609
phecode_708-8	Secondary osteoarthritis	14	20	22	37	40	36	36	48	60	56	369
phecode_708-9	Heberden's or Bouchard's nodes*	20	18	24	25	38	45	44	48	57	39	358
phecode_709	Acquired deformities of fingers and toes	128	179	173	208	213	251	262	311	324	357	2406
phecode_709-1	Acquired deformities of fingers	16	20	28	20	14	20	30	26	26	19	219
phecode_709-11	Mallet finger	13	12	15	14	14	16	17	22	14	17	154
phecode_709-2	Acquired deformities of toe	101	157	177	184	196	226	250	280	320	332	2223
phecode_709-21	Hallux valgus (Bunion)	70	102	143	133	171	160	210	234	242	276	1741
phecode_709-22	Hallux rigidus	17	22	30	34	30	46	40	36	36	49	340
phecode_709-24	Hammer toe	16	16	21	39	49	42	45	52	71	80	431

5 Phenome-wide prediction of disease onset from retinal fundus photographs

Table 21 continued from previous page

Endpoint	PheCode string	1	2	3	4	5	6	7	8	9	10	Total
phecode_710	Acquired deformities of limbs	47	58	53	81	88	107	105	120	118	132	909
phecode_710-3	Acquired deformities of the knee	8	13	14	13	19	16	26	31	35	37	212
phecode_710-31	Genu valgum (acquired)	5	7	9	11	6	5	10	21	21	8	103
phecode_710-32	Genu varum (acquired)	6	3	5	5	10	12	11	17	16	17	102
phecode_710-4	Acquired deformities of the ankle and foot	34	32	35	53	65	60	68	67	85	95	594
phecode_710-41	Flat foot [pes planus]	17	26	17	20	25	20	28	26	20	38	237
phecode_711	Disorder of patella	126	139	131	145	163	146	153	167	158	186	1514
phecode_711-1	Derangement of meniscus	110	134	126	120	118	137	129	129	146	159	1308
phecode_712	Other specific joint derangements	67	79	76	79	77	54	63	87	77	87	746
phecode_712-1	Loose body in joint	10	14	14	18	17	14	17	15	13	15	147
phecode_712-5	Disorder of ligament	9	8	8	12	9	10	8	13	18	9	104
phecode_712-6	Instability of joint	23	18	18	18	19	21	27	14	12	21	191
phecode_713	Symptoms related to joints	1265	1370	1460	1479	1447	1536	1470	1545	1540	1689	14801
phecode_713-2	Effusion of joint	121	151	192	210	219	245	304	318	330	384	2474
phecode_713-3	Pain in joint	1247	1377	1447	1439	1494	1460	1514	1465	1511	1660	14614
phecode_713-4	Stiffness of joint	43	43	59	47	40	49	47	48	53	56	485
phecode_714	Deforming dorsopathies	25	20	30	37	45	67	66	72	99	122	583
phecode_714-3	Scoliosis	15	11	19	29	31	48	52	66	68	98	437
phecode_715	Non-inflammatory spondylopathy	157	232	295	333	357	407	504	533	621	678	4117
phecode_715-1	Spondylosis	129	181	245	272	301	353	380	432	463	552	3308
phecode_715-3	Spondylolisthesis	13	21	30	29	23	47	60	71	49	83	426
phecode_715-4	Spinal stenosis	38	59	85	81	104	134	145	183	199	234	1262
phecode_716	Intervertebral disc disorder	169	195	195	213	206	239	251	250	266	311	2295
phecode_716-2	Degenerative disc disease	54	61	59	65	59	78	68	76	86	92	698
phecode_716-3	Spinal disc displacement (herniation)	111	120	149	157	155	162	141	184	198	220	1597
phecode_717	Other and unspecified dorsopathies	67	74	69	65	80	76	86	103	76	95	791
phecode_717-2	Sacrococcygeal disorders	34	36	30	49	51	39	54	56	66	75	490
phecode_718	Back pain	952	1013	1016	1059	1069	1065	1116	1154	1149	1345	10938
phecode_718-1	Radiculopathy	66	75	100	106	119	148	137	141	140	164	1196
phecode_718-2	Cervicalgia	307	375	379	412	437	432	455	498	510	628	4433
phecode_718-3	Mid back pain	50	56	59	72	59	71	76	72	69	88	672
phecode_718-4	Low back pain	718	775	758	737	847	796	770	846	894	1055	8196
phecode_718-5	Sciatica	283	290	305	323	326	364	371	429	434	468	3593
phecode_719	Disorders of muscle	130	160	157	205	232	239	231	281	313	351	2299
phecode_719-1	Cramp and spasm	88	102	110	155	165	161	165	183	202	270	1601
phecode_719-11	Spasm of muscle	17	8	13	16	15	11	20	17	15	10	142
phecode_719-3	Separation of muscle (nontraumatic)	7	5	2	6	15	13	22	34	37	37	178
phecode_719-7	Muscle weakness (generalized)	21	13	18	20	22	24	25	33	32	42	250
phecode_720	Spontaneous rupture of synovium and tendon	30	32	32	45	38	43	29	55	50	74	428
phecode_721	Synoviopathy and bursopathy	624	660	738	770	774	691	844	814	821	815	7551
phecode_721-1	Synovitis and tenosynovitis	266	304	358	331	382	380	407	404	397	465	3694
phecode_721-11	Trigger finger	74	96	148	128	140	147	186	167	177	174	1437
phecode_721-12	Radial styloid tenosynovitis [de Quervain]	26	17	36	31	33	44	45	54	54	54	394
phecode_721-2	Ganglion cyst	139	152	155	148	168	176	170	178	184	181	1651
phecode_721-4	Calcium deposits in tendon and bursa	9	16	8	20	19	10	13	10	18	17	140
phecode_721-5	Bursitis	237	262	262	315	352	311	355	345	353	322	3114
phecode_721-6	Baker's cyst [popliteal cyst]	52	53	77	63	83	85	97	89	102	96	797
phecode_722	Fasciopathy	390	380	394	418	411	429	447	455	476	550	4350
phecode_722-1	Plantar fascial fibromatosis [Plantar fasciitis]	279	271	292	268	318	317	324	349	378	474	3270
phecode_722-4	Palmar fascial fibromatosis [Dupuytren]	40	64	74	96	127	145	160	164	173	183	1226
phecode_723	Enthesopathy/Enthesitis/Tendinopathy	632	696	693	666	705	688	774	821	853	892	7420
phecode_723-1	Adhesive capsulitis of shoulder	92	125	128	133	132	143	113	125	144	158	1293
phecode_723-2	Rotator cuff tear or rupture	150	175	215	216	234	271	246	245	235	291	2278
phecode_723-3	Medial epicondylitis (Golfer's elbow)	47	57	62	70	74	81	75	62	79	63	670
phecode_723-4	Lateral epicondylitis (Tennis elbow)	102	131	134	152	194	223	244	283	373	386	2222
phecode_723-5	Tendinitis	137	150	138	156	130	143	162	151	158	170	1495
phecode_723-51	Achilles tendinitis	110	123	128	128	123	128	117	137	143	150	1287
phecode_723-6	Impingement syndrome of shoulder*	155	132	194	169	192	182	212	186	224	203	1849
phecode_724	Other symptoms and disorders of the soft tissue	470	502	541	654	653	725	731	787	834	897	6794
phecode_724-1	Myalgia	203	241	270	277	302	318	338	354	371	376	3050

Supplementary Tables

Table 21 continued from previous page

Endpoint	PheCode string	1	2	3	4	5	6	7	8	9	10	Total
phecode_724-3	Nontraumatic hematoma of soft tissue	13	16	19	22	23	20	22	31	34	30	230
phecode_724-5	Exostosis	47	58	62	52	66	64	88	89	99	89	714
phecode_724-51	Calcaneal spur	16	11	6	11	18	21	20	16	15	14	148
phecode_724-52	Osteophyte*	35	36	46	51	47	57	64	64	92	73	565
phecode_726	Osteoporosis and low bone density	49	92	138	155	198	243	354	452	491	661	2833
phecode_726-1	Osteoporosis	43	84	134	146	190	242	355	427	480	658	2759
phecode_726-2	Pathologic fracture	5	6	7	19	18	24	44	43	52	75	293
phecode_727	Other disorders of bone	117	205	204	275	349	414	509	592	713	860	4238
phecode_727-1	Osteonecrosis	2	15	10	4	12	7	14	11	20	10	105
phecode_728	Chondropathies	86	76	89	78	84	86	96	107	99	173	974
phecode_728-1	Chondromalacia	14	10	8	9	10	19	12	17	23	24	146
phecode_728-3	Costochondritis (Tietze's disease)	64	54	73	62	69	55	69	79	91	124	740
phecode_730	Other disorders and symptoms of the musculoskeletal system	37	42	50	42	48	37	38	54	57	49	454
phecode_731	Symptoms involving musculoskeletal systems	133	132	124	137	144	137	138	152	133	164	1394
phecode_732	Nonspecific abnormal findings on radiological and other examination of musculoskeletal system	67	76	88	126	143	170	161	203	201	260	1495
phecode_733	Dentofacial anomalies, including malocclusion	57	52	65	78	87	110	79	103	107	119	857
phecode_733-6	Temporomandibular joint disorders	62	50	51	72	90	112	79	105	108	115	844
phecode_733-62	Arthralgia of temporomandibular joint	14	8	11	15	17	18	21	28	28	26	186
phecode_734	Diseases of the jaws	38	31	37	37	36	36	47	27	43	62	394
phecode_734-9	Jaw pain	31	19	35	34	40	36	38	30	39	57	359
phecode_800	Chest pain	591	682	686	725	717	789	775	802	815	929	7511
phecode_800-1	Chest pain on breathing	86	105	84	116	106	116	104	116	112	127	1072
phecode_800-11	Pleurodynia*	86	100	88	108	99	104	104	121	110	106	1026
phecode_800-2	Precordial pain	72	78	81	104	105	97	101	99	107	130	974
phecode_801	Cough	1034	1076	1172	1198	1260	1273	1361	1396	1385	1463	12618
phecode_802	Throat pain	109	112	146	147	169	119	133	141	146	170	1392
phecode_803	Snoring*	33	39	44	43	43	59	74	45	70	95	545
phecode_804	Other symptoms and signs involving the circulatory and respiratory system	380	410	476	441	514	537	616	578	647	739	5338
phecode_805	Fever of unknown origin	109	142	133	162	158	195	203	206	233	255	1796
phecode_807	Malaise and fatigue	487	519	572	603	638	693	707	705	791	866	6581
phecode_807-1	Chronic fatigue syndrome	182	228	240	220	259	249	319	294	351	435	2777
phecode_807-11	Postviral fatigue syndrome*	14	9	11	11	13	10	15	17	26	28	154
phecode_808	Syncope and collapse	171	215	200	270	297	288	350	392	405	430	3018
phecode_809	Pain	924	1001	1024	1076	1077	1116	1114	1127	1186	1346	10991
phecode_809-1	Acute pain	6	11	10	6	16	9	13	11	16	12	110
phecode_809-3	Pain in limb	852	937	956	971	1004	983	1015	1035	1093	1219	10065
phecode_810	Shock	2	3	4	12	5	13	20	19	22	48	148
phecode_812	Edema	126	166	175	255	289	324	378	402	462	572	3149
phecode_812-2	Angioneurotic edema	14	18	12	16	19	16	16	18	15	21	165
phecode_814	Jaundice (not of newborn)	6	9	18	13	22	23	31	28	36	32	218
phecode_815	Symptoms and signs concerning food and fluid intake	41	41	62	57	48	50	65	56	63	77	560
phecode_817	Motion sickness	3	8	9	10	12	14	13	14	13	18	114
phecode_819	General symptoms and other findings	619	719	847	839	990	1142	1171	1239	1408	1620	10594
phecode_820	Elevated erythrocyte sedimentation rate and abnormality of plasma viscosity	5	3	9	5	14	13	12	20	13	23	117
phecode_821	Abnormality of red blood cells	22	23	14	30	19	22	28	32	36	49	275
phecode_823	Abnormal serum enzyme levels	197	215	248	257	255	279	270	294	309	366	2690
phecode_823-2	Abnormal levels of other serum enzymes	195	217	243	237	249	291	262	296	298	367	2655
phecode_824	Other abnormalities of plasma proteins*	50	47	57	44	61	58	49	54	61	94	575
phecode_826	Other abnormal immunological findings in serum	26	25	26	25	21	33	20	17	20	54	267
phecode_826-3	Raised antibody titer*	17	14	21	17	20	19	17	21	18	45	209
phecode_827	Toxicology findings	20	38	36	43	48	59	64	62	74	77	521
phecode_827-1	Finding of alcohol in blood	19	33	35	40	51	49	62	64	70	74	497
phecode_829	Nonspecific findings on examination of blood	218	220	273	327	346	430	494	542	660	811	4321
phecode_829-2	Abnormal level of blood mineral*	45	41	57	56	54	59	57	70	71	87	597

5 Phenome-wide prediction of disease onset from retinal fundus photographs

Table 21 continued from previous page

Endpoint	PheCode string	1	2	3	4	5	6	7	8	9	10	Total
phecode_830	Proteinuria	320	411	481	541	547	605	660	693	790	889	5937
phecode_831	Glycosuria	33	64	72	98	99	115	106	141	179	228	1135
phecode_832	Other abnormal findings in urine	345	532	611	707	762	788	874	1042	1082	1243	7986
phecode_832-5	Acetonuria	64	68	83	75	84	92	104	107	105	107	889
phecode_832-6	Pyuria*	27	48	55	51	71	65	65	81	97	95	655
phecode_835	Cytology and pathology findings	105	114	103	135	162	189	248	262	390	665	2373
phecode_840	Allergy	381	432	474	497	457	506	521	556	548	615	4987
phecode_840-1	Food allergy	34	28	29	36	43	56	35	51	48	59	419
phecode_840-2	Allergy to insects	101	122	124	140	148	155	141	140	173	176	1420
phecode_840-8	Allergies related to other diseases/symptoms	235	255	280	279	291	297	273	301	310	443	2964
phecode_840-9	Anaphylactic reaction	13	12	10	10	11	12	6	16	13	13	116
phecode_841	Drug and medical agent allergy	331	437	495	564	577	701	759	843	888	1050	6645
phecode_841-1	Allergy to other anti-infective agents	207	311	320	346	386	395	479	489	553	619	4105
phecode_841-11	Penicillin allergy	175	251	243	310	311	355	386	371	412	472	3286
phecode_841-12	Allergy to antibiotic agent (excluding penicillin)	63	75	73	94	104	151	124	167	153	236	1240
phecode_841-13	Allergy to sulfonamides	2	8	16	6	17	15	19	28	31	35	177
phecode_841-3	Allergy to narcotic agent	19	32	36	44	48	66	67	93	103	134	642
phecode_841-4	Allergy to analgesic agent	52	73	99	86	125	139	161	181	213	238	1367
phecode_841-5	Allergy to serum and vaccine	8	3	6	11	9	7	14	16	16	15	105
phecode_848	Nonspecific abnormal findings of other body structures	127	159	168	214	275	309	293	345	400	475	2765
phecode_848-2	Nonspecific abnormal findings on radiological and other examination of other intrathoracic organs (echo)	28	37	40	72	71	101	130	140	188	231	1038
phecode_969	Adverse effects of agents primarily affecting gastrointestinal system	224	251	287	349	391	412	434	487	488	568	3891
phecode_973	Adverse effect of other drug	7	6	8	6	16	18	21	25	31	52	190
phecode_977	Long term (current) drug therapy	350	388	416	457	506	520	554	602	683	913	5389
phecode_977-4	Long term (current) use of steroids	29	36	45	49	43	54	51	59	63	81	510
phecode_977-41	Long term (current) use of inhaled steroids*	29	39	40	47	47	49	63	47	70	79	510
phecode_977-5	Long term (current) use of agents affecting hormones	2	8	13	12	22	31	27	84	135	335	669
phecode_977-51	Long term (current) use of hormonal contraceptives	1	1	2	0	3	2	13	24	51	112	209
phecode_977-52	Hormone replacement therapy (postmenopausal)	11	15	17	19	34	42	50	81	114	137	520
phecode_977-7	Long term (current) use of insulin or oral hypoglycemic drugs	55	90	122	158	157	211	236	292	334	576	2231
phecode_977-71	Long term (current) use of insulin	12	14	25	24	34	36	39	55	72	161	472
phecode_977-72	Long term (current) use of oral hypoglycemic drugs	51	86	123	149	158	199	225	281	315	561	2148
phecode_979	Transplanted organ	16	11	13	9	9	16	18	16	24	26	158

Supplementary Tables

Table 22: Event frequencies per deciles and rate ratios for all endpoints of the retinal risk model.

Endpoint	PheCode string	1	2	3	4	5	6	7	8	9	10	Rate Ratio (RR)
OMOP_4306655	All-Cause Death	0.011	0.013	0.023	0.033	0.039	0.053	0.067	0.08	0.109	0.141	13.31
phecode_002	Staphylococcus	0.006	0.004	0.006	0.006	0.01	0.012	0.01	0.014	0.016	0.024	4.383
phecode_002-1	Staphylococcus aureus	0.004	0.004	0.004	0.007	0.007	0.007	0.008	0.011	0.011	0.018	4.28
phecode_003	Escherichia coli	0.007	0.01	0.009	0.011	0.015	0.017	0.016	0.02	0.022	0.03	4.257
phecode_004	Streptococcus	0.005	0.004	0.005	0.005	0.007	0.006	0.01	0.009	0.014	0.015	3.035
phecode_005	Mycobacteria	0.001	0.001	0.001	0.002	0.001	0.001	0.001	0.002	0.001	0.007	7.001
phecode_007	Hemophilus infection	0.001	0	0.001	0	0.002	0.002	0.002	0.003	0.003	0.003	2.714
phecode_007-1	Hemophilus influenzae	0.001	0.001	0.001	0.001	0.002	0.002	0.002	0.002	0.003	0.003	3
phecode_008	Helicobacter [H. pylori]	0.004	0.006	0.004	0.005	0.004	0.006	0.007	0.004	0.006	0.014	3.584
phecode_009	Pseudomonas	0.002	0.001	0.002	0.003	0.003	0.004	0.004	0.005	0.006	0.009	5
phecode_011	Klebsiella	0.001	0.001	0.001	0.002	0.002	0.001	0.003	0.005	0.005	0.005	8.251
phecode_015	Clostridium	0.002	0.001	0.002	0.003	0.002	0.003	0.005	0.005	0.005	0.006	3.084
phecode_015-2	Clostridium difficile	0.001	0.001	0.002	0.001	0.003	0.002	0.004	0.004	0.004	0.006	8.751
phecode_030	Campylobacter	0.006	0.006	0.006	0.008	0.007	0.006	0.008	0.006	0.006	0.007	1.25
phecode_052	Herpesvirus	0.047	0.058	0.054	0.069	0.073	0.076	0.074	0.083	0.075	0.081	1.714
phecode_052-1	Herpes simplex	0.01	0.01	0.011	0.013	0.012	0.013	0.013	0.015	0.015	0.022	2.204
phecode_052-3	Varicella zoster virus	0.034	0.047	0.049	0.054	0.06	0.066	0.069	0.065	0.065	0.068	2.015
phecode_052-32	Herpes zoster	0.031	0.044	0.046	0.052	0.062	0.064	0.062	0.065	0.067	0.067	2.16
phecode_054	Hepatovirus	0.003	0.002	0	0.002	0.002	0.001	0.002	0.002	0.001	0.003	1.177
phecode_056	Human papillomavirus	0.036	0.044	0.042	0.039	0.043	0.042	0.044	0.042	0.045	0.044	1.232
phecode_056-1	Plantar wart	0.011	0.012	0.013	0.015	0.016	0.012	0.017	0.015	0.017	0.014	1.269
phecode_059	Coronavirus	0.013	0.011	0.016	0.016	0.015	0.013	0.021	0.017	0.022	0.028	2.099
phecode_059-1	COVID-19*	0.013	0.011	0.016	0.014	0.016	0.012	0.018	0.02	0.019	0.028	2.205
phecode_061	Influenza virus	0.003	0.006	0.005	0.004	0.006	0.006	0.006	0.007	0.008	0.01	3.389
phecode_070	Candidiasis	0.037	0.039	0.047	0.056	0.054	0.063	0.062	0.066	0.068	0.082	2.217
phecode_084	Parasites	0.003	0.003	0.004	0.005	0.004	0.005	0.004	0.005	0.004	0.009	2.6
phecode_086	Pediculosis, acariasis and other infestations	0.004	0.004	0.003	0.004	0.003	0.003	0.004	0.005	0.005	0.005	1.111
phecode_088	Sexually transmitted disease	0.002	0.003	0.002	0.004	0.004	0.003	0.004	0.005	0.007	0.007	3.154
phecode_089	Infections	0.277	0.304	0.303	0.336	0.32	0.331	0.307	0.33	0.336	0.351	1.264
phecode_089-1	Bacterial infections	0.047	0.052	0.053	0.061	0.064	0.073	0.072	0.084	0.092	0.12	2.582
phecode_089-2	Viral infections	0.158	0.167	0.162	0.177	0.185	0.184	0.178	0.196	0.19	0.206	1.302
phecode_089-3	Fungal infections	0.119	0.129	0.134	0.144	0.144	0.141	0.156	0.152	0.151	0.172	1.442
phecode_092	Bacteremia, Sepsis, and SIRS	0.008	0.011	0.013	0.015	0.019	0.03	0.033	0.035	0.048	0.06	7.831
phecode_092-2	Sepsis	0.008	0.01	0.013	0.015	0.02	0.029	0.032	0.034	0.048	0.06	7.605
phecode_095	Sequela of infection	0.003	0.004	0.003	0.005	0.003	0.002	0.004	0.004	0.006	0.004	1.286
phecode_096	Contact or exposure to infectious agent	0.007	0.007	0.007	0.005	0.006	0.006	0.008	0.008	0.01	0.012	1.738
phecode_097	Drug resistant microorganisms	0.002	0.001	0.001	0.002	0.002	0.004	0.002	0.002	0.003	0.003	1.7
phecode_097-1	Methicillin resistant Staphylococcus aureus	0.001	0.001	0.001	0.001	0.004	0.002	0.003	0.001	0.002	0.002	1.5
phecode_098	Carrier or suspected carrier of infectious diseases	0.002	0.003	0.004	0.003	0.005	0.004	0.006	0.006	0.006	0.01	4.917
phecode_099	Lab findings related to infections	0.006	0.009	0.011	0.011	0.01	0.015	0.016	0.014	0.017	0.018	2.795
phecode_100	Malignant neoplasm of the head and neck	0.002	0.002	0.003	0.002	0.005	0.003	0.005	0.004	0.007	0.008	4.701
phecode_101	Malignant neoplasm of the digestive organs	0.008	0.009	0.016	0.023	0.021	0.028	0.031	0.035	0.044	0.057	7.342
phecode_101-1	Malignant neoplasm of the esophagus	0	0	0.001	0.003	0.002	0.004	0.003	0.004	0.004	0.007	45.007
phecode_101-2	Malignant neoplasm of stomach	0	0.001	0.002	0.002	0.002	0.002	0.003	0.002	0.004	0.004	24.004
phecode_101-4	Malignant neoplasm of the colon and rectum	0.004	0.005	0.008	0.011	0.012	0.017	0.017	0.02	0.021	0.029	7.121
phecode_101-41	Malignant neoplasm of the colon	0.003	0.004	0.006	0.008	0.01	0.013	0.014	0.015	0.016	0.022	8.001
phecode_101-42	Malignant neoplasm of the rectum	0.002	0.002	0.003	0.004	0.003	0.005	0.005	0.007	0.007	0.009	5.7
phecode_101-6	Malignant neoplasm of the liver and intrahepatic bile ducts	0	0	0.001	0.001	0.001	0.001	0.003	0.003	0.003	0.006	11.335
phecode_101-8	Malignant neoplasm of the pancreas	0	0.001	0.002	0.002	0.004	0.002	0.004	0.003	0.005	0.006	13.002
phecode_102	Malignant neoplasm of the thoracic and respiratory organs	0.001	0.004	0.004	0.006	0.009	0.012	0.015	0.016	0.019	0.027	24.004
phecode_102-1	Malignant neoplasm of the of bronchus and lung	0	0.003	0.004	0.005	0.008	0.01	0.012	0.014	0.016	0.026	52.675
phecode_103	Malignant neoplasm of the skin	0.01	0.03	0.036	0.041	0.058	0.069	0.08	0.084	0.095	0.107	11.177
phecode_103-1	Melanomas of skin	0.002	0.006	0.008	0.009	0.009	0.012	0.011	0.013	0.012	0.013	7.365

5 Phenome-wide prediction of disease onset from retinal fundus photographs

Table 22 continued from previous page

Endpoint	PheCode string	1	2	3	4	5	6	7	8	9	10	Rate Ratio (RR)
phecode_103-2	Keratinocyte carcinoma	0.005	0.018	0.016	0.024	0.03	0.04	0.04	0.049	0.053	0.056	10.502
phecode_103-21	Basal cell carcinoma	0.004	0.017	0.015	0.02	0.026	0.036	0.035	0.04	0.047	0.046	10.335
phecode_103-22	Squamous cell carcinoma of the skin	0.001	0.002	0.002	0.003	0.004	0.007	0.008	0.009	0.01	0.014	14.169
phecode_103-3	Carcinoma in situ of skin	0.001	0.003	0.004	0.007	0.007	0.013	0.012	0.017	0.016	0.018	12.446
phecode_104	Malignant sarcoma-related cancers	0.002	0.002	0.003	0.004	0.003	0.005	0.005	0.005	0.004	0.004	1.734
phecode_105	Malignant neoplasm of the breast	0.002	0.005	0.01	0.019	0.023	0.026	0.031	0.038	0.04	0.044	26.304
phecode_105-1	Malignant neoplasm of the breast, female	0.027	0.033	0.03	0.039	0.039	0.034	0.029	0.037	0.033	0.035	1.326
phecode_106	Gynecological malignant neoplasms	0.01	0.014	0.01	0.01	0.016	0.013	0.015	0.019	0.018	0.012	1.291
phecode_106-2	Malignant neoplasm of the uterus	0.002	0.004	0.008	0.006	0.006	0.008	0.008	0.01	0.008	0.008	3.126
phecode_106-21	Malignant neoplasm of endometrium	0.002	0.004	0.007	0.005	0.007	0.007	0.008	0.011	0.008	0.007	2.875
phecode_106-3	Malignant neoplasm of the ovary	0.004	0.003	0.002	0.005	0.006	0.005	0.005	0.007	0.008	0.007	1.917
phecode_107	Malignant neoplasm of male genitalia	0.01	0.021	0.045	0.06	0.07	0.078	0.084	0.085	0.096	0.089	8.753
phecode_107-2	Malignant neoplasm of the prostate	0.007	0.019	0.044	0.059	0.068	0.079	0.075	0.085	0.092	0.09	13.057
phecode_108	Malignant neoplasm of the urinary tract	0.002	0.003	0.006	0.006	0.008	0.011	0.012	0.015	0.019	0.024	12.335
phecode_108-4	Malignant neoplasm of the kidney	0.001	0.002	0.003	0.001	0.003	0.004	0.004	0.005	0.007	0.007	6.001
phecode_108-41	Malignant neoplasm of kidney, except pelvis	0.001	0.001	0.003	0.002	0.003	0.004	0.004	0.004	0.006	0.006	9.502
phecode_108-5	Malignant neoplasm of the bladder	0.001	0.002	0.002	0.004	0.003	0.006	0.006	0.007	0.01	0.015	15.169
phecode_109	Malignant neoplasm of the eye, brain and other parts of central nervous system	0.002	0.001	0.001	0.001	0.003	0.004	0.004	0.005	0.004	0.005	2
phecode_109-3	Malignant neoplasm of brain	0.002	0.001	0.001	0.002	0.002	0.003	0.004	0.002	0.003	0.005	2.75
phecode_110	Malignant neoplasm of the endocrine glands	0.001	0.002	0.001	0.001	0.002	0.002	0.002	0.001	0.001	0.003	2.572
phecode_112	Malignant neoplasm of other and ill-defined sites	0.025	0.036	0.041	0.053	0.068	0.08	0.086	0.105	0.125	0.134	5.377
phecode_114	Neuroendocrine tumors	0.001	0.001	0.001	0.002	0.002	0.003	0.002	0.003	0.002	0.004	6.751
phecode_116	Secondary malignant neoplasm	0.01	0.013	0.021	0.024	0.027	0.038	0.046	0.046	0.056	0.067	6.493
phecode_116-1	Secondary malignancy of lymph nodes	0.005	0.008	0.01	0.012	0.018	0.018	0.016	0.019	0.019	0.023	4.469
phecode_116-2	Secondary malignancy of respiratory organs	0.003	0.005	0.007	0.007	0.007	0.01	0.012	0.016	0.016	0.019	6.39
phecode_116-3	Secondary malignant neoplasm of digestive systems	0.002	0.002	0.005	0.007	0.008	0.007	0.008	0.006	0.01	0.013	6.001
phecode_116-4	Secondary malignant neoplasm of liver	0.002	0.004	0.008	0.009	0.011	0.013	0.015	0.015	0.019	0.025	10.268
phecode_116-5	Secondary malignancy of brain/spine	0.001	0.003	0.002	0.002	0.003	0.006	0.005	0.006	0.008	0.006	6
phecode_116-6	Secondary malignancy of bone	0.001	0.002	0.006	0.008	0.01	0.011	0.014	0.017	0.018	0.024	16.558
phecode_120	Hemo onc - by cell of origin	0.005	0.005	0.008	0.009	0.013	0.019	0.018	0.024	0.023	0.023	4.376
phecode_120-1	Myeloid	0.002	0.003	0.002	0.004	0.006	0.007	0.01	0.01	0.012	0.013	5.924
phecode_120-2	Lymphoid	0.002	0.003	0.003	0.005	0.007	0.008	0.009	0.013	0.011	0.01	4.201
phecode_120-21	Mature B-cell	0.001	0.002	0.003	0.005	0.005	0.008	0.008	0.011	0.01	0.009	6.334
phecode_121	Leukemia	0.001	0.001	0.004	0.003	0.003	0.002	0.005	0.006	0.008	0.009	7.716
phecode_121-2	Chronic leukemia	0	0.001	0.002	0.001	0.002	0.002	0.003	0.004	0.005	0.006	39.006
phecode_121-21	Chronic lymphoid leukemia	0	0.001	0.002	0.001	0.002	0.002	0.002	0.004	0.005	0.005	Inf
phecode_122	Lymphoma	0.002	0.002	0.003	0.003	0.006	0.008	0.007	0.007	0.01	0.007	3.215
phecode_122-2	Non-Hodgkin lymphoma	0.001	0.002	0.002	0.003	0.005	0.007	0.007	0.007	0.009	0.007	5.858
phecode_122-22	Diffuse large B-cell lymphoma*	0.001	0	0	0.001	0.003	0.004	0.003	0.002	0.003	0.002	3.251
phecode_123	Multiple myeloma and malignant plasma cell neoplasms	0.002	0	0.001	0.001	0.001	0.003	0.002	0.003	0.005	0.005	2
phecode_123-1	Multiple myeloma	0.002	0	0	0.001	0.002	0.002	0.002	0.004	0.003	0.005	2.067
phecode_124	Myeloproliferative disorder	0.001	0.001	0.001	0.003	0.004	0.004	0.006	0.005	0.008	0.006	4.751
phecode_124-5	Essential thrombocythemia	0.001	0	0.001	0.002	0.001	0.002	0.003	0.002	0.003	0.003	3.201
phecode_130	Cancer (solid tumor, excluding BCC)	0.05	0.075	0.086	0.109	0.126	0.15	0.165	0.181	0.21	0.225	4.478
phecode_132	Sequelae of cancer	0.001	0.002	0.004	0.011	0.009	0.012	0.016	0.013	0.019	0.022	17.128
phecode_135	Benign neoplasm of the head and neck	0.011	0.009	0.008	0.011	0.011	0.011	0.012	0.01	0.013	0.016	1.508
phecode_135-1	Benign neoplasm of the oral cavity	0.002	0.001	0.001	0.002	0.003	0.003	0.002	0.002	0.003	0.004	2.167
phecode_135-5	Benign neoplasm of the paranasal sinus and nasal cavity	0.007	0.008	0.007	0.006	0.009	0.007	0.008	0.007	0.009	0.012	1.667
phecode_136	Benign neoplasm of the digestive organs	0.051	0.075	0.082	0.083	0.092	0.109	0.118	0.121	0.128	0.144	2.803
phecode_136-2	Benign neoplasm of stomach	0.008	0.012	0.016	0.019	0.022	0.024	0.028	0.026	0.029	0.028	3.605
phecode_136-4	Benign neoplasm of colon, rectum, anus and anal canal	0.044	0.065	0.064	0.066	0.076	0.087	0.099	0.098	0.116	0.127	2.864
phecode_136-41	Benign neoplasm of the colon	0.036	0.045	0.049	0.058	0.065	0.076	0.083	0.091	0.097	0.106	2.963

Supplementary Tables

Table 22 continued from previous page

Endpoint	PheCode string	1	2	3	4	5	6	7	8	9	10	Rate Ratio (RR)
phecode_136-42	Benign neoplasm of rectum and anus	0.019	0.022	0.027	0.026	0.029	0.032	0.039	0.039	0.041	0.054	2.894
phecode_138	Benign neoplasm of the skin	0.055	0.07	0.075	0.072	0.079	0.085	0.08	0.087	0.092	0.097	1.775
phecode_138-1	Nevus, non-neoplastic	0.005	0.008	0.007	0.008	0.007	0.006	0.008	0.007	0.01	0.01	1.844
phecode_138-2	Melanocytic nevi*	0.042	0.049	0.05	0.053	0.062	0.066	0.069	0.064	0.073	0.083	1.971
phecode_139	Benign sarcoma-related cancers	0.035	0.035	0.036	0.041	0.035	0.037	0.035	0.036	0.042	0.04	1.137
phecode_139-3	Benign neoplasm of other connective and soft tissue	0.002	0.001	0.001	0.001	0.003	0.001	0.002	0.002	0.002	0.002	1.5
phecode_139-5	Lipoma	0.026	0.03	0.026	0.025	0.026	0.026	0.022	0.026	0.026	0.035	1.331
phecode_139-52	Lipoma of intrathoracic organs	0.002	0.001	0.001	0.002	0.002	0.001	0.002	0.003	0.002	0.003	1.308
phecode_139-53	Lipoma of other skin subcutaneous tissue	0.009	0.011	0.008	0.008	0.009	0.007	0.008	0.007	0.008	0.01	1.105
phecode_139-54	Testicular lipoma	0.004	0.003	0.005	0.004	0.006	0.008	0.003	0.005	0.004	0.004	1
phecode_139-6	Hemangioma and lymphangioma	0.007	0.007	0.008	0.008	0.009	0.007	0.008	0.009	0.01	0.008	1.167
phecode_139-61	Hemangioma	0.007	0.007	0.009	0.007	0.009	0.007	0.008	0.007	0.012	0.008	1.143
phecode_140	Benign neoplasm of the breast	0	0.001	0.001	0.001	0.002	0.003	0.003	0.004	0.006	0.006	12.335
phecode_142	Lump or mass in breast or nonspecific abnormal breast exam	0.006	0.014	0.017	0.032	0.044	0.049	0.064	0.075	0.088	0.111	17.364
phecode_142-1	Lump or mass in breast	0.006	0.011	0.013	0.023	0.025	0.028	0.039	0.04	0.048	0.07	11.767
phecode_144	Gynecological benign neoplasms	0.044	0.04	0.035	0.049	0.054	0.055	0.063	0.083	0.088	0.132	3.008
phecode_144-1	Benign neoplasms of external female genital organs and cervix	0.008	0.006	0.007	0.009	0.012	0.012	0.014	0.014	0.017	0.018	2.193
phecode_144-13	Benign neoplasms of the cervix	0.006	0.006	0.007	0.007	0.01	0.012	0.011	0.015	0.016	0.015	2.474
phecode_144-2	Benign neoplasms of the uterus	0.036	0.029	0.031	0.043	0.037	0.05	0.048	0.062	0.07	0.12	3.304
phecode_144-21	Leiomyoma of uterus	0.024	0.014	0.02	0.022	0.029	0.031	0.034	0.049	0.058	0.101	4.221
phecode_144-3	Benign neoplasms of the ovary	0.004	0.006	0.005	0.003	0.003	0.005	0.007	0.007	0.005	0.003	0.769
phecode_146	Benign neoplasm of the genitourinary system	0.001	0.003	0.004	0.006	0.009	0.011	0.014	0.023	0.029	0.037	32.434
phecode_146-2	Benign neoplasm of the prostate	0.012	0.015	0.023	0.023	0.028	0.032	0.035	0.036	0.04	0.045	3.844
phecode_148	Benign neoplasm of the eye, brain and other parts of central nervous system	0.003	0.003	0.005	0.007	0.007	0.007	0.008	0.008	0.007	0.009	2.667
phecode_148-1	Benign neoplasm of eye	0.002	0.002	0.002	0.003	0.004	0.004	0.004	0.004	0.006	0.005	3.301
phecode_148-16	Benign neoplasm of choroid	0.001	0.001	0.002	0.002	0.005	0.003	0.003	0.004	0.005	0.006	3.778
phecode_148-2	Benign neoplasm of meninges (Meningioma)	0.001	0.001	0.002	0.001	0.002	0.002	0.002	0.003	0.002	0.003	2.375
phecode_149	Benign neoplasm of the endocrine glands	0.002	0.002	0.003	0.002	0.004	0.005	0.005	0.003	0.005	0.005	2.637
phecode_149-3	Benign neoplasm of the parathyroid gland	0.001	0.001	0.001	0.001	0.001	0.002	0.003	0.002	0.003	0.002	3
phecode_160	Nutritional anemias	0.021	0.026	0.033	0.037	0.041	0.047	0.051	0.064	0.067	0.099	4.765
phecode_160-1	Iron deficiency anemia	0.019	0.025	0.028	0.033	0.035	0.047	0.047	0.055	0.06	0.09	4.76
phecode_160-2	Megaloblastic anemia	0.003	0.003	0.002	0.005	0.005	0.007	0.008	0.01	0.01	0.014	4.632
phecode_162	Aplastic anemia	0.001	0.001	0.001	0.001	0.003	0.003	0.003	0.002	0.004	0.005	7.751
phecode_164	Anemia	0.035	0.049	0.052	0.061	0.074	0.086	0.094	0.104	0.128	0.166	4.747
phecode_164-1	Microcytic anemia	0.019	0.025	0.027	0.036	0.035	0.045	0.047	0.054	0.06	0.09	4.846
phecode_164-2	Macrocytic anemia	0.003	0.003	0.003	0.005	0.005	0.008	0.009	0.01	0.01	0.016	4.572
phecode_164-6	Anemia secondary to chronic diseases and conditions	0	0.001	0.001	0.001	0.001	0.002	0.004	0.002	0.004	0.008	25.004
phecode_165	Hemoglobinopathies	0	0	0	0	0	0.001	0	0.001	0.002	0.017	103.017
phecode_168	Coagulation defects, purpura and other hemorrhagic conditions	0.008	0.008	0.009	0.008	0.012	0.014	0.017	0.017	0.018	0.023	2.915
phecode_168-1	Hypo-coagulability	0.003	0.004	0.004	0.005	0.007	0.008	0.008	0.008	0.011	0.015	4.451
phecode_168-19	Spontaneous ecchymoses	0.003	0.003	0.004	0.003	0.004	0.005	0.004	0.004	0.005	0.007	2.5
phecode_168-2	Hyper-coagulability	0.001	0.002	0.002	0.002	0.002	0.003	0.001	0.002	0.003	0.002	2.333
phecode_168-4	Abnormal coagulation profile	0.001	0	0.001	0.001	0.002	0.002	0.005	0.003	0.005	0.005	4.001
phecode_169	Platelet defects	0.004	0.005	0.006	0.007	0.008	0.011	0.016	0.015	0.02	0.02	5.262
phecode_169-1	Thrombocytopenia	0.004	0.005	0.006	0.006	0.008	0.011	0.015	0.016	0.019	0.02	5.637
phecode_169-11	Immune thrombocytopenic purpura [ITP]	0.001	0.001	0.001	0.001	0.001	0.001	0.002	0.003	0.002	0.003	4.751
phecode_170	Decreased white blood cell count	0.007	0.008	0.013	0.013	0.014	0.016	0.017	0.018	0.018	0.024	3.364
phecode_170-1	Neutropenia	0.007	0.009	0.011	0.013	0.014	0.015	0.017	0.018	0.018	0.023	3.415
phecode_170-19	Neutropenia NOS	0.003	0.003	0.004	0.003	0.003	0.004	0.005	0.005	0.007	0.008	2.381
phecode_171	Increased white blood cell count	0.002	0.001	0.002	0.002	0.003	0.004	0.003	0.004	0.004	0.003	1.5
phecode_171-1	Lymphocytosis (symptomatic)	0.001	0.001	0.001	0.001	0.002	0.002	0.002	0.003	0.002	0.002	2.167

5 Phenome-wide prediction of disease onset from retinal fundus photographs

Table 22 continued from previous page

Endpoint	PheCode string	1	2	3	4	5	6	7	8	9	10	Rate Ratio (RR)
phecode_172	Other disorders of white blood cells	0.002	0.003	0.003	0.004	0.004	0.004	0.005	0.006	0.006	0.007	3.308
phecode_174	Diseases of spleen	0.002	0.002	0.003	0.003	0.004	0.004	0.003	0.007	0.007	0.006	3.084
phecode_174-2	Splenomegaly	0.001	0.002	0.001	0.001	0.002	0.002	0.003	0.004	0.004	0.004	3.286
phecode_175	Polycythemias	0	0.001	0.002	0.001	0.002	0.003	0.003	0.003	0.004	0.005	28.005
phecode_175-2	Secondary polycythemia	0	0.001	0.001	0.002	0.001	0.002	0.003	0.002	0.004	0.004	25
phecode_176	Other diseases of blood and blood-forming organs	0.003	0.006	0.007	0.006	0.006	0.008	0.008	0.009	0.01	0.009	2.715
phecode_177	Abnormality of the lymph nodes	0.02	0.025	0.031	0.028	0.034	0.035	0.039	0.037	0.037	0.048	2.419
phecode_177-2	Enlargement of lymph nodes [Lymphadenopathy]	0.018	0.018	0.027	0.023	0.026	0.023	0.024	0.026	0.029	0.033	1.851
phecode_177-4	Lymphedema	0.002	0.004	0.004	0.005	0.005	0.007	0.009	0.01	0.011	0.01	5.168
phecode_179	Immunodeficiencies	0.003	0.003	0.003	0.002	0.004	0.003	0.003	0.003	0.003	0.003	0.95
phecode_179-9	Immunodeficiency NOS	0.003	0.001	0.001	0.002	0.002	0.002	0.002	0.001	0.003	0.002	0.938
phecode_180	Other disorders involving the immune mechanism	0.004	0.004	0.004	0.004	0.006	0.006	0.008	0.009	0.009	0.01	2.44
phecode_180-3	Paraproteinemias	0.001	0.001	0.002	0.002	0.003	0.005	0.006	0.005	0.007	0.008	7.668
phecode_180-31	Monoclonal gammopathy	0.001	0.001	0.002	0.002	0.003	0.004	0.005	0.005	0.007	0.007	7.168
phecode_181	Autoimmune disease	0.019	0.021	0.024	0.022	0.023	0.028	0.028	0.032	0.033	0.036	1.878
phecode_200	Disorders of thyroid gland	0.024	0.032	0.038	0.055	0.055	0.054	0.067	0.072	0.074	0.074	3.08
phecode_200-1	Hypothyroidism	0.019	0.024	0.035	0.038	0.049	0.044	0.048	0.055	0.055	0.067	3.587
phecode_200-13	Postprocedural hypothyroidism	0.002	0.001	0.002	0.004	0.005	0.005	0.004	0.005	0.004	0.007	4.301
phecode_200-14	Hypothyroidism, not specified as secondary	0.016	0.021	0.029	0.033	0.043	0.038	0.042	0.05	0.05	0.058	3.628
phecode_200-2	Goiter	0.005	0.008	0.009	0.011	0.011	0.011	0.015	0.011	0.015	0.019	3.563
phecode_200-21	Diffuse goiter	0.002	0.001	0.002	0.003	0.003	0.003	0.003	0.004	0.002	0.002	1.364
phecode_200-22	Uninodular goiter [single thyroid nodule]	0.004	0.003	0.004	0.004	0.004	0.006	0.006	0.006	0.008	0.007	1.667
phecode_200-23	Multinodular goiter	0.001	0.002	0.003	0.004	0.004	0.004	0.005	0.004	0.004	0.008	8.001
phecode_200-3	Thyrotoxicosis [hyperthyroidism]	0.005	0.006	0.007	0.008	0.011	0.009	0.009	0.009	0.008	0.009	1.728
phecode_200-31	Graves' disease [Toxic diffuse goiter]	0.001	0.001	0.002	0.002	0.002	0.002	0.002	0.003	0.002	0.001	1
phecode_200-4	Thyroiditis	0.001	0.001	0.001	0.002	0.001	0.001	0.003	0.001	0.003	0.003	2
phecode_200-7	Iodine-deficiency related thyroid disorders*	0.004	0.003	0.006	0.004	0.006	0.007	0.008	0.008	0.006	0.006	1.445
phecode_200-9	Abnormal thyroid function studies	0.003	0.004	0.004	0.004	0.003	0.005	0.004	0.004	0.005	0.006	1.857
phecode_202	Diabetes mellitus	0.022	0.033	0.044	0.057	0.066	0.079	0.081	0.084	0.113	0.156	7.056
phecode_202-1	Type 1 diabetes	0.001	0.003	0.005	0.004	0.007	0.007	0.007	0.008	0.009	0.022	14.778
phecode_202-2	Type 2 diabetes	0.022	0.032	0.045	0.055	0.065	0.076	0.078	0.081	0.113	0.15	6.922
phecode_202-4	Other specified diabetes*	0.018	0.03	0.046	0.056	0.066	0.08	0.091	0.098	0.13	0.213	12.039
phecode_204	Elevated blood glucose level	0.031	0.042	0.061	0.07	0.079	0.087	0.092	0.098	0.111	0.142	4.591
phecode_204-1	Impaired fasting glucose	0.003	0.007	0.009	0.01	0.014	0.012	0.014	0.016	0.016	0.02	6.668
phecode_204-2	Impaired glucose tolerance (oral)	0.009	0.016	0.019	0.023	0.027	0.03	0.034	0.036	0.037	0.049	5.176
phecode_204-4	Prediabetes*	0.001	0.001	0.001	0.002	0.002	0.001	0.002	0.003	0.003	0.002	1.667
phecode_205	Hypoglycemia	0.002	0.002	0.002	0.002	0.003	0.004	0.005	0.008	0.013	0.027	13.583
phecode_208	Disorders of parathyroid gland	0.002	0.003	0.003	0.005	0.006	0.007	0.007	0.01	0.007	0.011	6.801
phecode_208-2	Hyperparathyroidism	0.001	0.004	0.004	0.004	0.007	0.006	0.007	0.009	0.009	0.01	10.168
phecode_208-21	Primary hyperparathyroidism	0.001	0.002	0.002	0.003	0.004	0.003	0.005	0.006	0.006	0.005	6.601
phecode_209	Disorders of the pituitary gland and its hypothalamic control	0.002	0.002	0.002	0.003	0.003	0.003	0.003	0.003	0.006	0.005	2.067
phecode_209-1	Pituitary hyperfunction	0	0.001	0.001	0.001	0.001	0.002	0.002	0.002	0.002	0.003	7.001
phecode_211	Disorders of adrenal glands	0.002	0.001	0.001	0.001	0.002	0.004	0.004	0.003	0.005	0.006	3.801
phecode_215	Testicular dysfunction	0.003	0.004	0.006	0.003	0.002	0.001	0.004	0.004	0.004	0.005	1.667
phecode_230	Malnutrition and underweight	0.028	0.029	0.033	0.04	0.044	0.052	0.063	0.063	0.067	0.088	3.167
phecode_230-2	Abnormal loss of weight and underweight	0.021	0.02	0.03	0.032	0.035	0.041	0.052	0.048	0.056	0.071	3.416
phecode_230-21	Abnormal weight loss	0.018	0.02	0.029	0.028	0.033	0.041	0.048	0.049	0.051	0.07	3.802
phecode_230-22	Underweight	0.002	0.001	0.002	0.003	0.002	0.003	0.003	0.005	0.003	0.002	0.769
phecode_230-3	Anorexia	0.006	0.009	0.007	0.009	0.012	0.014	0.017	0.015	0.016	0.022	3.436
phecode_232	Vitamin deficiencies	0.049	0.057	0.058	0.067	0.071	0.083	0.078	0.099	0.108	0.177	3.61
phecode_232-2	Vitamin B group deficiency	0.012	0.013	0.015	0.02	0.021	0.03	0.032	0.031	0.036	0.051	4.401
phecode_232-27	Vitamin B12 deficiency	0.002	0.003	0.002	0.003	0.004	0.005	0.005	0.006	0.005	0.009	3.786
phecode_232-29	Folate deficiency [Vitamin B9]	0.001	0	0	0.001	0.001	0.002	0.004	0.004	0.004	0.007	6.668
phecode_232-4	Vitamin D deficiency	0.039	0.036	0.046	0.052	0.054	0.06	0.062	0.067	0.081	0.143	3.684
phecode_234	Other nutritional deficiencies	0.013	0.016	0.021	0.016	0.021	0.023	0.023	0.019	0.027	0.031	2.375

Supplementary Tables

Table 22 continued from previous page

Endpoint	PheCode string	1	2	3	4	5	6	7	8	9	10	Rate Ratio (RR)
phecode_236	Overweight and obesity	0.064	0.082	0.092	0.107	0.106	0.107	0.119	0.129	0.139	0.161	2.531
phecode_236-1	Obesity	0.062	0.083	0.091	0.108	0.103	0.111	0.12	0.125	0.139	0.161	2.581
phecode_236-11	Morbid obesity	0.001	0.003	0.002	0.003	0.003	0.006	0.007	0.005	0.008	0.008	6.251
phecode_237	Abnormal weight gain	0.003	0.004	0.005	0.007	0.006	0.007	0.006	0.006	0.007	0.008	2.55
phecode_239	Hyperlipidemia	0.086	0.112	0.148	0.161	0.18	0.204	0.223	0.238	0.266	0.3	3.474
phecode_239-1	Hypercholesterolemia	0.075	0.097	0.135	0.141	0.164	0.188	0.197	0.223	0.243	0.281	3.748
phecode_239-11	Pure hypercholesterolemia	0.074	0.097	0.13	0.141	0.164	0.186	0.204	0.22	0.244	0.283	3.819
phecode_239-12	Familial hypercholesterolemia*	0.002	0.002	0.001	0.002	0.002	0.004	0.003	0.005	0.004	0.004	2.167
phecode_239-2	Hyperglyceridemia	0.002	0.003	0.005	0.003	0.005	0.003	0.005	0.005	0.006	0.006	3.601
phecode_239-3	Mixed hyperlipidemia	0.001	0.003	0.004	0.002	0.003	0.003	0.004	0.004	0.004	0.004	4.6
phecode_244	Disorders of lipoprotein metabolism and other lipidemias	0.003	0.004	0.005	0.004	0.008	0.008	0.007	0.01	0.013	0.01	3.648
phecode_247	Disorders of mineral metabolism and mineral deficiencies	0.031	0.038	0.045	0.053	0.056	0.067	0.071	0.081	0.091	0.123	3.973
phecode_247-3	Disorder of phosphorus metabolism	0.002	0.001	0.002	0.002	0.003	0.004	0.004	0.004	0.005	0.008	4.701
phecode_247-4	Disorders of magnesium metabolism	0.002	0.001	0.002	0.004	0.005	0.005	0.006	0.007	0.011	0.015	9.101
phecode_247-5	Disorders of calcium metabolism	0.006	0.006	0.007	0.009	0.009	0.012	0.013	0.019	0.02	0.02	3.529
phecode_247-51	Hypocalcemia	0.002	0.002	0.001	0.001	0.001	0.002	0.003	0.003	0.002	0.003	1.417
phecode_247-52	Hypercalcemia	0.001	0.003	0.002	0.002	0.004	0.004	0.003	0.005	0.005	0.005	3.556
phecode_247-7	Disorders of iron metabolism	0.026	0.029	0.035	0.04	0.043	0.052	0.053	0.061	0.066	0.099	3.768
phecode_247-72	Iron deficiency	0.023	0.027	0.033	0.039	0.038	0.05	0.054	0.058	0.065	0.096	4.24
phecode_248	Disorders of plasma-protein metabolism, NEC	0.001	0.001	0.001	0.001	0.001	0.003	0.002	0.002	0.003	0.006	9.001
phecode_251	Disorders of bilirubin excretion	0.001	0.002	0.002	0.002	0.002	0.003	0.006	0.004	0.004	0.004	2.778
phecode_251-1	Gilbert syndrome*	0.001	0.002	0.002	0.002	0.002	0.003	0.005	0.004	0.002	0.005	5.001
phecode_256	Disorders of fluid, electrolyte and acid-base balance	0.043	0.045	0.06	0.063	0.072	0.078	0.086	0.102	0.109	0.138	3.192
phecode_256-1	Hyperosmolality and/or hyponatremia	0.003	0.005	0.005	0.005	0.005	0.006	0.005	0.005	0.005	0.009	2.524
phecode_256-2	Hyposmolality and/or hyponatremia	0.026	0.031	0.035	0.033	0.038	0.037	0.039	0.042	0.047	0.052	1.988
phecode_256-3	Mixed disorder of acid-base balance	0.002	0.003	0.003	0.005	0.006	0.007	0.008	0.011	0.014	0.022	11
phecode_256-31	Acidosis	0.002	0.002	0.004	0.003	0.005	0.006	0.007	0.012	0.013	0.018	11.302
phecode_256-4	Hyperkalemia [Hyperpotassemia]	0.001	0.003	0.003	0.005	0.008	0.008	0.014	0.016	0.016	0.027	18.225
phecode_256-5	Hypokalemia [Hypopotassemia]	0.006	0.005	0.009	0.01	0.015	0.013	0.022	0.021	0.026	0.031	5.429
phecode_256-6	Fluid overload	0	0.001	0.003	0.002	0.004	0.004	0.008	0.008	0.01	0.016	31.672
phecode_256-7	Volume depletion	0.005	0.007	0.009	0.012	0.015	0.019	0.023	0.024	0.033	0.044	8.968
phecode_257	Polydipsia	0.002	0.004	0.003	0.003	0.005	0.002	0.003	0.003	0.003	0.002	1
phecode_280	Substance related disorders	0.022	0.026	0.026	0.036	0.04	0.04	0.046	0.048	0.059	0.062	2.81
phecode_280-1	Alcohol use disorders	0.019	0.027	0.023	0.033	0.034	0.038	0.044	0.046	0.053	0.06	3.167
phecode_280-11	Alcohol abuse	0.009	0.011	0.011	0.015	0.021	0.02	0.024	0.025	0.028	0.036	4.001
phecode_280-12	Alcohol dependence	0.005	0.008	0.008	0.009	0.008	0.012	0.014	0.011	0.017	0.02	3.758
phecode_280-13	Alcoholic liver disease	0	0.001	0.002	0.002	0.003	0.001	0.004	0.003	0.006	0.008	16.336
phecode_280-8	Other psychoactive substance related disorders	0.002	0.002	0.001	0.002	0.001	0.003	0.001	0.003	0.003	0.005	2.385
phecode_280-82	Other psychoactive substance dependence	0.002	0.001	0.001	0.001	0.001	0.001	0.002	0.002	0.003	0.003	1.462
phecode_281	Substance abuse, dependence, and withdrawal	0.047	0.055	0.064	0.066	0.068	0.075	0.084	0.098	0.109	0.128	2.7
phecode_281-1	Substance abuse	0.01	0.011	0.013	0.015	0.023	0.021	0.025	0.025	0.03	0.037	3.751
phecode_281-2	Substance dependence	0.039	0.043	0.052	0.054	0.058	0.063	0.074	0.087	0.094	0.114	2.914
phecode_282-1	Current tobacco use and nicotine dependence	0.026	0.031	0.033	0.036	0.04	0.047	0.052	0.061	0.062	0.086	3.252
phecode_283	Other behavioral problems	0.174	0.191	0.188	0.198	0.208	0.212	0.218	0.225	0.235	0.243	1.398
phecode_283-4	Patient's noncompliance with medical treatment and regimen	0.027	0.028	0.035	0.035	0.033	0.038	0.043	0.043	0.052	0.064	2.314
phecode_283-8	Other problems related to lifestyle	0.15	0.157	0.168	0.166	0.177	0.18	0.188	0.178	0.195	0.205	1.359
phecode_284	Suicide attempt	0.008	0.007	0.007	0.009	0.013	0.01	0.011	0.012	0.017	0.015	1.898
phecode_284-1	Suicidal ideations	0.006	0.004	0.007	0.008	0.009	0.01	0.008	0.011	0.014	0.011	2.03
phecode_284-2	Suicide and self-inflicted harm	0.002	0.003	0.003	0.003	0.003	0.003	0.003	0.004	0.003	0.007	3.231
phecode_284-29	Intentional self-harm*	0.002	0.003	0.003	0.003	0.003	0.002	0.004	0.003	0.003	0.007	2.734
phecode_286	Mood [affective] disorders	0.069	0.07	0.083	0.085	0.079	0.086	0.092	0.103	0.115	0.136	1.958
phecode_286-1	Bipolar disorder	0.001	0.002	0.001	0.001	0.002	0.002	0.001	0.003	0.002	0.002	1.625
phecode_286-2	Major depressive disorder	0.069	0.069	0.083	0.079	0.082	0.089	0.093	0.099	0.118	0.131	1.893
phecode_286-21	Major depressive disorder, recurrent	0.004	0.006	0.002	0.005	0.003	0.003	0.004	0.005	0.005	0.007	1.667

5 Phenome-wide prediction of disease onset from retinal fundus photographs

Table 22 continued from previous page

Endpoint	PheCode string	1	2	3	4	5	6	7	8	9	10	Rate Ratio (RR)
phecode_287	Psychotic disorder	0.002	0.003	0.003	0.002	0.002	0.002	0.001	0.003	0.004	0.005	2.2
phecode_288	Anxiety disorders	0.055	0.071	0.074	0.078	0.081	0.092	0.095	0.107	0.11	0.127	2.315
phecode_288-2	Panic disorder [episodic paroxysmal anxiety]	0.007	0.008	0.007	0.008	0.007	0.011	0.009	0.011	0.008	0.014	1.931
phecode_288-3	Generalized anxiety disorder	0.013	0.014	0.017	0.018	0.017	0.016	0.02	0.022	0.022	0.028	2.154
phecode_288-4	Phobic disorders	0.007	0.007	0.011	0.011	0.008	0.011	0.011	0.012	0.014	0.02	2.75
phecode_290	Reaction to severe stress, and adjustment disorders	0.021	0.023	0.027	0.027	0.029	0.035	0.043	0.051	0.058	0.07	3.384
phecode_290-1	Posttraumatic stress disorder	0.001	0.001	0.001	0.002	0.001	0.002	0.001	0.002	0.002	0.004	3.858
phecode_292	Somatiform disorders	0.002	0.001	0.003	0.003	0.002	0.001	0.001	0.002	0.002	0.003	1.616
phecode_294	Sexual dysfunction and disorders	0.008	0.013	0.019	0.028	0.043	0.058	0.07	0.095	0.117	0.137	16.503
phecode_299	Mental disorder, not otherwise specified	0.004	0.004	0.003	0.006	0.004	0.005	0.005	0.006	0.006	0.008	1.778
phecode_308	Signs and symptoms involving emotional state	0.057	0.064	0.065	0.068	0.07	0.071	0.076	0.079	0.097	0.113	1.997
phecode_308-1	Irritability	0.001	0.002	0.001	0.002	0.002	0.003	0.003	0.003	0.002	0.003	2.858
phecode_308-3	Emotional lability	0.001	0.001	0.001	0.002	0.003	0.003	0.001	0.003	0.004	0.008	5.875
phecode_308-4	Demoralization and apathy	0.002	0.001	0.002	0.002	0.003	0.002	0.003	0.003	0.004	0.006	2.643
phecode_308-5	Nervousness	0.003	0.002	0.003	0.002	0.003	0.003	0.003	0.003	0.004	0.004	1.389
phecode_308-6	Excessive crying of child, adolescent, or adult	0.001	0.001	0.001	0.001	0.003	0.002	0.001	0.001	0.003	0.004	3.834
phecode_308-7	Restlessness and agitation*	0.002	0.002	0.003	0.002	0.003	0.003	0.004	0.005	0.004	0.006	2.5
phecode_323	Systemic atrophies primarily affecting the central nervous system	0.001	0.001	0.001	0.003	0.003	0.003	0.003	0.002	0.002	0.002	3
phecode_324	Extrapyramidal and movement disorders	0.022	0.025	0.033	0.036	0.04	0.046	0.052	0.054	0.054	0.063	2.884
phecode_324-1	Parkinsonism	0.001	0.002	0.004	0.005	0.008	0.009	0.012	0.012	0.015	0.014	21.003
phecode_324-11	Parkinson's disease	0.001	0.001	0.004	0.005	0.007	0.009	0.011	0.012	0.015	0.013	20.003
phecode_324-3	Dystonia	0.01	0.013	0.012	0.014	0.016	0.015	0.016	0.017	0.02	0.02	2.106
phecode_324-34	Torticollis	0.007	0.009	0.009	0.009	0.01	0.009	0.009	0.014	0.011	0.013	1.818
phecode_324-4	Tremor	0.006	0.007	0.01	0.014	0.017	0.016	0.021	0.02	0.024	0.024	3.821
phecode_324-41	Essential tremor*	0.001	0.003	0.003	0.003	0.005	0.005	0.007	0.008	0.006	0.009	7.858
phecode_324-8	Restless legs syndrome	0.003	0.003	0.008	0.005	0.007	0.008	0.006	0.008	0.011	0.009	3.167
phecode_325	Symptoms and signs related to movement disorders	0.017	0.016	0.022	0.026	0.031	0.04	0.047	0.057	0.068	0.088	5.095
phecode_325-1	Abnormal involuntary movements	0.003	0.002	0.001	0.003	0.003	0.002	0.003	0.002	0.001	0.003	0.85
phecode_325-2	Abnormality of gait and mobility	0.013	0.013	0.018	0.023	0.028	0.034	0.043	0.054	0.063	0.084	6.376
phecode_325-23	Unsteadiness on feet*	0.002	0.002	0.002	0.003	0.004	0.004	0.006	0.009	0.009	0.007	3.385
phecode_325-3	Lack of coordination	0.005	0.004	0.004	0.004	0.005	0.007	0.011	0.009	0.01	0.011	2.38
phecode_327	Other degenerative diseases of nervous system	0	0.001	0.002	0.003	0.003	0.004	0.008	0.007	0.008	0.013	26.671
phecode_328	Dementias and cerebral degeneration	0.001	0.002	0.003	0.006	0.009	0.017	0.02	0.025	0.034	0.041	41.507
phecode_328-1	Alzheimer's disease	0	0	0.001	0.005	0.004	0.009	0.013	0.014	0.019	0.022	133.022
phecode_328-7	Vascular dementia	0	0	0	0.001	0.002	0.003	0.004	0.005	0.007	0.009	19.003
phecode_328-8	Dementia in conditions classified elsewhere	0	0.001	0.002	0.005	0.006	0.01	0.011	0.014	0.017	0.02	Inf
phecode_328-9	Dementia NOS	0.001	0.002	0.001	0.002	0.004	0.007	0.011	0.012	0.019	0.021	25.204
phecode_329	Symptoms and signs involving cognitive functions and awareness	0.025	0.031	0.036	0.046	0.054	0.065	0.078	0.088	0.107	0.129	5.14
phecode_329-1	Memory loss	0.012	0.014	0.019	0.023	0.024	0.035	0.035	0.044	0.053	0.057	4.888
phecode_329-4	Other specified cognitive deficit	0.004	0.007	0.005	0.006	0.007	0.004	0.008	0.006	0.007	0.005	1.348
phecode_329-41	Attention and concentration deficit	0.001	0.002	0.001	0.002	0.001	0.002	0.002	0.002	0.003	0.004	4.601
phecode_329-42	Cognitive communication deficit	0.003	0.003	0.004	0.004	0.004	0.005	0.004	0.004	0.005	0.003	1.111
phecode_329-5	Mild cognitive impairment, so stated	0	0	0.002	0.003	0.004	0.006	0.008	0.01	0.011	0.012	37.506
phecode_329-6	Transient global amnesia	0.001	0.003	0.001	0.003	0.003	0.002	0.005	0.004	0.003	0.003	4.001
phecode_329-9	Delirium	0.001	0.001	0.002	0.003	0.008	0.008	0.014	0.014	0.021	0.028	43.507
phecode_330	Epilepsy, recurrent seizures, convulsions	0.007	0.006	0.008	0.007	0.01	0.011	0.014	0.013	0.013	0.017	2.396
phecode_330-1	Epilepsy	0.004	0.005	0.005	0.005	0.006	0.006	0.009	0.008	0.008	0.011	2.616
phecode_330-11	Generalized epilepsy	0.001	0.002	0.002	0.001	0.002	0.002	0.002	0.002	0.003	0.004	2.556
phecode_330-12	Partial epilepsy	0.001	0.001	0.002	0.001	0.001	0.004	0.001	0.002	0.003	0.003	4.251
phecode_330-3	Convulsions	0.006	0.005	0.005	0.005	0.009	0.008	0.01	0.01	0.009	0.012	2
phecode_331	Headache	0.085	0.08	0.086	0.099	0.092	0.115	0.109	0.121	0.124	0.157	1.853
phecode_331-1	Tension headache	0.008	0.009	0.009	0.012	0.011	0.011	0.012	0.011	0.014	0.022	2.7

Supplementary Tables

Table 22 continued from previous page

Endpoint	PheCode string	1	2	3	4	5	6	7	8	9	10	Rate Ratio (RR)
phecode_331-3	Headache syndromes, non migraine	0.003	0.002	0.002	0.001	0.001	0.001	0.003	0.003	0.002	0.003	1.05
phecode_331-6	Migraine	0.015	0.019	0.021	0.021	0.022	0.026	0.026	0.032	0.042	0.053	3.518
phecode_331-61	Migraine with aura	0.004	0.006	0.004	0.004	0.005	0.006	0.006	0.008	0.008	0.008	2.228
phecode_331-8	Headache NOS	0.065	0.067	0.073	0.082	0.085	0.088	0.096	0.091	0.099	0.141	2.168
phecode_333	Sleep disorders	0.071	0.072	0.077	0.082	0.079	0.083	0.084	0.092	0.101	0.109	1.543
phecode_333-1	Sleep apnea	0.009	0.01	0.016	0.018	0.018	0.019	0.022	0.025	0.03	0.035	3.84
phecode_333-11	Obstructive sleep apnea	0.005	0.006	0.008	0.008	0.008	0.01	0.014	0.014	0.014	0.021	4.094
phecode_333-2	Insomnia	0.028	0.028	0.03	0.036	0.037	0.038	0.039	0.043	0.047	0.048	1.697
phecode_333-4	Circadian rhythm sleep disorder	0.001	0.002	0.001	0.002	0.002	0.002	0.002	0.001	0.002	0.003	2.111
phecode_334	Disorders of other cranial nerves	0.009	0.011	0.013	0.012	0.013	0.013	0.015	0.015	0.014	0.017	1.962
phecode_334-1	Trigeminal nerve disorders [CNS]	0.004	0.005	0.004	0.006	0.006	0.006	0.008	0.007	0.009	0.007	1.818
phecode_334-11	Trigeminal neuralgia	0.002	0.004	0.005	0.005	0.005	0.005	0.006	0.005	0.006	0.007	3
phecode_334-2	Facial nerve disorders and weakness	0.004	0.003	0.004	0.005	0.005	0.004	0.006	0.006	0.007	0.008	1.846
phecode_334-21	Bell's palsy	0.004	0.002	0.004	0.005	0.003	0.004	0.005	0.006	0.005	0.007	1.667
phecode_335	Nerve root and plexus disorders	0.014	0.014	0.012	0.014	0.016	0.016	0.017	0.021	0.019	0.024	1.71
phecode_336	Mononeuropathies	0.036	0.048	0.048	0.054	0.053	0.059	0.056	0.058	0.066	0.062	1.723
phecode_336-1	Carpal tunnel syndrome	0.021	0.027	0.03	0.036	0.034	0.039	0.037	0.041	0.039	0.044	2.141
phecode_336-2	Lesion of median, ulnar, radial nerve	0.006	0.005	0.007	0.006	0.006	0.007	0.007	0.007	0.007	0.012	2.089
phecode_336-5	Mononeuritis of lower limb	0.011	0.014	0.016	0.012	0.018	0.015	0.015	0.015	0.019	0.019	1.667
phecode_336-52	Meralgia paresthetica	0.006	0.005	0.003	0.005	0.004	0.003	0.003	0.003	0.004	0.005	0.857
phecode_336-55	Lesion of plantar nerve	0.004	0.007	0.007	0.009	0.009	0.01	0.009	0.013	0.014	0.014	3.819
phecode_337	Polyneuropathies	0.007	0.01	0.015	0.012	0.013	0.015	0.016	0.024	0.026	0.034	5.001
phecode_337-8	Polyneuropathy in diseases classified elsewhere	0	0	0.001	0.001	0.001	0.001	0.002	0.003	0.005	0.014	41.507
phecode_341	Cerebral palsy and other paralytic syndromes	0.003	0.003	0.004	0.005	0.007	0.008	0.007	0.01	0.014	0.018	6.589
phecode_341-2	Hemiplegia and hemiparesis	0.002	0.003	0.003	0.003	0.005	0.007	0.005	0.009	0.012	0.017	7.287
phecode_342	Plegia and unspecified paralysis	0.002	0.003	0.004	0.004	0.006	0.006	0.003	0.005	0.006	0.006	2.924
phecode_342-4	Monoplegia	0.002	0.002	0.002	0.003	0.004	0.003	0.005	0.003	0.005	0.004	1.917
phecode_343	Disorders of autonomic nervous system	0.002	0.001	0.003	0.003	0.003	0.004	0.004	0.005	0.004	0.006	2.534
phecode_344	Disorders of the circulation of the cerebrospinal fluid	0.001	0.001	0.002	0.002	0.002	0.003	0.003	0.003	0.004	0.004	3.126
phecode_344-1	Hydrocephalus	0	0	0.002	0.001	0.001	0.002	0.002	0.002	0.003	0.004	13.502
phecode_346	Brain damage and brain death	0.001	0.002	0.002	0.002	0.004	0.003	0.002	0.003	0.003	0.004	3.251
phecode_347	Other disorders of the brain and CNS	0.002	0.002	0.003	0.003	0.004	0.005	0.006	0.007	0.006	0.007	2.929
phecode_348	Other diseases of spinal cord	0.006	0.005	0.005	0.008	0.008	0.011	0.011	0.011	0.01	0.013	2.25
phecode_348-2	Myelopathies	0.004	0.004	0.005	0.006	0.006	0.008	0.01	0.007	0.008	0.009	2.08
phecode_349	Disorder of nervous system	0.005	0.007	0.006	0.007	0.012	0.011	0.014	0.014	0.013	0.016	3.38
phecode_349-1	Abnormal findings on diagnostic test of central nervous system	0.003	0.003	0.003	0.004	0.005	0.007	0.006	0.01	0.008	0.01	3.389
phecode_349-13	Abnormal findings on diagnostic imaging of skull and head	0.002	0.002	0.002	0.003	0.004	0.004	0.005	0.008	0.006	0.009	5.2
phecode_349-2	Abnormal results of function studies of peripheral nervous system	0.002	0.003	0.003	0.002	0.003	0.005	0.004	0.005	0.004	0.004	1.846
phecode_350	Other symptoms involving nervous system	0.031	0.037	0.04	0.043	0.055	0.066	0.067	0.078	0.088	0.117	3.797
phecode_350-5	Repeated falls*	0.004	0.008	0.01	0.017	0.022	0.033	0.035	0.044	0.061	0.081	18.299
phecode_351	Disturbances of skin sensation	0.057	0.06	0.072	0.075	0.075	0.078	0.085	0.091	0.106	0.118	2.074
phecode_351-1	Anesthesia of skin*	0.017	0.016	0.016	0.017	0.02	0.019	0.018	0.021	0.015	0.023	1.376
phecode_351-2	Hypoesthesia of skin*	0	0.002	0.001	0.004	0.004	0.006	0.008	0.011	0.011	0.023	70.011
phecode_351-3	Paresthesia of skin*	0.034	0.034	0.038	0.036	0.036	0.04	0.04	0.038	0.04	0.046	1.368
phecode_352	Disturbances of sensation of smell and taste	0.007	0.007	0.009	0.008	0.008	0.009	0.009	0.009	0.01	0.01	1.4
phecode_352-1	Anosmia*	0.002	0.004	0.003	0.004	0.004	0.002	0.006	0.003	0.002	0.006	2.534
phecode_352-3	Parageusia*	0.001	0.003	0.002	0.004	0.003	0.002	0.003	0.003	0.003	0.002	1.667
phecode_353	Symptoms and signs involving general sensations and perceptions	0.01	0.014	0.012	0.014	0.019	0.018	0.019	0.021	0.02	0.024	2.4
phecode_353-1	Hallucinations	0.001	0.002	0.001	0.003	0.004	0.005	0.005	0.005	0.006	0.006	4.626
phecode_354	Dizziness and giddiness	0.067	0.088	0.092	0.105	0.116	0.114	0.122	0.133	0.131	0.164	2.443
phecode_355	Coma and other alteration of consciousness	0.005	0.008	0.007	0.008	0.008	0.011	0.013	0.013	0.015	0.022	4.786
phecode_355-1	Coma	0.003	0.003	0.003	0.003	0.002	0.003	0.004	0.003	0.004	0.005	1.65
phecode_355-2	Alteration of consciousness	0.003	0.003	0.005	0.004	0.005	0.009	0.01	0.009	0.011	0.017	5.317

5 Phenome-wide prediction of disease onset from retinal fundus photographs

Table 22 continued from previous page

Endpoint	PheCode string	1	2	3	4	5	6	7	8	9	10	Rate Ratio (RR)
phecode_356	Speech disturbance	0.003	0.004	0.005	0.007	0.01	0.012	0.01	0.011	0.015	0.016	5.707
phecode_356-1	Dysarthria	0	0	0	0.001	0.002	0.002	0.002	0.003	0.003	0.004	13.002
phecode_356-2	Aphasia and dysphasia	0.001	0.001	0.003	0.001	0.005	0.004	0.005	0.007	0.006	0.007	10.002
phecode_360	Inflammation of eyelids	0.051	0.053	0.056	0.056	0.061	0.064	0.067	0.072	0.082	0.077	1.507
phecode_360-1	Hordeolum	0.02	0.022	0.019	0.026	0.027	0.023	0.022	0.026	0.023	0.026	1.252
phecode_360-11	Hordeolum externum	0.021	0.021	0.021	0.025	0.025	0.023	0.021	0.027	0.023	0.026	1.238
phecode_360-12	Hordeolum internum	0.004	0.002	0.003	0.002	0.006	0.004	0.004	0.004	0.005	0.005	1.25
phecode_360-2	Chalazion	0.014	0.014	0.018	0.014	0.016	0.015	0.017	0.018	0.015	0.017	1.244
phecode_360-4	Blepharitis	0.02	0.02	0.027	0.029	0.036	0.033	0.035	0.046	0.042	0.045	2.213
phecode_360-5	Noninfectious dermatoses of eyelid	0.001	0.001	0.002	0.002	0.001	0.002	0.003	0.003	0.004	0.005	3.626
phecode_360-51	Eczematous dermatitis of eyelid	0.001	0.001	0.001	0.002	0.002	0.001	0.002	0.003	0.003	0.004	3.376
phecode_361	Disorders of eyelid function	0.005	0.005	0.006	0.006	0.01	0.011	0.014	0.016	0.019	0.024	4.901
phecode_361-1	Entropion and trichiasis of eyelid	0	0.001	0	0.001	0.002	0.004	0.004	0.003	0.004	0.006	12.002
phecode_361-3	Ptosis of eyelid	0.004	0.003	0.003	0.004	0.005	0.006	0.006	0.008	0.01	0.015	3.462
phecode_361-4	Blepharochalasis	0.001	0.001	0	0.001	0.001	0.002	0.003	0.003	0.004	0.003	5.001
phecode_361-9	Ectropion of eyelid	0.001	0	0.001	0.001	0.001	0.002	0.002	0.003	0.004	0.005	4.286
phecode_362	Other disorders of the eyelids	0.01	0.013	0.012	0.013	0.017	0.013	0.015	0.017	0.016	0.016	1.69
phecode_362-5	Cysts of eyelid	0.002	0.002	0.002	0.002	0.004	0.002	0.003	0.002	0.005	0.004	1.572
phecode_363	Disorders of lacrimal system	0.021	0.031	0.032	0.043	0.055	0.054	0.069	0.073	0.079	0.092	4.366
phecode_363-2	Dry eye syndrome [Tear film insufficiency]	0.016	0.025	0.026	0.035	0.045	0.045	0.058	0.062	0.062	0.078	4.948
phecode_363-5	Epiphora	0.005	0.005	0.006	0.009	0.008	0.01	0.011	0.011	0.014	0.011	2.125
phecode_363-7	Stenosis and insufficiency of lacrimal passages	0.001	0.002	0.003	0.001	0.005	0.003	0.004	0.005	0.004	0.004	3
phecode_366	Noninflammatory disorders of conjunctiva	0.005	0.005	0.006	0.006	0.007	0.009	0.009	0.009	0.009	0.01	2.067
phecode_366-1	Pterygium of eye	0.001	0.001	0.001	0.002	0.003	0.001	0.003	0.001	0.001	0.004	3.126
phecode_366-4	Vascular abnormalities of conjunctiva	0.003	0.002	0.003	0.003	0.004	0.005	0.004	0.003	0.005	0.004	1.5
phecode_366-42	Conjunctival hyperemia	0.002	0.001	0.002	0.002	0.003	0.002	0.003	0.001	0.003	0.003	1.231
phecode_367	Inflammation of the eye	0.049	0.061	0.062	0.064	0.07	0.072	0.076	0.09	0.096	0.098	2
phecode_367-1	Conjunctivitis	0.04	0.05	0.052	0.058	0.06	0.062	0.063	0.073	0.074	0.077	1.951
phecode_367-12	Allergic [atopic] conjunctivitis	0.004	0.006	0.008	0.008	0.007	0.008	0.009	0.008	0.009	0.013	3.637
phecode_367-13	Blepharoconjunctivitis	0.001	0.002	0.003	0.003	0.003	0.004	0.005	0.005	0.006	0.005	6.601
phecode_367-2	Keratitis	0.005	0.005	0.005	0.004	0.006	0.004	0.005	0.005	0.006	0.006	1.214
phecode_367-21	Corneal ulcer	0.002	0.001	0.001	0.002	0.002	0.001	0.002	0.002	0.002	0.001	0.9
phecode_367-5	Uveitis	0.005	0.004	0.004	0.003	0.004	0.005	0.006	0.006	0.009	0.013	2.634
phecode_367-52	Iridocyclitis	0.005	0.004	0.004	0.004	0.004	0.005	0.006	0.006	0.01	0.013	2.534
phecode_367-6	Episcleritis	0.003	0.001	0.002	0.002	0.002	0.002	0.003	0.004	0.004	0.004	1.53
phecode_369	Noninflammatory disorders of the cornea	0.002	0.003	0.005	0.006	0.008	0.008	0.011	0.014	0.016	0.032	13.069
phecode_369-1	Corneal scars and opacities	0.001	0.001	0.001	0.001	0.001	0.001	0.002	0.002	0.002	0.006	8.501
phecode_369-4	Corneal degenerations	0	0.001	0.003	0.002	0.004	0.004	0.005	0.007	0.009	0.012	23.667
phecode_369-5	Hereditary corneal dystrophies	0	0.001	0.001	0	0.003	0.001	0.002	0.003	0.004	0.012	38.006
phecode_370	Disorders of iris and ciliary body	0.001	0.001	0.001	0.001	0.002	0.003	0.004	0.004	0.005	0.015	18.203
phecode_371	Cataract	0.006	0.014	0.025	0.044	0.071	0.094	0.126	0.167	0.228	0.363	58.604
phecode_371-3	Nuclear cataract	0.001	0.002	0.004	0.009	0.018	0.02	0.036	0.049	0.07	0.121	185.03
phecode_371-31	Age-related nuclear cataract	0.001	0.002	0.004	0.008	0.018	0.02	0.036	0.048	0.071	0.122	186.03
phecode_373	Noninflammatory disorders of choroid	0.001	0.001	0.001	0.001	0.001	0.002	0.002	0.002	0.004	0.015	13.431
phecode_374	Disorders of the retina	0.014	0.029	0.04	0.051	0.062	0.072	0.087	0.098	0.122	0.182	13.274
phecode_374-1	Retinal detachments and breaks	0.004	0.004	0.006	0.008	0.011	0.01	0.013	0.012	0.018	0.024	6.683
phecode_374-11	Serous retinal detachment	0.001	0.003	0.002	0.004	0.005	0.004	0.006	0.006	0.009	0.014	16.603
phecode_374-3	Retinal vascular changes and occlusions	0.002	0.003	0.005	0.008	0.008	0.015	0.013	0.013	0.019	0.035	19.094
phecode_374-38	Retinal vein occlusions	0	0.001	0.003	0.003	0.004	0.005	0.007	0.007	0.009	0.012	25.337
phecode_374-39	Transient retinal arterial occlusion [Amaurosis fugax]	0	0	0.001	0.001	0.002	0.003	0.003	0.003	0.004	0.004	8.001
phecode_374-4	Retinal disorders in diseases classified elsewhere	0.005	0.008	0.016	0.02	0.023	0.03	0.029	0.037	0.051	0.084	18.11
phecode_374-42	Diabetic retinopathy	0.004	0.008	0.016	0.02	0.025	0.028	0.028	0.037	0.048	0.076	17.111
phecode_374-5	Macular degeneration	0.002	0.005	0.008	0.015	0.018	0.022	0.028	0.037	0.053	0.11	51.316
phecode_374-51	Age-related macular degeneration	0	0.001	0.001	0.003	0.006	0.006	0.01	0.013	0.016	0.057	116.686
phecode_374-511	Nonexudative (dry) age-related macular degeneration	0	0	0.001	0.001	0.002	0.003	0.003	0.005	0.006	0.019	115.019

Supplementary Tables

Table 22 continued from previous page

Endpoint	PheCode string	1	2	3	4	5	6	7	8	9	10	Rate Ratio (RR)
phecode_374-512	Exudative (wet) age-related macular degeneration	0	0	0	0	0.001	0.001	0.002	0.002	0.003	0.014	86.014
phecode_374-52	Macular cyst, hole, or pseudohole	0	0.001	0.002	0.003	0.004	0.005	0.005	0.005	0.007	0.008	15.336
phecode_374-55	Puckering of macula	0	0.002	0.002	0.005	0.007	0.005	0.008	0.01	0.012	0.017	35.339
phecode_374-8	Retinal edema	0	0.001	0.002	0.002	0.002	0.003	0.004	0.005	0.007	0.017	35.006
phecode_375	Abnormal intraocular pressure	0.013	0.016	0.028	0.03	0.037	0.037	0.048	0.055	0.06	0.086	6.655
phecode_375-1	Glaucoma	0.005	0.01	0.015	0.016	0.019	0.03	0.031	0.039	0.041	0.067	12.721
phecode_375-11	Open angle glaucoma	0.001	0.004	0.006	0.007	0.009	0.012	0.012	0.017	0.018	0.038	38.673
phecode_375-113	Primary open angle glaucoma	0	0.002	0.004	0.002	0.005	0.005	0.007	0.006	0.008	0.015	46.008
phecode_375-12	Angle-Closure Glaucoma	0.001	0	0.002	0.002	0.004	0.005	0.004	0.006	0.005	0.008	6.573
phecode_375-14	Low-tension glaucoma (Normal-tension glaucoma)	0	0.001	0.001	0.001	0.002	0.003	0.002	0.002	0.003	0.005	Inf
phecode_375-6	Ocular hypertension	0.005	0.006	0.011	0.01	0.012	0.014	0.013	0.016	0.017	0.02	4.139
phecode_376	Disorders of vitreous body	0.007	0.016	0.032	0.043	0.043	0.047	0.051	0.057	0.059	0.063	9.745
phecode_376-1	Vitreous degeneration	0.004	0.014	0.023	0.031	0.033	0.034	0.037	0.044	0.037	0.049	12.294
phecode_376-2	Vitreous opacities	0.006	0.016	0.032	0.042	0.042	0.05	0.05	0.054	0.059	0.064	10.38
phecode_376-21	Crystalline deposits in vitreous body	0.003	0.006	0.008	0.01	0.011	0.012	0.012	0.014	0.014	0.017	5.723
phecode_377	Hemorrhage of the eye	0.014	0.016	0.022	0.025	0.028	0.031	0.028	0.033	0.034	0.039	2.796
phecode_377-2	Conjunctival hemorrhage	0.013	0.015	0.018	0.023	0.023	0.025	0.026	0.023	0.027	0.026	2.039
phecode_377-4	Retinal hemorrhage	0.001	0.001	0.001	0.001	0.003	0.003	0.004	0.004	0.003	0.007	9.001
phecode_377-5	Vitreous hemorrhage	0.001	0.001	0.001	0.002	0.002	0.002	0.002	0.003	0.002	0.009	13.752
phecode_379	Infection of the eye	0.004	0.004	0.004	0.002	0.006	0.005	0.006	0.007	0.007	0.006	1.308
phecode_379-2	Eye infection, viral	0.004	0.005	0.003	0.003	0.005	0.005	0.005	0.005	0.007	0.005	1.218
phecode_379-21	Infection of the eye, herpes	0.001	0.001	0.001	0.002	0.002	0.003	0.003	0.002	0.004	0.004	2.556
phecode_380	Disorders of optic nerve and visual pathways	0.003	0.002	0.003	0.003	0.004	0.003	0.004	0.007	0.008	0.013	3.851
phecode_380-2	Disorders of optic disc	0.002	0.002	0.003	0.001	0.002	0.002	0.002	0.005	0.006	0.008	4.601
phecode_381	Strabismus	0.002	0.002	0.002	0.004	0.005	0.005	0.006	0.007	0.008	0.009	3.715
phecode_381-1	Paralytic strabismus [Neurogenic strabismus]	0.001	0	0.001	0.002	0.003	0.002	0.002	0.003	0.004	0.003	2.375
phecode_383	Irregular eye movements	0.002	0	0.001	0.002	0.001	0.002	0.001	0.003	0.002	0.003	1.8
phecode_384	Anomalies of pupillary function	0.001	0.001	0.001	0.002	0.001	0.002	0.001	0.002	0.002	0.004	3.667
phecode_385	Abnormal results of function studies of eye	0.001	0.002	0.002	0.002	0.002	0.003	0.003	0.003	0.004	0.008	10.402
phecode_386	Visual disturbances	0.017	0.024	0.028	0.032	0.032	0.032	0.041	0.044	0.052	0.067	3.855
phecode_386-1	Amblyopia	0.001	0.001	0.003	0.003	0.003	0.003	0.005	0.006	0.01	0.014	14.5
phecode_386-2	Diplopia	0.004	0.005	0.004	0.005	0.006	0.006	0.006	0.008	0.008	0.009	2.292
phecode_386-4	Visual field defects	0.002	0.002	0.003	0.004	0.004	0.007	0.004	0.007	0.007	0.011	5.667
phecode_386-9	Visual distortions and subjective visual disturbances	0.001	0.002	0.001	0.002	0.003	0.004	0.003	0.002	0.003	0.004	4.167
phecode_387	Disorders of refraction and accommodation	0.003	0.004	0.004	0.013	0.013	0.018	0.027	0.03	0.039	0.073	25.945
phecode_387-1	Hypermetropia	0	0	0	0	0.001	0.001	0.002	0.003	0.005	0.008	23.5
phecode_387-2	Myopia	0.001	0.002	0.002	0.003	0.004	0.005	0.005	0.01	0.014	0.046	47.174
phecode_387-3	Astigmatism	0.002	0.001	0.003	0.005	0.009	0.009	0.013	0.018	0.022	0.03	18.403
phecode_388	Blindness and low vision	0.003	0.004	0.006	0.007	0.009	0.013	0.016	0.02	0.028	0.039	14.877
phecode_389	Other disorders of eye	0.052	0.06	0.063	0.074	0.076	0.075	0.084	0.09	0.097	0.114	2.178
phecode_389-1	Ocular pain	0.006	0.005	0.006	0.007	0.006	0.007	0.005	0.007	0.008	0.006	0.895
phecode_390	Disorders of external ear	0.116	0.135	0.15	0.167	0.183	0.193	0.193	0.187	0.212	0.21	1.815
phecode_390-1	Otitis externa	0.058	0.065	0.067	0.071	0.074	0.082	0.075	0.078	0.078	0.091	1.55
phecode_390-4	Impacted cerumen	0.066	0.087	0.097	0.107	0.135	0.145	0.136	0.152	0.15	0.159	2.434
phecode_390-6	Perichondritis and chondritis of pinna	0	0.002	0.001	0.002	0.006	0.004	0.005	0.005	0.007	0.008	15.667
phecode_391	Disorders of the middle ear	0.059	0.071	0.075	0.075	0.074	0.075	0.074	0.076	0.08	0.082	1.374
phecode_391-1	Otitis media	0.027	0.039	0.035	0.038	0.042	0.033	0.04	0.042	0.043	0.043	1.586
phecode_391-11	Acute otitis media	0.002	0.003	0.003	0.003	0.003	0.004	0.003	0.003	0.004	0.003	1.417
phecode_391-12	Chronic otitis media	0.005	0.004	0.006	0.005	0.005	0.006	0.007	0.004	0.005	0.004	0.788
phecode_391-2	Eustachian tube disorders	0.025	0.029	0.028	0.027	0.03	0.032	0.033	0.037	0.037	0.038	1.49
phecode_391-21	Eustachian salpingitis	0.003	0.003	0.003	0.003	0.004	0.004	0.003	0.004	0.002	0.002	0.688
phecode_391-7	Perforation of tympanic membrane	0.003	0.004	0.005	0.004	0.006	0.006	0.006	0.006	0.006	0.005	1.45
phecode_391-9	Otorrhea	0.003	0.004	0.004	0.006	0.005	0.005	0.006	0.005	0.007	0.006	2.25
phecode_392	Otalgia and effusion of ear	0.039	0.042	0.039	0.04	0.046	0.05	0.056	0.054	0.053	0.064	1.632
phecode_394	Disorders of vestibular function	0.017	0.024	0.027	0.03	0.032	0.031	0.033	0.036	0.041	0.049	2.883
phecode_394-1	Meniere disease	0.002	0.002	0.004	0.003	0.003	0.004	0.005	0.004	0.005	0.004	1.917

5 Phenome-wide prediction of disease onset from retinal fundus photographs

Table 22 continued from previous page

Endpoint	PheCode string	1	2	3	4	5	6	7	8	9	10	Rate Ratio (RR)
phecode_394-2	Vertigo	0.014	0.021	0.023	0.026	0.023	0.025	0.023	0.035	0.029	0.043	2.978
phecode_394-21	Paroxysmal vertigo	0.011	0.018	0.019	0.022	0.019	0.02	0.022	0.028	0.026	0.037	3.279
phecode_394-22	Vestibular neuronitis	0.005	0.003	0.004	0.004	0.004	0.004	0.005	0.005	0.005	0.005	1.034
phecode_394-4	Abnormal vestibular function study	0.001	0.001	0.002	0.002	0.001	0.002	0.003	0.002	0.002	0.002	3.751
phecode_395	Other diseases of inner ear	0.02	0.021	0.025	0.028	0.028	0.03	0.027	0.029	0.03	0.034	1.661
phecode_395-1	Labyrinthitis	0.019	0.021	0.024	0.025	0.023	0.031	0.024	0.027	0.029	0.031	1.676
phecode_396	Hearing impairment	0.044	0.057	0.067	0.086	0.105	0.118	0.128	0.14	0.157	0.169	3.88
phecode_396-1	Conductive hearing loss	0.002	0.002	0.002	0.002	0.002	0.002	0.003	0.003	0.003	0.004	1.917
phecode_396-2	Sensorineural hearing loss	0.009	0.008	0.012	0.017	0.02	0.023	0.024	0.025	0.029	0.033	3.457
phecode_396-21	Sensorineural hearing loss, bilateral	0.002	0.003	0.005	0.005	0.009	0.008	0.007	0.008	0.01	0.012	4.734
phecode_396-22	Presbycusis	0.001	0.001	0.002	0.004	0.005	0.005	0.007	0.008	0.009	0.01	10.335
phecode_397	Other hearing abnormality	0.036	0.037	0.037	0.042	0.039	0.043	0.046	0.04	0.038	0.042	1.168
phecode_397-1	Tinnitus	0.035	0.038	0.036	0.04	0.039	0.043	0.042	0.041	0.039	0.042	1.187
phecode_398	Other disorders of ear	0.011	0.013	0.011	0.015	0.017	0.015	0.02	0.018	0.021	0.021	1.939
phecode_398-1	Abnormal auditory function study	0.002	0.003	0.003	0.002	0.004	0.005	0.006	0.004	0.005	0.004	2.455
phecode_400	Rheumatic heart disease	0.003	0.004	0.007	0.01	0.01	0.017	0.023	0.024	0.032	0.038	12.213
phecode_400-2	Chronic rheumatic heart diseases	0.003	0.004	0.006	0.01	0.01	0.017	0.022	0.025	0.031	0.038	11.652
phecode_401	Hypertension	0.087	0.125	0.148	0.188	0.202	0.241	0.276	0.309	0.334	0.38	4.381
phecode_401-1	Essential hypertension	0.086	0.124	0.148	0.188	0.202	0.242	0.274	0.309	0.334	0.378	4.372
phecode_401-3	Hypertensive chronic kidney disease	0	0.001	0	0.001	0.001	0.001	0.003	0.002	0.003	0.007	13.669
phecode_401-6	Secondary hypertension	0.004	0.002	0.005	0.006	0.008	0.009	0.011	0.01	0.01	0.009	2.546
phecode_402	Elevated blood pressure reading without diagnosis of hypertension	0.032	0.04	0.051	0.051	0.058	0.065	0.06	0.063	0.058	0.063	1.962
phecode_403	Angina pectoris	0.007	0.019	0.024	0.03	0.03	0.042	0.053	0.051	0.062	0.086	12.652
phecode_404	Coronary heart disease	0.011	0.028	0.039	0.047	0.056	0.064	0.083	0.098	0.112	0.151	13.241
phecode_404-1	Myocardial infarction	0.004	0.009	0.016	0.019	0.022	0.028	0.032	0.046	0.051	0.072	17.323
phecode_404-11	Acute myocardial infarction	0.003	0.007	0.012	0.017	0.018	0.021	0.022	0.032	0.033	0.045	15.114
phecode_404-2	Coronary atherosclerosis [Atherosclerotic heart disease]	0.007	0.019	0.028	0.035	0.039	0.05	0.061	0.074	0.085	0.115	15.5
phecode_406	Chronic pulmonary heart disease	0	0.001	0.003	0.002	0.002	0.005	0.005	0.01	0.009	0.015	90.015
phecode_406-1	Pulmonary hypertension	0	0.001	0.003	0.001	0.002	0.005	0.005	0.009	0.008	0.014	85.014
phecode_410	Inflammation of the heart [Carditis]	0.006	0.003	0.006	0.009	0.006	0.007	0.011	0.011	0.009	0.013	2.383
phecode_410-2	Endocarditis	0.003	0.002	0.004	0.007	0.002	0.007	0.009	0.008	0.01	0.01	3.53
phecode_411	Other diseases of pericardium	0.002	0.002	0.004	0.004	0.006	0.005	0.007	0.008	0.008	0.013	5.334
phecode_411-2	Pericardial effusion (noninflammatory)*	0.001	0.002	0.002	0.003	0.004	0.003	0.005	0.005	0.007	0.009	11.402
phecode_413	Heart valve disorders	0.009	0.011	0.016	0.021	0.031	0.039	0.047	0.055	0.07	0.093	9.844
phecode_413-1	Mitral valve disorders	0.007	0.006	0.01	0.014	0.019	0.024	0.03	0.034	0.045	0.053	7.806
phecode_413-11	Mitral insufficiency	0.005	0.005	0.005	0.009	0.012	0.014	0.017	0.023	0.025	0.032	7.001
phecode_413-12	Mitral valve prolapse*	0	0	0.001	0.002	0.002	0.003	0.002	0.002	0.002	0.003	6.334
phecode_413-2	Aortic valve disorders	0.004	0.005	0.007	0.01	0.013	0.021	0.023	0.032	0.038	0.051	13.567
phecode_413-21	Aortic stenosis	0.001	0.002	0.002	0.004	0.004	0.009	0.013	0.016	0.021	0.028	21.504
phecode_413-22	Aortic insufficiency	0.002	0.003	0.004	0.005	0.007	0.007	0.01	0.012	0.011	0.016	8.729
phecode_413-3	Tricuspid valve disorders	0.002	0.002	0.005	0.008	0.008	0.012	0.016	0.022	0.026	0.03	16.912
phecode_413-32	Tricuspid valve insufficiency*	0	0	0.001	0.002	0.004	0.003	0.003	0.005	0.007	0.007	43.007
phecode_413-4	Pulmonary valve disorders	0.001	0.001	0	0.001	0.001	0.003	0.003	0.003	0.004	0.003	2
phecode_413-42	Pulmonary valve insufficiency*	0.001	0	0	0.001	0.001	0.003	0.002	0.003	0.004	0.002	1.667
phecode_413-6	Heart valve replaced	0.001	0.001	0.002	0.003	0.005	0.004	0.006	0.008	0.009	0.012	10.573
phecode_414	Cardiomyopathy	0.001	0.003	0.006	0.003	0.005	0.007	0.008	0.008	0.011	0.015	11.502
phecode_414-2	Dilated cardiomyopathy*	0.001	0	0.002	0.001	0.002	0.002	0.003	0.003	0.003	0.005	4.001
phecode_416	Cardiac arrhythmia and conduction disorders	0.026	0.036	0.049	0.062	0.079	0.106	0.13	0.139	0.169	0.211	8.119
phecode_416-1	Paroxysmal tachycardia	0.004	0.006	0.01	0.011	0.012	0.015	0.015	0.016	0.018	0.021	5.251
phecode_416-11	Supraventricular tachycardia	0.004	0.002	0.008	0.008	0.009	0.012	0.011	0.012	0.015	0.014	3.458
phecode_416-12	Ventricular tachycardia	0.001	0.002	0.003	0.002	0.004	0.003	0.004	0.005	0.007	0.008	8.501
phecode_416-2	Atrial fibrillation and flutter	0.008	0.015	0.024	0.032	0.043	0.063	0.08	0.089	0.115	0.145	18.128
phecode_416-21	Atrial fibrillation	0.004	0.011	0.018	0.023	0.034	0.046	0.055	0.069	0.084	0.104	23.3
phecode_416-211	Paroxysmal atrial fibrillation*	0.004	0.005	0.008	0.013	0.018	0.023	0.025	0.027	0.03	0.038	10.502
phecode_416-212	Persistent atrial fibrillation*	0	0.001	0.001	0.002	0.002	0.003	0.005	0.003	0.008	0.007	14.002
phecode_416-22	Atrial flutter	0	0.002	0.003	0.005	0.006	0.009	0.01	0.009	0.016	0.016	95.016
phecode_416-4	Heart block	0.007	0.009	0.014	0.021	0.024	0.03	0.047	0.049	0.057	0.076	10.746
phecode_416-41	Atrioventricular block	0.002	0.005	0.004	0.007	0.012	0.015	0.017	0.023	0.024	0.036	16.926
phecode_416-42	Left bundle branch block	0.002	0.002	0.004	0.006	0.005	0.012	0.014	0.015	0.019	0.023	9.93

Supplementary Tables

Table 22 continued from previous page

Endpoint	PheCode string	1	2	3	4	5	6	7	8	9	10	Rate Ratio (RR)
phecode_416-43	Right bundle branch block	0.003	0.004	0.006	0.009	0.01	0.012	0.017	0.013	0.023	0.023	6.951
phecode_416-5	Premature depolarization [Premature beats]	0.006	0.009	0.008	0.011	0.013	0.017	0.019	0.017	0.018	0.023	3.811
phecode_416-51	Atrial premature depolarization [Supraventricular premature beats]	0.001	0.001	0.001	0.002	0.003	0.003	0.006	0.004	0.005	0.004	3.143
phecode_416-52	Ventricular premature depolarization*	0.004	0.005	0.006	0.007	0.007	0.011	0.011	0.009	0.014	0.015	3.87
phecode_416-7	Sinoatrial node dysfunction	0.001	0	0	0.001	0.002	0.003	0.002	0.003	0.003	0.004	3.858
phecode_416-71	Sick sinus syndrome*	0.001	0	0	0.001	0.002	0.003	0.002	0.003	0.003	0.004	3.858
phecode_417	Abnormalities of heart beat	0.077	0.085	0.091	0.101	0.116	0.122	0.14	0.142	0.155	0.176	2.296
phecode_417-1	Palpitations	0.037	0.041	0.047	0.053	0.055	0.064	0.058	0.068	0.069	0.076	2.042
phecode_417-2	Tachycardia	0.008	0.009	0.013	0.01	0.013	0.02	0.019	0.018	0.02	0.029	3.481
phecode_417-3	Bradycardia*	0.013	0.012	0.017	0.018	0.023	0.024	0.025	0.03	0.036	0.04	3.013
phecode_418	Abnormal results of cardiovascular function studies	0.014	0.019	0.024	0.026	0.031	0.036	0.045	0.047	0.056	0.076	5.471
phecode_418-1	Abnormal electrocardiogram [ECG] [EKG]	0.012	0.012	0.017	0.019	0.023	0.031	0.031	0.033	0.042	0.053	4.494
phecode_419	Presence of cardiac device	0.002	0.002	0.002	0.006	0.008	0.011	0.014	0.019	0.02	0.027	13.502
phecode_420	Cardiac arrest	0.001	0.002	0.002	0.004	0.002	0.006	0.006	0.008	0.01	0.015	23.504
phecode_423	Abnormal cardiac sounds	0.004	0.006	0.006	0.008	0.01	0.013	0.014	0.014	0.02	0.022	5.039
phecode_423-1	Cardiac murmurs	0.004	0.004	0.005	0.006	0.009	0.01	0.013	0.013	0.016	0.018	4.27
phecode_424	Heart failure	0.004	0.008	0.009	0.016	0.02	0.031	0.04	0.045	0.058	0.098	24.754
phecode_424-1	Left heart failure	0.002	0.004	0.006	0.008	0.009	0.013	0.018	0.022	0.022	0.042	23.277
phecode_424-2	Systolic heart failure	0.001	0.002	0.002	0.003	0.002	0.005	0.006	0.008	0.009	0.015	23.254
phecode_424-3	Diastolic heart failure	0	0	0.001	0.001	0.002	0.003	0.003	0.004	0.004	0.006	12.002
phecode_425	Cardiomegaly	0.004	0.006	0.012	0.013	0.016	0.021	0.028	0.032	0.039	0.047	12.437
phecode_426	Other heart disorders in diseases NEC	0.004	0.005	0.009	0.01	0.014	0.018	0.023	0.029	0.031	0.047	13.093
phecode_430	Nontraumatic Intracranial hemorrhage	0.002	0.003	0.004	0.006	0.007	0.006	0.006	0.013	0.011	0.015	9.101
phecode_430-1	Nontraumatic subarachnoid hemorrhage	0.001	0.001	0.001	0.003	0.003	0.003	0.003	0.004	0.003	0.004	5.001
phecode_430-2	Nontraumatic intracerebral hemorrhage	0.001	0.001	0.001	0.003	0.003	0.003	0.003	0.006	0.007	0.011	16.253
phecode_430-3	Nontraumatic subdural hemorrhage	0.001	0.001	0.001	0.002	0.002	0.001	0.002	0.004	0.003	0.003	4.251
phecode_431	Stroke and transient cerebral ischemic attacks	0.008	0.013	0.017	0.022	0.032	0.033	0.041	0.054	0.064	0.08	10.106
phecode_431-1	Stroke	0.005	0.008	0.008	0.013	0.019	0.022	0.024	0.037	0.041	0.053	11.105
phecode_431-11	Ischemic stroke	0.004	0.006	0.006	0.01	0.015	0.019	0.021	0.03	0.035	0.043	10.156
phecode_431-12	Hemorrhagic stroke	0.001	0.003	0.002	0.005	0.005	0.004	0.004	0.009	0.009	0.012	9.126
phecode_431-2	Transient cerebral ischemic attacks	0.003	0.008	0.011	0.011	0.015	0.022	0.023	0.027	0.03	0.035	11.212
phecode_433	Other cerebrovascular disease	0.004	0.006	0.009	0.011	0.017	0.026	0.032	0.038	0.048	0.062	14.425
phecode_433-1	Occlusion and stenosis of cerebral arteries	0.001	0.002	0.001	0.002	0.003	0.005	0.005	0.008	0.008	0.012	17.753
phecode_433-2	Occlusion and stenosis of precerebral arteries	0.001	0	0.001	0.002	0.002	0.005	0.007	0.005	0.008	0.012	18.753
phecode_433-21	Carotid artery stenosis	0.001	0	0.001	0.002	0.003	0.004	0.006	0.005	0.008	0.011	17.503
phecode_436	Atherosclerosis [ASCVD]	0.007	0.02	0.027	0.034	0.04	0.05	0.063	0.07	0.083	0.114	16.193
phecode_437	Vascular insufficiency of intestine	0.001	0.001	0.001	0.002	0.002	0.004	0.004	0.004	0.005	0.005	6.601
phecode_438	Aneurysm or ectasia	0.003	0.006	0.009	0.009	0.014	0.015	0.014	0.021	0.022	0.033	11.168
phecode_438-1	Aortic aneurysm and ectasia	0	0.002	0.003	0.005	0.01	0.011	0.011	0.012	0.017	0.026	54.009
phecode_438-11	Abdominal aortic aneurysm	0	0.001	0.001	0.004	0.006	0.007	0.006	0.007	0.012	0.019	59.01
phecode_438-12	Thoracic aneurysm	0	0.001	0.001	0.003	0.002	0.005	0.004	0.004	0.004	0.004	13.5
phecode_439	Hemorrhoids	0.087	0.091	0.084	0.088	0.09	0.092	0.092	0.094	0.093	0.089	1.031
phecode_440	Embolism and thrombosis	0.026	0.031	0.038	0.039	0.05	0.055	0.064	0.069	0.067	0.084	3.25
phecode_440-1	Venous thromboembolism	0.023	0.023	0.027	0.032	0.031	0.036	0.044	0.039	0.041	0.049	2.124
phecode_440-11	Deep vein thrombosis [DVT]	0.005	0.006	0.008	0.009	0.012	0.013	0.011	0.013	0.014	0.017	3.367
phecode_440-13	Phlebitis and thrombophlebitis	0.014	0.017	0.02	0.023	0.025	0.033	0.035	0.032	0.035	0.04	2.88
phecode_440-2	Arterial embolism and thrombosis	0.002	0.002	0.003	0.003	0.008	0.005	0.009	0.01	0.013	0.016	7.694
phecode_440-3	Pulmonary embolism	0.006	0.009	0.009	0.013	0.016	0.018	0.018	0.021	0.025	0.028	4.474
phecode_443	Other specified disorders of arteries and arterioles	0.002	0.005	0.006	0.009	0.009	0.014	0.016	0.018	0.025	0.027	10.735
phecode_443-1	Stricture of artery [Arterial stenosis]	0	0	0.001	0.001	0.001	0.002	0.003	0.004	0.003	0.008	48
phecode_444	Venous insufficiency	0.025	0.029	0.03	0.042	0.043	0.052	0.052	0.058	0.057	0.064	2.545
phecode_444-1	Varicose veins	0.027	0.027	0.03	0.04	0.044	0.046	0.054	0.051	0.057	0.059	2.237
phecode_444-11	Varicose veins of lower extremities	0.021	0.022	0.029	0.035	0.042	0.041	0.047	0.047	0.05	0.056	2.698
phecode_444-15	Scrotal varices [Varicocele]	0	0.001	0.001	0.001	0.002	0.003	0.005	0.004	0.007	0.008	16.003

5 Phenome-wide prediction of disease onset from retinal fundus photographs

Table 22 continued from previous page

Endpoint	PheCode string	1	2	3	4	5	6	7	8	9	10	Rate Ratio (RR)
phecode_444-5	Venous insufficiency (chronic) (peripheral)	0.004	0.005	0.008	0.012	0.016	0.015	0.022	0.024	0.026	0.03	7.914
phecode_446	Hypotension	0.013	0.012	0.017	0.024	0.037	0.041	0.052	0.052	0.064	0.085	6.371
phecode_446-2	Orthostatic hypotension	0.004	0.004	0.006	0.01	0.015	0.018	0.024	0.027	0.028	0.038	8.668
phecode_447	Nonspecific low blood-pressure reading	0.002	0.003	0.004	0.005	0.006	0.006	0.009	0.008	0.012	0.015	7.232
phecode_448	Peripheral vascular disease	0.009	0.01	0.015	0.015	0.018	0.021	0.021	0.027	0.029	0.04	4.464
phecode_448-1	Raynaud's syndrome	0.007	0.007	0.008	0.01	0.012	0.012	0.01	0.013	0.013	0.014	2.075
phecode_448-9	Peripheral vascular disease NOS [includes PAD]	0.001	0.002	0.002	0.005	0.007	0.009	0.011	0.014	0.019	0.031	37.206
phecode_449	Other disorders of the circulatory system	0.001	0.001	0.002	0.002	0.001	0.002	0.002	0.005	0.005	0.007	6.668
phecode_452	Hemorrhage, NOS	0.002	0.003	0.004	0.004	0.003	0.004	0.005	0.005	0.007	0.006	3
phecode_460	Acute respiratory infection	0.254	0.281	0.294	0.313	0.323	0.306	0.337	0.338	0.35	0.362	1.425
phecode_460-1	Acute upper respiratory infection	0.181	0.194	0.199	0.211	0.206	0.221	0.237	0.233	0.228	0.276	1.523
phecode_460-2	Acute lower respiratory infection	0.111	0.14	0.15	0.165	0.178	0.189	0.204	0.221	0.243	0.257	2.319
phecode_462	Sinusitis	0.065	0.075	0.072	0.074	0.077	0.082	0.079	0.091	0.092	0.097	1.494
phecode_462-1	Acute sinusitis	0.029	0.037	0.033	0.037	0.035	0.038	0.04	0.042	0.045	0.052	1.785
phecode_462-2	Chronic sinusitis	0.051	0.05	0.055	0.056	0.062	0.06	0.067	0.067	0.072	0.073	1.45
phecode_463	Rhinitis and nasal congestion	0.074	0.079	0.083	0.087	0.089	0.096	0.093	0.097	0.093	0.105	1.426
phecode_463-1	Chronic rhinitis	0.021	0.02	0.02	0.025	0.022	0.028	0.029	0.029	0.028	0.03	1.405
phecode_463-2	Allergic rhinitis	0.04	0.039	0.042	0.045	0.04	0.046	0.046	0.044	0.046	0.068	1.701
phecode_463-21	Seasonal allergic rhinitis	0.024	0.027	0.025	0.028	0.031	0.034	0.027	0.03	0.029	0.05	2.086
phecode_463-4	Nasal congestion*	0.023	0.024	0.024	0.026	0.026	0.024	0.027	0.027	0.027	0.028	1.216
phecode_463-5	Postnasal drip	0.011	0.013	0.016	0.015	0.016	0.019	0.019	0.021	0.019	0.02	1.809
phecode_464	Nasopharyngitis	0.014	0.014	0.019	0.018	0.017	0.02	0.018	0.016	0.022	0.021	1.512
phecode_464-1	Acute nasopharyngitis	0.013	0.015	0.018	0.018	0.018	0.017	0.018	0.018	0.019	0.022	1.65
phecode_465	Pharyngitis	0.06	0.065	0.072	0.074	0.076	0.087	0.078	0.091	0.102	0.113	1.884
phecode_465-1	Acute Pharyngitis	0.058	0.068	0.071	0.069	0.076	0.086	0.084	0.088	0.098	0.116	1.994
phecode_465-2	Chronic Pharyngitis	0.047	0.052	0.056	0.057	0.057	0.065	0.064	0.072	0.071	0.09	1.945
phecode_466	Tonsillitis and adenoiditis	0.011	0.015	0.017	0.016	0.017	0.017	0.02	0.028	0.03	0.044	3.985
phecode_466-1	Acute tonsillitis and adenoiditis	0.01	0.012	0.013	0.013	0.015	0.014	0.016	0.022	0.026	0.04	4.141
phecode_466-4	Hypertrophy of tonsils and adenoids	0.001	0.002	0.001	0.002	0.001	0.001	0.003	0.001	0.004	0.003	2.25
phecode_467	Laryngitis and tracheitis	0.01	0.01	0.017	0.014	0.014	0.016	0.017	0.017	0.019	0.022	2.225
phecode_467-1	Acute laryngitis and tracheitis	0.009	0.011	0.014	0.013	0.014	0.018	0.015	0.017	0.019	0.022	2.434
phecode_468	Pneumonia	0.017	0.023	0.026	0.038	0.037	0.052	0.058	0.073	0.088	0.124	7.401
phecode_468-1	Viral pneumonia	0.004	0.004	0.005	0.005	0.005	0.008	0.007	0.012	0.011	0.015	4.092
phecode_468-2	Bacterial pneumonia	0.001	0.002	0.003	0.003	0.004	0.005	0.005	0.006	0.006	0.011	11.002
phecode_468-8	Bronchopneumonia	0.001	0	0	0.002	0.002	0.002	0.001	0.004	0.006	0.007	4.667
phecode_468-9	Lobar pneumonia*	0.007	0.012	0.014	0.016	0.022	0.028	0.034	0.04	0.051	0.073	9.889
phecode_469	Bronchitis	0.026	0.027	0.029	0.031	0.037	0.037	0.04	0.037	0.051	0.048	1.856
phecode_469-1	Acute bronchitis	0.003	0.003	0.004	0.004	0.003	0.006	0.005	0.004	0.008	0.005	1.737
phecode_469-2	Chronic bronchitis	0.001	0.001	0.001	0.004	0.005	0.004	0.005	0.007	0.009	0.011	7.779
phecode_471	Other disorders of nose and nasal sinuses	0.054	0.06	0.061	0.061	0.061	0.059	0.066	0.061	0.062	0.069	1.278
phecode_471-2	Deviated nasal septum	0.006	0.005	0.006	0.005	0.006	0.007	0.008	0.007	0.011	0.01	1.658
phecode_471-3	Hypertrophy of nasal turbinates	0.002	0.001	0.002	0.002	0.002	0.001	0.002	0.002	0.003	0.003	1.546
phecode_471-5	Nasal polyps	0.007	0.008	0.006	0.007	0.01	0.006	0.007	0.008	0.009	0.012	1.786
phecode_472	Diseases of vocal cords and larynx, not elsewhere classified	0.002	0.003	0.003	0.004	0.004	0.005	0.005	0.006	0.005	0.006	3.901
phecode_473	Other diseases of upper respiratory tract	0.016	0.02	0.017	0.019	0.02	0.019	0.016	0.02	0.019	0.022	1.364
phecode_474	Chronic obstructive pulmonary disease	0.007	0.01	0.016	0.024	0.031	0.04	0.045	0.056	0.07	0.1	14.783
phecode_474-1	Emphysema	0.002	0.002	0.003	0.003	0.006	0.008	0.009	0.014	0.018	0.027	16.503
phecode_475	Asthma	0.036	0.037	0.041	0.045	0.044	0.044	0.051	0.051	0.056	0.055	1.513
phecode_475-5	Exercise induced bronchospasm	0.002	0.002	0.001	0.001	0.001	0.002	0.002	0.002	0.003	0.002	0.786
phecode_476	Bronchiectasis	0.004	0.005	0.006	0.008	0.013	0.015	0.016	0.018	0.022	0.024	6.592
phecode_477	Inhalation lung injury	0.001	0.001	0	0.001	0	0.001	0.003	0.002	0.004	0.007	5.501
phecode_478	Aspiration pneumonia	0	0.002	0.001	0.003	0.004	0.005	0.006	0.008	0.009	0.014	28.338
phecode_479	Pulmonary insufficiency and acute respiratory distress syndrome	0.007	0.007	0.012	0.012	0.019	0.02	0.024	0.027	0.032	0.048	6.838
phecode_479-3	Respiratory failure	0.002	0.002	0.001	0.004	0.005	0.004	0.006	0.005	0.009	0.016	9.092
phecode_479-6	Pulmonary collapse [Atelectasis]	0.005	0.007	0.009	0.01	0.014	0.017	0.018	0.019	0.026	0.032	6.434
phecode_480	Pulmonary edema	0.001	0.001	0	0.001	0.002	0.002	0.002	0.004	0.003	0.008	9.402
phecode_481	Interstitial pulmonary diseases	0.001	0.003	0.002	0.005	0.006	0.008	0.006	0.011	0.016	0.021	16.128

Supplementary Tables

Table 22 continued from previous page

Endpoint	PheCode string	1	2	3	4	5	6	7	8	9	10	Rate Ratio (RR)
phecode_481-4	Pulmonary fibrosis	0	0	0.001	0.001	0.002	0.003	0.003	0.003	0.006	0.007	21.003
phecode_483	Pleural effusion	0.006	0.01	0.014	0.015	0.021	0.028	0.03	0.037	0.049	0.064	10.344
phecode_484	Pneumothorax and air leak	0.002	0.002	0.002	0.003	0.003	0.004	0.004	0.006	0.005	0.009	3.786
phecode_486	Other respiratory disorders	0.073	0.089	0.095	0.101	0.109	0.118	0.123	0.132	0.141	0.165	2.265
phecode_486-2	Other diseases of bronchus	0.003	0.002	0.003	0.002	0.004	0.003	0.003	0.005	0.002	0.005	1.45
phecode_486-21	Bronchospasm	0.003	0.002	0.002	0.002	0.002	0.002	0.001	0.003	0.003	0.003	1.177
phecode_486-5	Abnormal sputum	0.027	0.031	0.039	0.033	0.045	0.042	0.048	0.05	0.061	0.062	2.313
phecode_487	Hemorrhage from respiratory passages	0.019	0.02	0.027	0.027	0.03	0.032	0.039	0.041	0.045	0.05	2.629
phecode_487-1	Epistaxis	0.015	0.015	0.02	0.019	0.024	0.027	0.03	0.028	0.03	0.037	2.484
phecode_487-3	Hemoptysis	0.005	0.006	0.006	0.007	0.008	0.01	0.009	0.013	0.013	0.017	3.091
phecode_488	Abnormalities of breathing	0.101	0.108	0.123	0.131	0.159	0.16	0.177	0.179	0.215	0.243	2.407
phecode_488-1	Dyspnea [Shortness of breath]	0.067	0.077	0.088	0.099	0.123	0.132	0.14	0.15	0.18	0.215	3.214
phecode_488-6	Wheezing	0.02	0.027	0.023	0.025	0.026	0.029	0.034	0.031	0.039	0.043	2.15
phecode_488-8	Mouth breathing*	0.009	0.013	0.011	0.013	0.011	0.012	0.014	0.012	0.016	0.02	2.327
phecode_491	Pleurisy	0.004	0.005	0.004	0.006	0.006	0.006	0.01	0.01	0.012	0.016	4.262
phecode_494	Voice disturbance	0.01	0.015	0.019	0.021	0.022	0.022	0.023	0.024	0.025	0.025	2.404
phecode_495	Abnormal findings on diagnostic imaging of lung	0.006	0.008	0.01	0.012	0.018	0.022	0.023	0.029	0.031	0.039	6.942
phecode_495-1	Solitary pulmonary nodule	0.001	0.001	0.001	0.002	0.004	0.004	0.005	0.005	0.005	0.004	4.801
phecode_496	Abnormal results of pulmonary function studies	0.013	0.012	0.018	0.02	0.018	0.025	0.022	0.026	0.029	0.038	2.975
phecode_498	Asphyxia and hypoxemia	0.001	0.001	0.001	0.001	0.001	0.002	0.001	0.003	0.003	0.005	6.601
phecode_500	Disorders of tooth development	0.002	0.002	0.002	0.003	0.002	0.002	0.003	0.003	0.004	0.003	1.214
phecode_500-4	Disturbances in tooth eruption	0.003	0.001	0.001	0.003	0.002	0.002	0.003	0.002	0.004	0.003	1.177
phecode_500-41	Impacted teeth*	0.001	0.001	0.001	0.002	0.001	0.003	0.002	0.003	0.003	0.002	2.143
phecode_501	Dental caries	0.006	0.009	0.009	0.007	0.01	0.01	0.01	0.011	0.011	0.012	1.898
phecode_502	Other diseases of teeth and supporting structures	0.013	0.014	0.016	0.014	0.016	0.015	0.015	0.015	0.019	0.019	1.553
phecode_503	Diseases of pulp and periapical tissues	0.011	0.011	0.014	0.015	0.014	0.013	0.012	0.009	0.015	0.018	1.708
phecode_503-5	Periapical abscess	0.01	0.012	0.014	0.014	0.013	0.013	0.011	0.008	0.015	0.018	1.814
phecode_504	Periodontal diseases	0.006	0.007	0.005	0.006	0.007	0.005	0.008	0.006	0.009	0.011	1.883
phecode_504-1	Gingivitis	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.003	1.455
phecode_504-3	Periodontitis	0.001	0.003	0.004	0.002	0.001	0.002	0.003	0.002	0.002	0.005	3.626
phecode_504-32	Chronic periodontitis	0.002	0.003	0.002	0.001	0.002	0.002	0.002	0.001	0.002	0.003	1.429
phecode_506	Diseases of salivary glands	0.009	0.012	0.014	0.014	0.014	0.019	0.019	0.017	0.022	0.023	2.604
phecode_506-3	Sialoadenitis	0.002	0.003	0.004	0.003	0.003	0.004	0.004	0.004	0.005	0.005	2
phecode_506-5	Disturbances of salivary secretion	0.004	0.006	0.006	0.008	0.01	0.009	0.012	0.012	0.012	0.014	3.667
phecode_507	Lesions of mouth	0.023	0.023	0.024	0.027	0.025	0.032	0.03	0.036	0.037	0.039	1.716
phecode_507-1	Stomatitis	0.012	0.01	0.012	0.012	0.016	0.015	0.016	0.017	0.017	0.017	1.389
phecode_507-11	Recurrent oral aphthae [Recurrent aphthous stomatitis]	0.002	0.002	0.002	0.003	0.003	0.003	0.003	0.003	0.005	0.005	2.143
phecode_508	Diseases of lips	0.006	0.007	0.01	0.01	0.01	0.01	0.012	0.01	0.013	0.015	2.411
phecode_509	Diseases of tongue	0.015	0.017	0.012	0.014	0.017	0.017	0.021	0.017	0.02	0.019	1.226
phecode_509-1	Glossitis	0.004	0.004	0.005	0.004	0.005	0.006	0.005	0.007	0.006	0.007	1.667
phecode_509-11	Glossodynia	0.002	0.002	0.002	0.003	0.003	0.002	0.003	0.004	0.003	0.004	2.3
phecode_509-3	Hypertrophy of tongue papillae	0.002	0.001	0.002	0.002	0.002	0.002	0.003	0.002	0.002	0.002	1.364
phecode_510	Diseases of esophagus	0.04	0.052	0.052	0.064	0.072	0.077	0.078	0.076	0.086	0.099	2.491
phecode_510-2	Esophagitis	0.032	0.034	0.037	0.049	0.049	0.06	0.055	0.06	0.058	0.072	2.248
phecode_510-5	Dyskinesia of esophagus	0.002	0.002	0.003	0.003	0.001	0.004	0.004	0.003	0.004	0.006	2.467
phecode_510-8	Barrett's esophagus	0.004	0.005	0.008	0.011	0.012	0.014	0.013	0.015	0.016	0.016	3.841
phecode_511	Gastro-esophageal reflux disease	0.09	0.102	0.11	0.124	0.132	0.143	0.134	0.149	0.159	0.171	1.909
phecode_512	Dysphagia	0.021	0.025	0.03	0.03	0.038	0.039	0.043	0.045	0.051	0.053	2.532
phecode_513	Peptic ulcer	0.01	0.013	0.018	0.017	0.024	0.023	0.029	0.028	0.029	0.037	3.683
phecode_513-1	Esophageal ulcer	0.004	0.002	0.005	0.005	0.006	0.007	0.008	0.011	0.01	0.011	3.137
phecode_513-2	Gastric ulcer	0.005	0.005	0.012	0.009	0.012	0.014	0.015	0.014	0.016	0.02	3.728
phecode_513-3	Duodenal ulcer	0.004	0.004	0.004	0.006	0.006	0.006	0.006	0.008	0.009	0.012	3.182
phecode_514	Gastrointestinal obstruction	0.01	0.008	0.011	0.018	0.019	0.02	0.028	0.024	0.028	0.033	3.449
phecode_514-1	Esophageal obstruction (Stricture and stenosis of esophagus)	0.004	0.003	0.004	0.007	0.008	0.01	0.009	0.009	0.011	0.011	2.616
phecode_514-2	Intestinal obstruction	0.005	0.005	0.007	0.008	0.01	0.011	0.015	0.013	0.014	0.019	3.807
phecode_514-21	Impaction of intestine	0.002	0.002	0.001	0.002	0.002	0.002	0.003	0.004	0.001	0.003	1.455
phecode_514-3	Ileus	0.001	0.001	0.001	0.002	0.002	0.003	0.002	0.004	0.005	0.005	4.572
phecode_515	Heartburn and epigastric pain	0.028	0.03	0.036	0.04	0.039	0.047	0.041	0.041	0.046	0.056	2.012

5 Phenome-wide prediction of disease onset from retinal fundus photographs

Table 22 continued from previous page

Endpoint	PheCode string	1	2	3	4	5	6	7	8	9	10	Rate Ratio (RR)
phecode_516	Other diseases of stomach and duodenum	0.007	0.007	0.006	0.009	0.009	0.011	0.014	0.013	0.015	0.02	2.796
phecode_517	Gastrointestinal angiodysplasia	0	0	0.002	0.001	0.003	0.002	0.004	0.004	0.005	0.004	8.001
phecode_518	Appendicitis	0.006	0.007	0.005	0.005	0.006	0.006	0.005	0.008	0.008	0.008	1.342
phecode_520	Hernia	0.063	0.074	0.086	0.107	0.119	0.135	0.15	0.158	0.172	0.189	3.001
phecode_520-1	Hernia of the abdominal wall	0.023	0.026	0.032	0.041	0.045	0.057	0.068	0.088	0.107	0.119	5.19
phecode_520-11	Inguinal hernia	0.006	0.012	0.014	0.02	0.028	0.035	0.052	0.055	0.082	0.093	16.15
phecode_520-12	Femoral hernia	0.001	0.002	0.001	0.001	0.001	0.002	0.002	0.001	0.003	0.002	2
phecode_520-13	Umbilical hernia	0.006	0.009	0.009	0.01	0.012	0.012	0.016	0.017	0.018	0.022	3.829
phecode_520-14	Ventral hernia	0.009	0.006	0.009	0.007	0.008	0.011	0.012	0.014	0.013	0.018	2.077
phecode_520-15	Incisional hernia	0.005	0.003	0.003	0.003	0.005	0.005	0.006	0.008	0.007	0.009	2.072
phecode_520-2	Diaphragmatic hernia [Hiatal hernia]	0.041	0.042	0.055	0.057	0.08	0.075	0.089	0.092	0.097	0.112	2.76
phecode_522	Gastrointestinal inflammation	0.081	0.094	0.116	0.115	0.122	0.131	0.132	0.152	0.152	0.176	2.164
phecode_522-1	Inflammatory bowel disease	0.006	0.009	0.007	0.008	0.007	0.008	0.01	0.011	0.01	0.013	1.975
phecode_522-11	Crohn's disease	0.004	0.002	0.002	0.003	0.002	0.003	0.003	0.003	0.003	0.003	0.875
phecode_522-12	Ulcerative colitis	0.003	0.003	0.004	0.003	0.003	0.005	0.004	0.003	0.006	0.004	1.588
phecode_522-14	Microscopic colitis*	0.001	0.001	0.003	0.002	0.002	0.002	0.002	0.004	0.004	0.004	4.334
phecode_522-7	Ulceration of the lower GI tract	0.003	0.003	0.002	0.002	0.004	0.004	0.004	0.004	0.005	0.005	1.778
phecode_522-8	Duodenitis	0.01	0.011	0.016	0.02	0.017	0.021	0.024	0.022	0.024	0.031	3.065
phecode_522-9	Gastritis	0.044	0.058	0.069	0.074	0.078	0.081	0.083	0.093	0.102	0.114	2.604
phecode_523	Diverticular disease [Bowel diverticulosis]	0.042	0.067	0.073	0.086	0.099	0.124	0.13	0.149	0.147	0.178	4.274
phecode_523-1	Diverticula of small intestine	0.001	0.001	0.001	0.001	0.002	0.003	0.003	0.004	0.003	0.002	3
phecode_523-2	Diverticula of colon	0.037	0.052	0.064	0.075	0.09	0.109	0.119	0.13	0.14	0.153	4.164
phecode_523-4	Diverticulitis	0.006	0.011	0.011	0.013	0.014	0.017	0.017	0.021	0.019	0.023	3.944
phecode_524	Functional intestinal disorder	0.012	0.017	0.018	0.021	0.02	0.022	0.022	0.024	0.028	0.03	2.606
phecode_524-1	Irritable bowel syndrome	0.01	0.014	0.015	0.018	0.02	0.019	0.022	0.019	0.024	0.028	2.742
phecode_525	Intestinal malabsorption	0.007	0.007	0.007	0.009	0.007	0.007	0.006	0.008	0.008	0.01	1.439
phecode_525-1	Celiac disease	0.002	0.003	0.003	0.004	0.005	0.004	0.003	0.004	0.004	0.005	2.215
phecode_525-3	Disorders of intestinal carbohydrate absorption	0.002	0.001	0.002	0.001	0.001	0.001	0.001	0.003	0.003	0.003	1.637
phecode_526	Intestinal infection	0.04	0.043	0.046	0.052	0.058	0.059	0.064	0.074	0.076	0.094	2.339
phecode_526-1	Bacterial enteritis	0.016	0.014	0.017	0.015	0.017	0.019	0.02	0.018	0.02	0.02	1.243
phecode_526-11	Intestinal e.coli	0.003	0.003	0.001	0.003	0.003	0.002	0.004	0.002	0.005	0.004	1.5
phecode_526-12	Intestinal infection due to C. difficile	0.001	0.001	0.002	0.001	0.003	0.002	0.005	0.003	0.005	0.006	8.501
phecode_526-2	Viral enteritis	0.003	0.002	0.002	0.002	0.002	0.002	0.004	0.003	0.004	0.007	2.5
phecode_527	Abdominal pain	0.169	0.179	0.184	0.198	0.201	0.219	0.215	0.221	0.241	0.25	1.473
phecode_528	Nausea and vomiting	0.049	0.059	0.063	0.066	0.077	0.07	0.091	0.096	0.103	0.12	2.442
phecode_528-1	Nausea	0.049	0.058	0.065	0.065	0.076	0.071	0.092	0.095	0.101	0.122	2.472
phecode_528-2	Vomiting	0.049	0.059	0.062	0.068	0.075	0.073	0.09	0.095	0.102	0.121	2.474
phecode_529	Symptoms involving digestive system	0.159	0.177	0.189	0.22	0.218	0.239	0.246	0.262	0.278	0.296	1.862
phecode_529-1	Diarrhea	0.059	0.058	0.058	0.071	0.068	0.068	0.076	0.082	0.086	0.094	1.589
phecode_529-2	Abdominal distension and flatulence	0.025	0.027	0.029	0.041	0.036	0.041	0.041	0.042	0.043	0.053	2.114
phecode_529-3	Fecal incontinence	0.003	0.007	0.009	0.009	0.014	0.015	0.018	0.016	0.02	0.029	9.001
phecode_529-5	Constipation	0.04	0.05	0.071	0.071	0.086	0.092	0.11	0.134	0.132	0.153	3.833
phecode_529-6	Halitosis*	0.001	0.002	0.002	0.002	0.002	0.002	0.002	0.001	0.002	0.003	2.286
phecode_530	Disease of anus and rectum	0.027	0.026	0.028	0.029	0.032	0.031	0.033	0.034	0.033	0.036	1.331
phecode_530-1	Anal fissure	0.01	0.012	0.013	0.013	0.01	0.012	0.01	0.013	0.014	0.013	1.404
phecode_530-2	Anorectal abscess	0.002	0.002	0.002	0.002	0.003	0.001	0.002	0.003	0.005	0.005	2.819
phecode_530-3	Rectal prolapse	0.001	0.002	0.002	0.002	0.003	0.003	0.004	0.003	0.004	0.008	7.667
phecode_532	Other disorders of the intestines	0.014	0.011	0.013	0.012	0.017	0.017	0.017	0.022	0.019	0.023	1.651
phecode_532-1	Intestinal fistula	0.002	0.003	0.002	0.002	0.003	0.003	0.004	0.004	0.002	0.005	2
phecode_532-4	Volvulus	0.001	0.001	0.001	0.001	0.002	0.003	0.003	0.004	0.003	0.003	2.125
phecode_535	Intestinal or peritoneal adhesions	0.007	0.01	0.011	0.013	0.014	0.014	0.015	0.014	0.016	0.023	3.089
phecode_537	Abnormality of the peritoneum	0.005	0.005	0.004	0.005	0.006	0.005	0.005	0.007	0.008	0.01	1.818
phecode_537-1	Peritonitis	0.005	0.004	0.004	0.005	0.003	0.004	0.005	0.005	0.006	0.008	1.621
phecode_540	Hepatitis	0.002	0.002	0.003	0.003	0.002	0.003	0.003	0.004	0.004	0.003	1.4
phecode_540-1	Chronic hepatitis	0.001	0.001	0.001	0.001	0.003	0.002	0.001	0.002	0.004	0.003	2.667
phecode_540-3	Viral hepatitis	0.003	0.002	0	0.002	0.002	0.002	0.002	0.002	0.002	0.003	1.176
phecode_542	Chronic liver disease and sequelae	0.02	0.026	0.028	0.03	0.033	0.037	0.033	0.043	0.046	0.055	2.777
phecode_542-1	Fibrosis and cirrhosis of liver	0.001	0.001	0.002	0.002	0.002	0.003	0.004	0.004	0.006	0.012	17.753
phecode_542-2	Fatty liver disease (FLD)	0.019	0.021	0.025	0.028	0.031	0.032	0.028	0.036	0.042	0.044	2.331
phecode_542-3	Hepatic failure	0.001	0.001	0.001	0.002	0.002	0.002	0.002	0.003	0.004	0.007	11.252

Supplementary Tables

Table 22 continued from previous page

Endpoint	PheCode string	1	2	3	4	5	6	7	8	9	10	Rate Ratio (RR)
phecode_542-4	Portal hypertension	0.001	0	0	0.002	0.001	0.001	0.003	0.001	0.005	0.005	4.144
phecode_545	Nonspecific abnormal results of function study of liver	0.02	0.023	0.027	0.024	0.026	0.033	0.033	0.035	0.035	0.049	2.407
phecode_546	Other disorders of liver	0.009	0.01	0.01	0.012	0.016	0.016	0.021	0.021	0.024	0.024	2.607
phecode_546-3	Hepatomegaly	0.001	0.001	0.001	0.001	0.002	0.002	0.002	0.002	0.003	0.004	3.286
phecode_550	Disorders of the gallbladder	0.03	0.03	0.038	0.042	0.046	0.047	0.053	0.053	0.055	0.063	2.125
phecode_550-1	Gallstones [Cholelithiasis]	0.024	0.026	0.03	0.037	0.04	0.044	0.044	0.045	0.049	0.057	2.374
phecode_550-2	Cholecystitis	0.008	0.007	0.008	0.011	0.009	0.01	0.009	0.012	0.011	0.016	2.043
phecode_550-4	Cholesterosis of gallbladder	0.003	0.002	0.003	0.003	0.003	0.002	0.004	0.003	0.004	0.005	1.875
phecode_552	Other diseases of biliary tract	0.003	0.004	0.006	0.007	0.006	0.008	0.01	0.008	0.012	0.014	4.401
phecode_552-1	Cholangitis	0.001	0.001	0.001	0.002	0.004	0.004	0.002	0.003	0.003	0.006	4.501
phecode_552-2	Obstruction of bile duct	0.001	0.001	0.002	0.002	0.004	0.002	0.004	0.005	0.005	0.007	7
phecode_554	Diseases of the pancreas	0.004	0.004	0.007	0.008	0.01	0.009	0.012	0.011	0.017	0.014	3.955
phecode_554-1	Pancreatitis	0.002	0.003	0.004	0.007	0.004	0.006	0.006	0.006	0.008	0.009	4.155
phecode_554-11	Acute pancreatitis	0.002	0.003	0.004	0.006	0.004	0.005	0.006	0.005	0.007	0.007	3.584
phecode_554-2	Cyst and pseudocyst of pancreas	0.001	0.001	0.001	0.002	0.002	0.003	0.003	0.004	0.003	0.005	7.5
phecode_555	Ascites	0.002	0.002	0.003	0.004	0.004	0.006	0.009	0.008	0.011	0.013	6.078
phecode_556	Other symptoms involving the digestive system and abdomen	0.037	0.042	0.048	0.049	0.054	0.056	0.062	0.071	0.061	0.077	2.102
phecode_556-3	Abdominal or pelvic swelling, mass, or lump	0.01	0.009	0.014	0.013	0.013	0.015	0.013	0.015	0.016	0.015	1.5
phecode_556-8	Nonspecific abnormal findings in stool contents	0.03	0.034	0.036	0.037	0.041	0.047	0.05	0.052	0.056	0.061	2.017
phecode_557	Gastrointestinal hemorrhage	0.067	0.067	0.072	0.08	0.079	0.082	0.09	0.091	0.086	0.107	1.614
phecode_557-1	Hematemesis	0.003	0.003	0.003	0.003	0.005	0.004	0.005	0.007	0.008	0.013	4.707
phecode_557-2	Blood in stool	0.005	0.006	0.009	0.008	0.009	0.01	0.013	0.015	0.016	0.021	4.536
phecode_557-8	Hemorrhage of rectum and anus	0.053	0.052	0.053	0.051	0.054	0.057	0.059	0.055	0.054	0.061	1.141
phecode_558	Abnormal findings on diagnostic imaging of the digestive tract	0.012	0.015	0.019	0.021	0.023	0.027	0.028	0.032	0.036	0.038	3.122
phecode_559	Other disease of digestive system, NOS	0.002	0.003	0.004	0.005	0.005	0.004	0.005	0.005	0.006	0.005	2.333
phecode_580	Glomerular diseases	0.002	0.002	0.003	0.004	0.004	0.004	0.004	0.007	0.007	0.014	6.538
phecode_580-3	Nephrotic syndrome	0.002	0.002	0.003	0.003	0.003	0.002	0.004	0.004	0.003	0.004	1.572
phecode_581	Renal tubulo-interstitial diseases	0.013	0.017	0.016	0.016	0.018	0.019	0.023	0.021	0.023	0.027	2.051
phecode_581-1	Pyelonephritis	0.003	0.004	0.005	0.004	0.005	0.005	0.004	0.005	0.006	0.006	1.9
phecode_581-11	Acute pyelonephritis	0.002	0.003	0.002	0.005	0.004	0.005	0.004	0.005	0.004	0.006	2.267
phecode_581-3	Obstructive and reflux uropathy	0.007	0.007	0.007	0.011	0.014	0.013	0.014	0.017	0.016	0.019	2.675
phecode_581-31	Hydronephrosis	0.007	0.006	0.006	0.01	0.013	0.012	0.01	0.014	0.014	0.017	2.513
phecode_581-33	Stricture or kinking of ureter	0.002	0.002	0.001	0.002	0.003	0.002	0.003	0.003	0.004	0.004	1.8
phecode_582	Acute kidney failure	0.006	0.009	0.017	0.017	0.024	0.034	0.042	0.05	0.066	0.104	18.709
phecode_583	Chronic kidney disease	0.013	0.018	0.026	0.028	0.042	0.056	0.072	0.073	0.091	0.122	9.628
phecode_583-1	End stage renal disease [CDK, stage 5]	0.001	0.001	0.001	0.001	0.001	0.003	0.003	0.004	0.004	0.012	10.143
phecode_584	Renal failure	0.003	0.002	0.004	0.003	0.005	0.007	0.009	0.01	0.013	0.018	6.412
phecode_585	Kidney stone disease	0.012	0.011	0.012	0.015	0.019	0.02	0.022	0.02	0.025	0.025	2.158
phecode_585-1	Renal colic	0.004	0.003	0.006	0.005	0.007	0.005	0.008	0.008	0.008	0.013	3.479
phecode_586	Other disorders of the kidney and ureters	0.007	0.01	0.013	0.016	0.021	0.021	0.026	0.027	0.033	0.043	5.779
phecode_586-2	Cyst of kidney	0.004	0.006	0.008	0.01	0.013	0.013	0.013	0.018	0.022	0.024	5.371
phecode_588	Disorders and findings from impaired renal function	0	0.001	0.001	0.001	0.001	0.002	0.003	0.004	0.004	0.005	15.503
phecode_588-2	Abnormal results of function study of kidney	0	0.001	0.001	0.001	0.002	0.001	0.003	0.003	0.004	0.005	31.005
phecode_591	Urinary tract infection	0.046	0.066	0.081	0.093	0.1	0.104	0.121	0.131	0.147	0.16	3.512
phecode_592	Cystitis and urethritis	0.01	0.015	0.019	0.023	0.026	0.032	0.028	0.037	0.051	0.048	4.634
phecode_592-1	Cystitis	0.009	0.013	0.017	0.022	0.025	0.029	0.029	0.037	0.051	0.046	5.295
phecode_592-11	Acute cystitis	0.002	0.002	0.002	0.002	0.004	0.006	0.006	0.009	0.007	0.009	4.584
phecode_592-12	Chronic cystitis	0.001	0.001	0.003	0.002	0.002	0.003	0.003	0.002	0.005	0.004	4.334
phecode_593	Hematuria	0.076	0.101	0.127	0.13	0.129	0.154	0.169	0.172	0.173	0.191	2.509
phecode_593-1	Gross hematuria	0.003	0.003	0.004	0.003	0.006	0.005	0.007	0.008	0.012	0.01	3.445
phecode_593-2	Microscopic hematuria	0.007	0.009	0.009	0.013	0.014	0.011	0.013	0.014	0.014	0.017	2.65
phecode_593-3	Recurrent and persistent hematuria*	0.001	0.001	0.002	0.002	0.002	0.002	0.003	0.002	0.003	0.001	1.286
phecode_594	Abnormality of urination	0.085	0.114	0.12	0.129	0.15	0.17	0.179	0.182	0.194	0.216	2.557
phecode_594-1	Retention of urine	0.007	0.013	0.02	0.024	0.032	0.035	0.045	0.06	0.071	0.086	11.646
phecode_594-11	Urinary hesitancy	0.001	0.001	0.001	0.002	0.001	0.002	0.003	0.003	0.004	0.005	7.751
phecode_594-2	Dysuria	0.016	0.027	0.035	0.036	0.037	0.045	0.043	0.051	0.052	0.057	3.501

5 Phenome-wide prediction of disease onset from retinal fundus photographs

Table 22 continued from previous page

Endpoint	PheCode string	1	2	3	4	5	6	7	8	9	10	Rate Ratio (RR)
phecode_594-3	Urinary incontinence	0.018	0.032	0.038	0.047	0.052	0.052	0.054	0.067	0.062	0.079	4.288
phecode_594-31	Urge incontinence	0.005	0.011	0.011	0.013	0.015	0.017	0.019	0.017	0.019	0.022	4.061
phecode_594-32	Stress incontinence	0.003	0.004	0.01	0.009	0.015	0.015	0.021	0.022	0.024	0.026	8.054
phecode_594-4	Frequency of urination and polyuria	0.041	0.046	0.055	0.054	0.067	0.08	0.078	0.083	0.082	0.084	2.034
phecode_594-41	Nocturia	0.007	0.01	0.012	0.017	0.018	0.022	0.026	0.033	0.035	0.044	6.406
phecode_594-6	Urinary urgency	0.013	0.017	0.021	0.025	0.026	0.026	0.032	0.028	0.034	0.027	2.077
phecode_596	Other disorders of bladder	0.02	0.026	0.029	0.032	0.034	0.042	0.043	0.049	0.063	0.058	2.924
phecode_596-1	Bladder neck obstruction	0.001	0.001	0.001	0.002	0.003	0.004	0.006	0.008	0.01	0.012	8.876
phecode_596-2	Overactive bladder	0.012	0.016	0.019	0.018	0.018	0.021	0.024	0.025	0.024	0.025	2.081
phecode_596-3	Diverticulum of bladder	0.001	0	0.002	0.002	0.002	0.002	0.005	0.004	0.004	0.006	8.751
phecode_596-5	Neuromuscular dysfunction of bladder	0.001	0.001	0.001	0.002	0.002	0.001	0.002	0.003	0.003	0.004	3.834
phecode_597	Other disorders of urethra and urinary tract	0.004	0.005	0.007	0.009	0.011	0.011	0.012	0.011	0.014	0.013	3.637
phecode_597-1	Urethral stricture	0.003	0.003	0.003	0.005	0.006	0.008	0.008	0.01	0.01	0.011	3.89
phecode_597-5	Urethral caruncle	0	0.001	0.001	0.001	0.002	0.001	0.003	0.003	0.003	0.004	12.502
phecode_599	Other symptoms/disorders or the urinary system	0.074	0.084	0.089	0.103	0.106	0.123	0.123	0.127	0.138	0.136	1.84
phecode_600	Benign prostatic hyperplasia	0.046	0.065	0.096	0.126	0.149	0.157	0.164	0.181	0.179	0.189	4.128
phecode_601	Inflammatory diseases of prostate	0.012	0.016	0.021	0.021	0.02	0.019	0.017	0.013	0.015	0.021	1.736
phecode_601-1	Prostatitis	0.013	0.016	0.02	0.022	0.02	0.019	0.017	0.013	0.015	0.022	1.715
phecode_601-11	Acute prostatitis	0.003	0.005	0.005	0.005	0.006	0.003	0.004	0.008	0.005	0.006	2.001
phecode_601-12	Chronic prostatitis	0.003	0.004	0.008	0.008	0.005	0.008	0.007	0.005	0.006	0.008	3.287
phecode_602	Other disorders of prostate	0.02	0.029	0.046	0.05	0.051	0.054	0.064	0.065	0.067	0.074	3.71
phecode_602-4	Elevated prostate specific antigen [PSA]	0.014	0.022	0.038	0.039	0.041	0.044	0.046	0.047	0.053	0.059	4.181
phecode_603	Disorders and symptoms of testis	0.051	0.052	0.053	0.043	0.054	0.053	0.051	0.062	0.059	0.054	1.07
phecode_603-1	Hydrocele	0.011	0.012	0.011	0.012	0.009	0.008	0.016	0.014	0.01	0.011	0.968
phecode_603-2	Spermatocele	0.013	0.014	0.017	0.015	0.012	0.017	0.012	0.016	0.019	0.019	1.515
phecode_603-5	Orchitis and epididymitis	0.024	0.02	0.021	0.016	0.021	0.023	0.02	0.025	0.024	0.021	0.907
phecode_603-6	Scrotal pain*	0	0.001	0.002	0.004	0.005	0.009	0.01	0.012	0.017	0.02	41.34
phecode_604	Disorders of penis	0.048	0.035	0.039	0.033	0.037	0.037	0.035	0.035	0.041	0.04	0.838
phecode_604-1	Redundant prepuce and phimosis	0.01	0.005	0.005	0.006	0.006	0.007	0.007	0.009	0.005	0.006	0.63
phecode_604-3	Peyronie's disease	0.009	0.007	0.006	0.007	0.008	0.005	0.006	0.008	0.006	0.008	0.84
phecode_604-5	Balanoposthitis	0.016	0.016	0.016	0.015	0.016	0.015	0.015	0.018	0.016	0.017	1.046
phecode_605	Male sexual dysfunction	0.084	0.102	0.119	0.121	0.136	0.127	0.14	0.159	0.147	0.158	1.873
phecode_605-1	Male erectile dysfunction	0.084	0.102	0.117	0.124	0.13	0.13	0.137	0.158	0.151	0.157	1.882
phecode_608	Other disorders of male genital organs	0.083	0.104	0.105	0.096	0.097	0.1	0.105	0.094	0.098	0.103	1.239
phecode_608-1	Abnormal findings in semen	0.001	0.004	0.007	0.01	0.017	0.019	0.03	0.038	0.045	0.06	72.212
phecode_610	Benign mammary dysplasias	0.011	0.01	0.006	0.007	0.011	0.01	0.012	0.013	0.018	0.024	2.177
phecode_610-1	Cystic mastopathy	0.006	0.004	0.006	0.003	0.006	0.006	0.009	0.011	0.015	0.021	3.779
phecode_613	Other nonmalignant breast conditions	0.072	0.073	0.069	0.069	0.068	0.072	0.08	0.08	0.069	0.094	1.293
phecode_613-1	Inflammatory disease of breast	0.01	0.006	0.007	0.006	0.01	0.005	0.007	0.007	0.006	0.015	1.5
phecode_613-5	Mastodynia	0.051	0.045	0.045	0.051	0.056	0.05	0.053	0.055	0.059	0.069	1.354
phecode_613-7	Other signs and symptoms in breast	0.015	0.021	0.019	0.02	0.017	0.02	0.017	0.019	0.016	0.014	0.959
phecode_614	Inflammatory diseases of female pelvic organs	0.049	0.053	0.049	0.052	0.053	0.051	0.051	0.059	0.071	0.083	1.681
phecode_614-1	Pelvic peritoneal adhesions, female (postoperative) (postinfection)	0.007	0.005	0.004	0.006	0.005	0.005	0.005	0.006	0.005	0.01	1.546
phecode_614-5	Inflammatory disease of cervix, vagina, and vulva	0.043	0.048	0.043	0.047	0.044	0.045	0.047	0.054	0.061	0.072	1.659
phecode_614-52	Vaginitis and vulvovaginitis	0.036	0.04	0.042	0.036	0.04	0.033	0.045	0.043	0.051	0.058	1.624
phecode_614-53	Cyst or abscess of Bartholin's gland	0.003	0.003	0.004	0.002	0.004	0.004	0.005	0.005	0.007	0.006	2.334
phecode_614-54	Abscess or ulceration of vulva	0.004	0.002	0.005	0.002	0.004	0.003	0.002	0.003	0.003	0.003	0.715
phecode_614-55	Candidiasis of vulva and vagina	0.001	0.003	0.006	0.011	0.015	0.018	0.019	0.022	0.027	0.036	35.506
phecode_615	Endometriosis	0.003	0.002	0.004	0.004	0.003	0.008	0.008	0.009	0.018	0.02	6.602
phecode_618	Genital prolapse	0.04	0.051	0.053	0.052	0.06	0.067	0.063	0.064	0.071	0.066	1.646
phecode_618-1	Prolapse of vaginal walls	0.034	0.048	0.043	0.04	0.051	0.055	0.056	0.051	0.061	0.054	1.608
phecode_618-11	Cystocele	0.023	0.027	0.035	0.03	0.038	0.034	0.04	0.042	0.044	0.04	1.795
phecode_618-12	Rectocele	0.016	0.022	0.02	0.021	0.023	0.026	0.021	0.023	0.028	0.024	1.538
phecode_618-2	Uterine/Uterovaginal prolapse	0.013	0.024	0.022	0.026	0.03	0.028	0.03	0.027	0.03	0.022	1.781
phecode_619	Noninflammatory female genital disorders	0.111	0.118	0.118	0.113	0.121	0.12	0.143	0.153	0.167	0.19	1.707
phecode_619-2	Disorders of uterus, NEC	0.011	0.011	0.012	0.008	0.013	0.013	0.013	0.014	0.013	0.017	1.514
phecode_619-3	Noninflammatory disorders of cervix	0.006	0.01	0.008	0.011	0.008	0.013	0.017	0.021	0.025	0.034	6.113

Supplementary Tables

Table 22 continued from previous page

Endpoint	PheCode string	1	2	3	4	5	6	7	8	9	10	Rate Ratio (RR)
phecode_619-4	Noninflammatory disorders of vagina	0.073	0.07	0.089	0.073	0.085	0.085	0.107	0.109	0.125	0.14	1.922
phecode_619-5	Noninflammatory disorders of vulva and perineum	0.033	0.027	0.031	0.028	0.025	0.03	0.035	0.03	0.026	0.031	0.926
phecode_620	Dysplasia of female genital organs	0.002	0.002	0.002	0.002	0.003	0.003	0.006	0.008	0.01	0.012	6.502
phecode_621	Endometrial hyperplasia	0.002	0.003	0.005	0.001	0.004	0.002	0.003	0.004	0.005	0.005	2.251
phecode_622	Polyp of female genital organs	0.023	0.022	0.024	0.024	0.029	0.028	0.028	0.032	0.033	0.037	1.568
phecode_622-1	Polyp of corpus uteri	0.017	0.019	0.02	0.021	0.02	0.02	0.019	0.018	0.02	0.021	1.233
phecode_622-2	Mucous polyp of cervix	0.006	0.005	0.006	0.007	0.008	0.012	0.011	0.015	0.015	0.015	2.351
phecode_623	Hypertrophy of female genital organs	0.03	0.026	0.029	0.025	0.025	0.026	0.032	0.024	0.027	0.028	0.939
phecode_624	Symptoms involving female genital tract	0.01	0.011	0.013	0.013	0.013	0.014	0.014	0.013	0.015	0.017	1.781
phecode_624-1	Dystrophy of female genital tract	0.005	0.005	0.011	0.006	0.006	0.007	0.009	0.009	0.008	0.012	2.438
phecode_624-2	Atrophy of female genital tract	0.004	0.006	0.006	0.004	0.008	0.008	0.008	0.009	0.005	0.005	1.215
phecode_625	Pain and other symptoms associated with female genital organs	0.035	0.03	0.036	0.037	0.043	0.039	0.05	0.056	0.064	0.074	2.092
phecode_625-1	Dyspareunia	0.006	0.008	0.008	0.011	0.009	0.014	0.016	0.017	0.022	0.024	4.054
phecode_625-2	Postcoital bleeding	0.003	0.002	0.002	0.003	0.003	0.007	0.013	0.015	0.017	0.02	7.224
phecode_626	Disorders of menstruation and other abnormal bleeding from female genital tract	0.018	0.015	0.017	0.015	0.017	0.025	0.042	0.072	0.117	0.167	9.371
phecode_626-1	Irregular menstrual cycle/bleeding	0.001	0.002	0.004	0.004	0.008	0.012	0.021	0.051	0.081	0.125	91.281
phecode_626-11	Absent or infrequent menstruation	0.001	0.001	0.002	0.002	0.002	0.005	0.009	0.019	0.023	0.036	38.667
phecode_626-13	Irregular menstrual cycle	0	0.001	0.001	0	0.002	0.005	0.01	0.021	0.041	0.064	203.064
phecode_626-14	Irregular menstrual bleeding	0	0	0.001	0.002	0.005	0.006	0.008	0.018	0.027	0.05	Inf
phecode_626-2	Dysmenorrhea	0	0	0	0	0	0.002	0.002	0.007	0.008	0.018	Inf
phecode_627	Menopausal and postmenopausal disorders	0.103	0.094	0.108	0.112	0.139	0.158	0.205	0.279	0.326	0.375	3.639
phecode_627-1	Postmenopausal bleeding	0.05	0.044	0.043	0.058	0.051	0.057	0.059	0.06	0.073	0.072	1.44
phecode_627-2	Symptomatic menopause	0.02	0.021	0.024	0.035	0.04	0.065	0.107	0.147	0.191	0.227	11.556
phecode_627-3	Postmenopausal atrophic vaginitis	0.034	0.043	0.047	0.045	0.051	0.05	0.054	0.059	0.059	0.055	1.634
phecode_627-4	Menorrhagia/Excessive and frequent menstruation	0.001	0.001	0.004	0.005	0.007	0.016	0.024	0.058	0.099	0.15	211.575
phecode_628	Ovarian cyst	0.018	0.022	0.015	0.02	0.019	0.023	0.024	0.019	0.025	0.038	2.104
phecode_628-2	Corpus luteum cyst or hematoma	0.016	0.022	0.016	0.021	0.019	0.019	0.023	0.021	0.02	0.037	2.334
phecode_660	Infection of the skin	0.2	0.209	0.216	0.227	0.236	0.249	0.258	0.257	0.264	0.276	1.378
phecode_660-1	Fungal infection of the skin	0.072	0.09	0.098	0.093	0.093	0.114	0.107	0.111	0.116	0.127	1.756
phecode_660-11	Candidiasis of skin and nails	0.005	0.006	0.005	0.012	0.009	0.008	0.008	0.011	0.009	0.011	2.242
phecode_660-12	Dermatophytosis	0.057	0.064	0.068	0.078	0.075	0.078	0.083	0.091	0.098	0.098	1.713
phecode_660-13	Pityriasis versicolor	0.001	0.001	0.003	0.002	0.002	0.002	0.005	0.005	0.004	0.008	5.334
phecode_660-2	Bacterial infection of the skin	0.007	0.007	0.011	0.008	0.007	0.009	0.01	0.008	0.008	0.009	1.156
phecode_660-21	Impetigo	0.006	0.007	0.008	0.005	0.008	0.007	0.009	0.007	0.007	0.007	1.051
phecode_660-4	Carbuncle and furuncle	0.011	0.012	0.013	0.011	0.013	0.013	0.014	0.016	0.013	0.021	1.812
phecode_660-6	Cellulitis and abscess	0.084	0.081	0.089	0.097	0.101	0.111	0.119	0.12	0.128	0.149	1.776
phecode_662	Rosacea	0.014	0.022	0.022	0.026	0.022	0.029	0.027	0.03	0.03	0.036	2.591
phecode_664	Papulosquamous disorders	0.011	0.011	0.012	0.012	0.013	0.01	0.01	0.008	0.012	0.015	1.349
phecode_664-1	Lichen planus, nitidus, or striatus	0.007	0.007	0.008	0.008	0.007	0.009	0.008	0.007	0.008	0.014	2.125
phecode_664-2	Pityriasis	0.003	0.003	0.003	0.003	0.002	0.003	0.001	0.004	0.004	0.003	0.905
phecode_664-21	Pityriasis rosea	0.003	0.001	0.002	0.002	0.002	0.002	0.002	0.003	0.004	0.003	1.053
phecode_665	Psoriasis	0.014	0.017	0.017	0.018	0.018	0.019	0.019	0.021	0.023	0.022	1.506
phecode_665-2	Psoriatic arthropathy	0.001	0.002	0.003	0.002	0.002	0.001	0.001	0.001	0.003	0.002	2
phecode_665-3	Other psoriasis	0.002	0.002	0.002	0.002	0.002	0.003	0.002	0.001	0.003	0.002	1.1
phecode_666	Urticaria	0.017	0.017	0.017	0.018	0.024	0.024	0.022	0.021	0.025	0.025	1.485
phecode_666-1	Allergic urticaria	0.002	0.002	0.002	0.003	0.003	0.003	0.003	0.003	0.004	0.006	3.455
phecode_666-2	Idiopathic urticaria	0.001	0.001	0.001	0.001	0.001	0.002	0.002	0.003	0.002	0.003	2.125
phecode_667	Erythematous conditions	0.006	0.006	0.005	0.006	0.007	0.009	0.007	0.008	0.008	0.009	1.46
phecode_668	Dermatitis [eczema]	0.087	0.091	0.104	0.108	0.108	0.112	0.129	0.123	0.127	0.133	1.537
phecode_668-1	Atopic dermatitis	0.019	0.022	0.022	0.019	0.018	0.018	0.022	0.021	0.022	0.025	1.313
phecode_668-2	Seborrheic dermatitis	0.017	0.018	0.019	0.023	0.022	0.026	0.026	0.029	0.027	0.029	1.687
phecode_668-3	Contact dermatitis	0.017	0.017	0.02	0.021	0.023	0.022	0.023	0.022	0.024	0.026	1.546
phecode_668-4	Dermatitis due to substances taken internally	0.002	0	0.002	0.002	0.003	0.002	0.003	0.002	0.003	0.003	1.616
phecode_668-5	Lichen simplex chronicus	0.003	0.003	0.003	0.003	0.004	0.004	0.004	0.005	0.005	0.007	2.75
phecode_668-6	Prurigo	0.002	0.001	0.002	0.003	0.002	0.002	0.004	0.003	0.003	0.003	1.214

5 Phenome-wide prediction of disease onset from retinal fundus photographs

Table 22 continued from previous page

Endpoint	PheCode string	1	2	3	4	5	6	7	8	9	10	Rate Ratio (RR)
phecode_670	Seborrheic keratosis	0.003	0.007	0.008	0.008	0.013	0.013	0.012	0.014	0.01	0.011	3.095
phecode_672	Other acute skin changes due to ultraviolet radiation	0.002	0.002	0.002	0.002	0.004	0.003	0.002	0.004	0.003	0.003	1.462
phecode_673	Skin changes due to chronic exposure to nonionizing radiation	0.058	0.117	0.143	0.163	0.186	0.206	0.212	0.236	0.237	0.236	4.073
phecode_673-1	Actinic keratosis	0.058	0.116	0.136	0.166	0.182	0.202	0.218	0.229	0.23	0.236	4.063
phecode_674	Disorders of pigmentation	0.018	0.016	0.021	0.02	0.02	0.023	0.023	0.019	0.025	0.029	1.593
phecode_674-1	Hypopigmentation	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.001	0.002	0.007	3.727
phecode_674-11	Vitiligo	0.002	0.003	0.001	0.002	0.001	0.001	0.001	0.001	0.002	0.005	3.201
phecode_674-2	Hyperpigmentation	0.006	0.007	0.006	0.006	0.007	0.008	0.009	0.008	0.011	0.012	1.821
phecode_675	Atrophic conditions of skin	0.004	0.005	0.005	0.009	0.011	0.012	0.013	0.014	0.02	0.023	5.751
phecode_675-1	Circumscribed scleroderma	0.004	0.005	0.005	0.008	0.011	0.012	0.013	0.014	0.02	0.022	5.668
phecode_676	Hypertrophic conditions of skin	0.036	0.034	0.034	0.034	0.036	0.036	0.039	0.039	0.039	0.043	1.192
phecode_676-1	Hypertrophic scar [Keloid scar]	0.001	0.001	0.001	0.002	0.002	0.002	0.001	0.002	0.002	0.006	4.876
phecode_676-2	Scar conditions and fibrosis of skin	0.004	0.007	0.005	0.007	0.006	0.009	0.007	0.008	0.007	0.007	1.519
phecode_678	Other skin and connective tissue disorders	0.131	0.173	0.191	0.208	0.213	0.222	0.227	0.242	0.239	0.258	1.973
phecode_679	Skin symptoms	0.16	0.184	0.197	0.205	0.209	0.219	0.224	0.236	0.236	0.253	1.577
phecode_679-1	Rash and other nonspecific skin eruption	0.117	0.117	0.131	0.127	0.135	0.139	0.148	0.143	0.149	0.157	1.344
phecode_679-2	Pallor and flushing	0.003	0.003	0.004	0.005	0.004	0.006	0.007	0.006	0.004	0.005	1.765
phecode_679-21	Pallor	0.002	0	0.002	0.001	0.002	0.002	0.001	0.003	0.002	0.003	1.455
phecode_679-22	Flushing	0.002	0.002	0.003	0.004	0.003	0.003	0.003	0.004	0.003	0.001	0.818
phecode_679-3	Changes in skin texture	0.004	0.007	0.006	0.008	0.009	0.011	0.012	0.01	0.011	0.017	4.392
phecode_679-4	Pruritus	0.042	0.044	0.06	0.058	0.066	0.068	0.07	0.077	0.077	0.093	2.24
phecode_679-7	Abnormal granulation tissue, NOS	0.001	0.002	0.001	0.003	0.004	0.004	0.005	0.007	0.007	0.006	5.571
phecode_680	Epidermal thickening	0.015	0.018	0.024	0.029	0.033	0.034	0.039	0.038	0.041	0.051	3.363
phecode_680-1	Corns and callosities	0.007	0.009	0.013	0.01	0.015	0.014	0.016	0.017	0.017	0.021	2.841
phecode_680-3	Xerosis cutis*	0.004	0.008	0.009	0.01	0.012	0.011	0.017	0.012	0.02	0.022	5.361
phecode_681	Localized swelling, mass and lump of skin and subcutaneous tissue	0.071	0.083	0.09	0.094	0.104	0.109	0.112	0.118	0.135	0.145	2.035
phecode_682	Other follicular disorders	0.082	0.074	0.07	0.078	0.071	0.078	0.073	0.071	0.083	0.081	0.987
phecode_682-1	Cutaneous cyst	0.047	0.053	0.053	0.051	0.051	0.048	0.05	0.052	0.053	0.051	1.092
phecode_682-11	Sebaceous cyst [Epidermal cyst]	0.037	0.035	0.038	0.035	0.04	0.035	0.036	0.04	0.04	0.039	1.046
phecode_682-12	Pilar and trichodermal cyst	0.045	0.049	0.047	0.041	0.048	0.048	0.045	0.045	0.046	0.046	1.023
phecode_682-4	Acne	0.005	0.003	0.003	0.003	0.003	0.005	0.005	0.006	0.007	0.011	2.207
phecode_683	Nail disorders	0.033	0.035	0.037	0.042	0.044	0.043	0.043	0.044	0.046	0.045	1.384
phecode_683-1	Ingrowing nail	0.011	0.014	0.014	0.015	0.014	0.016	0.014	0.019	0.016	0.017	1.616
phecode_683-2	Nail dystrophy*	0.002	0.003	0.002	0.002	0.001	0.002	0.004	0.004	0.005	0.005	2.819
phecode_684	Diseases of hair and hair follicles	0.003	0.007	0.008	0.012	0.016	0.018	0.021	0.02	0.022	0.035	12.237
phecode_684-1	Alopecia	0.003	0.007	0.007	0.013	0.014	0.017	0.023	0.016	0.023	0.034	10.633
phecode_684-11	Alopecia Areata	0.002	0.003	0.003	0.005	0.005	0.004	0.006	0.008	0.008	0.009	4.501
phecode_685	Disorders of sweat glands	0.017	0.022	0.025	0.017	0.024	0.024	0.031	0.025	0.028	0.032	1.91
phecode_685-1	Dyshidrosis	0.006	0.003	0.004	0.004	0.004	0.007	0.004	0.006	0.005	0.005	0.829
phecode_685-4	Prickly heat and miliaria	0.002	0.002	0.002	0.003	0.002	0.003	0.002	0.003	0.002	0.002	0.833
phecode_685-8	Hyperhidrosis	0.01	0.014	0.017	0.016	0.016	0.019	0.023	0.018	0.019	0.023	2.191
phecode_685-82	Generalized hyperhidrosis	0.01	0.015	0.016	0.015	0.018	0.02	0.021	0.019	0.019	0.023	2.238
phecode_686	Chronic ulcer of skin	0.004	0.004	0.009	0.01	0.015	0.017	0.024	0.03	0.039	0.056	15.412
phecode_686-1	Pressure ulcer	0.001	0.002	0.003	0.004	0.007	0.009	0.012	0.015	0.02	0.03	26.147
phecode_686-2	Non-pressure chronic ulcer	0.003	0.004	0.005	0.006	0.007	0.01	0.016	0.017	0.022	0.034	12.061
phecode_688	Granulomatous disorder of the skin	0.006	0.008	0.007	0.006	0.008	0.008	0.006	0.007	0.008	0.008	1.378
phecode_688-1	Sarcoidosis	0.002	0.001	0.001	0.002	0.001	0.002	0.001	0.002	0.002	0.003	1.5
phecode_688-3	Pyogenic granuloma of skin and subcutaneous tissue	0.001	0.002	0.003	0.002	0.004	0.001	0.003	0.003	0.004	0.003	2.5
phecode_700	Diffuse diseases of connective tissue	0.003	0.004	0.004	0.005	0.004	0.006	0.008	0.007	0.008	0.01	3.471
phecode_700-2	Sicca syndrome [Sjogren syndrome]	0.001	0.001	0.002	0.002	0.002	0.003	0.003	0.002	0.003	0.004	4.601
phecode_701	Osteomyelitis, periostitis, and other infections involving bone	0.002	0.002	0.001	0.001	0.001	0.002	0.002	0.002	0.003	0.007	4.1
phecode_701-1	Osteomyelitis	0.002	0.002	0.001	0.001	0.001	0.002	0.003	0.001	0.003	0.006	3.364
phecode_702	Infective and reactive arthropathies	0.003	0.005	0.004	0.003	0.004	0.003	0.004	0.006	0.004	0.006	1.619
phecode_702-3	Enteropathic arthropathies	0.002	0.001	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.003	1.455
phecode_703	Crystal arthropathies	0.008	0.014	0.022	0.027	0.031	0.037	0.037	0.05	0.057	0.077	9.349
phecode_703-1	Hyperuricemia	0.008	0.013	0.017	0.027	0.029	0.031	0.037	0.047	0.054	0.075	9.512

Supplementary Tables

Table 22 continued from previous page

Endpoint	PheCode string	1	2	3	4	5	6	7	8	9	10	Rate Ratio (RR)
phecode_703-11	Gout	0.007	0.012	0.018	0.025	0.029	0.03	0.036	0.048	0.054	0.072	10.465
phecode_703-2	Chondrocalcinosis	0.001	0	0.001	0.002	0.003	0.003	0.003	0.003	0.005	0.006	8.751
phecode_704	Systemic vasculitis	0.002	0.002	0.002	0.003	0.004	0.006	0.007	0.007	0.008	0.009	4.818
phecode_704-5	Giant cell arteritis	0	0.001	0.001	0.001	0.001	0.003	0.003	0.003	0.006	0.005	31
phecode_705	Rheumatoid arthritis and other inflammatory polyarthropathies	0.012	0.015	0.02	0.023	0.022	0.028	0.032	0.037	0.04	0.046	3.654
phecode_705-1	Rheumatoid arthritis	0.007	0.008	0.01	0.012	0.012	0.011	0.018	0.019	0.017	0.02	2.819
phecode_705-3	Polymyalgia rheumatica	0.002	0.004	0.006	0.008	0.01	0.013	0.016	0.019	0.021	0.024	13.457
phecode_705-5	Rheumatism, unspecified	0.002	0.002	0.003	0.003	0.002	0.002	0.003	0.002	0.003	0.003	1.134
phecode_706	Other inflammatory spondylopathies	0.004	0.006	0.006	0.007	0.008	0.011	0.01	0.011	0.012	0.014	3.584
phecode_706-1	Sacroiliitis NEC	0.001	0.002	0.002	0.001	0.002	0.003	0.002	0.003	0.003	0.006	4.626
phecode_707	Other arthropathies	0.021	0.031	0.039	0.052	0.062	0.069	0.077	0.094	0.11	0.127	6.123
phecode_707-8	Polyarthrits	0.002	0.003	0.003	0.003	0.003	0.007	0.003	0.006	0.006	0.006	2.643
phecode_708	Osteoarthritis	0.098	0.136	0.167	0.2	0.222	0.233	0.258	0.281	0.3	0.318	3.246
phecode_708-1	Primary osteoarthritis	0.028	0.043	0.047	0.055	0.061	0.071	0.078	0.081	0.098	0.096	3.466
phecode_708-11	Primary osteoarthritis of hip, pelvic region and thigh	0.003	0.003	0.007	0.007	0.008	0.012	0.009	0.014	0.013	0.015	4.451
phecode_708-12	Primary osteoarthritis of knee, lower leg	0.007	0.01	0.013	0.015	0.017	0.021	0.027	0.025	0.024	0.025	3.535
phecode_708-13	Primary osteoarthritis of the hand	0.004	0.009	0.014	0.013	0.012	0.02	0.021	0.023	0.024	0.023	5.561
phecode_708-14	Primary osteoarthritis of the shoulder, upper arm	0.004	0.007	0.009	0.005	0.011	0.01	0.011	0.013	0.016	0.014	3.91
phecode_708-15	Primary osteoarthritis of the wrist, forearm	0.001	0.001	0.002	0.002	0.002	0.002	0.003	0.004	0.004	0.004	5.501
phecode_708-16	Primary osteoarthritis ankle and foot	0.007	0.01	0.012	0.012	0.01	0.014	0.015	0.017	0.014	0.016	2.228
phecode_708-7	Generalized osteoarthritis	0.007	0.014	0.015	0.021	0.024	0.028	0.032	0.042	0.04	0.044	6.676
phecode_708-8	Secondary osteoarthritis	0.002	0.003	0.004	0.006	0.007	0.006	0.006	0.008	0.01	0.009	4.001
phecode_708-9	Heberden's or Bouchard's nodes*	0.003	0.003	0.004	0.004	0.006	0.007	0.007	0.008	0.009	0.006	1.95
phecode_709	Acquired deformities of fingers and toes	0.022	0.031	0.029	0.035	0.036	0.043	0.045	0.053	0.055	0.061	2.79
phecode_709-1	Acquired deformities of fingers	0.003	0.003	0.005	0.003	0.002	0.003	0.005	0.004	0.004	0.003	1.188
phecode_709-11	Mallet finger	0.002	0.002	0.002	0.002	0.002	0.003	0.003	0.004	0.002	0.003	1.308
phecode_709-2	Acquired deformities of toe	0.017	0.027	0.03	0.031	0.033	0.038	0.042	0.048	0.054	0.056	3.288
phecode_709-21	Hallux valgus (Bunion)	0.012	0.017	0.024	0.022	0.029	0.027	0.035	0.039	0.041	0.046	3.944
phecode_709-22	Hallux rigidus	0.003	0.004	0.005	0.006	0.005	0.008	0.007	0.006	0.006	0.008	2.882
phecode_709-24	Hammer toe	0.003	0.003	0.003	0.006	0.008	0.007	0.007	0.009	0.012	0.013	5.001
phecode_710	Acquired deformities of limbs	0.008	0.01	0.009	0.013	0.015	0.018	0.017	0.02	0.019	0.022	2.809
phecode_710-3	Acquired deformities of the knee	0.001	0.002	0.002	0.002	0.003	0.003	0.004	0.005	0.006	0.006	4.626
phecode_710-31	Genu valgum (acquired)	0.001	0.001	0.001	0.002	0.001	0.001	0.002	0.003	0.003	0.001	1.6
phecode_710-32	Genu varum (acquired)	0.001	0	0.001	0.001	0.002	0.002	0.002	0.003	0.003	0.003	2.834
phecode_710-4	Acquired deformities of the ankle and foot	0.006	0.005	0.006	0.009	0.011	0.01	0.011	0.011	0.014	0.016	2.795
phecode_710-41	Flat foot [pes planus]	0.003	0.004	0.003	0.003	0.004	0.003	0.005	0.004	0.003	0.006	2.236
phecode_711	Disorder of patella	0.021	0.023	0.022	0.024	0.027	0.025	0.026	0.028	0.027	0.031	1.476
phecode_711-1	Derangement of meniscus	0.018	0.022	0.021	0.02	0.02	0.023	0.021	0.021	0.024	0.026	1.446
phecode_712	Other specific joint derangements	0.011	0.013	0.013	0.013	0.013	0.009	0.01	0.014	0.013	0.014	1.299
phecode_712-1	Loose body in joint	0.002	0.002	0.002	0.003	0.003	0.002	0.003	0.002	0.002	0.002	1.5
phecode_712-5	Disorder of ligament	0.001	0.001	0.001	0.002	0.001	0.002	0.001	0.002	0.003	0.001	1
phecode_712-6	Instability of joint	0.004	0.003	0.003	0.003	0.003	0.003	0.004	0.002	0.002	0.003	0.913
phecode_713	Symptoms related to joints	0.304	0.329	0.35	0.355	0.347	0.369	0.353	0.371	0.37	0.405	1.335
phecode_713-2	Effusion of joint	0.02	0.025	0.032	0.035	0.037	0.041	0.051	0.053	0.055	0.065	3.174
phecode_713-3	Pain in joint	0.291	0.322	0.338	0.336	0.349	0.341	0.354	0.342	0.353	0.388	1.332
phecode_713-4	Stiffness of joint	0.007	0.007	0.01	0.008	0.007	0.008	0.008	0.008	0.009	0.009	1.303
phecode_714	Deforming dorsopathies	0.004	0.003	0.005	0.006	0.007	0.011	0.011	0.012	0.016	0.02	4.881
phecode_714-3	Scoliosis	0.002	0.002	0.003	0.005	0.005	0.008	0.009	0.011	0.011	0.016	6.534
phecode_715	Non-inflammatory spondylopathy	0.027	0.04	0.051	0.058	0.062	0.071	0.088	0.093	0.108	0.118	4.319
phecode_715-1	Spondylolysis	0.022	0.031	0.042	0.047	0.052	0.061	0.066	0.075	0.08	0.095	4.279
phecode_715-3	Spondylolisthesis	0.002	0.003	0.005	0.005	0.004	0.008	0.01	0.012	0.008	0.014	6.386
phecode_715-4	Spinal stenosis	0.006	0.01	0.014	0.013	0.017	0.022	0.024	0.03	0.033	0.038	6.159
phecode_716	Intervertebral disc disorder	0.029	0.033	0.033	0.036	0.035	0.041	0.043	0.042	0.045	0.053	1.841
phecode_716-2	Degenerative disc disease	0.009	0.01	0.01	0.011	0.01	0.013	0.011	0.013	0.014	0.015	1.704
phecode_716-3	Spinal disc displacement (herniation)	0.019	0.02	0.025	0.026	0.026	0.027	0.024	0.031	0.033	0.037	1.982
phecode_717	Other and unspecified dorsopathies	0.011	0.012	0.011	0.011	0.013	0.013	0.014	0.017	0.013	0.016	1.418
phecode_717-2	Sacrococcygeal disorders	0.006	0.006	0.005	0.008	0.008	0.006	0.009	0.009	0.011	0.012	2.206
phecode_718	Back pain	0.235	0.25	0.251	0.262	0.264	0.263	0.276	0.285	0.284	0.332	1.413

5 Phenome-wide prediction of disease onset from retinal fundus photographs

Table 22 continued from previous page

Endpoint	PheCode string	1	2	3	4	5	6	7	8	9	10	Rate Ratio (RR)
phecode_718-1	Radiculopathy	0.011	0.012	0.017	0.018	0.02	0.025	0.023	0.023	0.023	0.027	2.485
phecode_718-2	Cervicalgia	0.055	0.067	0.067	0.073	0.078	0.077	0.081	0.089	0.091	0.112	2.046
phecode_718-3	Mid back pain	0.008	0.009	0.01	0.012	0.01	0.012	0.013	0.012	0.011	0.015	1.76
phecode_718-4	Low back pain	0.147	0.159	0.155	0.151	0.174	0.163	0.158	0.173	0.183	0.216	1.47
phecode_718-5	Sciatica	0.05	0.051	0.053	0.057	0.057	0.064	0.065	0.075	0.076	0.082	1.654
phecode_719	Disorders of muscle	0.022	0.027	0.026	0.034	0.039	0.04	0.039	0.047	0.052	0.059	2.7
phecode_719-1	Cramp and spasm	0.015	0.017	0.018	0.026	0.027	0.027	0.027	0.03	0.034	0.045	3.069
phecode_719-11	Spasm of muscle	0.003	0.001	0.002	0.003	0.002	0.002	0.003	0.003	0.002	0.002	0.588
phecode_719-3	Separation of muscle (nontraumatic)	0.001	0.001	0	0.001	0.002	0.002	0.004	0.006	0.006	0.006	5.287
phecode_719-7	Muscle weakness (generalized)	0.003	0.002	0.003	0.003	0.004	0.004	0.004	0.005	0.005	0.007	2
phecode_720	Spontaneous rupture of synovium and tendon	0.005	0.005	0.005	0.007	0.006	0.007	0.005	0.009	0.008	0.012	2.467
phecode_721	Synoviopathy and bursopathy	0.115	0.122	0.136	0.142	0.143	0.127	0.156	0.15	0.151	0.15	1.306
phecode_721-1	Synovitis and tenosynovitis	0.046	0.052	0.062	0.057	0.066	0.065	0.07	0.07	0.068	0.08	1.748
phecode_721-11	Trigger finger	0.012	0.016	0.024	0.021	0.023	0.024	0.031	0.028	0.029	0.029	2.352
phecode_721-12	Radial styloid tenosynovitis [de Quervain]	0.004	0.003	0.006	0.005	0.005	0.007	0.007	0.009	0.009	0.009	2.077
phecode_721-2	Ganglion cyst	0.023	0.026	0.026	0.025	0.028	0.03	0.029	0.03	0.031	0.03	1.302
phecode_721-4	Calcium deposits in tendon and bursa	0.001	0.003	0.001	0.003	0.003	0.002	0.002	0.002	0.003	0.003	1.889
phecode_721-5	Bursitis	0.04	0.044	0.044	0.053	0.059	0.053	0.06	0.058	0.06	0.054	1.359
phecode_721-6	Baker's cyst [popliteal cyst]	0.009	0.009	0.013	0.01	0.014	0.014	0.016	0.015	0.017	0.016	1.846
phecode_722	Fasciopathy	0.067	0.065	0.068	0.072	0.07	0.074	0.077	0.078	0.082	0.094	1.41
phecode_722-1	Plantar fascial fibromatosis [Plantar fasciitis]	0.047	0.046	0.049	0.045	0.054	0.054	0.055	0.059	0.064	0.08	1.699
phecode_722-4	Palmar fascial fibromatosis [Dupuytren]	0.007	0.011	0.012	0.016	0.021	0.024	0.026	0.027	0.029	0.03	4.576
phecode_723	Enthesopathy/Enthesitis/Tendinopathy	0.122	0.134	0.133	0.128	0.136	0.132	0.149	0.158	0.164	0.172	1.412
phecode_723-1	Adhesive capsulitis of shoulder	0.015	0.021	0.021	0.022	0.022	0.024	0.019	0.021	0.024	0.026	1.718
phecode_723-2	Rotator cuff tear or rupture	0.025	0.029	0.036	0.036	0.039	0.046	0.041	0.041	0.039	0.049	1.94
phecode_723-3	Medial epicondylitis (Golfer's elbow)	0.008	0.009	0.01	0.012	0.012	0.013	0.012	0.01	0.013	0.01	1.341
phecode_723-4	Lateral epicondylitis (Tennis elbow)	0.018	0.023	0.023	0.027	0.034	0.039	0.043	0.049	0.065	0.067	3.785
phecode_723-5	Tendinitis	0.023	0.025	0.023	0.026	0.022	0.024	0.027	0.025	0.026	0.028	1.241
phecode_723-51	Achilles tendinitis	0.018	0.02	0.021	0.021	0.02	0.021	0.019	0.023	0.024	0.025	1.364
phecode_723-6	Impingement syndrome of shoulder*	0.026	0.022	0.032	0.028	0.032	0.03	0.035	0.031	0.037	0.034	1.31
phecode_724	Other symptoms and disorders of the soft tissue	0.083	0.088	0.095	0.115	0.115	0.127	0.128	0.138	0.147	0.158	1.909
phecode_724-1	Myalgia	0.034	0.041	0.046	0.047	0.051	0.054	0.057	0.06	0.063	0.064	1.853
phecode_724-3	Nontraumatic hematoma of soft tissue	0.002	0.003	0.003	0.004	0.004	0.003	0.004	0.005	0.006	0.005	2.308
phecode_724-5	Exostosis	0.008	0.01	0.01	0.009	0.011	0.011	0.014	0.015	0.016	0.015	1.894
phecode_724-51	Calcaneal spur	0.003	0.002	0.001	0.002	0.003	0.003	0.003	0.003	0.002	0.002	0.875
phecode_724-52	Osteophyte*	0.006	0.006	0.008	0.008	0.008	0.009	0.01	0.01	0.015	0.012	2.086
phecode_726	Osteoporosis and low bone density	0.008	0.015	0.023	0.026	0.033	0.041	0.059	0.076	0.082	0.11	13.492
phecode_726-1	Osteoporosis	0.007	0.014	0.022	0.024	0.032	0.04	0.059	0.071	0.08	0.11	15.305
phecode_726-2	Pathologic fracture	0.001	0.001	0.001	0.003	0.003	0.004	0.007	0.007	0.009	0.012	15.002
phecode_727	Other disorders of bone	0.02	0.035	0.035	0.047	0.059	0.07	0.087	0.101	0.121	0.146	7.352
phecode_727-1	Osteonecrosis	0	0.002	0.002	0.001	0.002	0.001	0.002	0.002	0.003	0.002	5.001
phecode_728	Chondropathies	0.014	0.013	0.015	0.013	0.014	0.014	0.016	0.018	0.017	0.029	2.012
phecode_728-1	Chondromalacia	0.002	0.002	0.001	0.001	0.002	0.003	0.002	0.003	0.004	0.004	1.715
phecode_728-3	Costochondritis (Tietze's disease)	0.011	0.009	0.012	0.01	0.011	0.009	0.011	0.013	0.015	0.02	1.938
phecode_730	Other disorders and symptoms of the musculoskeletal system	0.006	0.007	0.008	0.007	0.008	0.006	0.006	0.009	0.009	0.008	1.325
phecode_731	Symptoms involving musculoskeletal systems	0.022	0.022	0.021	0.023	0.024	0.023	0.023	0.025	0.022	0.027	1.233
phecode_732	Nonspecific abnormal findings on radiological and other examination of musculoskeletal system	0.011	0.013	0.015	0.021	0.024	0.028	0.027	0.034	0.034	0.043	3.881
phecode_733	Dentofacial anomalies, including malocclusion	0.009	0.009	0.011	0.013	0.014	0.018	0.013	0.017	0.018	0.02	2.088
phecode_733-6	Temporomandibular joint disorders	0.01	0.008	0.008	0.012	0.015	0.019	0.013	0.017	0.018	0.019	1.855
phecode_733-62	Arthralgia of temporomandibular joint	0.002	0.001	0.002	0.002	0.003	0.003	0.003	0.005	0.005	0.004	1.857
phecode_734	Diseases of the jaws	0.006	0.005	0.006	0.006	0.006	0.006	0.008	0.004	0.007	0.01	1.632
phecode_734-9	Jaw pain	0.005	0.003	0.006	0.006	0.007	0.006	0.006	0.005	0.006	0.009	1.839
phecode_800	Chest pain	0.114	0.132	0.133	0.14	0.139	0.152	0.15	0.155	0.157	0.179	1.572
phecode_800-1	Chest pain on breathing	0.014	0.017	0.014	0.019	0.018	0.019	0.017	0.019	0.019	0.021	1.477

Supplementary Tables

Table 22 continued from previous page

Endpoint	PheCode string	1	2	3	4	5	6	7	8	9	10	Rate Ratio (RR)
phecode_800-11	Pleurodynia*	0.014	0.017	0.015	0.018	0.016	0.017	0.017	0.02	0.018	0.018	1.233
phecode_800-2	Precordial pain	0.012	0.013	0.013	0.017	0.017	0.016	0.017	0.016	0.018	0.021	1.806
phecode_801	Cough	0.217	0.226	0.246	0.251	0.264	0.267	0.286	0.293	0.291	0.307	1.415
phecode_802	Throat pain	0.019	0.019	0.025	0.025	0.029	0.02	0.023	0.024	0.025	0.029	1.56
phecode_803	Snoring*	0.005	0.006	0.007	0.007	0.007	0.01	0.012	0.007	0.012	0.016	2.879
phecode_804	Other symptoms and signs involving the circulatory and respiratory system	0.068	0.073	0.085	0.079	0.092	0.096	0.11	0.103	0.116	0.132	1.945
phecode_805	Fever of unknown origin	0.018	0.023	0.022	0.027	0.026	0.032	0.034	0.034	0.039	0.042	2.34
phecode_807	Malaise and fatigue	0.089	0.095	0.105	0.111	0.117	0.127	0.13	0.129	0.145	0.159	1.779
phecode_807-1	Chronic fatigue syndrome	0.031	0.039	0.041	0.038	0.044	0.043	0.055	0.05	0.06	0.074	2.391
phecode_807-11	Postviral fatigue syndrome*	0.002	0.001	0.002	0.002	0.002	0.002	0.002	0.003	0.004	0.005	2
phecode_808	Syncope and collapse	0.029	0.036	0.034	0.046	0.05	0.049	0.059	0.066	0.068	0.073	2.515
phecode_809	Pain	0.185	0.2	0.205	0.215	0.215	0.223	0.223	0.225	0.237	0.269	1.457
phecode_809-1	Acute pain	0.001	0.002	0.002	0.001	0.003	0.001	0.002	0.002	0.003	0.002	2
phecode_809-3	Pain in limb	0.163	0.179	0.183	0.186	0.192	0.188	0.194	0.198	0.209	0.233	1.431
phecode_810	Shock	0	0	0.001	0.002	0.001	0.002	0.003	0.003	0.004	0.008	24.004
phecode_812	Edema	0.021	0.028	0.03	0.043	0.049	0.055	0.064	0.068	0.078	0.097	4.54
phecode_812-2	Angioneurotic edema	0.002	0.003	0.002	0.003	0.003	0.003	0.003	0.003	0.002	0.003	1.5
phecode_814	Jaundice (not of newborn)	0.001	0.001	0.003	0.002	0.004	0.004	0.005	0.005	0.006	0.005	5.334
phecode_815	Symptoms and signs concerning food and fluid intake	0.007	0.007	0.01	0.009	0.008	0.008	0.011	0.009	0.01	0.013	1.878
phecode_817	Motion sickness	0	0.001	0.001	0.002	0.002	0.002	0.002	0.002	0.002	0.003	6.001
phecode_819	General symptoms and other findings	0.113	0.131	0.154	0.153	0.18	0.208	0.213	0.225	0.256	0.295	2.618
phecode_820	Elevated erythrocyte sedimentation rate and abnormality of plasma viscosity	0.001	0	0.001	0.001	0.002	0.002	0.002	0.003	0.002	0.004	4.601
phecode_821	Abnormality of red blood cells	0.004	0.004	0.002	0.005	0.003	0.004	0.005	0.005	0.006	0.008	2.228
phecode_823	Abnormal serum enzyme levels	0.034	0.037	0.043	0.044	0.044	0.048	0.047	0.051	0.053	0.063	1.858
phecode_823-2	Abnormal levels of other serum enzymes	0.034	0.038	0.042	0.041	0.043	0.05	0.045	0.051	0.052	0.063	1.882
phecode_824	Other abnormalities of plasma proteins*	0.008	0.008	0.009	0.007	0.01	0.01	0.008	0.009	0.01	0.015	1.88
phecode_826	Other abnormal immunological findings in serum	0.004	0.004	0.004	0.004	0.004	0.006	0.003	0.003	0.003	0.009	2.077
phecode_826-3	Raised antibody titer*	0.003	0.002	0.004	0.003	0.003	0.003	0.003	0.004	0.003	0.008	2.648
phecode_827	Toxicology findings	0.003	0.006	0.006	0.007	0.008	0.01	0.01	0.01	0.012	0.013	3.851
phecode_827-1	Finding of alcohol in blood	0.003	0.005	0.006	0.007	0.008	0.008	0.01	0.01	0.011	0.012	3.895
phecode_829	Nonspecific findings on examination of blood	0.037	0.038	0.047	0.056	0.059	0.074	0.085	0.093	0.113	0.139	3.721
phecode_829-2	Abnormal level of blood mineral*	0.007	0.007	0.009	0.009	0.009	0.01	0.009	0.011	0.012	0.014	1.934
phecode_830	Proteinuria	0.055	0.071	0.083	0.093	0.094	0.104	0.114	0.119	0.136	0.153	2.779
phecode_831	Glycosuria	0.005	0.011	0.012	0.016	0.016	0.019	0.018	0.023	0.03	0.038	6.909
phecode_832	Other abnormal findings in urine	0.059	0.091	0.104	0.121	0.13	0.135	0.149	0.178	0.185	0.212	3.604
phecode_832-5	Acetonuria	0.01	0.011	0.014	0.012	0.014	0.015	0.017	0.018	0.017	0.018	1.672
phecode_832-6	Pyuria*	0.004	0.008	0.009	0.008	0.012	0.011	0.011	0.013	0.016	0.016	3.519
phecode_835	Cytology and pathology findings	0.022	0.024	0.021	0.028	0.034	0.039	0.052	0.055	0.081	0.138	6.335
phecode_840	Allergy	0.073	0.082	0.09	0.095	0.087	0.097	0.099	0.106	0.105	0.117	1.614
phecode_840-1	Food allergy	0.006	0.005	0.005	0.006	0.007	0.009	0.006	0.008	0.008	0.01	1.736
phecode_840-2	Allergy to insects	0.017	0.02	0.021	0.023	0.025	0.026	0.024	0.023	0.029	0.029	1.743
phecode_840-8	Allergies related to other diseases/symptoms	0.043	0.046	0.051	0.051	0.053	0.054	0.05	0.055	0.056	0.081	1.885
phecode_840-9	Anaphylactic reaction	0.002	0.002	0.002	0.002	0.002	0.002	0.001	0.003	0.002	0.002	1
phecode_841	Drug and medical agent allergy	0.058	0.076	0.086	0.098	0.101	0.122	0.132	0.147	0.155	0.183	3.173
phecode_841-1	Allergy to other anti-infective agents	0.035	0.052	0.054	0.058	0.065	0.066	0.081	0.082	0.093	0.104	2.991
phecode_841-11	Penicillin allergy	0.029	0.042	0.041	0.052	0.052	0.059	0.065	0.062	0.069	0.079	2.698
phecode_841-12	Allergy to antibiotic agent (excluding penicillin)	0.01	0.012	0.012	0.015	0.017	0.025	0.02	0.027	0.025	0.039	3.747
phecode_841-13	Allergy to sulfonamides	0	0.001	0.003	0.001	0.003	0.002	0.003	0.005	0.005	0.006	17.503
phecode_841-3	Allergy to narcotic agent	0.003	0.005	0.006	0.007	0.008	0.011	0.011	0.015	0.017	0.022	7.053
phecode_841-4	Allergy to analgesic agent	0.009	0.012	0.016	0.014	0.02	0.023	0.026	0.03	0.035	0.039	4.577
phecode_841-5	Allergy to serum and vaccine	0.001	0	0.001	0.002	0.001	0.001	0.002	0.003	0.003	0.002	1.875
phecode_848	Nonspecific abnormal findings of other body structures	0.022	0.027	0.029	0.036	0.047	0.052	0.05	0.059	0.068	0.081	3.741

5 Phenome-wide prediction of disease onset from retinal fundus photographs

Table 22 continued from previous page

Endpoint	PheCode string	1	2	3	4	5	6	7	8	9	10	Rate Ratio (RR)
phecode_848-2	Nonspecific abnormal findings on radiological and other examination of other intrathoracic organs (echo)	0.005	0.006	0.007	0.012	0.012	0.017	0.021	0.023	0.031	0.038	8.251
phecode_969	Adverse effects of agents primarily affecting gastrointestinal system	0.04	0.044	0.051	0.062	0.069	0.073	0.077	0.086	0.086	0.101	2.536
phecode_973	Adverse effect of other drug	0.001	0.001	0.001	0.001	0.003	0.003	0.003	0.004	0.005	0.008	7.43
phecode_977	Long term (current) drug therapy	0.065	0.072	0.077	0.084	0.093	0.096	0.102	0.111	0.126	0.168	2.609
phecode_977-4	Long term (current) use of steroids	0.005	0.006	0.007	0.008	0.007	0.009	0.008	0.01	0.01	0.013	2.794
phecode_977-41	Long term (current) use of inhaled steroids*	0.005	0.006	0.007	0.008	0.008	0.008	0.01	0.008	0.011	0.013	2.725
phecode_977-5	Long term (current) use of agents affecting hormones	0	0.001	0.002	0.002	0.004	0.005	0.005	0.014	0.023	0.057	167.5
phecode_977-51	Long term (current) use of hormonal contraceptives	0	0	0.001	0	0.001	0.001	0.004	0.008	0.016	0.035	112.035
phecode_977-52	Hormone replacement therapy (postmenopausal)	0.003	0.005	0.005	0.006	0.011	0.013	0.016	0.026	0.036	0.043	12.458
phecode_977-7	Long term (current) use of insulin or oral hypoglycemic drugs	0.009	0.015	0.02	0.026	0.026	0.035	0.039	0.049	0.056	0.096	10.474
phecode_977-71	Long term (current) use of insulin	0.002	0.002	0.004	0.004	0.006	0.006	0.006	0.009	0.012	0.026	13.419
phecode_977-72	Long term (current) use of oral hypoglycemic drugs	0.008	0.014	0.02	0.025	0.026	0.033	0.037	0.047	0.052	0.093	11.002
phecode_979	Transplated organ	0.003	0.002	0.002	0.001	0.001	0.003	0.003	0.003	0.004	0.004	1.625

Supplementary Tables

Table 23: Discriminative Performances for all endpoints of the retinal risk model.

Endpoint	Phecode String	Age+Sex	Age+Sex+Retina	Delta C-Index
phecode_979	Transplated organ	0.547 (0.517, 0.585)	0.406 (0.368, 0.442)	-0.139 (-0.203, -0.096)
phecode_701	Osteomyelitis, periostitis, and other infections involving bone	0.602 (0.56, 0.642)	0.47 (0.425, 0.514)	-0.128 (-0.192, -0.06)
phecode_604-1	Redundant prepuce and phimosis	0.487 (0.459, 0.53)	0.477 (0.441, 0.516)	-0.014 (-0.044, 0.014)
phecode_604-3	Peyronie's disease	0.499 (0.468, 0.525)	0.489 (0.461, 0.515)	-0.012 (-0.043, 0.023)
phecode_391-12	Chronic otitis media	0.496 (0.478, 0.52)	0.492 (0.47, 0.512)	-0.007 (-0.029, 0.021)
phecode_604	Disorders of penis	0.51 (0.496, 0.526)	0.499 (0.483, 0.513)	-0.012 (-0.022, -0.001)
phecode_329-4	Other specified cognitive deficit	0.515 (0.489, 0.542)	0.5 (0.481, 0.522)	-0.016 (-0.042, 0.016)
phecode_623	Hypertrophy of female genital organs	0.505 (0.493, 0.52)	0.501 (0.484, 0.515)	-0.005 (-0.019, 0.007)
phecode_701-1	Osteomyelitis	0.574 (0.547, 0.616)	0.508 (0.464, 0.561)	-0.072 (-0.136, 0.003)
phecode_682	Other follicular disorders	0.512 (0.505, 0.519)	0.508 (0.502, 0.514)	-0.004 (-0.008, 0.001)
phecode_685-4	Prickly heat and miliaria	0.525 (0.491, 0.56)	0.509 (0.476, 0.544)	-0.018 (-0.038, 0.01)
phecode_215	Testicular dysfunction	0.515 (0.465, 0.567)	0.509 (0.458, 0.563)	-0.004 (-0.05, 0.04)
phecode_391-21	Eustachian salpingitis	0.521 (0.501, 0.562)	0.511 (0.481, 0.541)	-0.014 (-0.037, 0.009)
phecode_486-21	Bronchospasm	0.543 (0.507, 0.576)	0.518 (0.478, 0.552)	-0.028 (-0.058, 0.009)
phecode_139-53	Lipoma of other skin subcutaneous tissue	0.526 (0.508, 0.54)	0.517 (0.492, 0.537)	-0.01 (-0.025, 0.011)
phecode_603-5	Orchitis and epididymitis	0.518 (0.502, 0.539)	0.517 (0.498, 0.532)	-0.002 (-0.019, 0.014)
phecode_723-5	Tendinitis	0.523 (0.512, 0.534)	0.517 (0.506, 0.53)	-0.006 (-0.013, 0.002)
phecode_331-3	Headache syndromes, non migraine	0.552 (0.506, 0.583)	0.519 (0.478, 0.557)	-0.033 (-0.07, -0.003)
phecode_604-5	Balanoposthitis	0.526 (0.503, 0.547)	0.519 (0.499, 0.542)	-0.006 (-0.018, 0.006)
phecode_619-5	Noninflammatory disorders of vulva and perineum	0.531 (0.519, 0.544)	0.52 (0.505, 0.537)	-0.011 (-0.018, -0.001)
phecode_522-11	Crohn's disease	0.537 (0.507, 0.58)	0.518 (0.491, 0.553)	-0.018 (-0.046, 0.008)
phecode_179-9	Immunodeficiency NOS	0.543 (0.502, 0.579)	0.52 (0.477, 0.571)	-0.021 (-0.057, 0.025)
phecode_685-1	Dyshidrosis	0.537 (0.502, 0.563)	0.523 (0.492, 0.549)	-0.013 (-0.031, 0.003)
phecode_475-5	Exercise induced bronchospasm	0.54 (0.5, 0.582)	0.524 (0.484, 0.559)	-0.02 (-0.063, 0.032)
phecode_329-42	Cognitive communication deficit	0.549 (0.526, 0.58)	0.523 (0.501, 0.555)	-0.027 (-0.05, -0.006)
phecode_525	Intestinal malabsorption	0.534 (0.515, 0.553)	0.525 (0.504, 0.542)	-0.009 (-0.026, 0.008)
phecode_105-1	Malignant neoplasm of the breast, female	0.533 (0.518, 0.546)	0.525 (0.513, 0.537)	-0.009 (-0.017, -0.001)
phecode_721-4	Calcium deposits in tendon and bursa	0.543 (0.508, 0.586)	0.527 (0.491, 0.569)	-0.019 (-0.051, 0.026)
phecode_352-3	Parageusia*	0.537 (0.505, 0.569)	0.523 (0.495, 0.562)	-0.01 (-0.04, 0.013)
phecode_622-1	Polyp of corpus uteri	0.536 (0.52, 0.557)	0.525 (0.508, 0.549)	-0.011 (-0.021, 0.003)
phecode_500-4	Disturbances in tooth eruption	0.569 (0.532, 0.599)	0.525 (0.49, 0.565)	-0.042 (-0.064, -0.019)
phecode_389-1	Ocular pain	0.538 (0.518, 0.559)	0.528 (0.5, 0.548)	-0.011 (-0.034, 0.009)
phecode_463-4	Nasal congestion*	0.536 (0.525, 0.546)	0.527 (0.518, 0.537)	-0.009 (-0.016, 0.002)
phecode_683	Nail disorders	0.533 (0.527, 0.541)	0.528 (0.521, 0.536)	-0.007 (-0.011, -0.001)
phecode_500	Disorders of tooth development	0.573 (0.54, 0.619)	0.527 (0.499, 0.57)	-0.046 (-0.063, -0.029)
phecode_504-32	Chronic periodontitis	0.548 (0.509, 0.581)	0.527 (0.492, 0.569)	-0.02 (-0.051, 0.007)
phecode_367-2	Keratitis	0.536 (0.503, 0.556)	0.53 (0.502, 0.552)	-0.006 (-0.029, 0.016)
phecode_608	Other disorders of male genital organs	0.53 (0.524, 0.539)	0.529 (0.524, 0.538)	-0.001 (-0.003, 0.002)
phecode_683-1	Ingrowing nail	0.534 (0.523, 0.548)	0.531 (0.517, 0.545)	-0.004 (-0.012, 0.009)
phecode_379-2	Eye infection, viral	0.546 (0.517, 0.571)	0.533 (0.512, 0.558)	-0.011 (-0.031, 0.009)
phecode_721	Synoviopathy and bursopathy	0.535 (0.53, 0.54)	0.535 (0.53, 0.54)	0 (-0.003, 0.002)
phecode_624-2	Atrophy of female genital tract	0.562 (0.535, 0.591)	0.532 (0.507, 0.567)	-0.027 (-0.049, -0.009)
phecode_394-22	Vestibular neuronitis	0.549 (0.526, 0.574)	0.537 (0.509, 0.56)	-0.012 (-0.039, 0.007)
phecode_712	Other specific joint derangements	0.539 (0.521, 0.552)	0.536 (0.519, 0.549)	-0.003 (-0.014, 0.007)
phecode_723-51	Achilles tendinitis	0.539 (0.525, 0.549)	0.536 (0.523, 0.547)	-0.002 (-0.008, 0.004)
phecode_056-1	Plantar wart	0.542 (0.527, 0.555)	0.538 (0.522, 0.551)	-0.005 (-0.01, 0)
phecode_504-3	Periodontitis	0.54 (0.513, 0.57)	0.539 (0.501, 0.574)	-0.003 (-0.036, 0.027)
phecode_522-1	Inflammatory bowel disease	0.546 (0.53, 0.574)	0.538 (0.522, 0.558)	-0.009 (-0.02, 0.007)
phecode_530-1	Anal fissure	0.555 (0.536, 0.57)	0.54 (0.52, 0.554)	-0.015 (-0.024, -0.005)

5 Phenome-wide prediction of disease onset from retinal fundus photographs

Table 23 continued from previous page

Endpoint	Phecode String	Age+Sex	Age+Sex+Retina	Delta C-Index
phecode_379	Infection of the eye	0.546 (0.527, 0.567)	0.543 (0.519, 0.563)	-0.005 (-0.025, 0.018)
phecode_613	Other nonmalignant breast conditions	0.544 (0.537, 0.554)	0.542 (0.533, 0.552)	-0.003 (-0.006, 0.001)
phecode_580-3	Nephrotic syndrome	0.553 (0.523, 0.58)	0.543 (0.516, 0.577)	-0.011 (-0.036, 0.021)
phecode_106	Gynecological malignant neoplasms	0.551 (0.526, 0.57)	0.543 (0.519, 0.564)	-0.007 (-0.024, 0.009)
phecode_702	Infective and reactive arthropathies	0.556 (0.526, 0.581)	0.543 (0.518, 0.576)	-0.011 (-0.029, 0.007)
phecode_514-21	Impaction of intestine	0.571 (0.531, 0.62)	0.542 (0.506, 0.591)	-0.032 (-0.067, 0.008)
phecode_721-2	Ganglion cyst	0.547 (0.537, 0.557)	0.544 (0.534, 0.554)	-0.002 (-0.007, 0.003)
phecode_463-1	Chronic rhinitis	0.549 (0.537, 0.561)	0.545 (0.534, 0.556)	-0.004 (-0.009, 0.003)
phecode_362	Other disorders of the eyelids	0.553 (0.538, 0.567)	0.545 (0.53, 0.564)	-0.006 (-0.019, 0.004)
phecode_619-2	Disorders of uterus, NEC	0.551 (0.531, 0.573)	0.546 (0.525, 0.572)	-0.004 (-0.019, 0.011)
phecode_334	Disorders of other cranial nerves	0.55 (0.536, 0.568)	0.547 (0.533, 0.563)	-0.003 (-0.014, 0.014)
phecode_603-1	Hydrocele	0.555 (0.53, 0.584)	0.547 (0.523, 0.573)	-0.008 (-0.021, 0.005)
phecode_104	Malignant sarcoma-related cancers	0.552 (0.524, 0.582)	0.546 (0.519, 0.575)	-0.005 (-0.029, 0.028)
phecode_518	Appendicitis	0.557 (0.532, 0.585)	0.549 (0.524, 0.579)	-0.01 (-0.029, 0.008)
phecode_668-4	Dermatitis due to substances taken internally	0.564 (0.531, 0.601)	0.554 (0.505, 0.584)	-0.015 (-0.058, 0.019)
phecode_613-5	Mastodynia	0.554 (0.543, 0.564)	0.55 (0.54, 0.563)	-0.003 (-0.007, 0.001)
phecode_601-11	Acute prostatitis	0.559 (0.532, 0.591)	0.553 (0.514, 0.588)	-0.009 (-0.046, 0.019)
phecode_355-1	Coma	0.577 (0.548, 0.608)	0.553 (0.517, 0.594)	-0.024 (-0.044, -0.003)
phecode_106-21	Malignant neoplasm of endometrium	0.566 (0.533, 0.592)	0.554 (0.529, 0.576)	-0.011 (-0.035, 0.007)
phecode_349-2	Abnormal results of function studies of peripheral nervous system	0.58 (0.556, 0.611)	0.555 (0.531, 0.586)	-0.027 (-0.045, -0.008)
phecode_362-5	Cysts of eyelid	0.56 (0.526, 0.592)	0.553 (0.526, 0.59)	-0.004 (-0.038, 0.018)
phecode_462	Sinusitis	0.558 (0.551, 0.565)	0.556 (0.549, 0.562)	-0.002 (-0.005, 0.001)
phecode_664-2	Pityriasis	0.588 (0.556, 0.626)	0.556 (0.527, 0.594)	-0.033 (-0.05, -0.014)
phecode_308-5	Nervousness	0.561 (0.531, 0.591)	0.553 (0.526, 0.591)	-0.002 (-0.035, 0.022)
phecode_618-12	Rectocele	0.558 (0.538, 0.578)	0.558 (0.54, 0.573)	0 (-0.011, 0.01)
phecode_500-41	Impacted teeth*	0.593 (0.553, 0.632)	0.559 (0.516, 0.599)	-0.034 (-0.054, -0.009)
phecode_624	Symptoms involving female genital tract	0.57 (0.553, 0.589)	0.558 (0.537, 0.581)	-0.011 (-0.029, 0.002)
phecode_052	Herpesvirus	0.561 (0.555, 0.566)	0.558 (0.553, 0.565)	-0.002 (-0.005, 0)
phecode_342-4	Monoplegia	0.561 (0.526, 0.596)	0.56 (0.53, 0.591)	-0.003 (-0.03, 0.025)
phecode_723-3	Medial epicondylitis (Golfer's elbow)	0.562 (0.546, 0.578)	0.559 (0.545, 0.577)	-0.003 (-0.012, 0.008)
phecode_614-52	Vaginitis and vulvovaginitis	0.561 (0.548, 0.574)	0.56 (0.549, 0.573)	-0.002 (-0.009, 0.008)
phecode_360	Inflammation of eyelids	0.564 (0.557, 0.571)	0.561 (0.554, 0.568)	-0.003 (-0.006, -0.001)
phecode_614-1	Pelvic peritoneal adhesions, female (postoperative) (postinfection)	0.575 (0.539, 0.61)	0.563 (0.53, 0.595)	-0.012 (-0.037, 0.009)
phecode_614-5	Inflammatory disease of cervix, vagina, and vulva	0.562 (0.548, 0.577)	0.562 (0.55, 0.572)	0 (-0.007, 0.007)
phecode_601-1	Prostatitis	0.576 (0.555, 0.592)	0.564 (0.544, 0.581)	-0.012 (-0.019, -0.005)
phecode_724-3	Nontraumatic hematoma of soft tissue	0.569 (0.535, 0.597)	0.565 (0.528, 0.593)	-0.003 (-0.026, 0.015)
phecode_601	Inflammatory diseases of prostate	0.575 (0.556, 0.593)	0.565 (0.542, 0.582)	-0.01 (-0.017, -0.003)
phecode_622	Polyp of female genital organs	0.571 (0.558, 0.583)	0.565 (0.551, 0.578)	-0.006 (-0.013, 0.002)
phecode_346	Brain damage and brain death	0.568 (0.532, 0.602)	0.565 (0.526, 0.6)	-0.003 (-0.036, 0.025)
phecode_308-1	Irritability	0.575 (0.535, 0.615)	0.568 (0.532, 0.603)	-0.01 (-0.038, 0.015)
phecode_613-1	Inflammatory disease of breast	0.582 (0.556, 0.616)	0.567 (0.537, 0.597)	-0.015 (-0.027, -0.005)
phecode_672	Other acute skin changes due to ultraviolet radiation	0.59 (0.551, 0.627)	0.565 (0.533, 0.603)	-0.022 (-0.047, 0.001)
phecode_621	Endometrial hyperplasia	0.588 (0.548, 0.627)	0.568 (0.53, 0.615)	-0.019 (-0.051, 0.013)
phecode_052-3	Varicella zoster virus	0.57 (0.563, 0.578)	0.569 (0.562, 0.575)	-0.001 (-0.004, 0)
phecode_395	Other diseases of inner ear	0.572 (0.562, 0.583)	0.569 (0.56, 0.579)	-0.004 (-0.007, -0.001)
phecode_510-2	Esophagitis	0.57 (0.563, 0.581)	0.569 (0.561, 0.58)	-0.001 (-0.007, 0.005)

Supplementary Tables

Table 23 continued from previous page

Endpoint	Phecode String	Age+Sex	Age+Sex+Retina	Delta C-Index
phecode_624-1	Dystrophy of female genital tract	0.576 (0.551, 0.602)	0.57 (0.544, 0.597)	-0.005 (-0.034, 0.019)
phecode_395-1	Labyrinthitis	0.576 (0.564, 0.587)	0.571 (0.561, 0.582)	-0.005 (-0.008, 0)
phecode_353	Symptoms and signs involving general sensations and perceptions	0.577 (0.564, 0.595)	0.572 (0.559, 0.588)	-0.006 (-0.014, 0)
phecode_230-22	Underweight	0.593 (0.563, 0.631)	0.572 (0.538, 0.606)	-0.023 (-0.049, 0.003)
phecode_200-4	Thyroiditis	0.607 (0.566, 0.642)	0.574 (0.523, 0.611)	-0.036 (-0.068, 0.001)
phecode_710-41	Flat foot [pes planus]	0.584 (0.559, 0.616)	0.572 (0.537, 0.608)	-0.01 (-0.03, 0.009)
phecode_509-11	Glossodynia	0.591 (0.561, 0.625)	0.573 (0.548, 0.606)	-0.018 (-0.037, 0.002)
phecode_599	Other symptoms/disorders or the urinary system	0.577 (0.571, 0.582)	0.574 (0.568, 0.58)	-0.002 (-0.004, -0.001)
phecode_510	Diseases of esophagus	0.575 (0.568, 0.581)	0.575 (0.568, 0.582)	-0.001 (-0.005, 0.004)
phecode_618	Genital prolapse	0.577 (0.569, 0.587)	0.575 (0.566, 0.586)	-0.003 (-0.006, 0.001)
phecode_463-5	Postnasal drip	0.579 (0.566, 0.591)	0.575 (0.561, 0.589)	-0.004 (-0.011, 0.003)
phecode_509	Diseases of tongue	0.58 (0.568, 0.594)	0.575 (0.563, 0.59)	-0.005 (-0.011, 0.001)
phecode_532-4	Volvulus	0.598 (0.557, 0.644)	0.572 (0.53, 0.622)	-0.023 (-0.044, -0.003)
phecode_520-12	Femoral hernia	0.585 (0.546, 0.627)	0.575 (0.535, 0.614)	-0.011 (-0.039, 0.023)
phecode_462-1	Acute sinusitis	0.579 (0.57, 0.588)	0.576 (0.567, 0.586)	-0.003 (-0.004, -0.001)
phecode_052-32	Herpes zoster	0.578 (0.57, 0.585)	0.576 (0.568, 0.583)	-0.002 (-0.004, 0)
phecode_581-33	Stricture or kinking of ureter	0.609 (0.574, 0.644)	0.579 (0.539, 0.61)	-0.033 (-0.058, -0.015)
phecode_471-3	Hypertrophy of nasal turbinates	0.59 (0.55, 0.625)	0.577 (0.532, 0.613)	-0.012 (-0.036, 0.013)
phecode_618-1	Prolapse of vaginal walls	0.581 (0.57, 0.593)	0.577 (0.567, 0.588)	-0.004 (-0.008, 0.001)
phecode_717	Other and unspecified dorsoopathies	0.578 (0.563, 0.594)	0.578 (0.561, 0.592)	0 (-0.009, 0.009)
phecode_286-21	Major depressive disorder, recurrent	0.587 (0.561, 0.611)	0.577 (0.55, 0.606)	-0.008 (-0.024, 0.008)
phecode_171	Increased white blood cell count	0.595 (0.565, 0.627)	0.579 (0.545, 0.61)	-0.014 (-0.044, 0.006)
phecode_721-11	Trigger finger	0.58 (0.569, 0.593)	0.578 (0.568, 0.59)	-0.002 (-0.008, 0.004)
phecode_721-6	Baker's cyst [popliteal cyst]	0.584 (0.57, 0.598)	0.579 (0.562, 0.593)	-0.004 (-0.011, 0.003)
phecode_520-14	Ventral hernia	0.585 (0.562, 0.599)	0.58 (0.564, 0.594)	-0.004 (-0.016, 0.009)
phecode_619	Noninflammatory female genital disorders	0.581 (0.573, 0.588)	0.58 (0.574, 0.588)	0 (-0.002, 0.002)
phecode_377-2	Conjunctival hemorrhage	0.584 (0.573, 0.595)	0.581 (0.569, 0.592)	-0.004 (-0.006, -0.001)
phecode_712-5	Disorder of ligament	0.61 (0.565, 0.657)	0.582 (0.541, 0.62)	-0.029 (-0.06, -0.006)
phecode_106-3	Malignant neoplasm of the ovary	0.585 (0.558, 0.619)	0.582 (0.546, 0.609)	-0.003 (-0.034, 0.026)
phecode_628-2	Corpus luteum cyst or hematoma	0.589 (0.573, 0.606)	0.583 (0.566, 0.599)	-0.006 (-0.013, -0.001)
phecode_379-21	Infection of the eye, herpes	0.584 (0.54, 0.61)	0.581 (0.548, 0.611)	-0.001 (-0.022, 0.023)
phecode_448-1	Raynaud's syndrome	0.585 (0.568, 0.603)	0.582 (0.565, 0.603)	-0.003 (-0.013, 0.005)
phecode_618-2	Uterine/Uterovaginal prolapse	0.587 (0.572, 0.6)	0.582 (0.57, 0.595)	-0.005 (-0.01, 0.001)
phecode_594	Abnormality of urination	0.584 (0.579, 0.589)	0.583 (0.577, 0.588)	-0.002 (-0.004, 0.001)
phecode_394-4	Abnormal vestibular function study	0.612 (0.565, 0.648)	0.584 (0.546, 0.621)	-0.026 (-0.048, -0.007)
phecode_367-21	Corneal ulcer	0.601 (0.562, 0.655)	0.585 (0.54, 0.627)	-0.018 (-0.035, 0.011)
phecode_444-1	Varicose veins	0.586 (0.578, 0.595)	0.584 (0.576, 0.593)	-0.002 (-0.007, 0.002)
phecode_398-1	Abnormal auditory function study	0.606 (0.581, 0.631)	0.585 (0.558, 0.616)	-0.02 (-0.033, -0.005)
phecode_417	Abnormalities of heart beat	0.588 (0.584, 0.593)	0.586 (0.581, 0.59)	-0.002 (-0.005, 0)
phecode_593	Hematuria	0.588 (0.583, 0.592)	0.586 (0.582, 0.59)	-0.002 (-0.004, 0)
phecode_594-4	Frequency of urination and polyuria	0.587 (0.581, 0.594)	0.586 (0.58, 0.593)	-0.001 (-0.003, 0.001)
phecode_628	Ovarian cyst	0.592 (0.577, 0.612)	0.585 (0.571, 0.605)	-0.007 (-0.012, -0.001)
phecode_509-1	Glossitis	0.602 (0.581, 0.624)	0.585 (0.565, 0.61)	-0.015 (-0.03, -0.002)
phecode_487-1	Epistaxis	0.592 (0.58, 0.605)	0.586 (0.575, 0.599)	-0.006 (-0.011, 0)
phecode_386-2	Diplopia	0.595 (0.576, 0.617)	0.587 (0.563, 0.608)	-0.01 (-0.019, 0.004)
phecode_377	Hemorrhage of the eye	0.587 (0.575, 0.597)	0.587 (0.575, 0.597)	0 (-0.008, 0.007)
phecode_841-5	Allergy to serum and vaccine	0.623 (0.583, 0.677)	0.587 (0.536, 0.638)	-0.034 (-0.066, -0.005)
phecode_618-11	Cystocele	0.59 (0.58, 0.602)	0.588 (0.576, 0.601)	-0.002 (-0.006, 0.002)

5 Phenome-wide prediction of disease onset from retinal fundus photographs

Table 23 continued from previous page

Endpoint	Phecode String	Age+Sex	Age+Sex+Retina	Delta C-Index
phecode_596-2	Overactive bladder	0.597 (0.585, 0.612)	0.587 (0.576, 0.602)	-0.009 (-0.014, -0.003)
phecode_323	Systemic atrophies primarily affecting the central nervous system	0.623 (0.593, 0.662)	0.591 (0.559, 0.627)	-0.036 (-0.055, -0.017)
phecode_280-82	Other psychoactive substance dependence	0.59 (0.548, 0.629)	0.589 (0.545, 0.63)	-0.001 (-0.043, 0.049)
phecode_363-5	Epiphora	0.592 (0.576, 0.607)	0.589 (0.574, 0.605)	-0.003 (-0.015, 0.01)
phecode_444	Venous insufficiency	0.591 (0.585, 0.6)	0.589 (0.582, 0.597)	-0.002 (-0.005, 0.001)
phecode_488-8	Mouth breathing*	0.598 (0.583, 0.615)	0.589 (0.574, 0.605)	-0.009 (-0.013, -0.002)
phecode_506	Diseases of salivary glands	0.59 (0.577, 0.605)	0.59 (0.574, 0.606)	-0.001 (-0.008, 0.006)
phecode_334-1	Trigeminal nerve disorders [CNS]	0.6 (0.575, 0.618)	0.59 (0.57, 0.611)	-0.009 (-0.021, 0.005)
phecode_324	Extrapyramidal and movement disorders	0.596 (0.586, 0.602)	0.59 (0.582, 0.599)	-0.005 (-0.011, 0)
phecode_594-6	Urinary urgency	0.594 (0.582, 0.604)	0.591 (0.579, 0.603)	-0.004 (-0.007, 0.001)
phecode_171-1	Lymphocytosis (symptomatic)	0.607 (0.569, 0.648)	0.589 (0.546, 0.636)	-0.018 (-0.054, 0.019)
phecode_664-21	Pityriasis rosea	0.608 (0.573, 0.647)	0.592 (0.553, 0.626)	-0.019 (-0.031, -0.007)
phecode_593-2	Microscopic hematuria	0.598 (0.582, 0.611)	0.592 (0.576, 0.606)	-0.007 (-0.015, 0.001)
phecode_109	Malignant neoplasm of the eye, brain and other parts of central nervous system	0.603 (0.573, 0.637)	0.592 (0.557, 0.627)	-0.012 (-0.044, 0.015)
phecode_200-21	Diffuse goiter	0.627 (0.591, 0.654)	0.591 (0.56, 0.62)	-0.036 (-0.05, -0.018)
phecode_708-8	Secondary osteoarthritis	0.595 (0.58, 0.622)	0.592 (0.574, 0.613)	-0.005 (-0.02, 0.01)
phecode_808	Syncope and collapse	0.6 (0.592, 0.606)	0.593 (0.587, 0.6)	-0.006 (-0.01, -0.002)
phecode_360-4	Blepharitis	0.595 (0.586, 0.606)	0.593 (0.584, 0.603)	-0.003 (-0.005, 0)
phecode_444-11	Varicose veins of lower extremities	0.596 (0.588, 0.606)	0.594 (0.585, 0.603)	-0.002 (-0.005, 0.001)
phecode_334-11	Trigeminal neuralgia	0.596 (0.574, 0.619)	0.592 (0.571, 0.617)	-0.001 (-0.014, 0.011)
phecode_325-3	Lack of coordination	0.603 (0.582, 0.619)	0.594 (0.575, 0.613)	-0.008 (-0.021, 0.007)
phecode_139-54	Testicular lipoma	0.609 (0.571, 0.65)	0.593 (0.551, 0.635)	-0.016 (-0.027, -0.004)
phecode_597	Other disorders of urethra and urinary tract	0.6 (0.581, 0.615)	0.595 (0.578, 0.61)	-0.005 (-0.017, 0.005)
phecode_288-4	Phobic disorders	0.602 (0.581, 0.613)	0.596 (0.574, 0.609)	-0.006 (-0.012, 0.002)
phecode_494	Voice disturbance	0.595 (0.584, 0.608)	0.595 (0.585, 0.605)	0 (-0.007, 0.005)
phecode_510-5	Dyskinesia of esophagus	0.612 (0.58, 0.645)	0.595 (0.554, 0.633)	-0.018 (-0.04, 0.002)
phecode_284	Suicide ideation and attempt or self harm	0.599 (0.585, 0.616)	0.594 (0.58, 0.613)	-0.005 (-0.018, 0.009)
phecode_592-12	Chronic cystitis	0.603 (0.565, 0.633)	0.595 (0.559, 0.628)	-0.006 (-0.034, 0.019)
phecode_466-4	Hypertrophy of tonsils and adenoids	0.611 (0.577, 0.657)	0.598 (0.558, 0.635)	-0.016 (-0.041, 0.017)
phecode_329-6	Transient global amnesia	0.622 (0.594, 0.655)	0.597 (0.561, 0.631)	-0.026 (-0.043, -0.011)
phecode_807-1	Chronic fatigue syndrome	0.597 (0.59, 0.604)	0.596 (0.589, 0.604)	-0.001 (-0.003, 0.003)
phecode_052-1	Herpes simplex	0.598 (0.585, 0.613)	0.597 (0.583, 0.611)	-0.002 (-0.008, 0.004)
phecode_200-31	Graves' disease [Toxic diffuse goiter]	0.634 (0.603, 0.661)	0.597 (0.56, 0.622)	-0.035 (-0.052, -0.016)
phecode_394-1	Meniere disease	0.612 (0.581, 0.634)	0.599 (0.569, 0.623)	-0.013 (-0.027, 0.001)
phecode_522-8	Duodenitis	0.598 (0.584, 0.613)	0.598 (0.586, 0.609)	0 (-0.009, 0.008)
phecode_581-31	Hydronephrosis	0.606 (0.589, 0.629)	0.598 (0.578, 0.617)	-0.008 (-0.021, 0.004)
phecode_440-11	Deep vein thrombosis [DVT]	0.603 (0.588, 0.62)	0.599 (0.581, 0.619)	-0.003 (-0.016, 0.008)
phecode_284-1	Suicidal ideations	0.611 (0.596, 0.635)	0.599 (0.581, 0.622)	-0.012 (-0.024, 0)
phecode_552-1	Cholangitis	0.605 (0.571, 0.641)	0.599 (0.56, 0.641)	-0.005 (-0.036, 0.026)
phecode_144-1	Benign neoplasms of external female genital organs and cervix	0.612 (0.59, 0.635)	0.601 (0.58, 0.623)	-0.011 (-0.027, 0)
phecode_394	Disorders of vestibular function	0.603 (0.593, 0.613)	0.602 (0.592, 0.612)	-0.002 (-0.005, 0)
phecode_180	Other disorders involving the immune mechanism	0.608 (0.589, 0.63)	0.603 (0.586, 0.622)	-0.006 (-0.021, 0.008)
phecode_601-12	Chronic prostatitis	0.616 (0.59, 0.657)	0.602 (0.575, 0.639)	-0.018 (-0.027, -0.001)
phecode_148	Benign neoplasm of the eye, brain and other parts of central nervous system	0.602 (0.582, 0.626)	0.603 (0.583, 0.625)	-0.001 (-0.012, 0.016)

Supplementary Tables

Table 23 continued from previous page

Endpoint	Phecode String	Age+Sex	Age+Sex+Retina	Delta C-Index
phecode_172	Other disorders of white blood cells	0.624 (0.598, 0.649)	0.602 (0.576, 0.634)	-0.022 (-0.043, 0)
phecode_109-3	Malignant neoplasm of brain	0.608 (0.57, 0.645)	0.604 (0.563, 0.643)	-0.005 (-0.032, 0.029)
phecode_209-1	Pituitary hyperfunction	0.615 (0.555, 0.653)	0.605 (0.554, 0.649)	-0.009 (-0.047, 0.034)
phecode_596	Other disorders of bladder	0.607 (0.599, 0.617)	0.604 (0.596, 0.613)	-0.003 (-0.007, 0.001)
phecode_622-2	Mucous polyp of cervix	0.618 (0.598, 0.64)	0.605 (0.584, 0.625)	-0.014 (-0.03, 0)
phecode_394-2	Vertigo	0.607 (0.597, 0.618)	0.606 (0.594, 0.616)	-0.002 (-0.004, 0.001)
phecode_728	Chondropathies	0.608 (0.596, 0.623)	0.608 (0.592, 0.623)	-0.002 (-0.007, 0.005)
phecode_144-13	Benign neoplasms of the cervix	0.624 (0.597, 0.648)	0.61 (0.582, 0.628)	-0.015 (-0.03, 0)
phecode_200-3	Thyrototoxicosis [hyperthyroidism]	0.613 (0.597, 0.63)	0.609 (0.59, 0.626)	-0.004 (-0.012, 0.002)
phecode_513-3	Duodenal ulcer	0.62 (0.602, 0.641)	0.609 (0.592, 0.628)	-0.01 (-0.026, 0.001)
phecode_625	Pain and other symptoms associated with female genital organs	0.611 (0.6, 0.622)	0.609 (0.597, 0.621)	-0.002 (-0.006, 0.002)
phecode_523-4	Diverticulitis	0.612 (0.598, 0.627)	0.608 (0.596, 0.623)	-0.004 (-0.011, 0.007)
phecode_728-3	Costochondritis (Tietze's disease)	0.611 (0.595, 0.626)	0.61 (0.595, 0.627)	-0.001 (-0.009, 0.006)
phecode_581-3	Obstructive and reflux uropathy	0.62 (0.604, 0.635)	0.61 (0.593, 0.625)	-0.01 (-0.021, 0.001)
phecode_430-1	Nontraumatic subarachnoid hemorrhage	0.61 (0.575, 0.644)	0.613 (0.57, 0.641)	-0.002 (-0.026, 0.034)
phecode_124-5	Essential thrombocythemia	0.639 (0.584, 0.685)	0.61 (0.571, 0.65)	-0.026 (-0.059, 0.008)
phecode_467	Laryngitis and tracheitis	0.616 (0.606, 0.632)	0.611 (0.6, 0.629)	-0.005 (-0.011, 0)
phecode_814	Jaundice (not of newborn)	0.621 (0.591, 0.653)	0.613 (0.588, 0.636)	-0.009 (-0.031, 0.015)
phecode_594-3	Urinary incontinence	0.618 (0.609, 0.624)	0.613 (0.604, 0.62)	-0.005 (-0.009, -0.002)
phecode_416-11	Supraventricular tachycardia	0.62 (0.602, 0.635)	0.613 (0.595, 0.63)	-0.008 (-0.018, 0.003)
phecode_514-3	Ileus	0.69 (0.655, 0.718)	0.612 (0.576, 0.652)	-0.076 (-0.105, -0.042)
phecode_394-21	Paroxysmal vertigo	0.618 (0.606, 0.629)	0.614 (0.603, 0.627)	-0.003 (-0.005, 0)
phecode_331-1	Tension headache	0.615 (0.603, 0.628)	0.615 (0.602, 0.629)	0 (-0.006, 0.007)
phecode_410	Inflammation of the heart [Carditis]	0.623 (0.605, 0.648)	0.615 (0.595, 0.638)	-0.009 (-0.021, 0.002)
phecode_363-7	Stenosis and insufficiency of lacrimal passages	0.622 (0.595, 0.649)	0.615 (0.588, 0.642)	-0.008 (-0.022, 0.007)
phecode_116-1	Secondary malignancy of lymph nodes	0.617 (0.601, 0.63)	0.616 (0.599, 0.63)	-0.001 (-0.012, 0.008)
phecode_717-2	Sacrococcygeal disorders	0.621 (0.602, 0.639)	0.616 (0.594, 0.635)	-0.006 (-0.013, 0.004)
phecode_710-31	Genu valgum (acquired)	0.645 (0.599, 0.687)	0.616 (0.566, 0.668)	-0.029 (-0.048, -0.013)
phecode_848	Nonspecific abnormal findings of other body structures	0.617 (0.609, 0.624)	0.617 (0.609, 0.624)	-0.001 (-0.006, 0.004)
phecode_552	Other diseases of biliary tract	0.62 (0.596, 0.638)	0.616 (0.598, 0.642)	-0.002 (-0.015, 0.009)
phecode_708-12	Primary osteoarthritis of knee, lower leg	0.618 (0.605, 0.628)	0.618 (0.605, 0.627)	0 (-0.008, 0.006)
phecode_591	Urinary tract infection	0.62 (0.615, 0.625)	0.618 (0.612, 0.622)	-0.003 (-0.006, 0)
phecode_381	Strabismus	0.631 (0.6, 0.655)	0.617 (0.592, 0.64)	-0.012 (-0.025, 0.006)
phecode_552-2	Obstruction of bile duct	0.635 (0.607, 0.674)	0.617 (0.587, 0.653)	-0.019 (-0.048, 0.009)
phecode_386-9	Visual distortions and subjective visual disturbances	0.626 (0.597, 0.664)	0.618 (0.583, 0.647)	-0.01 (-0.031, 0.016)
phecode_467-1	Acute laryngitis and tracheitis	0.623 (0.612, 0.635)	0.619 (0.607, 0.631)	-0.005 (-0.01, -0.001)
phecode_247-3	Disorder of phosphorus metabolism	0.637 (0.611, 0.663)	0.619 (0.58, 0.653)	-0.022 (-0.057, 0.017)
phecode_381-1	Paralytic strabismus [Neurogenic strabismus]	0.649 (0.617, 0.683)	0.62 (0.578, 0.648)	-0.031 (-0.056, -0.003)
phecode_803	Snoring*	0.631 (0.615, 0.648)	0.62 (0.605, 0.637)	-0.011 (-0.015, -0.007)
phecode_324-4	Tremor	0.629 (0.616, 0.641)	0.621 (0.607, 0.634)	-0.008 (-0.016, -0.001)
phecode_580	Glomerular diseases	0.628 (0.607, 0.652)	0.622 (0.599, 0.646)	-0.007 (-0.032, 0.02)
phecode_670	Seborrheic keratosis	0.622 (0.597, 0.64)	0.623 (0.599, 0.642)	0 (-0.009, 0.008)
phecode_416-1	Paroxysmal tachycardia	0.63 (0.617, 0.647)	0.621 (0.608, 0.637)	-0.009 (-0.019, -0.001)
phecode_416-51	Atrial premature depolarization [Supraventricular premature beats]	0.629 (0.604, 0.66)	0.623 (0.597, 0.648)	-0.008 (-0.031, 0.023)
phecode_614-53	Cyst or abscess of Bartholin's gland	0.627 (0.598, 0.668)	0.621 (0.592, 0.656)	-0.005 (-0.022, 0.016)

5 Phenome-wide prediction of disease onset from retinal fundus photographs

Table 23 continued from previous page

Endpoint	Phecode String	Age+Sex	Age+Sex+Retina	Delta C-Index
phecode_116-3	Secondary malignant neoplasm of digestive systems	0.633 (0.614, 0.654)	0.623 (0.605, 0.642)	-0.011 (-0.024, 0.003)
phecode_416-5	Premature depolarization [Premature beats]	0.626 (0.61, 0.638)	0.623 (0.606, 0.636)	-0.003 (-0.009, 0.006)
phecode_520	Hernia	0.624 (0.619, 0.629)	0.624 (0.619, 0.629)	0 (-0.001, 0.001)
phecode_506-5	Disturbances of salivary secretion	0.629 (0.61, 0.645)	0.625 (0.607, 0.642)	-0.004 (-0.013, 0.006)
phecode_401-6	Secondary hypertension	0.626 (0.609, 0.647)	0.625 (0.607, 0.646)	-0.001 (-0.005, 0.004)
phecode_440-3	Pulmonary embolism	0.63 (0.616, 0.642)	0.626 (0.614, 0.641)	-0.003 (-0.014, 0.005)
phecode_610	Benign mammary dysplasias	0.634 (0.608, 0.652)	0.628 (0.603, 0.645)	-0.007 (-0.015, 0.001)
phecode_514-1	Esophageal obstruction (Stricture and stenosis of esophagus)	0.633 (0.615, 0.648)	0.626 (0.61, 0.645)	-0.006 (-0.018, 0.007)
phecode_413-12	Mitral valve prolapse*	0.639 (0.605, 0.674)	0.628 (0.595, 0.667)	-0.01 (-0.039, 0.011)
phecode_136-2	Benign neoplasm of stomach	0.633 (0.623, 0.642)	0.627 (0.617, 0.64)	-0.005 (-0.01, 0)
phecode_602-4	Elevated prostate specific antigen [PSA]	0.63 (0.62, 0.64)	0.628 (0.618, 0.638)	-0.002 (-0.007, 0.001)
phecode_522-14	Microscopic colitis*	0.64 (0.604, 0.665)	0.63 (0.591, 0.659)	-0.011 (-0.031, 0.016)
phecode_708-9	Heberden's or Bouchard's nodes*	0.634 (0.612, 0.655)	0.63 (0.606, 0.65)	-0.004 (-0.013, 0.003)
phecode_200	Disorders of thyroid gland	0.632 (0.626, 0.638)	0.629 (0.624, 0.636)	-0.002 (-0.004, 0)
phecode_682-4	Acne	0.635 (0.608, 0.658)	0.631 (0.604, 0.657)	-0.005 (-0.012, 0.003)
phecode_122-22	Diffuse large B-cell lymphoma*	0.669 (0.637, 0.705)	0.632 (0.597, 0.671)	-0.037 (-0.059, -0.009)
phecode_602	Other disorders of prostate	0.636 (0.625, 0.646)	0.632 (0.623, 0.643)	-0.003 (-0.006, 0)
phecode_200-1	Hypothyroidism	0.635 (0.629, 0.643)	0.633 (0.626, 0.639)	-0.002 (-0.005, 0.001)
phecode_356	Speech disturbance	0.635 (0.617, 0.655)	0.633 (0.619, 0.651)	-0.002 (-0.018, 0.012)
phecode_523-1	Diverticula of small intestine	0.643 (0.608, 0.691)	0.632 (0.595, 0.673)	-0.008 (-0.031, 0.023)
phecode_416-52	Ventricular premature depolarization*	0.636 (0.621, 0.655)	0.633 (0.617, 0.654)	-0.002 (-0.01, 0.005)
phecode_510-8	Barrett's esophagus	0.641 (0.627, 0.657)	0.634 (0.617, 0.649)	-0.007 (-0.014, 0)
phecode_709	Acquired deformities of fingers and toes	0.636 (0.629, 0.645)	0.635 (0.627, 0.643)	-0.001 (-0.003, 0)
phecode_396-2	Sensorineural hearing loss	0.637 (0.626, 0.649)	0.635 (0.625, 0.646)	-0.001 (-0.005, 0.002)
phecode_410-2	Endocarditis	0.646 (0.627, 0.669)	0.635 (0.618, 0.665)	-0.011 (-0.024, 0.003)
phecode_200-14	Hypothyroidism, not specified as secondary	0.638 (0.631, 0.646)	0.637 (0.629, 0.645)	-0.001 (-0.004, 0.002)
phecode_088	Sexually transmitted disease	0.645 (0.62, 0.674)	0.638 (0.612, 0.667)	-0.007 (-0.013, -0.001)
phecode_247-52	Hypercalcemia	0.66 (0.636, 0.69)	0.639 (0.606, 0.671)	-0.022 (-0.035, -0.005)
phecode_396-21	Sensorineural hearing loss, bilateral	0.647 (0.629, 0.665)	0.64 (0.622, 0.656)	-0.008 (-0.012, -0.003)
phecode_324-41	Essential tremor*	0.64 (0.619, 0.669)	0.639 (0.618, 0.66)	-0.002 (-0.015, 0.015)
phecode_325-23	Unsteadiness on feet*	0.651 (0.625, 0.676)	0.641 (0.615, 0.666)	-0.01 (-0.024, 0)
phecode_148-2	Benign neoplasm of meninges (Meningioma)	0.655 (0.622, 0.686)	0.639 (0.606, 0.679)	-0.016 (-0.037, 0.008)
phecode_308-3	Emotional lability	0.653 (0.618, 0.704)	0.642 (0.61, 0.682)	-0.013 (-0.036, 0.007)
phecode_099	Lab findings related to infections	0.648 (0.638, 0.663)	0.64 (0.631, 0.656)	-0.007 (-0.01, -0.003)
phecode_594-31	Urge incontinence	0.648 (0.636, 0.66)	0.642 (0.628, 0.654)	-0.006 (-0.012, -0.001)
phecode_331-6	Migraine	0.643 (0.635, 0.651)	0.643 (0.635, 0.65)	0 (-0.001, 0.001)
phecode_130	Cancer (solid tumor, excluding BCC)	0.644 (0.639, 0.649)	0.643 (0.638, 0.648)	-0.001 (-0.002, 0.001)
phecode_208	Disorders of parathyroid gland	0.648 (0.627, 0.668)	0.645 (0.622, 0.665)	-0.004 (-0.017, 0.008)
phecode_710-3	Acquired deformities of the knee	0.651 (0.624, 0.676)	0.646 (0.616, 0.667)	-0.006 (-0.016, 0.006)
phecode_832	Other abnormal findings in urine	0.645 (0.641, 0.65)	0.645 (0.641, 0.649)	0 (-0.001, 0.001)
phecode_124	Myeloproliferative disorder	0.674 (0.648, 0.696)	0.647 (0.618, 0.665)	-0.027 (-0.043, -0.009)
phecode_728-1	Chondromalacia	0.659 (0.629, 0.7)	0.646 (0.616, 0.684)	-0.013 (-0.023, -0.005)
phecode_122	Lymphoma	0.658 (0.631, 0.683)	0.646 (0.626, 0.668)	-0.012 (-0.026, 0.005)
phecode_708-15	Primary osteoarthritis of the wrist, forearm	0.652 (0.614, 0.677)	0.646 (0.611, 0.673)	-0.005 (-0.025, 0.019)
phecode_200-13	Postprocedural hypothyroidism	0.659 (0.643, 0.681)	0.647 (0.626, 0.665)	-0.014 (-0.025, -0.002)

Supplementary Tables

Table 23 continued from previous page

Endpoint	Phecode String	Age+Sex	Age+Sex+Retina	Delta C-Index
phecode_709-2	Acquired deformities of toe	0.649 (0.639, 0.657)	0.647 (0.638, 0.655)	-0.001 (-0.003, -0.001)
phecode_466	Tonsillitis and adenoiditis	0.649 (0.636, 0.66)	0.648 (0.635, 0.659)	-0.001 (-0.002, 0)
phecode_208-2	Hyperparathyroidism	0.654 (0.631, 0.676)	0.648 (0.621, 0.669)	-0.004 (-0.024, 0.013)
phecode_625-1	Dyspareunia	0.65 (0.632, 0.664)	0.649 (0.633, 0.666)	-0.002 (-0.009, 0.008)
phecode_679-7	Abnormal granulation tissue, NOS	0.649 (0.626, 0.671)	0.649 (0.62, 0.673)	0 (-0.017, 0.018)
phecode_414-2	Dilated cardiomyopathy*	0.662 (0.633, 0.703)	0.65 (0.61, 0.684)	-0.014 (-0.046, 0.007)
phecode_586-2	Cyst of kidney	0.652 (0.636, 0.671)	0.65 (0.634, 0.668)	-0.003 (-0.013, 0.009)
phecode_430-3	Nontraumatic subdural hemorrhage	0.664 (0.624, 0.71)	0.65 (0.613, 0.692)	-0.015 (-0.032, 0.004)
phecode_817	Motion sickness	0.651 (0.627, 0.687)	0.651 (0.619, 0.682)	-0.002 (-0.026, 0.02)
phecode_108-4	Malignant neoplasm of the kidney	0.661 (0.634, 0.687)	0.653 (0.626, 0.678)	-0.008 (-0.027, 0.01)
phecode_101-8	Malignant neoplasm of the pancreas	0.669 (0.643, 0.692)	0.653 (0.621, 0.675)	-0.017 (-0.039, 0.006)
phecode_438-12	Thoracic aneurysm	0.675 (0.65, 0.702)	0.652 (0.625, 0.687)	-0.022 (-0.043, -0.001)
phecode_120	Hemo onc - by cell of origin	0.667 (0.653, 0.679)	0.652 (0.637, 0.665)	-0.015 (-0.023, -0.006)
phecode_108-41	Malignant neoplasm of kidney, except pelvis	0.663 (0.63, 0.699)	0.652 (0.624, 0.683)	-0.009 (-0.029, 0.008)
phecode_120-2	Lymphoid	0.666 (0.649, 0.685)	0.653 (0.635, 0.672)	-0.013 (-0.026, -0.001)
phecode_116	Secondary malignant neoplasm	0.655 (0.646, 0.664)	0.653 (0.644, 0.663)	-0.002 (-0.007, 0.004)
phecode_430	Nontraumatic Intracranial hemorrhage	0.656 (0.638, 0.68)	0.652 (0.634, 0.674)	-0.002 (-0.017, 0.01)
phecode_116-4	Secondary malignant neoplasm of liver	0.661 (0.646, 0.673)	0.654 (0.638, 0.668)	-0.006 (-0.017, 0.005)
phecode_290-1	Posttraumatic stress disorder	0.657 (0.621, 0.709)	0.65 (0.62, 0.704)	-0.003 (-0.034, 0.023)
phecode_122-2	Non-Hodgkin lymphoma	0.667 (0.643, 0.691)	0.654 (0.629, 0.68)	-0.013 (-0.03, 0.004)
phecode_413-42	Pulmonary valve insufficiency*	0.69 (0.64, 0.734)	0.655 (0.618, 0.696)	-0.036 (-0.059, -0.001)
phecode_417-3	Bradycardia*	0.658 (0.647, 0.67)	0.654 (0.644, 0.666)	-0.004 (-0.007, -0.001)
phecode_177-4	Lymphedema	0.656 (0.641, 0.673)	0.655 (0.639, 0.671)	-0.003 (-0.016, 0.013)
phecode_832-6	Pyuria*	0.657 (0.642, 0.672)	0.655 (0.638, 0.67)	-0.003 (-0.007, 0.002)
phecode_123-1	Multiple myeloma	0.672 (0.637, 0.715)	0.654 (0.613, 0.707)	-0.019 (-0.041, 0.003)
phecode_841-12	Allergy to antibiotic agent (excluding penicillin)	0.661 (0.649, 0.672)	0.658 (0.647, 0.669)	-0.003 (-0.01, 0.003)
phecode_413-4	Pulmonary valve disorders	0.693 (0.652, 0.729)	0.658 (0.618, 0.702)	-0.032 (-0.055, -0.004)
phecode_583-1	End stage renal disease [CDK, stage 5]	0.708 (0.674, 0.738)	0.659 (0.623, 0.697)	-0.048 (-0.09, -0.002)
phecode_011	Klebsiella	0.667 (0.625, 0.701)	0.662 (0.616, 0.695)	-0.005 (-0.031, 0.027)
phecode_144-2	Benign neoplasms of the uterus	0.662 (0.65, 0.672)	0.661 (0.649, 0.671)	-0.001 (-0.004, 0.003)
phecode_418-1	Abnormal electrocardiogram [ECG] [EKG]	0.664 (0.655, 0.672)	0.661 (0.654, 0.672)	-0.002 (-0.006, 0.002)
phecode_208-21	Primary hyperparathyroidism	0.678 (0.651, 0.7)	0.663 (0.638, 0.684)	-0.013 (-0.029, 0.002)
phecode_466-1	Acute tonsillitis and adenoiditis	0.663 (0.653, 0.675)	0.662 (0.652, 0.675)	-0.001 (-0.002, 0.001)
phecode_600	Benign prostatic hyperplasia	0.664 (0.659, 0.669)	0.664 (0.659, 0.669)	0 (-0.001, 0.001)
phecode_121	Leukemia	0.683 (0.659, 0.709)	0.664 (0.628, 0.691)	-0.019 (-0.039, -0.004)
phecode_704	Systemic vasculitis	0.677 (0.653, 0.7)	0.665 (0.643, 0.686)	-0.012 (-0.024, 0.003)
phecode_709-21	Hallux valgus (Bunion)	0.667 (0.66, 0.675)	0.665 (0.658, 0.674)	-0.002 (-0.003, 0)
phecode_356-2	Aphasia and dysphasia	0.692 (0.664, 0.723)	0.664 (0.636, 0.694)	-0.027 (-0.044, -0.011)
phecode_660-13	Pityriasis versicolor	0.666 (0.635, 0.696)	0.666 (0.632, 0.692)	-0.001 (-0.013, 0.013)
phecode_123	Multiple myeloma and malignant plasma cell neoplasms	0.683 (0.647, 0.722)	0.665 (0.631, 0.707)	-0.018 (-0.04, 0.001)
phecode_308-6	Excessive crying of child, adolescent, or adult	0.674 (0.646, 0.71)	0.665 (0.634, 0.702)	-0.007 (-0.035, 0.015)
phecode_120-21	Mature B-cell	0.68 (0.66, 0.703)	0.665 (0.646, 0.687)	-0.014 (-0.026, -0.002)
phecode_374-39	Transient retinal arterial occlusion [Amaurosis fugax]	0.683 (0.651, 0.714)	0.666 (0.637, 0.696)	-0.015 (-0.031, 0)
phecode_413-22	Aortic insufficiency	0.677 (0.662, 0.692)	0.668 (0.651, 0.684)	-0.009 (-0.021, 0.002)
phecode_101	Malignant neoplasm of the digestive organs	0.672 (0.663, 0.683)	0.669 (0.659, 0.68)	-0.003 (-0.009, 0.002)

5 Phenome-wide prediction of disease onset from retinal fundus photographs

Table 23 continued from previous page

Endpoint	Phecode String	Age+Sex	Age+Sex+Retina	Delta C-Index
phecode_120-1	Myeloid	0.689 (0.668, 0.709)	0.67 (0.649, 0.686)	-0.018 (-0.031, -0.008)
phecode_329-41	Attention and concentration deficit	0.677 (0.647, 0.714)	0.67 (0.642, 0.708)	-0.007 (-0.015, 0)
phecode_710-32	Genu varum (acquired)	0.682 (0.646, 0.721)	0.672 (0.636, 0.715)	-0.011 (-0.022, 0.001)
phecode_714	Deforming dorsopathies	0.677 (0.659, 0.696)	0.672 (0.656, 0.692)	-0.006 (-0.013, 0.003)
phecode_101-42	Malignant neoplasm of the rectum	0.682 (0.657, 0.705)	0.675 (0.65, 0.693)	-0.008 (-0.019, 0.003)
phecode_361-9	Ectropion of eyelid	0.684 (0.64, 0.722)	0.676 (0.629, 0.714)	-0.009 (-0.023, 0.008)
phecode_101-2	Malignant neoplasm of stomach	0.688 (0.656, 0.729)	0.673 (0.65, 0.708)	-0.014 (-0.036, 0.006)
phecode_476	Bronchiectasis	0.677 (0.667, 0.692)	0.676 (0.664, 0.692)	-0.001 (-0.008, 0.005)
phecode_180-3	Paraproteinemias	0.687 (0.66, 0.714)	0.675 (0.654, 0.704)	-0.01 (-0.024, 0.001)
phecode_121-2	Chronic leukemia	0.698 (0.673, 0.726)	0.679 (0.65, 0.709)	-0.019 (-0.043, 0)
phecode_101-41	Malignant neoplasm of the colon	0.687 (0.667, 0.702)	0.682 (0.664, 0.699)	-0.005 (-0.01, 0.001)
phecode_200-23	Multinodular goiter	0.69 (0.671, 0.715)	0.682 (0.66, 0.705)	-0.009 (-0.016, -0.003)
phecode_592	Cystitis and urethritis	0.685 (0.674, 0.693)	0.682 (0.672, 0.69)	-0.003 (-0.005, -0.001)
phecode_101-4	Malignant neoplasm of the colon and rectum	0.685 (0.673, 0.699)	0.682 (0.672, 0.696)	-0.003 (-0.008, 0.002)
phecode_180-31	Monoclonal gammopathy	0.689 (0.668, 0.713)	0.683 (0.657, 0.704)	-0.007 (-0.026, 0.009)
phecode_416-211	Paroxysmal atrial fibrillation*	0.683 (0.674, 0.695)	0.683 (0.674, 0.694)	0 (-0.007, 0.004)
phecode_619-3	Noninflammatory disorders of cervix	0.686 (0.664, 0.703)	0.684 (0.663, 0.702)	-0.002 (-0.007, 0.003)
phecode_107	Malignant neoplasm of male genitalia	0.685 (0.675, 0.693)	0.684 (0.674, 0.692)	-0.001 (-0.003, 0)
phecode_416-12	Ventricular tachycardia	0.709 (0.685, 0.734)	0.683 (0.66, 0.714)	-0.024 (-0.037, -0.015)
phecode_593-1	Gross hematuria	0.69 (0.668, 0.711)	0.683 (0.663, 0.705)	-0.006 (-0.009, -0.002)
phecode_610-1	Cystic mastopathy	0.69 (0.664, 0.713)	0.685 (0.66, 0.71)	-0.004 (-0.013, 0.006)
phecode_709-24	Hammer toe	0.687 (0.671, 0.707)	0.687 (0.67, 0.705)	0 (-0.005, 0.006)
phecode_107-2	Malignant neoplasm of the prostate	0.689 (0.683, 0.697)	0.688 (0.681, 0.696)	-0.001 (-0.003, 0)
phecode_121-21	Chronic lymphoid leukemia	0.706 (0.673, 0.735)	0.689 (0.657, 0.718)	-0.016 (-0.035, 0.003)
phecode_413-11	Mitral valve insufficiency	0.696 (0.684, 0.713)	0.691 (0.679, 0.705)	-0.005 (-0.013, 0.002)
phecode_592-1	Cystitis	0.694 (0.686, 0.703)	0.691 (0.683, 0.701)	-0.003 (-0.005, -0.001)
phecode_416-43	Right bundle branch block	0.706 (0.69, 0.721)	0.695 (0.678, 0.711)	-0.011 (-0.016, -0.005)
phecode_430-2	Nontraumatic intracerebral hemorrhage	0.697 (0.678, 0.72)	0.695 (0.676, 0.719)	-0.003 (-0.029, 0.016)
phecode_416	Cardiac arrhythmia and conduction disorders	0.697 (0.692, 0.702)	0.698 (0.692, 0.701)	0 (-0.002, 0.002)
phecode_530-3	Rectal prolapse	0.717 (0.694, 0.744)	0.697 (0.676, 0.728)	-0.02 (-0.031, -0.009)
phecode_413-1	Mitral valve disorders	0.706 (0.694, 0.717)	0.702 (0.689, 0.714)	-0.005 (-0.011, 0.002)
phecode_705-3	Polymyalgia rheumatica	0.709 (0.692, 0.719)	0.708 (0.692, 0.717)	-0.001 (-0.006, 0.003)
phecode_108	Malignant neoplasm of the urinary tract	0.71 (0.693, 0.724)	0.709 (0.692, 0.724)	-0.001 (-0.008, 0.007)
phecode_413	Heart valve disorders	0.712 (0.705, 0.72)	0.709 (0.702, 0.718)	-0.003 (-0.007, 0.002)
phecode_592-11	Acute cystitis	0.712 (0.698, 0.731)	0.71 (0.695, 0.729)	-0.003 (-0.009, 0.006)
phecode_413-6	Heart valve replaced	0.726 (0.706, 0.751)	0.711 (0.692, 0.732)	-0.014 (-0.027, -0.004)
phecode_841-13	Allergy to sulfonamides	0.718 (0.693, 0.743)	0.712 (0.684, 0.735)	-0.006 (-0.021, 0.007)
phecode_168-4	Abnormal coagulation profile	0.729 (0.694, 0.763)	0.712 (0.679, 0.743)	-0.016 (-0.028, -0.002)
phecode_324-1	Parkinsonism	0.722 (0.71, 0.737)	0.713 (0.7, 0.729)	-0.009 (-0.017, -0.004)
phecode_416-212	Persistent atrial fibrillation*	0.717 (0.689, 0.742)	0.713 (0.689, 0.741)	-0.001 (-0.018, 0.014)
phecode_594-32	Stress incontinence	0.716 (0.705, 0.724)	0.714 (0.706, 0.723)	-0.001 (-0.008, 0.006)
phecode_594-41	Nocturia	0.716 (0.704, 0.723)	0.715 (0.704, 0.723)	0 (-0.001, 0.001)
phecode_481-4	Pulmonary fibrosis	0.726 (0.702, 0.748)	0.716 (0.681, 0.738)	-0.011 (-0.032, 0.009)
phecode_324-11	Parkinson's disease (Primary)	0.724 (0.709, 0.74)	0.716 (0.698, 0.731)	-0.008 (-0.015, -0.002)

Supplementary Tables

Table 23 continued from previous page

Endpoint	Phecode String	Age+Sex	Age+Sex+Retina	Delta C-Index
phecode_400	Rheumatic fever and chronic rheumatic heart diseases	0.723 (0.712, 0.733)	0.717 (0.707, 0.728)	-0.005 (-0.013, 0.003)
phecode_400-2	Chronic rheumatic heart diseases	0.723 (0.713, 0.737)	0.717 (0.701, 0.73)	-0.007 (-0.014, 0)
phecode_416-71	Sick sinus syndrome*	0.733 (0.7, 0.771)	0.717 (0.678, 0.75)	-0.016 (-0.035, -0.002)
phecode_700-2	Sicca syndrome [Sjogren syndrome]	0.731 (0.707, 0.758)	0.718 (0.689, 0.751)	-0.012 (-0.026, 0.003)
phecode_416-4	Heart block	0.722 (0.715, 0.73)	0.718 (0.71, 0.726)	-0.004 (-0.007, 0)
phecode_620	Dysplasia of female genital organs	0.72 (0.691, 0.746)	0.72 (0.692, 0.746)	-0.001 (-0.009, 0.009)
phecode_416-7	Sinoatrial node dysfunction	0.735 (0.695, 0.775)	0.718 (0.683, 0.766)	-0.016 (-0.035, -0.001)
phecode_416-42	Left bundle branch block	0.722 (0.706, 0.735)	0.719 (0.706, 0.734)	-0.002 (-0.009, 0.007)
phecode_101-1	Malignant neoplasm of the esophagus	0.735 (0.71, 0.759)	0.721 (0.697, 0.746)	-0.012 (-0.028, 0)
phecode_413-2	Aortic valve disorders	0.727 (0.716, 0.734)	0.726 (0.714, 0.734)	-0.001 (-0.008, 0.003)
phecode_108-5	Malignant neoplasm of the bladder	0.734 (0.718, 0.756)	0.73 (0.711, 0.753)	-0.004 (-0.011, 0.005)
phecode_416-22	Atrial flutter	0.739 (0.725, 0.753)	0.732 (0.718, 0.747)	-0.006 (-0.013, 0)
phecode_615	Endometriosis	0.734 (0.71, 0.756)	0.733 (0.711, 0.754)	0 (-0.007, 0.005)
phecode_594-1	Retention of urine	0.735 (0.726, 0.744)	0.733 (0.724, 0.741)	-0.002 (-0.004, 0.001)
phecode_132	Sequelae of cancer	0.734 (0.725, 0.745)	0.733 (0.722, 0.745)	-0.002 (-0.007, 0.003)
phecode_596-3	Diverticulum of bladder	0.738 (0.709, 0.771)	0.734 (0.704, 0.767)	-0.005 (-0.013, 0.005)
phecode_433-2	Occlusion and stenosis of pre-cerebral arteries	0.738 (0.718, 0.76)	0.737 (0.715, 0.754)	0 (-0.015, 0.015)
phecode_704-5	Giant cell arteritis	0.752 (0.725, 0.784)	0.741 (0.713, 0.775)	-0.013 (-0.029, 0.003)
phecode_594-11	Urinary hesitancy	0.75 (0.724, 0.781)	0.745 (0.721, 0.771)	-0.006 (-0.016, 0.008)
phecode_142-1	Lump or mass in breast	0.748 (0.74, 0.755)	0.747 (0.739, 0.753)	-0.001 (-0.002, -0.001)
phecode_416-41	Atrioventricular block	0.75 (0.74, 0.763)	0.747 (0.735, 0.76)	-0.004 (-0.01, 0.001)
phecode_625-2	Postcoital bleeding	0.754 (0.732, 0.774)	0.748 (0.727, 0.767)	-0.005 (-0.013, 0.001)
phecode_105	Malignant neoplasm of the breast	0.753 (0.747, 0.76)	0.75 (0.743, 0.756)	-0.003 (-0.006, -0.001)
phecode_329-5	Mild cognitive impairment, so stated	0.764 (0.751, 0.782)	0.756 (0.743, 0.775)	-0.007 (-0.014, -0.001)
phecode_142	Lump or mass in breast or non-specific abnormal breast exam	0.761 (0.756, 0.765)	0.76 (0.756, 0.765)	0 (-0.001, 0)
phecode_328-9	Dementia NOS	0.779 (0.761, 0.794)	0.768 (0.751, 0.783)	-0.009 (-0.018, -0.002)
phecode_597-5	Urethral caruncle	0.787 (0.766, 0.81)	0.778 (0.754, 0.801)	-0.01 (-0.015, -0.004)
phecode_328	Dementias and cerebral degeneration	0.788 (0.776, 0.802)	0.783 (0.772, 0.796)	-0.005 (-0.01, -0.001)
phecode_438-1	Aortic aneurysm and ectasia	0.788 (0.777, 0.797)	0.788 (0.778, 0.796)	0 (-0.005, 0.004)
phecode_626	Disorders of menstruation and other abnormal bleeding from female genital tract	0.792 (0.783, 0.801)	0.791 (0.782, 0.8)	-0.001 (-0.002, 0)
phecode_444-15	Scrotal varices [Varicocele]	0.802 (0.788, 0.815)	0.794 (0.778, 0.807)	-0.009 (-0.016, 0)
phecode_140	Benign neoplasm of the breast	0.797 (0.779, 0.812)	0.795 (0.779, 0.809)	-0.003 (-0.008, 0.001)
phecode_328-7	Vascular dementia	0.8 (0.774, 0.823)	0.795 (0.771, 0.819)	-0.004 (-0.018, 0.009)
phecode_328-8	Dementia in conditions classified elsewhere	0.801 (0.79, 0.813)	0.796 (0.783, 0.808)	-0.005 (-0.01, -0.002)
phecode_596-1	Bladder neck obstruction	0.801 (0.788, 0.818)	0.799 (0.787, 0.816)	-0.002 (-0.006, 0.002)
phecode_328-1	Alzheimer's disease	0.812 (0.798, 0.826)	0.809 (0.796, 0.822)	-0.003 (-0.007, 0.002)
phecode_608-1	Abnormal findings in semen	0.824 (0.819, 0.829)	0.823 (0.818, 0.828)	-0.001 (-0.002, 0)
phecode_626-11	Absent or infrequent menstruation	0.847 (0.836, 0.857)	0.846 (0.835, 0.859)	-0.001 (-0.004, 0.002)
phecode_626-14	Irregular menstrual bleeding	0.871 (0.863, 0.878)	0.87 (0.863, 0.879)	0 (-0.003, 0.002)
phecode_626-2	Dysmenorrhea	0.894 (0.882, 0.908)	0.891 (0.879, 0.902)	-0.004 (-0.01, 0.004)
phecode_520-2	Diaphragmatic hernia [Hiatal hernia]	0.603 (0.597, 0.612)	0.603 (0.596, 0.612)	0 (-0.003, 0.003)
phecode_290	Reaction to severe stress, and adjustment disorders	0.642 (0.634, 0.65)	0.642 (0.634, 0.65)	0 (-0.001, 0.001)
phecode_396-22	Presbycusis	0.7 (0.684, 0.72)	0.698 (0.683, 0.721)	0 (-0.01, 0.007)
phecode_532-1	Intestinal fistula	0.606 (0.579, 0.637)	0.606 (0.582, 0.63)	0 (-0.02, 0.015)

5 Phenome-wide prediction of disease onset from retinal fundus photographs

Table 23 continued from previous page

Endpoint	Phecode String	Age+Sex	Age+Sex+Retina	Delta C-Index
phecode_146	Benign neoplasm of the genitourinary system	0.834 (0.827, 0.841)	0.834 (0.828, 0.841)	0 (-0.002, 0.002)
phecode_144	Gynecological benign neoplasms	0.649 (0.64, 0.661)	0.649 (0.638, 0.661)	0 (-0.004, 0.003)
phecode_342	Plegia and unspecified paralysis	0.56 (0.533, 0.58)	0.559 (0.534, 0.582)	0 (-0.019, 0.022)
phecode_603	Disorders and symptoms of testis	0.526 (0.514, 0.539)	0.527 (0.511, 0.538)	0 (-0.009, 0.009)
phecode_529-1	Diarrhea	0.544 (0.538, 0.551)	0.545 (0.538, 0.55)	0 (-0.005, 0.004)
phecode_520-11	Inguinal hernia	0.788 (0.783, 0.794)	0.788 (0.783, 0.794)	0 (0, 0.001)
phecode_520-1	Hernia of the abdominal wall	0.715 (0.71, 0.72)	0.715 (0.71, 0.72)	0 (-0.001, 0.001)
phecode_433-21	Carotid artery stenosis	0.734 (0.708, 0.758)	0.733 (0.712, 0.753)	0 (-0.018, 0.014)
phecode_520-15	Incisional hernia	0.59 (0.565, 0.618)	0.59 (0.568, 0.614)	0 (-0.022, 0.022)
phecode_723-4	Lateral epicondylitis (Tennis elbow)	0.638 (0.629, 0.645)	0.638 (0.629, 0.645)	0 (-0.001, 0.002)
phecode_593-3	Recurrent and persistent hematuria*	0.547 (0.512, 0.59)	0.547 (0.517, 0.575)	0 (-0.037, 0.035)
phecode_144-21	Leiomyoma of uterus	0.709 (0.697, 0.72)	0.71 (0.697, 0.72)	0 (-0.004, 0.004)
phecode_977-5	Long term (current) use of agents affecting hormones	0.896 (0.886, 0.901)	0.896 (0.886, 0.902)	0 (-0.001, 0.001)
phecode_418	Abnormal results of cardiovascular function studies	0.66 (0.651, 0.666)	0.659 (0.651, 0.668)	0 (-0.004, 0.005)
phecode_146-2	Benign neoplasm of the prostate	0.647 (0.634, 0.66)	0.648 (0.634, 0.661)	0 (-0.005, 0.005)
phecode_626-13	Irregular menstrual cycle	0.889 (0.883, 0.895)	0.889 (0.883, 0.896)	0 (-0.001, 0.002)
phecode_174-2	Splenomegaly	0.617 (0.566, 0.665)	0.618 (0.582, 0.646)	0 (-0.032, 0.035)
phecode_139-5	Lipoma	0.519 (0.51, 0.531)	0.52 (0.508, 0.532)	0 (-0.009, 0.01)
phecode_708-11	Primary osteoarthritis of hip, pelvic region and thigh	0.657 (0.641, 0.674)	0.657 (0.642, 0.675)	0 (-0.004, 0.006)
phecode_614	Inflammatory diseases of female pelvic organs	0.564 (0.551, 0.573)	0.565 (0.553, 0.574)	0 (-0.005, 0.007)
phecode_626-1	Irregular menstrual cycle/bleeding	0.877 (0.873, 0.885)	0.878 (0.874, 0.884)	0 (-0.001, 0.002)
phecode_614-55	Candidiasis of vulva and vagina	0.763 (0.756, 0.773)	0.764 (0.757, 0.772)	0 (-0.006, 0.008)
phecode_200-7	Iodine-deficiency related thyroid disorders*	0.585 (0.569, 0.606)	0.586 (0.567, 0.602)	0 (-0.015, 0.014)
phecode_529	Symptoms involving digestive system	0.562 (0.558, 0.566)	0.562 (0.559, 0.566)	0 (-0.002, 0.003)
phecode_594-2	Dysuria	0.611 (0.603, 0.621)	0.611 (0.605, 0.62)	0 (-0.004, 0.004)
phecode_679-1	Rash and other nonspecific skin eruption	0.533 (0.529, 0.538)	0.534 (0.53, 0.538)	0 (-0.003, 0.004)
phecode_627-2	Symptomatic menopause	0.766 (0.761, 0.771)	0.767 (0.761, 0.772)	0 (-0.001, 0.001)
phecode_977-52	Hormone replacement therapy (postmenopausal)	0.749 (0.733, 0.764)	0.749 (0.734, 0.766)	0 (-0.003, 0.004)
phecode_627-4	Menorrhagia/Excessive and frequent menstruation	0.878 (0.873, 0.883)	0.879 (0.873, 0.883)	0 (-0.001, 0.002)
phecode_009	Pseudomonas	0.658 (0.63, 0.688)	0.661 (0.632, 0.689)	0.001 (-0.018, 0.024)
phecode_977-51	Long term (current) use of hormonal contraceptives	0.922 (0.916, 0.935)	0.923 (0.917, 0.936)	0.001 (-0.001, 0.002)
phecode_413-3	Tricuspid valve disorders	0.738 (0.722, 0.748)	0.738 (0.723, 0.748)	0.001 (-0.01, 0.01)
phecode_106-2	Malignant neoplasm of the uterus	0.556 (0.528, 0.578)	0.554 (0.529, 0.582)	0.001 (-0.026, 0.028)
phecode_676-2	Scar conditions and fibrosis of skin	0.539 (0.518, 0.561)	0.539 (0.519, 0.566)	0.001 (-0.014, 0.021)
phecode_438-11	Abdominal aortic aneurysm	0.822 (0.811, 0.834)	0.824 (0.81, 0.836)	0.001 (-0.005, 0.007)
phecode_721-5	Bursitis	0.535 (0.529, 0.542)	0.536 (0.529, 0.543)	0.001 (-0.004, 0.005)
phecode_734	Diseases of the jaws	0.584 (0.565, 0.605)	0.585 (0.565, 0.604)	0.001 (-0.011, 0.013)
phecode_603-6	Scrotal pain*	0.804 (0.795, 0.813)	0.805 (0.796, 0.813)	0.001 (-0.004, 0.006)
phecode_396	Hearing impairment	0.638 (0.633, 0.643)	0.638 (0.634, 0.643)	0.001 (0, 0.002)
phecode_546	Other disorders of liver	0.592 (0.576, 0.605)	0.592 (0.579, 0.607)	0.001 (-0.011, 0.015)
phecode_390-4	Impacted cerumen	0.59 (0.585, 0.595)	0.591 (0.586, 0.596)	0.001 (0, 0.001)
phecode_176	Other diseases of blood and blood-forming organs	0.569 (0.552, 0.59)	0.571 (0.548, 0.588)	0.001 (-0.02, 0.021)
phecode_820	Elevated erythrocyte sedimentation rate and abnormality of plasma viscosity	0.656 (0.628, 0.702)	0.657 (0.619, 0.698)	0.001 (-0.025, 0.022)

Supplementary Tables

Table 23 continued from previous page

Endpoint	Phecode String	Age+Sex	Age+Sex+Retina	Delta C-Index
phecode_471-2	Deviated nasal septum	0.583 (0.566, 0.605)	0.584 (0.567, 0.607)	0.001 (-0.012, 0.013)
phecode_627	Menopausal and postmenopausal disorders	0.686 (0.681, 0.691)	0.687 (0.682, 0.692)	0.001 (0, 0.002)
phecode_135	Benign neoplasm of the head and neck	0.558 (0.542, 0.577)	0.559 (0.542, 0.573)	0.001 (-0.012, 0.014)
phecode_848-2	Nonspecific abnormal findings on radiological and other examination of other intrathoracic organs (echo)	0.697 (0.685, 0.709)	0.697 (0.685, 0.709)	0.001 (-0.006, 0.007)
phecode_440-1	Venous thromboembolism	0.566 (0.554, 0.575)	0.566 (0.555, 0.574)	0.001 (-0.007, 0.006)
phecode_511	Gastro-esophageal reflux disease	0.561 (0.556, 0.566)	0.562 (0.556, 0.567)	0.001 (-0.003, 0.006)
phecode_707-8	Polyarthrits	0.603 (0.58, 0.629)	0.602 (0.578, 0.631)	0.001 (-0.014, 0.015)
phecode_705	Rheumatoid arthritis and other inflammatory polyarthropathies	0.619 (0.608, 0.628)	0.62 (0.61, 0.629)	0.001 (-0.004, 0.007)
phecode_714-3	Scoliosis	0.693 (0.67, 0.712)	0.694 (0.676, 0.712)	0.001 (-0.009, 0.01)
phecode_487	Hemorrhage from respiratory passages	0.587 (0.576, 0.597)	0.587 (0.576, 0.596)	0.001 (-0.004, 0.005)
phecode_431-2	Transient cerebral ischemic attacks	0.677 (0.666, 0.69)	0.678 (0.667, 0.69)	0.001 (-0.005, 0.006)
phecode_391-9	Otorrhea	0.553 (0.529, 0.58)	0.553 (0.534, 0.58)	0.001 (-0.016, 0.023)
phecode_329-1	Memory loss	0.651 (0.643, 0.66)	0.652 (0.644, 0.663)	0.001 (-0.004, 0.005)
phecode_200-2	Goiter	0.639 (0.625, 0.655)	0.641 (0.627, 0.655)	0.001 (-0.005, 0.008)
phecode_331-61	Migraine with aura	0.595 (0.576, 0.619)	0.598 (0.575, 0.62)	0.001 (-0.009, 0.017)
phecode_112	Malignant neoplasm of other and ill-defined sites	0.662 (0.655, 0.667)	0.663 (0.657, 0.669)	0.001 (-0.001, 0.003)
phecode_540	Hepatitis	0.567 (0.531, 0.595)	0.569 (0.535, 0.599)	0.001 (-0.023, 0.022)
phecode_710	Acquired deformities of limbs	0.599 (0.583, 0.611)	0.6 (0.585, 0.613)	0.001 (-0.01, 0.012)
phecode_390	Disorders of external ear	0.559 (0.556, 0.564)	0.561 (0.557, 0.565)	0.001 (0, 0.002)
phecode_619-4	Noninflammatory disorders of vagina	0.597 (0.589, 0.605)	0.598 (0.589, 0.607)	0.001 (0, 0.003)
phecode_835	Cytology and pathology findings	0.722 (0.713, 0.729)	0.723 (0.714, 0.73)	0.002 (0, 0.003)
phecode_392	Otalgia and effusion of ear	0.56 (0.552, 0.568)	0.562 (0.555, 0.569)	0.002 (-0.003, 0.007)
phecode_708-14	Primary osteoarthritis of the shoulder, upper arm	0.605 (0.589, 0.62)	0.605 (0.59, 0.625)	0.002 (-0.011, 0.012)
phecode_708-16	Primary osteoarthritis ankle and foot	0.578 (0.566, 0.592)	0.579 (0.564, 0.595)	0.002 (-0.01, 0.01)
phecode_514	Gastrointestinal obstruction	0.625 (0.613, 0.64)	0.626 (0.616, 0.642)	0.002 (-0.005, 0.01)
phecode_556	Other symptoms involving the digestive system and abdomen	0.557 (0.551, 0.564)	0.559 (0.552, 0.566)	0.002 (-0.003, 0.006)
phecode_703-2	Chondrocalcinosis	0.678 (0.649, 0.707)	0.68 (0.645, 0.711)	0.002 (-0.013, 0.015)
phecode_507-1	Stomatitis	0.558 (0.545, 0.572)	0.558 (0.543, 0.573)	0.002 (-0.008, 0.011)
phecode_481	Interstitial pulmonary diseases	0.729 (0.711, 0.744)	0.731 (0.715, 0.748)	0.002 (-0.011, 0.017)
phecode_734-9	Jaw pain	0.589 (0.569, 0.61)	0.591 (0.57, 0.615)	0.002 (-0.012, 0.017)
phecode_367-1	Conjunctivitis	0.57 (0.563, 0.578)	0.571 (0.564, 0.579)	0.002 (-0.001, 0.005)
phecode_969	Adverse effects of agents primarily affecting gastrointestinal system	0.593 (0.587, 0.6)	0.595 (0.589, 0.601)	0.002 (-0.001, 0.005)
phecode_708-13	Primary osteoarthritis of the hand	0.644 (0.633, 0.659)	0.646 (0.636, 0.66)	0.002 (-0.004, 0.009)
phecode_523-2	Diverticula of colon	0.63 (0.625, 0.636)	0.632 (0.627, 0.638)	0.002 (0, 0.005)
phecode_330-3	Convulsions	0.567 (0.546, 0.589)	0.568 (0.546, 0.586)	0.002 (-0.023, 0.018)
phecode_727	Other disorders of bone	0.71 (0.706, 0.716)	0.712 (0.708, 0.718)	0.002 (0.001, 0.003)
phecode_367-13	Blepharoconjunctivitis	0.637 (0.615, 0.673)	0.641 (0.617, 0.67)	0.002 (-0.01, 0.011)
phecode_667	Erythematous conditions	0.555 (0.534, 0.57)	0.556 (0.54, 0.571)	0.002 (-0.017, 0.024)
phecode_404-11	Acute myocardial infarction	0.702 (0.693, 0.711)	0.704 (0.694, 0.713)	0.002 (-0.004, 0.008)
phecode_416-2	Atrial fibrillation and flutter	0.731 (0.726, 0.736)	0.733 (0.728, 0.739)	0.002 (-0.001, 0.005)
phecode_581	Renal tubulo-interstitial diseases	0.547 (0.533, 0.559)	0.549 (0.537, 0.562)	0.002 (-0.01, 0.015)
phecode_841-1	Allergy to other anti-infective agents	0.602 (0.595, 0.609)	0.604 (0.596, 0.611)	0.002 (-0.002, 0.006)
phecode_726	Osteoporosis and low bone density	0.741 (0.734, 0.747)	0.743 (0.737, 0.748)	0.002 (0, 0.004)
phecode_715-1	Spondylosis	0.62 (0.611, 0.627)	0.622 (0.613, 0.629)	0.002 (-0.002, 0.006)
phecode_416-21	Atrial fibrillation	0.731 (0.722, 0.737)	0.733 (0.725, 0.74)	0.002 (0, 0.006)

5 Phenome-wide prediction of disease onset from retinal fundus photographs

Table 23 continued from previous page

Endpoint	Phecode String	Age+Sex	Age+Sex+Retina	Delta C-Index
phecode_700	Diffuse diseases of connective tissue	0.63 (0.613, 0.652)	0.635 (0.613, 0.652)	0.002 (-0.012, 0.02)
phecode_556-8	Nonspecific abnormal findings in stool contents	0.56 (0.552, 0.568)	0.562 (0.554, 0.571)	0.002 (-0.003, 0.009)
phecode_478	Aspiration pneumonia	0.738 (0.719, 0.757)	0.74 (0.722, 0.758)	0.002 (-0.012, 0.018)
phecode_331	Headache	0.579 (0.571, 0.584)	0.581 (0.575, 0.585)	0.002 (0, 0.005)
phecode_438	Aneurysm or ectasia	0.718 (0.705, 0.73)	0.721 (0.708, 0.732)	0.002 (-0.005, 0.008)
phecode_726-1	Osteoporosis	0.746 (0.74, 0.753)	0.748 (0.743, 0.755)	0.003 (0.001, 0.005)
phecode_433-1	Occlusion and stenosis of cerebral arteries	0.716 (0.693, 0.746)	0.719 (0.695, 0.739)	0.003 (-0.02, 0.019)
phecode_465-2	Chronic Pharyngitis	0.564 (0.558, 0.571)	0.567 (0.56, 0.574)	0.003 (-0.001, 0.008)
phecode_627-1	Postmenopausal bleeding	0.551 (0.542, 0.559)	0.553 (0.545, 0.562)	0.003 (-0.003, 0.008)
phecode_288	Anxiety disorders	0.59 (0.584, 0.596)	0.593 (0.587, 0.598)	0.003 (0, 0.006)
phecode_586	Other disorders of the kidney and ureters	0.644 (0.632, 0.653)	0.645 (0.637, 0.658)	0.003 (-0.008, 0.01)
phecode_447	Nonspecific low blood-pressure reading	0.642 (0.623, 0.664)	0.646 (0.625, 0.668)	0.003 (-0.012, 0.016)
phecode_708	Osteoarthritis	0.612 (0.609, 0.616)	0.615 (0.612, 0.619)	0.003 (0.002, 0.004)
phecode_230-2	Abnormal loss of weight and underweight	0.608 (0.598, 0.617)	0.61 (0.602, 0.62)	0.003 (-0.003, 0.007)
phecode_417-1	Palpitations	0.577 (0.571, 0.584)	0.58 (0.574, 0.587)	0.003 (-0.001, 0.007)
phecode_507	Lesions of mouth	0.565 (0.555, 0.576)	0.569 (0.556, 0.578)	0.003 (-0.003, 0.009)
phecode_528-2	Vomiting	0.585 (0.576, 0.591)	0.588 (0.579, 0.595)	0.003 (-0.001, 0.008)
phecode_116-2	Secondary malignancy of respiratory organs	0.632 (0.618, 0.646)	0.634 (0.616, 0.65)	0.003 (-0.01, 0.014)
phecode_329	Symptoms and signs involving cognitive functions and awareness	0.644 (0.637, 0.652)	0.648 (0.64, 0.654)	0.003 (-0.002, 0.007)
phecode_406	Chronic pulmonary heart disease	0.741 (0.723, 0.758)	0.745 (0.726, 0.764)	0.003 (-0.012, 0.02)
phecode_723	Enthesopathy/Enthesitis/Tendinopathy	0.535 (0.529, 0.538)	0.538 (0.533, 0.542)	0.003 (-0.001, 0.006)
phecode_523	Diverticular disease [Bowel diverticulosis]	0.622 (0.616, 0.627)	0.625 (0.62, 0.63)	0.003 (0.001, 0.005)
phecode_089-2	Viral infections	0.528 (0.522, 0.532)	0.531 (0.525, 0.536)	0.003 (-0.001, 0.006)
phecode_030	Campylobacter	0.532 (0.516, 0.552)	0.535 (0.518, 0.555)	0.003 (-0.022, 0.019)
phecode_479-6	Pulmonary collapse [Atelectasis]	0.646 (0.631, 0.666)	0.649 (0.635, 0.664)	0.003 (-0.009, 0.012)
phecode_446-2	Orthostatic hypotension	0.69 (0.677, 0.703)	0.692 (0.679, 0.706)	0.003 (-0.004, 0.01)
phecode_715-3	Spondylolisthesis	0.663 (0.638, 0.679)	0.665 (0.644, 0.682)	0.003 (-0.012, 0.015)
phecode_528-1	Nausea	0.585 (0.576, 0.59)	0.588 (0.579, 0.594)	0.003 (0, 0.008)
phecode_007	Hemophilus infection	0.611 (0.572, 0.669)	0.615 (0.584, 0.66)	0.003 (-0.025, 0.035)
phecode_136	Benign neoplasm of the digestive organs	0.589 (0.583, 0.593)	0.592 (0.586, 0.597)	0.003 (0.001, 0.006)
phecode_506-3	Sialoadenitis	0.553 (0.53, 0.575)	0.555 (0.527, 0.582)	0.003 (-0.019, 0.023)
phecode_440-13	Phlebitis and thrombophlebitis	0.583 (0.575, 0.593)	0.587 (0.575, 0.596)	0.003 (-0.003, 0.01)
phecode_354	Dizziness and giddiness	0.571 (0.566, 0.576)	0.575 (0.569, 0.58)	0.003 (0, 0.006)
phecode_597-1	Urethral stricture	0.655 (0.631, 0.675)	0.657 (0.638, 0.678)	0.003 (-0.01, 0.013)
phecode_230-21	Abnormal weight loss	0.612 (0.602, 0.621)	0.616 (0.606, 0.624)	0.003 (-0.002, 0.009)
phecode_715	Non-inflammatory spondylopathy	0.622 (0.615, 0.629)	0.625 (0.618, 0.633)	0.003 (-0.001, 0.007)
phecode_708-7	Generalized osteoarthritis	0.65 (0.643, 0.659)	0.654 (0.646, 0.661)	0.003 (-0.003, 0.008)
phecode_682-12	Pilar and trichodermal cyst	0.502 (0.491, 0.511)	0.506 (0.496, 0.513)	0.003 (-0.004, 0.013)
phecode_583	Chronic kidney disease	0.684 (0.677, 0.692)	0.688 (0.681, 0.693)	0.003 (0, 0.007)
phecode_708-1	Primary osteoarthritis	0.611 (0.603, 0.617)	0.614 (0.607, 0.62)	0.003 (0.001, 0.005)
phecode_169-1	Thrombocytopenia	0.667 (0.652, 0.683)	0.67 (0.656, 0.692)	0.003 (-0.008, 0.016)
phecode_423-1	Cardiac murmurs	0.641 (0.623, 0.655)	0.644 (0.628, 0.66)	0.003 (-0.007, 0.015)
phecode_720	Spontaneous rupture of synovium and tendon	0.605 (0.589, 0.63)	0.608 (0.593, 0.633)	0.004 (-0.006, 0.011)
phecode_529-3	Fecal incontinence	0.639 (0.626, 0.652)	0.642 (0.627, 0.657)	0.004 (-0.005, 0.014)
phecode_528	Nausea and vomiting	0.584 (0.577, 0.592)	0.588 (0.581, 0.594)	0.004 (-0.001, 0.007)
phecode_550-4	Cholesterosis of gallbladder	0.582 (0.545, 0.613)	0.586 (0.553, 0.615)	0.004 (-0.009, 0.018)
phecode_170-1	Neutropenia	0.585 (0.573, 0.599)	0.589 (0.577, 0.603)	0.004 (-0.008, 0.013)
phecode_363	Disorders of lacrimal system	0.634 (0.628, 0.64)	0.638 (0.632, 0.644)	0.004 (0.001, 0.007)

Supplementary Tables

Table 23 continued from previous page

Endpoint	Phecode String	Age+Sex	Age+Sex+Retina	Delta C-Index
phecode_280-8	Other psychoactive substance related disorders	0.598 (0.553, 0.638)	0.598 (0.565, 0.643)	0.004 (-0.029, 0.042)
phecode_819	General symptoms and other findings	0.582 (0.577, 0.586)	0.586 (0.581, 0.59)	0.004 (0.001, 0.007)
phecode_512	Dysphagia	0.575 (0.567, 0.585)	0.578 (0.57, 0.59)	0.004 (-0.004, 0.012)
phecode_530	Disease of anus and rectum	0.527 (0.518, 0.54)	0.53 (0.52, 0.544)	0.004 (-0.005, 0.013)
phecode_200-22	Uninodular goiter [single thyroid nodule]	0.628 (0.603, 0.651)	0.631 (0.604, 0.651)	0.004 (-0.01, 0.012)
phecode_446	Hypotension	0.663 (0.655, 0.671)	0.667 (0.66, 0.674)	0.004 (-0.001, 0.011)
phecode_423	Abnormal cardiac sounds	0.642 (0.626, 0.656)	0.645 (0.63, 0.661)	0.004 (-0.008, 0.015)
phecode_288-3	Generalized anxiety disorder	0.593 (0.58, 0.609)	0.597 (0.585, 0.61)	0.004 (-0.001, 0.008)
phecode_406-1	Pulmonary hypertension	0.745 (0.724, 0.761)	0.747 (0.725, 0.768)	0.004 (-0.013, 0.023)
phecode_363-2	Dry eye syndrome [Tear film insufficiency]	0.64 (0.634, 0.648)	0.644 (0.638, 0.653)	0.004 (0, 0.007)
phecode_829	Nonspecific findings on examination of blood	0.652 (0.646, 0.659)	0.656 (0.65, 0.664)	0.004 (0.001, 0.006)
phecode_324-8	Restless legs syndrome	0.592 (0.572, 0.613)	0.597 (0.58, 0.618)	0.004 (-0.01, 0.023)
phecode_703	Chrystal arthropathies	0.698 (0.691, 0.704)	0.702 (0.695, 0.71)	0.004 (0.001, 0.008)
phecode_444-5	Venous insufficiency (chronic) (peripheral)	0.662 (0.648, 0.676)	0.666 (0.654, 0.68)	0.004 (-0.002, 0.011)
phecode_101-6	Malignant neoplasm of the liver and intrahepatic bile ducts	0.719 (0.685, 0.752)	0.725 (0.678, 0.76)	0.004 (-0.022, 0.028)
phecode_397-1	Tinnitus	0.503 (0.493, 0.514)	0.507 (0.497, 0.515)	0.004 (-0.007, 0.016)
phecode_557-2	Blood in stool	0.621 (0.608, 0.64)	0.626 (0.611, 0.642)	0.004 (-0.007, 0.015)
phecode_440	Embolism and thrombosis	0.597 (0.588, 0.605)	0.601 (0.594, 0.608)	0.004 (-0.001, 0.008)
phecode_514-2	Intestinal obstruction	0.621 (0.606, 0.636)	0.625 (0.608, 0.643)	0.004 (-0.005, 0.015)
phecode_465-1	Acute Pharyngitis	0.561 (0.555, 0.567)	0.566 (0.558, 0.573)	0.004 (0, 0.008)
phecode_398	Other disorders of ear	0.575 (0.563, 0.588)	0.579 (0.567, 0.595)	0.004 (-0.004, 0.011)
phecode_294	Sexual dysfunction and disorders	0.786 (0.782, 0.791)	0.791 (0.786, 0.795)	0.004 (0.003, 0.006)
phecode_349	Disorder of nervous system	0.603 (0.587, 0.616)	0.607 (0.59, 0.62)	0.004 (-0.008, 0.016)
phecode_169	Platelet defects	0.666 (0.649, 0.68)	0.669 (0.654, 0.684)	0.004 (-0.006, 0.016)
phecode_675	Atrophic conditions of skin	0.716 (0.704, 0.726)	0.72 (0.709, 0.73)	0.004 (0, 0.008)
phecode_557	Gastrointestinal hemorrhage	0.532 (0.525, 0.538)	0.536 (0.53, 0.544)	0.004 (-0.001, 0.01)
phecode_404	Ischemic heart disease	0.69 (0.683, 0.695)	0.694 (0.687, 0.699)	0.004 (-0.001, 0.008)
phecode_136-41	Benign neoplasm of the colon	0.601 (0.595, 0.607)	0.606 (0.599, 0.612)	0.004 (0.001, 0.008)
phecode_675-1	Circumscribed scleroderma	0.716 (0.704, 0.727)	0.72 (0.708, 0.734)	0.004 (0, 0.008)
phecode_419	Presence of cardiac device	0.738 (0.723, 0.755)	0.743 (0.727, 0.759)	0.004 (-0.003, 0.011)
phecode_136-42	Benign neoplasm of rectum and anus	0.594 (0.586, 0.605)	0.599 (0.589, 0.611)	0.004 (-0.001, 0.009)
phecode_679	Skin symptoms	0.536 (0.532, 0.541)	0.541 (0.537, 0.545)	0.004 (0.002, 0.007)
phecode_390-6	Perichondritis and chondritis of pinna	0.678 (0.651, 0.7)	0.68 (0.656, 0.703)	0.004 (-0.008, 0.016)
phecode_674-2	Hypepigmentation	0.587 (0.569, 0.609)	0.593 (0.573, 0.613)	0.004 (-0.006, 0.014)
phecode_431	Stroke and transient cerebral ischemic attacks	0.689 (0.679, 0.696)	0.693 (0.686, 0.701)	0.005 (-0.001, 0.01)
phecode_815	Symptoms and signs concerning food and fluid intake	0.538 (0.522, 0.56)	0.542 (0.524, 0.566)	0.005 (-0.011, 0.021)
phecode_436	Atherosclerosis [ASCVD]	0.704 (0.697, 0.711)	0.709 (0.701, 0.715)	0.005 (0.001, 0.009)
phecode_404-2	Coronary atherosclerosis [Atherosclerotic heart disease]	0.703 (0.696, 0.71)	0.707 (0.701, 0.714)	0.005 (0, 0.008)
phecode_472	Diseases of vocal cords and larynx, not elsewhere classified	0.57 (0.543, 0.594)	0.573 (0.551, 0.608)	0.005 (-0.025, 0.03)
phecode_613-7	Other signs and symptoms in breast	0.458 (0.434, 0.477)	0.463 (0.447, 0.482)	0.005 (-0.013, 0.03)
phecode_330	Epilepsy, recurrent seizures, convulsions	0.578 (0.561, 0.597)	0.583 (0.563, 0.602)	0.005 (-0.012, 0.019)
phecode_465	Pharyngitis	0.561 (0.554, 0.566)	0.566 (0.558, 0.571)	0.005 (0, 0.008)
phecode_170	Decreased white blood cell count	0.586 (0.574, 0.597)	0.591 (0.579, 0.605)	0.005 (-0.004, 0.019)
phecode_103-22	Squamous cell carcinoma of the skin	0.716 (0.695, 0.743)	0.721 (0.699, 0.744)	0.005 (-0.004, 0.011)
phecode_678	Other skin and connective tissue disorders	0.553 (0.549, 0.557)	0.558 (0.554, 0.563)	0.005 (0.003, 0.007)
phecode_477	Inhalation lung injury	0.719 (0.677, 0.749)	0.723 (0.683, 0.769)	0.005 (-0.012, 0.03)

5 Phenome-wide prediction of disease onset from retinal fundus photographs

Table 23 continued from previous page

Endpoint	Phecode String	Age+Sex	Age+Sex+Retina	Delta C-Index
phecode_308	Signs and symptoms involving emotional state	0.564 (0.557, 0.571)	0.569 (0.561, 0.576)	0.005 (-0.002, 0.01)
phecode_723-2	Rotator cuff tear or rupture	0.536 (0.528, 0.547)	0.541 (0.533, 0.551)	0.005 (-0.004, 0.013)
phecode_718-3	Mid back pain	0.574 (0.558, 0.591)	0.58 (0.567, 0.593)	0.005 (-0.005, 0.016)
phecode_529-5	Constipation	0.609 (0.604, 0.616)	0.614 (0.609, 0.621)	0.005 (0.002, 0.009)
phecode_391-1	Otitis media	0.527 (0.516, 0.536)	0.532 (0.522, 0.54)	0.005 (-0.002, 0.014)
phecode_414	Cardiomyopathy	0.658 (0.635, 0.679)	0.665 (0.644, 0.684)	0.005 (-0.005, 0.019)
phecode_103-3	Carcinoma in situ of skin	0.718 (0.701, 0.729)	0.724 (0.706, 0.734)	0.005 (0.001, 0.009)
phecode_367	Inflammation of the eye	0.564 (0.558, 0.572)	0.569 (0.564, 0.576)	0.005 (0.001, 0.009)
phecode_688-3	Pyogenic granuloma of skin and subcutaneous tissue	0.51 (0.48, 0.543)	0.519 (0.486, 0.546)	0.005 (-0.021, 0.041)
phecode_527	Abdominal pain	0.548 (0.544, 0.553)	0.553 (0.549, 0.558)	0.005 (0.002, 0.008)
phecode_841	Drug and medical agent allergy	0.604 (0.6, 0.608)	0.609 (0.604, 0.614)	0.005 (0.002, 0.008)
phecode_529-2	Abdominal distension and flatulence	0.569 (0.561, 0.575)	0.573 (0.566, 0.582)	0.005 (0, 0.011)
phecode_413-21	Aortic stenosis	0.755 (0.741, 0.765)	0.758 (0.747, 0.772)	0.005 (-0.002, 0.012)
phecode_136-4	Benign neoplasm of colon, rectum, anus and anal canal	0.59 (0.585, 0.597)	0.596 (0.59, 0.602)	0.005 (0.002, 0.009)
phecode_350	Other symptoms involving nervous system	0.61 (0.602, 0.617)	0.615 (0.608, 0.622)	0.005 (0.001, 0.01)
phecode_286	Mood [affective] disorders	0.569 (0.562, 0.576)	0.574 (0.569, 0.582)	0.005 (0.002, 0.01)
phecode_230	Malnutrition and underweight	0.602 (0.593, 0.61)	0.607 (0.598, 0.614)	0.005 (0, 0.011)
phecode_660-12	Dermatophytosis	0.55 (0.543, 0.558)	0.556 (0.549, 0.564)	0.005 (0.002, 0.009)
phecode_256-2	Hyposmolality and/or hyponatremia	0.534 (0.524, 0.546)	0.539 (0.531, 0.55)	0.005 (-0.001, 0.013)
phecode_144-3	Benign neoplasms of the ovary	0.492 (0.456, 0.524)	0.498 (0.464, 0.539)	0.005 (-0.031, 0.057)
phecode_557-8	Hemorrhage of rectum and anus	0.51 (0.502, 0.521)	0.516 (0.507, 0.525)	0.006 (-0.002, 0.015)
phecode_522-7	Ulceration of the lower GI tract	0.541 (0.516, 0.57)	0.548 (0.52, 0.581)	0.006 (-0.019, 0.038)
phecode_550	Disorders of the gallbladder	0.572 (0.566, 0.582)	0.579 (0.572, 0.587)	0.006 (0.001, 0.012)
phecode_668-3	Contact dermatitis	0.539 (0.525, 0.549)	0.545 (0.529, 0.557)	0.006 (-0.003, 0.013)
phecode_383	Irregular eye movements	0.541 (0.497, 0.591)	0.546 (0.501, 0.589)	0.006 (-0.031, 0.041)
phecode_413-32	Tricuspid valve insufficiency*	0.724 (0.695, 0.752)	0.73 (0.705, 0.758)	0.006 (-0.014, 0.029)
phecode_732	Nonspecific abnormal findings on radiological and other examination of musculoskeletal system	0.632 (0.62, 0.641)	0.638 (0.626, 0.647)	0.006 (0.001, 0.009)
phecode_420	Cardiac arrest	0.728 (0.709, 0.748)	0.733 (0.713, 0.756)	0.006 (-0.009, 0.022)
phecode_550-1	Gallstones [Cholelithiasis]	0.579 (0.57, 0.587)	0.583 (0.577, 0.593)	0.006 (-0.001, 0.012)
phecode_391	Disorders of the middle ear	0.522 (0.516, 0.53)	0.528 (0.522, 0.535)	0.006 (0, 0.012)
phecode_800-11	Pleurodynia*	0.546 (0.535, 0.559)	0.552 (0.54, 0.564)	0.006 (-0.006, 0.017)
phecode_174	Diseases of spleen	0.586 (0.558, 0.62)	0.594 (0.567, 0.617)	0.006 (-0.018, 0.031)
phecode_724-52	Osteophyte*	0.582 (0.565, 0.601)	0.588 (0.567, 0.607)	0.006 (-0.008, 0.015)
phecode_330-11	Generalized epilepsy	0.536 (0.501, 0.578)	0.54 (0.498, 0.59)	0.006 (-0.042, 0.055)
phecode_288-2	Panic disorder [episodic paroxysmal anxiety]	0.567 (0.55, 0.586)	0.575 (0.557, 0.591)	0.006 (-0.009, 0.024)
phecode_703-11	Gout	0.706 (0.698, 0.716)	0.712 (0.702, 0.722)	0.006 (0.002, 0.01)
phecode_722-4	Palmar fascial fibromatosis [Dupuytren]	0.645 (0.634, 0.658)	0.651 (0.641, 0.663)	0.006 (0.002, 0.011)
phecode_116-5	Secondary malignancy of brain/spine	0.625 (0.599, 0.651)	0.628 (0.608, 0.66)	0.006 (-0.022, 0.028)
phecode_116-6	Secondary malignancy of bone	0.687 (0.677, 0.699)	0.694 (0.682, 0.706)	0.006 (-0.001, 0.016)
phecode_471	Other disorders of nose and nasal sinuses	0.507 (0.499, 0.513)	0.513 (0.507, 0.519)	0.006 (-0.002, 0.016)
phecode_668	Dermatitis [eczema]	0.537 (0.531, 0.542)	0.543 (0.538, 0.55)	0.006 (0.002, 0.01)
phecode_286-2	Major depressive disorder	0.569 (0.562, 0.574)	0.574 (0.569, 0.581)	0.006 (0.002, 0.011)
phecode_603-2	Spermatocele	0.531 (0.505, 0.552)	0.538 (0.519, 0.562)	0.006 (-0.008, 0.027)
phecode_841-11	Penicillin allergy	0.584 (0.577, 0.593)	0.59 (0.581, 0.597)	0.006 (0.001, 0.01)
phecode_532	Other disorders of the intestines	0.553 (0.54, 0.566)	0.559 (0.543, 0.574)	0.006 (-0.002, 0.017)
phecode_668-1	Atopic dermatitis	0.519 (0.505, 0.528)	0.524 (0.511, 0.535)	0.006 (-0.007, 0.018)
phecode_508	Diseases of lips	0.584 (0.569, 0.6)	0.591 (0.575, 0.606)	0.006 (-0.006, 0.019)
phecode_840-2	Allergy to insects	0.572 (0.56, 0.583)	0.579 (0.568, 0.59)	0.006 (0, 0.014)

Supplementary Tables

Table 23 continued from previous page

Endpoint	Phecode String	Age+Sex	Age+Sex+Retina	Delta C-Index
phecode_431-12	Hemorrhagic stroke	0.65 (0.626, 0.676)	0.658 (0.635, 0.675)	0.006 (-0.012, 0.026)
phecode_827-1	Finding of alcohol in blood	0.627 (0.609, 0.652)	0.635 (0.616, 0.657)	0.006 (-0.002, 0.015)
phecode_802	Throat pain	0.549 (0.537, 0.559)	0.555 (0.544, 0.567)	0.007 (0, 0.013)
phecode_403	Angina pectoris	0.672 (0.665, 0.682)	0.679 (0.671, 0.686)	0.007 (0.003, 0.011)
phecode_713-3	Pain in joint	0.526 (0.523, 0.53)	0.533 (0.529, 0.536)	0.007 (0.004, 0.009)
phecode_431-1	Stroke	0.695 (0.683, 0.706)	0.702 (0.692, 0.712)	0.007 (-0.001, 0.017)
phecode_703-1	Hyperuricemia	0.701 (0.693, 0.708)	0.707 (0.7, 0.715)	0.007 (0.003, 0.01)
phecode_397	Other hearing abnormality	0.502 (0.494, 0.511)	0.509 (0.5, 0.517)	0.007 (-0.005, 0.017)
phecode_483	Pleural effusion	0.682 (0.672, 0.691)	0.688 (0.679, 0.696)	0.007 (0.002, 0.014)
phecode_827	Toxicology findings	0.622 (0.604, 0.647)	0.629 (0.613, 0.65)	0.007 (-0.003, 0.015)
phecode_713	Symptoms related to joints	0.526 (0.523, 0.53)	0.533 (0.53, 0.536)	0.007 (0.004, 0.009)
phecode_439	Hemorrhoids	0.504 (0.499, 0.509)	0.511 (0.506, 0.517)	0.007 (0, 0.012)
phecode_721-1	Synovitis and tenosynovitis	0.539 (0.534, 0.547)	0.546 (0.54, 0.553)	0.007 (0.001, 0.012)
phecode_411	Other diseases of pericardium	0.642 (0.619, 0.662)	0.65 (0.626, 0.67)	0.007 (-0.013, 0.026)
phecode_807-11	Postviral fatigue syndrome*	0.605 (0.574, 0.644)	0.611 (0.576, 0.646)	0.007 (-0.015, 0.023)
phecode_089	Infections	0.521 (0.517, 0.525)	0.528 (0.524, 0.531)	0.007 (0.003, 0.01)
phecode_092	Bacteremia, Sepsis, and SIRS	0.666 (0.655, 0.676)	0.673 (0.663, 0.683)	0.007 (-0.001, 0.014)
phecode_431-11	Cerebral infarction [Ischemic stroke]	0.704 (0.692, 0.715)	0.711 (0.7, 0.721)	0.007 (0, 0.014)
phecode_247-5	Disorders of calcium metabolism	0.609 (0.595, 0.621)	0.615 (0.6, 0.632)	0.007 (-0.005, 0.019)
phecode_230-3	Anorexia	0.604 (0.588, 0.623)	0.611 (0.593, 0.629)	0.007 (-0.006, 0.019)
phecode_554-2	Cyst and pseudocyst of pancreas	0.626 (0.591, 0.658)	0.634 (0.6, 0.663)	0.007 (-0.016, 0.033)
phecode_404-1	Myocardial infarction [Heart attack]	0.707 (0.699, 0.715)	0.715 (0.707, 0.723)	0.007 (0.001, 0.012)
phecode_424-3	Diastolic heart failure	0.697 (0.667, 0.722)	0.703 (0.677, 0.73)	0.007 (-0.017, 0.026)
phecode_426	Other heart disorders in diseases NEC	0.709 (0.697, 0.72)	0.715 (0.705, 0.729)	0.007 (-0.002, 0.015)
phecode_092-2	Sepsis	0.663 (0.652, 0.672)	0.671 (0.66, 0.682)	0.007 (0.001, 0.013)
phecode_443	Other specified disorders of arteries and arterioles	0.687 (0.677, 0.702)	0.696 (0.682, 0.708)	0.008 (-0.001, 0.017)
phecode_513	Peptic ulcer	0.594 (0.585, 0.603)	0.601 (0.595, 0.613)	0.008 (-0.002, 0.018)
phecode_726-2	Pathologic fracture	0.724 (0.706, 0.745)	0.732 (0.714, 0.751)	0.008 (-0.005, 0.021)
phecode_424-2	Styolic heart failure	0.721 (0.696, 0.743)	0.73 (0.702, 0.751)	0.008 (-0.003, 0.019)
phecode_715-4	Spinal stenosis	0.645 (0.633, 0.657)	0.653 (0.641, 0.666)	0.008 (0.001, 0.015)
phecode_473	Other diseases of upper respiratory tract	0.537 (0.525, 0.551)	0.545 (0.532, 0.558)	0.008 (0, 0.014)
phecode_522	Gastrointestinal inflammation	0.555 (0.549, 0.56)	0.563 (0.558, 0.568)	0.008 (0.003, 0.012)
phecode_830	Proteinuria	0.584 (0.579, 0.59)	0.591 (0.587, 0.597)	0.008 (0.005, 0.012)
phecode_469-1	Acute bronchitis	0.564 (0.542, 0.584)	0.571 (0.545, 0.594)	0.008 (-0.016, 0.031)
phecode_070	Candidiasis	0.586 (0.579, 0.594)	0.594 (0.586, 0.603)	0.008 (0.003, 0.013)
phecode_239-11	Pure hypercholesterolemia	0.614 (0.611, 0.618)	0.623 (0.619, 0.627)	0.008 (0.006, 0.011)
phecode_239-1	Hypercholesterolemia	0.613 (0.609, 0.617)	0.621 (0.616, 0.626)	0.008 (0.006, 0.01)
phecode_239	Hyperlipidemia	0.606 (0.602, 0.612)	0.615 (0.61, 0.62)	0.008 (0.005, 0.011)
phecode_800-1	Chest pain on breathing	0.54 (0.53, 0.557)	0.55 (0.54, 0.563)	0.008 (-0.002, 0.02)
phecode_440-2	Arterial embolism and thrombosis	0.689 (0.668, 0.709)	0.697 (0.679, 0.715)	0.008 (-0.008, 0.02)
phecode_526-1	Bacterial enteritis	0.528 (0.518, 0.543)	0.538 (0.525, 0.55)	0.008 (-0.002, 0.022)
phecode_660-1	Fungal infection of the skin	0.541 (0.535, 0.548)	0.55 (0.544, 0.556)	0.008 (0.005, 0.012)
phecode_713-4	Stiffness of joint	0.529 (0.512, 0.545)	0.537 (0.511, 0.557)	0.008 (-0.008, 0.024)
phecode_375-12	Angle-Closure Glaucoma	0.675 (0.65, 0.698)	0.683 (0.659, 0.712)	0.008 (-0.008, 0.027)
phecode_724-1	Myalgia	0.555 (0.547, 0.564)	0.563 (0.553, 0.571)	0.009 (0.001, 0.015)
phecode_605	Male sexual dysfunction	0.558 (0.552, 0.564)	0.567 (0.56, 0.575)	0.009 (0.005, 0.013)
phecode_705-1	Rheumatoid arthritis	0.591 (0.576, 0.609)	0.601 (0.583, 0.617)	0.009 (-0.001, 0.023)
phecode_360-2	Chalazion	0.52 (0.503, 0.535)	0.531 (0.514, 0.541)	0.009 (-0.002, 0.018)
phecode_437	Vascular insufficiency of intestine	0.665 (0.634, 0.687)	0.673 (0.645, 0.707)	0.009 (-0.008, 0.03)
phecode_823	Abnormal serum enzyme levels	0.56 (0.55, 0.569)	0.569 (0.56, 0.578)	0.009 (0.003, 0.014)
phecode_361-1	Entropion and trichiasis of eyelid	0.682 (0.651, 0.712)	0.691 (0.662, 0.72)	0.009 (-0.013, 0.024)
phecode_103-2	Keratinocyte carcinoma	0.652 (0.646, 0.659)	0.66 (0.654, 0.668)	0.009 (0.006, 0.011)
phecode_605-1	Male erectile dysfunction	0.558 (0.551, 0.564)	0.567 (0.559, 0.573)	0.009 (0.004, 0.014)
phecode_557-1	Hematemesis	0.596 (0.578, 0.619)	0.605 (0.578, 0.627)	0.009 (-0.017, 0.031)
phecode_015-2	Clostridium difficile	0.646 (0.622, 0.673)	0.656 (0.629, 0.684)	0.009 (-0.017, 0.033)

5 Phenome-wide prediction of disease onset from retinal fundus photographs

Table 23 continued from previous page

Endpoint	Phecode String	Age+Sex	Age+Sex+Retina	Delta C-Index
phecode_823-2	Abnormal levels of other serum enzymes	0.56 (0.552, 0.568)	0.569 (0.56, 0.578)	0.009 (0.004, 0.013)
phecode_460-1	Acute upper respiratory infection	0.534 (0.529, 0.54)	0.543 (0.539, 0.548)	0.009 (0.005, 0.013)
phecode_809-3	Pain in limb	0.524 (0.52, 0.529)	0.533 (0.529, 0.537)	0.009 (0.005, 0.013)
phecode_468-8	Bronchopneumonia	0.686 (0.653, 0.731)	0.695 (0.663, 0.739)	0.009 (-0.012, 0.029)
phecode_433	Other cerebrovascular disease	0.716 (0.705, 0.726)	0.725 (0.717, 0.734)	0.009 (0.003, 0.016)
phecode_809	Pain	0.528 (0.525, 0.531)	0.537 (0.533, 0.541)	0.009 (0.005, 0.013)
phecode_522-9	Gastritis	0.565 (0.558, 0.571)	0.574 (0.568, 0.58)	0.009 (0.004, 0.014)
phecode_331-8	Headache NOS	0.57 (0.564, 0.574)	0.578 (0.573, 0.584)	0.009 (0.004, 0.014)
phecode_103-21	Basal cell carcinoma	0.643 (0.634, 0.652)	0.652 (0.643, 0.661)	0.009 (0.006, 0.013)
phecode_168	Coagulation defects, purpura and other hemorrhagic conditions	0.595 (0.579, 0.612)	0.604 (0.589, 0.618)	0.009 (-0.004, 0.018)
phecode_716-3	Spinal disc displacement (herniation)	0.542 (0.531, 0.556)	0.552 (0.541, 0.563)	0.009 (0.002, 0.019)
phecode_452	Hemorrhage, NOS	0.575 (0.546, 0.6)	0.586 (0.554, 0.611)	0.01 (-0.014, 0.034)
phecode_526-12	Intestinal infection due to C. difficile	0.646 (0.619, 0.681)	0.657 (0.628, 0.692)	0.01 (-0.017, 0.038)
phecode_424-1	Left heart failure	0.73 (0.716, 0.743)	0.739 (0.726, 0.752)	0.01 (0.004, 0.015)
phecode_424	Heart failure	0.737 (0.731, 0.748)	0.747 (0.739, 0.756)	0.01 (0.005, 0.016)
phecode_463	Rhinitis and nasal congestion	0.522 (0.517, 0.529)	0.532 (0.525, 0.538)	0.01 (0.004, 0.015)
phecode_366-4	Vascular abnormalities of conjunctiva	0.524 (0.495, 0.56)	0.533 (0.51, 0.567)	0.01 (-0.02, 0.042)
phecode_660	Infection of the skin	0.524 (0.52, 0.528)	0.534 (0.53, 0.539)	0.01 (0.007, 0.013)
phecode_840-1	Food allergy	0.547 (0.523, 0.573)	0.556 (0.531, 0.584)	0.01 (-0.008, 0.028)
phecode_162	Aplastic anemia	0.657 (0.631, 0.689)	0.667 (0.636, 0.699)	0.01 (-0.018, 0.033)
phecode_509-3	Hypertrophy of tongue papillae	0.533 (0.497, 0.576)	0.543 (0.507, 0.576)	0.01 (-0.024, 0.039)
phecode_585-1	Renal colic	0.608 (0.589, 0.629)	0.62 (0.598, 0.641)	0.01 (-0.002, 0.026)
phecode_673-1	Actinic keratosis	0.599 (0.595, 0.603)	0.609 (0.605, 0.614)	0.01 (0.008, 0.012)
phecode_336-5	Mononeuritis of lower limb	0.563 (0.549, 0.578)	0.574 (0.56, 0.589)	0.01 (0.001, 0.02)
phecode_234	Other nutritional deficiencies	0.58 (0.57, 0.593)	0.59 (0.58, 0.604)	0.01 (0.002, 0.02)
phecode_513-1	Esophageal ulcer	0.599 (0.577, 0.618)	0.609 (0.593, 0.632)	0.01 (-0.005, 0.026)
OMOP_4306655	All-Cause Death	0.707 (0.702, 0.714)	0.718 (0.712, 0.723)	0.01 (0.006, 0.014)
phecode_103	Malignant neoplasm of the skin	0.655 (0.648, 0.664)	0.666 (0.658, 0.674)	0.01 (0.008, 0.013)
phecode_181	Autoimmune disease	0.547 (0.536, 0.556)	0.557 (0.548, 0.568)	0.01 (0.001, 0.022)
phecode_682-11	Sebaceous cyst [Epidermal cyst]	0.486 (0.479, 0.496)	0.497 (0.489, 0.506)	0.01 (-0.003, 0.021)
phecode_707	Other arthropathies	0.641 (0.636, 0.647)	0.652 (0.646, 0.659)	0.011 (0.008, 0.014)
phecode_350-5	Repeated falls*	0.717 (0.707, 0.726)	0.727 (0.72, 0.735)	0.011 (0.004, 0.016)
phecode_681	Localized swelling, mass and lump of skin and subcutaneous tissue	0.556 (0.552, 0.561)	0.567 (0.563, 0.573)	0.011 (0.007, 0.015)
phecode_484	Pneumothorax and air leak	0.611 (0.583, 0.641)	0.624 (0.591, 0.652)	0.011 (-0.01, 0.04)
phecode_103-1	Melanomas of skin	0.603 (0.588, 0.621)	0.615 (0.598, 0.633)	0.011 (-0.002, 0.021)
phecode_554	Diseases of the pancreas	0.598 (0.578, 0.616)	0.606 (0.588, 0.626)	0.011 (-0.006, 0.024)
phecode_673	Skin changes due to chronic exposure to nonionizing radiation	0.598 (0.593, 0.602)	0.609 (0.604, 0.613)	0.011 (0.009, 0.013)
phecode_682-1	Cutaneous cyst	0.496 (0.489, 0.503)	0.506 (0.499, 0.514)	0.011 (0.002, 0.02)
phecode_095	Sequela of infection	0.546 (0.522, 0.574)	0.558 (0.534, 0.592)	0.011 (-0.011, 0.032)
phecode_812	Edema	0.61 (0.603, 0.617)	0.621 (0.614, 0.631)	0.011 (0.006, 0.017)
phecode_391-7	Perforation of tympanic membrane	0.532 (0.506, 0.555)	0.542 (0.517, 0.564)	0.011 (-0.015, 0.03)
phecode_495-1	Solitary pulmonary nodule	0.632 (0.594, 0.657)	0.642 (0.609, 0.674)	0.011 (-0.011, 0.041)
phecode_709-22	Hallux rigidus	0.573 (0.552, 0.592)	0.585 (0.561, 0.605)	0.011 (-0.003, 0.027)
phecode_718-2	Cervicalgia	0.55 (0.545, 0.556)	0.561 (0.556, 0.566)	0.011 (0.006, 0.017)
phecode_526-2	Viral enteritis	0.57 (0.537, 0.602)	0.585 (0.552, 0.616)	0.011 (-0.016, 0.04)
phecode_139-6	Hemangioma and lymphangioma	0.501 (0.485, 0.517)	0.513 (0.498, 0.534)	0.011 (-0.004, 0.034)
phecode_139	Benign sarcoma-related cancers	0.5 (0.49, 0.508)	0.511 (0.501, 0.521)	0.011 (0, 0.021)
phecode_584	Renal failure	0.683 (0.663, 0.699)	0.692 (0.676, 0.712)	0.011 (-0.001, 0.024)
phecode_256	Disorders of fluid, electrolyte and acid-base balance	0.581 (0.575, 0.588)	0.592 (0.585, 0.599)	0.011 (0.007, 0.016)
phecode_360-1	Hordeolum	0.519 (0.507, 0.533)	0.531 (0.52, 0.541)	0.011 (0.003, 0.022)
phecode_526	Intestinal infection	0.568 (0.561, 0.575)	0.579 (0.572, 0.588)	0.011 (0.005, 0.016)
phecode_524	Functional intestinal disorder	0.6 (0.589, 0.612)	0.612 (0.599, 0.623)	0.011 (0.003, 0.02)
phecode_718-5	Sciatica	0.546 (0.54, 0.553)	0.557 (0.55, 0.563)	0.011 (0.006, 0.016)

Supplementary Tables

Table 23 continued from previous page

Endpoint	Phecode String	Age+Sex	Age+Sex+Retina	Delta C-Index
phecode_004	Streptococcus	0.578 (0.557, 0.599)	0.59 (0.565, 0.609)	0.011 (-0.006, 0.03)
phecode_801	Cough	0.527 (0.524, 0.531)	0.539 (0.535, 0.543)	0.012 (0.009, 0.015)
phecode_360-11	Hordeolum externum	0.519 (0.51, 0.533)	0.531 (0.519, 0.543)	0.012 (0.002, 0.022)
phecode_353-1	Hallucinations	0.615 (0.583, 0.654)	0.626 (0.601, 0.651)	0.012 (-0.012, 0.039)
phecode_807	Malaise and fatigue	0.547 (0.543, 0.551)	0.559 (0.554, 0.565)	0.012 (0.007, 0.016)
phecode_325	Symptoms and signs related to movement disorders	0.636 (0.627, 0.643)	0.648 (0.64, 0.656)	0.012 (0.006, 0.019)
phecode_389	Other disorders of eye	0.561 (0.555, 0.569)	0.573 (0.566, 0.579)	0.012 (0.007, 0.016)
phecode_256-7	Volume depletion	0.662 (0.648, 0.675)	0.674 (0.661, 0.684)	0.012 (0.003, 0.021)
phecode_488-1	Dyspnea [Shortness of breath]	0.599 (0.594, 0.605)	0.612 (0.606, 0.616)	0.012 (0.009, 0.016)
phecode_169-11	Immune thrombocytopenic purpura [ITP]	0.599 (0.561, 0.639)	0.614 (0.567, 0.653)	0.012 (-0.013, 0.042)
phecode_089-3	Fungal infections	0.52 (0.516, 0.525)	0.532 (0.527, 0.538)	0.013 (0.009, 0.016)
phecode_558	Abnormal findings on diagnostic imaging of the digestive tract	0.588 (0.578, 0.601)	0.601 (0.586, 0.611)	0.013 (0.004, 0.02)
phecode_722-1	Plantar fascial fibromatosis [Plantar fasciitis]	0.536 (0.529, 0.544)	0.549 (0.541, 0.556)	0.013 (0.007, 0.018)
phecode_554-11	Acute pancreatitis	0.571 (0.547, 0.601)	0.585 (0.563, 0.607)	0.013 (-0.01, 0.038)
phecode_718-1	Radiculopathy	0.573 (0.561, 0.587)	0.586 (0.573, 0.6)	0.013 (0.004, 0.022)
phecode_491	Pleurisy	0.623 (0.602, 0.643)	0.637 (0.616, 0.653)	0.013 (0.003, 0.025)
phecode_324-3	Dystonia	0.548 (0.529, 0.566)	0.561 (0.55, 0.579)	0.013 (0, 0.028)
phecode_425	Cardiomegaly	0.679 (0.666, 0.694)	0.692 (0.682, 0.706)	0.013 (0.005, 0.022)
phecode_554-1	Pancreatitis	0.584 (0.558, 0.605)	0.595 (0.575, 0.617)	0.013 (-0.008, 0.03)
phecode_100	Malignant neoplasm of the head and neck	0.607 (0.578, 0.638)	0.622 (0.597, 0.651)	0.013 (-0.005, 0.037)
phecode_460	Acute respiratory infection	0.527 (0.522, 0.531)	0.54 (0.537, 0.544)	0.013 (0.01, 0.016)
phecode_841-3	Allergy to narcotic agent	0.657 (0.641, 0.68)	0.672 (0.654, 0.689)	0.013 (0.002, 0.024)
phecode_479	Pulmonary insufficiency and acute respiratory distress syndrome	0.657 (0.646, 0.672)	0.67 (0.657, 0.685)	0.013 (0.005, 0.022)
phecode_149	Benign neoplasm of the endocrine glands	0.581 (0.558, 0.611)	0.597 (0.574, 0.619)	0.013 (-0.01, 0.032)
phecode_520-13	Umbilical hernia	0.625 (0.614, 0.639)	0.637 (0.625, 0.652)	0.013 (0.003, 0.023)
phecode_084	Parasites	0.58 (0.563, 0.605)	0.593 (0.568, 0.619)	0.013 (-0.005, 0.029)
phecode_308-7	Restlessness and agitation*	0.569 (0.536, 0.6)	0.583 (0.554, 0.617)	0.013 (-0.019, 0.046)
phecode_724-5	Exostosis	0.559 (0.543, 0.577)	0.574 (0.554, 0.587)	0.013 (0.003, 0.024)
phecode_344-1	Hydrocephalus	0.627 (0.587, 0.67)	0.638 (0.604, 0.692)	0.013 (-0.036, 0.054)
phecode_015	Clostridium	0.602 (0.574, 0.632)	0.615 (0.589, 0.649)	0.013 (-0.011, 0.041)
phecode_348-2	Myelopathies	0.555 (0.537, 0.58)	0.57 (0.546, 0.594)	0.013 (-0.012, 0.038)
phecode_711	Disorder of patella	0.523 (0.51, 0.536)	0.536 (0.525, 0.549)	0.013 (0.002, 0.029)
phecode_716	Intervertebral disc disorder	0.535 (0.527, 0.544)	0.548 (0.539, 0.556)	0.014 (0.005, 0.02)
phecode_680	Epidermal thickening	0.586 (0.576, 0.595)	0.598 (0.589, 0.61)	0.014 (0.005, 0.019)
phecode_007-1	Hemophilus influenzae	0.595 (0.552, 0.627)	0.611 (0.566, 0.659)	0.014 (-0.012, 0.056)
phecode_840	Allergy	0.541 (0.534, 0.545)	0.554 (0.547, 0.56)	0.014 (0.009, 0.017)
phecode_209	Disorders of the pituitary gland and its hypothalamic control	0.609 (0.582, 0.648)	0.621 (0.584, 0.657)	0.014 (-0.022, 0.038)
phecode_537	Abnormality of the peritoneum	0.54 (0.521, 0.57)	0.557 (0.533, 0.583)	0.014 (-0.006, 0.036)
phecode_674	Disorders of pigmentation	0.545 (0.533, 0.556)	0.558 (0.548, 0.57)	0.014 (0.003, 0.022)
phecode_668-2	Seborrheic dermatitis	0.541 (0.53, 0.551)	0.555 (0.544, 0.566)	0.014 (0.006, 0.022)
phecode_559	Other disease of digestive system, NOS	0.523 (0.493, 0.548)	0.537 (0.508, 0.561)	0.014 (-0.011, 0.036)
phecode_711-1	Derangement of meniscus	0.524 (0.51, 0.535)	0.538 (0.525, 0.548)	0.014 (-0.001, 0.028)
phecode_327	Other degenerative diseases of nervous system	0.692 (0.667, 0.72)	0.707 (0.683, 0.729)	0.014 (-0.003, 0.036)
phecode_138-2	Melanocytic nevi*	0.544 (0.537, 0.553)	0.559 (0.551, 0.567)	0.014 (0.008, 0.019)
phecode_177	Abnormality of the lymph nodes	0.555 (0.547, 0.565)	0.57 (0.56, 0.58)	0.014 (0.006, 0.024)
phecode_517	Gastrointestinal angiodysplasia	0.68 (0.649, 0.712)	0.692 (0.663, 0.725)	0.014 (-0.01, 0.04)
phecode_724	Other symptoms and disorders of the soft tissue	0.55 (0.543, 0.555)	0.565 (0.56, 0.57)	0.014 (0.01, 0.019)
phecode_138-1	Nevus, non-neoplastic	0.532 (0.508, 0.552)	0.546 (0.527, 0.564)	0.015 (-0.005, 0.034)
phecode_361	Disorders of eyelid function	0.64 (0.629, 0.655)	0.655 (0.642, 0.672)	0.015 (0.004, 0.026)
phecode_556-3	Abdominal or pelvic swelling, mass, or lump	0.543 (0.53, 0.559)	0.559 (0.542, 0.572)	0.015 (-0.001, 0.029)
phecode_716-2	Degenerative disc disease	0.525 (0.508, 0.539)	0.54 (0.524, 0.551)	0.015 (0, 0.03)

5 Phenome-wide prediction of disease onset from retinal fundus photographs

Table 23 continued from previous page

Endpoint	Phecode String	Age+Sex	Age+Sex+Retina	Delta C-Index
phecode_688	Granulomatous disorder of the skin	0.507 (0.478, 0.527)	0.522 (0.501, 0.543)	0.015 (-0.005, 0.034)
phecode_355	Coma and other alteration of consciousness	0.61 (0.593, 0.631)	0.626 (0.61, 0.641)	0.015 (0.001, 0.029)
phecode_679-4	Pruritus	0.554 (0.547, 0.562)	0.57 (0.562, 0.576)	0.015 (0.01, 0.02)
phecode_468	Pneumonia	0.665 (0.657, 0.671)	0.68 (0.672, 0.687)	0.015 (0.01, 0.021)
phecode_468-9	Lobar pneumonia*	0.675 (0.667, 0.683)	0.689 (0.681, 0.699)	0.015 (0.008, 0.022)
phecode_841-4	Allergy to analgesic agent	0.624 (0.612, 0.633)	0.638 (0.626, 0.648)	0.015 (0.009, 0.021)
phecode_685	Disorders of sweat glands	0.55 (0.536, 0.561)	0.565 (0.553, 0.578)	0.015 (0.003, 0.028)
phecode_448	Peripheral vascular disease	0.613 (0.601, 0.628)	0.629 (0.616, 0.643)	0.015 (0.009, 0.024)
phecode_515	Heartburn and epigastric pain	0.553 (0.545, 0.563)	0.569 (0.562, 0.578)	0.015 (0.01, 0.023)
phecode_464	Nasopharyngitis	0.534 (0.521, 0.547)	0.549 (0.539, 0.562)	0.016 (0.003, 0.028)
phecode_488	Abnormalities of breathing	0.57 (0.565, 0.575)	0.586 (0.581, 0.591)	0.016 (0.013, 0.02)
phecode_524-1	Irritable bowel syndrome	0.606 (0.595, 0.617)	0.621 (0.611, 0.635)	0.016 (0.007, 0.026)
phecode_718	Back pain	0.518 (0.513, 0.522)	0.534 (0.529, 0.537)	0.016 (0.012, 0.02)
phecode_012	Proteus	0.675 (0.627, 0.723)	0.689 (0.64, 0.742)	0.016 (-0.034, 0.061)
phecode_325-2	Abnormality of gait and mobility	0.647 (0.639, 0.656)	0.662 (0.655, 0.671)	0.016 (0.009, 0.023)
phecode_486-5	Abnormal sputum	0.555 (0.547, 0.562)	0.571 (0.563, 0.58)	0.016 (0.009, 0.022)
phecode_391-11	Acute otitis media	0.506 (0.474, 0.536)	0.521 (0.491, 0.552)	0.016 (-0.017, 0.053)
phecode_256-5	Hypokalemia [Hypopotassemia]	0.638 (0.624, 0.655)	0.655 (0.643, 0.668)	0.016 (0.005, 0.029)
phecode_138	Benign neoplasm of the skin	0.528 (0.522, 0.535)	0.544 (0.537, 0.55)	0.016 (0.01, 0.021)
phecode_464-1	Acute nasopharyngitis	0.534 (0.523, 0.543)	0.55 (0.538, 0.563)	0.016 (0.005, 0.027)
phecode_679-22	Flushing	0.562 (0.535, 0.592)	0.578 (0.556, 0.602)	0.016 (-0.004, 0.04)
phecode_401	Hypertension	0.626 (0.623, 0.63)	0.642 (0.638, 0.646)	0.016 (0.014, 0.019)
phecode_684	Diseases of hair and hair follicles	0.695 (0.684, 0.704)	0.711 (0.7, 0.721)	0.016 (0.009, 0.025)
phecode_401-1	Essential hypertension	0.626 (0.622, 0.63)	0.642 (0.638, 0.647)	0.016 (0.014, 0.019)
phecode_386	Visual disturbances	0.588 (0.579, 0.595)	0.605 (0.596, 0.613)	0.016 (0.011, 0.024)
phecode_336	Mononeuropathies	0.544 (0.537, 0.552)	0.56 (0.554, 0.567)	0.016 (0.009, 0.022)
phecode_496	Abnormal results of pulmonary function studies	0.58 (0.57, 0.591)	0.597 (0.587, 0.608)	0.016 (0.008, 0.024)
phecode_460-2	Acute lower respiratory infection	0.559 (0.555, 0.563)	0.575 (0.572, 0.579)	0.016 (0.013, 0.02)
phecode_516	Other diseases of stomach and duodenum	0.586 (0.566, 0.601)	0.602 (0.589, 0.62)	0.017 (0.003, 0.029)
phecode_718-4	Low back pain	0.517 (0.512, 0.522)	0.534 (0.529, 0.538)	0.017 (0.011, 0.022)
phecode_148-16	Benign neoplasm of choroid	0.58 (0.557, 0.613)	0.596 (0.569, 0.626)	0.017 (-0.011, 0.044)
phecode_480	Pulmonary edema	0.709 (0.681, 0.737)	0.726 (0.696, 0.759)	0.017 (-0.002, 0.04)
phecode_375-6	Ocular hypertension	0.591 (0.577, 0.604)	0.607 (0.594, 0.624)	0.017 (0.003, 0.03)
phecode_204-1	Impaired fasting glucose	0.608 (0.591, 0.618)	0.625 (0.609, 0.641)	0.017 (0.006, 0.028)
phecode_002	Staphylococcus	0.605 (0.583, 0.623)	0.622 (0.601, 0.639)	0.017 (-0.002, 0.036)
phecode_003	Escherichia coli	0.599 (0.588, 0.611)	0.617 (0.601, 0.63)	0.017 (0.008, 0.027)
phecode_348	Other diseases of spinal cord	0.563 (0.548, 0.582)	0.58 (0.562, 0.603)	0.017 (0.002, 0.033)
phecode_179	Immunodeficiencies	0.465 (0.437, 0.497)	0.483 (0.455, 0.51)	0.017 (-0.019, 0.051)
phecode_360-12	Hordeolum internum	0.544 (0.513, 0.57)	0.56 (0.525, 0.586)	0.017 (-0.015, 0.041)
phecode_679-3	Changes in skin texture	0.581 (0.57, 0.596)	0.599 (0.585, 0.614)	0.018 (0.003, 0.03)
phecode_826	Other abnormal immunological findings in serum	0.521 (0.499, 0.55)	0.537 (0.507, 0.562)	0.018 (-0.016, 0.045)
phecode_555	Ascites	0.642 (0.622, 0.66)	0.66 (0.635, 0.679)	0.018 (0.002, 0.037)
phecode_545	Nonspecific abnormal results of function study of liver	0.544 (0.533, 0.553)	0.561 (0.553, 0.572)	0.018 (0.008, 0.029)
phecode_349-13	Abnormal findings on diagnostic imaging of skull and head	0.627 (0.607, 0.65)	0.647 (0.622, 0.672)	0.018 (0.004, 0.039)
phecode_349-1	Abnormal findings on diagnostic test of central nervous system	0.612 (0.594, 0.638)	0.633 (0.609, 0.654)	0.018 (0.001, 0.039)
phecode_719	Disorders of muscle	0.565 (0.554, 0.573)	0.583 (0.575, 0.593)	0.018 (0.011, 0.027)
phecode_402	Elevated blood pressure reading without diagnosis of hypertension	0.536 (0.529, 0.543)	0.554 (0.546, 0.561)	0.018 (0.01, 0.025)
phecode_706	Other inflammatory spondylopathies	0.585 (0.565, 0.603)	0.604 (0.586, 0.62)	0.018 (0.006, 0.032)
phecode_713-2	Effusion of joint	0.578 (0.57, 0.586)	0.596 (0.587, 0.604)	0.018 (0.011, 0.026)
phecode_417-2	Tachycardia	0.596 (0.583, 0.613)	0.615 (0.602, 0.628)	0.018 (0.004, 0.03)

Supplementary Tables

Table 23 continued from previous page

Endpoint	Phecode String	Age+Sex	Age+Sex+Retina	Delta C-Index
phecode_324-34	Torticollis	0.53 (0.511, 0.546)	0.549 (0.528, 0.567)	0.019 (-0.001, 0.038)
phecode_056	Human papillomavirus	0.492 (0.482, 0.5)	0.511 (0.501, 0.518)	0.019 (0.007, 0.029)
phecode_660-11	Candidiasis of skin and nails	0.543 (0.527, 0.559)	0.563 (0.545, 0.575)	0.019 (0.001, 0.04)
phecode_719-7	Muscle weakness (generalized)	0.55 (0.524, 0.575)	0.568 (0.545, 0.599)	0.019 (-0.01, 0.048)
phecode_329-9	Delirium	0.75 (0.732, 0.763)	0.769 (0.753, 0.783)	0.019 (0.007, 0.029)
phecode_664-1	Lichen planus, nitidus, or striatus	0.579 (0.56, 0.596)	0.597 (0.58, 0.612)	0.019 (0.002, 0.031)
phecode_468-1	Viral pneumonia	0.632 (0.596, 0.666)	0.652 (0.617, 0.685)	0.02 (-0.009, 0.055)
phecode_513-2	Gastric ulcer	0.589 (0.572, 0.603)	0.608 (0.591, 0.623)	0.02 (0.003, 0.033)
phecode_139-61	Hemangioma	0.498 (0.48, 0.516)	0.516 (0.496, 0.537)	0.02 (-0.006, 0.04)
phecode_719-1	Cramp and spasm	0.577 (0.567, 0.588)	0.598 (0.587, 0.608)	0.02 (0.012, 0.028)
phecode_462-2	Chronic sinusitis	0.54 (0.531, 0.547)	0.559 (0.551, 0.567)	0.02 (0.015, 0.025)
phecode_333-2	Insomnia	0.541 (0.531, 0.551)	0.561 (0.551, 0.569)	0.02 (0.01, 0.029)
phecode_585	Kidney stone disease	0.592 (0.58, 0.604)	0.612 (0.598, 0.624)	0.02 (0.01, 0.027)
phecode_002-1	Staphylococcus aureus	0.581 (0.561, 0.606)	0.601 (0.577, 0.621)	0.02 (0.002, 0.041)
phecode_723-1	Adhesive capsulitis of shoulder	0.536 (0.527, 0.548)	0.557 (0.546, 0.567)	0.02 (0.011, 0.033)
phecode_684-1	Alopecia	0.687 (0.677, 0.697)	0.708 (0.697, 0.717)	0.02 (0.011, 0.029)
phecode_330-1	Epilepsy	0.563 (0.543, 0.584)	0.584 (0.558, 0.606)	0.02 (-0.009, 0.042)
phecode_495	Abnormal findings on diagnostic imaging of lung	0.645 (0.634, 0.658)	0.666 (0.655, 0.678)	0.02 (0.009, 0.031)
phecode_148-1	Benign neoplasm of eye	0.572 (0.539, 0.603)	0.594 (0.565, 0.617)	0.02 (-0.006, 0.046)
phecode_582	Acute kidney failure	0.705 (0.696, 0.711)	0.725 (0.718, 0.731)	0.021 (0.014, 0.028)
phecode_530-2	Anorectal abscess	0.596 (0.566, 0.632)	0.615 (0.58, 0.652)	0.021 (-0.007, 0.046)
phecode_200-9	Abnormal thyroid function studies	0.593 (0.57, 0.616)	0.614 (0.587, 0.638)	0.021 (0.005, 0.033)
phecode_731	Symptoms involving musculoskeletal systems	0.496 (0.485, 0.507)	0.516 (0.503, 0.526)	0.021 (0.004, 0.032)
phecode_660-6	Cellulitis and abscess	0.538 (0.533, 0.543)	0.559 (0.554, 0.564)	0.021 (0.016, 0.026)
phecode_374-52	Macular cyst, hole, or pseudohole	0.663 (0.643, 0.685)	0.686 (0.663, 0.707)	0.021 (-0.002, 0.045)
phecode_283-8	Other problems related to lifestyle	0.513 (0.508, 0.518)	0.535 (0.529, 0.538)	0.022 (0.017, 0.027)
phecode_061	Influenza virus	0.56 (0.538, 0.586)	0.582 (0.559, 0.611)	0.022 (0.004, 0.041)
phecode_733-6	Temporomandibular joint disorders	0.587 (0.575, 0.599)	0.609 (0.595, 0.621)	0.022 (0.01, 0.031)
phecode_361-4	Blepharochalasis	0.623 (0.59, 0.653)	0.643 (0.603, 0.68)	0.022 (-0.01, 0.052)
phecode_149-3	Benign neoplasm of the parathyroid gland	0.64 (0.61, 0.671)	0.661 (0.633, 0.694)	0.022 (-0.009, 0.047)
phecode_733	Dentofacial anomalies, including malocclusion	0.587 (0.572, 0.605)	0.607 (0.591, 0.624)	0.022 (0.012, 0.032)
phecode_135-5	Benign neoplasm of the paranasal sinus and nasal cavity	0.555 (0.535, 0.569)	0.577 (0.559, 0.597)	0.022 (0.008, 0.041)
phecode_335	Nerve root and plexus disorders	0.543 (0.527, 0.558)	0.565 (0.553, 0.578)	0.022 (0.009, 0.035)
phecode_376-21	Crystalline deposits in vitreous body	0.592 (0.571, 0.608)	0.614 (0.601, 0.629)	0.022 (0.01, 0.035)
phecode_550-2	Cholecystitis	0.535 (0.52, 0.55)	0.559 (0.539, 0.577)	0.022 (0.006, 0.045)
phecode_376	Disorders of vitreous body	0.612 (0.604, 0.619)	0.635 (0.627, 0.642)	0.022 (0.017, 0.029)
phecode_588	Disorders and findings from impaired renal function	0.655 (0.624, 0.692)	0.678 (0.647, 0.716)	0.022 (0.002, 0.046)
phecode_391-2	Eustachian tube disorders	0.53 (0.52, 0.538)	0.553 (0.543, 0.563)	0.023 (0.016, 0.029)
phecode_366-42	Conjunctival hyperemia	0.516 (0.473, 0.552)	0.538 (0.502, 0.576)	0.023 (-0.008, 0.061)
phecode_668-5	Lichen simplex chronicus	0.566 (0.542, 0.596)	0.591 (0.558, 0.617)	0.023 (0.002, 0.049)
phecode_376-2	Vitreous opacities	0.612 (0.604, 0.618)	0.634 (0.626, 0.643)	0.023 (0.016, 0.028)
phecode_168-1	Hypo-coagulability	0.592 (0.566, 0.617)	0.614 (0.594, 0.64)	0.024 (0.008, 0.039)
phecode_710-4	Acquired deformities of the ankle and foot	0.582 (0.563, 0.597)	0.605 (0.586, 0.621)	0.024 (0.009, 0.042)
phecode_341-2	Hemiplegia and hemiparesis	0.667 (0.65, 0.685)	0.692 (0.669, 0.712)	0.024 (0.01, 0.041)
phecode_089-1	Bacterial infections	0.552 (0.546, 0.559)	0.577 (0.57, 0.582)	0.024 (0.017, 0.029)
phecode_336-1	Carpal tunnel syndrome	0.542 (0.534, 0.552)	0.566 (0.555, 0.577)	0.024 (0.013, 0.034)
phecode_351-1	Anesthesia of skin*	0.5 (0.487, 0.513)	0.524 (0.512, 0.537)	0.024 (0.008, 0.038)
phecode_355-2	Alteration of consciousness	0.621 (0.603, 0.646)	0.647 (0.625, 0.668)	0.024 (0.008, 0.039)
phecode_366	Noninflammatory disorders of conjunctiva	0.536 (0.52, 0.554)	0.559 (0.539, 0.58)	0.024 (0.003, 0.044)

5 Phenome-wide prediction of disease onset from retinal fundus photographs

Table 23 continued from previous page

Endpoint	Phecode String	Age+Sex	Age+Sex+Retina	Delta C-Index
phecode_239-12	Familial hypercholesterolemia*	0.598 (0.571, 0.624)	0.625 (0.594, 0.646)	0.024 (0.006, 0.045)
phecode_588-2	Abnormal results of function study of kidney	0.687 (0.653, 0.712)	0.71 (0.676, 0.739)	0.024 (0.002, 0.047)
phecode_361-3	Ptosis of eyelid	0.627 (0.606, 0.648)	0.652 (0.628, 0.673)	0.024 (0.009, 0.041)
phecode_810	Shock	0.722 (0.687, 0.754)	0.745 (0.71, 0.781)	0.024 (-0.004, 0.048)
phecode_804	Other symptoms and signs involving the circulatory and respiratory system	0.533 (0.527, 0.54)	0.558 (0.553, 0.564)	0.025 (0.02, 0.031)
phecode_507-11	Recurrent oral aphthae [Recurrent aphthous stomatitis]	0.559 (0.53, 0.584)	0.583 (0.55, 0.61)	0.025 (0.001, 0.052)
phecode_330-12	Partial epilepsy	0.587 (0.559, 0.62)	0.612 (0.573, 0.656)	0.025 (-0.017, 0.057)
phecode_256-6	Fluid overload	0.721 (0.697, 0.74)	0.744 (0.725, 0.766)	0.025 (0.008, 0.041)
phecode_525-1	Celiac disease	0.508 (0.482, 0.537)	0.534 (0.506, 0.563)	0.025 (-0.016, 0.055)
phecode_666-1	Allergic urticaria	0.599 (0.575, 0.627)	0.621 (0.587, 0.654)	0.025 (-0.007, 0.054)
phecode_800-2	Precordial pain	0.536 (0.521, 0.551)	0.561 (0.545, 0.576)	0.025 (0.012, 0.037)
phecode_660-2	Bacterial infection of the skin	0.503 (0.486, 0.52)	0.53 (0.506, 0.545)	0.025 (0.012, 0.04)
phecode_832-5	Acetonuria	0.547 (0.529, 0.563)	0.572 (0.556, 0.586)	0.025 (0.011, 0.038)
phecode_336-52	Meralgia paresthetica	0.472 (0.441, 0.5)	0.496 (0.471, 0.524)	0.025 (-0.01, 0.057)
phecode_469	Bronchitis	0.543 (0.534, 0.553)	0.568 (0.561, 0.578)	0.025 (0.016, 0.034)
phecode_535	Intestinal or peritoneal adhesions	0.554 (0.539, 0.569)	0.579 (0.566, 0.601)	0.026 (0.012, 0.041)
phecode_660-21	Impetigo	0.513 (0.493, 0.534)	0.537 (0.522, 0.563)	0.026 (0.01, 0.04)
phecode_800	Chest pain	0.513 (0.508, 0.519)	0.539 (0.533, 0.543)	0.026 (0.02, 0.031)
phecode_686-1	Pressure ulcer	0.721 (0.705, 0.737)	0.746 (0.733, 0.761)	0.026 (0.011, 0.038)
phecode_333-11	Obstructive sleep apnea	0.612 (0.593, 0.629)	0.638 (0.622, 0.654)	0.026 (0.015, 0.041)
phecode_284-29	Intentional self-harm*	0.566 (0.531, 0.599)	0.595 (0.563, 0.632)	0.026 (-0.003, 0.052)
phecode_683-2	Nail dystrophy*	0.593 (0.56, 0.623)	0.618 (0.585, 0.651)	0.026 (-0.002, 0.051)
phecode_840-9	Anaphylactic reaction	0.496 (0.455, 0.552)	0.524 (0.476, 0.566)	0.026 (-0.021, 0.073)
phecode_248	Disorders of plasma-protein metabolism, NEC	0.601 (0.56, 0.653)	0.629 (0.573, 0.676)	0.026 (-0.003, 0.056)
phecode_390-1	Otitis externa	0.502 (0.496, 0.508)	0.529 (0.523, 0.536)	0.026 (0.019, 0.034)
phecode_668-6	Prurigo	0.528 (0.497, 0.565)	0.556 (0.523, 0.589)	0.026 (-0.009, 0.061)
phecode_280-1	Alcohol use disorders	0.578 (0.568, 0.585)	0.604 (0.595, 0.614)	0.026 (0.018, 0.034)
phecode_486	Other respiratory disorders	0.544 (0.54, 0.55)	0.571 (0.565, 0.576)	0.026 (0.021, 0.03)
phecode_487-3	Hemoptysis	0.585 (0.566, 0.602)	0.611 (0.594, 0.627)	0.026 (0.012, 0.041)
phecode_341	Cerebral palsy and other paralytic syndromes	0.65 (0.626, 0.674)	0.675 (0.655, 0.695)	0.027 (0.011, 0.042)
phecode_377-4	Retinal hemorrhage	0.641 (0.611, 0.675)	0.668 (0.633, 0.702)	0.027 (0.01, 0.046)
phecode_479-3	Respiratory failure	0.672 (0.653, 0.695)	0.701 (0.679, 0.722)	0.027 (0.012, 0.042)
phecode_257	Polydipsia	0.512 (0.483, 0.545)	0.537 (0.509, 0.57)	0.027 (0, 0.049)
phecode_177-2	Enlargement of lymph nodes [Lymphadenopathy]	0.531 (0.52, 0.541)	0.558 (0.547, 0.569)	0.027 (0.014, 0.038)
phecode_411-2	Pericardial effusion (noninflammatory)*	0.635 (0.608, 0.663)	0.662 (0.635, 0.693)	0.027 (0.004, 0.049)
phecode_337	Polyneuropathies	0.597 (0.583, 0.609)	0.624 (0.608, 0.64)	0.027 (0.018, 0.037)
phecode_284-2	Suicide and self-inflicted harm	0.566 (0.533, 0.597)	0.594 (0.563, 0.63)	0.027 (-0.005, 0.066)
phecode_281-1	Substance abuse	0.622 (0.608, 0.633)	0.648 (0.636, 0.658)	0.027 (0.019, 0.036)
phecode_160-2	Megaloblastic anemia	0.624 (0.605, 0.652)	0.651 (0.631, 0.671)	0.027 (0.013, 0.042)
phecode_475	Asthma	0.52 (0.512, 0.53)	0.548 (0.537, 0.556)	0.027 (0.018, 0.035)
phecode_468-2	Bacterial pneumonia	0.619 (0.599, 0.644)	0.647 (0.624, 0.671)	0.027 (0.011, 0.047)
phecode_102	Malignant neoplasm of the thoracic and respiratory organs	0.677 (0.657, 0.689)	0.704 (0.689, 0.714)	0.028 (0.013, 0.04)
phecode_581-1	Pyelonephritis	0.542 (0.518, 0.562)	0.568 (0.539, 0.592)	0.028 (0.005, 0.044)
phecode_204-4	Prediabetes*	0.544 (0.509, 0.589)	0.57 (0.535, 0.611)	0.028 (-0.014, 0.063)
phecode_369-4	Corneal degenerations	0.676 (0.651, 0.708)	0.705 (0.682, 0.738)	0.028 (0.009, 0.046)
phecode_352	Disturbances of sensation of smell and taste	0.502 (0.484, 0.522)	0.531 (0.508, 0.548)	0.028 (0.004, 0.047)
phecode_344	Disorders of the circulation of the cerebrospinal fluid	0.566 (0.533, 0.602)	0.595 (0.564, 0.629)	0.028 (0, 0.07)
phecode_977-4	Long term (current) use of steroids	0.545 (0.528, 0.563)	0.575 (0.552, 0.594)	0.028 (0.011, 0.052)

Supplementary Tables

Table 23 continued from previous page

Endpoint	Phecode String	Age+Sex	Age+Sex+Retina	Delta C-Index
phecode_666	Urticaria	0.55 (0.54, 0.56)	0.578 (0.569, 0.589)	0.028 (0.019, 0.037)
phecode_727-1	Osteonecrosis	0.525 (0.476, 0.568)	0.549 (0.518, 0.595)	0.028 (-0.03, 0.077)
phecode_719-11	Spasm of muscle	0.472 (0.442, 0.512)	0.501 (0.461, 0.532)	0.028 (-0.024, 0.066)
phecode_448-9	Peripheral vascular disease NOS [includes PAD]	0.73 (0.717, 0.744)	0.759 (0.745, 0.773)	0.029 (0.019, 0.04)
phecode_723-6	Impingement syndrome of shoulder*	0.5 (0.489, 0.51)	0.528 (0.517, 0.537)	0.029 (0.011, 0.04)
phecode_256-31	Acidosis	0.647 (0.624, 0.673)	0.678 (0.658, 0.697)	0.029 (0.008, 0.05)
phecode_280	Substance related disorders	0.573 (0.563, 0.582)	0.602 (0.593, 0.611)	0.029 (0.021, 0.037)
phecode_977-41	Long term (current) use of inhaled steroids*	0.547 (0.523, 0.566)	0.576 (0.554, 0.596)	0.029 (0.01, 0.055)
phecode_232-29	Folate deficiency [Vitamin B9]	0.72 (0.69, 0.745)	0.748 (0.712, 0.773)	0.029 (0.011, 0.046)
phecode_722	Fasciopathy	0.505 (0.499, 0.512)	0.534 (0.529, 0.541)	0.029 (0.023, 0.036)
phecode_232-27	Vitamin B12 deficiency	0.576 (0.552, 0.603)	0.604 (0.58, 0.629)	0.029 (0.004, 0.053)
phecode_283	Other behavioral problems	0.506 (0.502, 0.511)	0.536 (0.532, 0.542)	0.029 (0.024, 0.035)
phecode_244	Disorders of lipoprotein metabolism and other lipidemias	0.594 (0.577, 0.617)	0.625 (0.602, 0.643)	0.03 (0.01, 0.045)
phecode_525-3	Disorders of intestinal carbohydrate absorption	0.55 (0.5, 0.589)	0.581 (0.546, 0.622)	0.03 (-0.004, 0.075)
phecode_471-5	Nasal polyps	0.546 (0.528, 0.565)	0.577 (0.558, 0.596)	0.03 (0.017, 0.049)
phecode_360-51	Eczematous dermatitis of eyelid	0.619 (0.591, 0.645)	0.647 (0.618, 0.673)	0.03 (0.007, 0.053)
phecode_239-2	Hyperglyceridemia	0.541 (0.511, 0.564)	0.567 (0.546, 0.598)	0.03 (0.003, 0.061)
phecode_367-12	Allergic [atopic] conjunctivitis	0.567 (0.545, 0.586)	0.598 (0.578, 0.617)	0.03 (0.015, 0.048)
phecode_333	Sleep disorders	0.521 (0.514, 0.528)	0.551 (0.545, 0.557)	0.03 (0.023, 0.037)
phecode_686	Chronic ulcer of skin	0.695 (0.685, 0.708)	0.726 (0.716, 0.737)	0.03 (0.024, 0.041)
phecode_336-2	Lesion of median, ulnar, radial nerve	0.52 (0.501, 0.541)	0.55 (0.528, 0.572)	0.03 (0.007, 0.047)
phecode_676	Hypertrophic conditions of skin	0.482 (0.474, 0.491)	0.514 (0.504, 0.523)	0.03 (0.021, 0.043)
phecode_347	Other disorders of the brain and CNS	0.574 (0.547, 0.593)	0.603 (0.569, 0.631)	0.031 (0.006, 0.052)
phecode_502	Other diseases of teeth and supporting structures	0.505 (0.491, 0.521)	0.535 (0.52, 0.549)	0.031 (0.012, 0.05)
phecode_702-3	Enteropathic arthropathies	0.508 (0.466, 0.556)	0.54 (0.502, 0.574)	0.031 (-0.006, 0.071)
phecode_680-1	Corns and callosities	0.557 (0.542, 0.57)	0.587 (0.575, 0.602)	0.031 (0.018, 0.044)
phecode_664	Papulosquamous disorders	0.532 (0.512, 0.55)	0.564 (0.544, 0.577)	0.031 (0.018, 0.049)
phecode_730	Other disorders and symptoms of the musculoskeletal system	0.519 (0.498, 0.538)	0.55 (0.532, 0.567)	0.031 (0.008, 0.052)
phecode_351-3	Paresthesia of skin*	0.488 (0.477, 0.497)	0.52 (0.51, 0.53)	0.031 (0.023, 0.042)
phecode_401-3	Hypertensive chronic kidney disease	0.7 (0.67, 0.737)	0.732 (0.699, 0.768)	0.032 (0.009, 0.059)
phecode_680-3	Xerosis cutis*	0.592 (0.578, 0.606)	0.622 (0.609, 0.638)	0.032 (0.019, 0.042)
phecode_336-55	Lesion of plantar nerve	0.599 (0.581, 0.611)	0.631 (0.613, 0.647)	0.032 (0.021, 0.047)
phecode_375	Abnormal intraocular pressure	0.619 (0.61, 0.629)	0.651 (0.641, 0.66)	0.032 (0.026, 0.038)
phecode_164-2	Macrocytic anemia	0.624 (0.599, 0.643)	0.656 (0.632, 0.675)	0.032 (0.017, 0.046)
phecode_386-4	Visual field defects	0.608 (0.585, 0.63)	0.639 (0.611, 0.662)	0.032 (0.005, 0.054)
phecode_660-4	Carbuncle and furuncle	0.524 (0.511, 0.541)	0.557 (0.543, 0.573)	0.032 (0.014, 0.046)
phecode_280-11	Alcohol abuse	0.616 (0.603, 0.631)	0.648 (0.636, 0.66)	0.032 (0.023, 0.042)
phecode_486-2	Other diseases of bronchus	0.506 (0.477, 0.538)	0.54 (0.499, 0.57)	0.032 (-0.009, 0.078)
phecode_168-19	Spontaneous ecchymoses	0.533 (0.507, 0.558)	0.566 (0.536, 0.591)	0.033 (0.004, 0.064)
phecode_376-1	Vitreous degeneration	0.606 (0.598, 0.615)	0.638 (0.631, 0.647)	0.033 (0.022, 0.041)
phecode_102-1	Malignant neoplasm of the of bronchus and lung	0.679 (0.665, 0.697)	0.714 (0.697, 0.726)	0.033 (0.021, 0.045)
phecode_351	Disturbances of skin sensation	0.535 (0.529, 0.54)	0.568 (0.563, 0.573)	0.034 (0.027, 0.039)
phecode_375-113	Primary open angle glaucoma	0.661 (0.636, 0.684)	0.694 (0.666, 0.714)	0.034 (0.016, 0.052)
phecode_463-21	Seasonal allergic rhinitis	0.51 (0.499, 0.52)	0.545 (0.532, 0.555)	0.034 (0.019, 0.048)
phecode_375-14	Low-tension glaucoma (Normal-tension glaucoma)	0.682 (0.657, 0.717)	0.72 (0.683, 0.749)	0.035 (0.009, 0.061)

5 Phenome-wide prediction of disease onset from retinal fundus photographs

Table 23 continued from previous page

Endpoint	Phecode String	Age+Sex	Age+Sex+Retina	Delta C-Index
phecode_356-1	Dysarthria	0.646 (0.604, 0.683)	0.684 (0.646, 0.721)	0.035 (0.007, 0.089)
phecode_685-82	Generalized hyperhidrosis	0.542 (0.527, 0.558)	0.578 (0.565, 0.591)	0.035 (0.022, 0.051)
phecode_384	Anomalies of pupillary function	0.558 (0.528, 0.598)	0.592 (0.555, 0.647)	0.035 (-0.013, 0.078)
phecode_581-11	Acute pyelonephritis	0.559 (0.533, 0.585)	0.595 (0.568, 0.616)	0.036 (0.012, 0.052)
phecode_375-1	Glaucoma	0.653 (0.643, 0.661)	0.689 (0.678, 0.697)	0.036 (0.028, 0.043)
phecode_733-62	Arthralgia of temporomandibular joint	0.594 (0.57, 0.623)	0.629 (0.604, 0.664)	0.036 (0.016, 0.059)
phecode_685-8	Hyperhidrosis	0.54 (0.528, 0.554)	0.577 (0.565, 0.59)	0.036 (0.025, 0.05)
phecode_374-38	Retinal vein occlusions	0.67 (0.649, 0.691)	0.703 (0.688, 0.724)	0.036 (0.016, 0.05)
phecode_665	Psoriasis	0.512 (0.498, 0.528)	0.549 (0.537, 0.562)	0.037 (0.022, 0.054)
phecode_503-5	Periapical abscess	0.512 (0.497, 0.526)	0.548 (0.533, 0.563)	0.037 (0.022, 0.055)
phecode_204	Elevated blood glucose level	0.581 (0.575, 0.587)	0.618 (0.612, 0.624)	0.037 (0.031, 0.044)
phecode_542-3	Hepatic failure	0.654 (0.622, 0.692)	0.694 (0.66, 0.724)	0.037 (0.008, 0.072)
phecode_686-2	Non-pressure chronic ulcer	0.683 (0.667, 0.699)	0.72 (0.707, 0.737)	0.038 (0.027, 0.052)
phecode_662	Rosacea	0.53 (0.52, 0.542)	0.567 (0.559, 0.576)	0.038 (0.029, 0.045)
phecode_474	Chronic obstructive pulmonary disease [COPD]	0.666 (0.658, 0.674)	0.704 (0.697, 0.711)	0.038 (0.033, 0.046)
phecode_388	Blindness and low vision	0.681 (0.666, 0.697)	0.72 (0.705, 0.735)	0.038 (0.027, 0.048)
phecode_469-2	Chronic bronchitis	0.651 (0.626, 0.676)	0.689 (0.668, 0.712)	0.038 (0.016, 0.064)
phecode_463-2	Allergic rhinitis	0.501 (0.493, 0.51)	0.539 (0.532, 0.548)	0.038 (0.028, 0.048)
phecode_333-1	Sleep apnea	0.598 (0.587, 0.611)	0.637 (0.627, 0.649)	0.038 (0.029, 0.05)
phecode_526-11	Intestinal e.coli	0.504 (0.466, 0.538)	0.543 (0.514, 0.58)	0.038 (0.008, 0.074)
phecode_374-3	Retinal vascular changes and occlusions	0.668 (0.656, 0.683)	0.707 (0.693, 0.72)	0.039 (0.027, 0.048)
phecode_443-1	Stricture of artery [Arterial stenosis]	0.726 (0.697, 0.76)	0.764 (0.739, 0.795)	0.039 (0.015, 0.057)
phecode_334-2	Facial nerve disorders and weakness	0.527 (0.5, 0.546)	0.565 (0.541, 0.591)	0.039 (0.014, 0.066)
phecode_299	Mental disorder, not otherwise specified	0.532 (0.501, 0.559)	0.568 (0.542, 0.596)	0.039 (0.013, 0.058)
phecode_627-3	Postmenopausal atrophic vaginitis	0.513 (0.504, 0.522)	0.552 (0.54, 0.56)	0.039 (0.027, 0.051)
phecode_840-8	Allergies related to other diseases/symptoms	0.508 (0.5, 0.517)	0.547 (0.54, 0.556)	0.039 (0.031, 0.047)
phecode_286-1	Bipolar disorder	0.512 (0.458, 0.55)	0.554 (0.505, 0.585)	0.039 (-0.001, 0.092)
phecode_721-12	Radial styloid tenosynovitis [de Quervain]	0.544 (0.524, 0.563)	0.584 (0.566, 0.602)	0.04 (0.026, 0.057)
phecode_204-2	Impaired glucose tolerance (oral)	0.581 (0.57, 0.59)	0.621 (0.613, 0.629)	0.04 (0.031, 0.048)
phecode_706-1	Sacroiliitis NEC	0.559 (0.531, 0.597)	0.598 (0.567, 0.633)	0.04 (0.014, 0.07)
phecode_256-4	Hyperkalemia [Hyperpotassemia]	0.667 (0.651, 0.687)	0.708 (0.691, 0.724)	0.04 (0.025, 0.052)
phecode_175-2	Secondary polycythemia	0.627 (0.587, 0.671)	0.668 (0.627, 0.712)	0.04 (-0.01, 0.081)
phecode_712-6	Instability of joint	0.475 (0.437, 0.499)	0.513 (0.481, 0.558)	0.04 (0.014, 0.068)
phecode_719-3	Separation of muscle (nontraumatic)	0.672 (0.643, 0.697)	0.71 (0.684, 0.735)	0.041 (0.02, 0.057)
phecode_540-3	Viral hepatitis	0.544 (0.506, 0.584)	0.587 (0.535, 0.622)	0.041 (0.011, 0.077)
phecode_387-3	Astigmatism	0.717 (0.702, 0.739)	0.759 (0.741, 0.779)	0.041 (0.028, 0.055)
phecode_054	Hepatovirus	0.555 (0.507, 0.603)	0.594 (0.559, 0.642)	0.041 (0.011, 0.07)
phecode_247-4	Disorders of magnesium metabolism	0.67 (0.649, 0.691)	0.712 (0.69, 0.735)	0.041 (0.024, 0.058)
phecode_826-3	Raised antibody titer*	0.517 (0.491, 0.555)	0.562 (0.531, 0.588)	0.042 (0.012, 0.068)
phecode_474-1	Emphysema	0.697 (0.677, 0.712)	0.738 (0.719, 0.755)	0.043 (0.026, 0.055)
phecode_334-21	Bell's palsy	0.527 (0.501, 0.557)	0.57 (0.544, 0.59)	0.043 (0.012, 0.069)
phecode_256-1	Hyperosmolality and/or hypernatremia	0.502 (0.475, 0.531)	0.547 (0.523, 0.568)	0.044 (0.021, 0.074)
phecode_396-1	Conductive hearing loss	0.537 (0.501, 0.572)	0.579 (0.548, 0.615)	0.044 (0.011, 0.076)
phecode_164-6	Anemia secondary to chronic diseases and conditions	0.648 (0.618, 0.688)	0.697 (0.662, 0.726)	0.046 (0.013, 0.08)
phecode_666-2	Idiopathic urticaria	0.557 (0.515, 0.604)	0.603 (0.558, 0.643)	0.046 (0.001, 0.088)
phecode_449	Other disorders of the circulatory system	0.633 (0.604, 0.668)	0.679 (0.643, 0.723)	0.046 (0.021, 0.075)

Supplementary Tables

Table 23 continued from previous page

Endpoint	Phecode String	Age+Sex	Age+Sex+Retina	Delta C-Index
phecode_503	Diseases of pulp and periapical tissues	0.509 (0.494, 0.525)	0.557 (0.54, 0.569)	0.046 (0.033, 0.064)
phecode_282-1	Current tobacco use and nicotine dependence	0.568 (0.559, 0.576)	0.614 (0.608, 0.622)	0.046 (0.039, 0.055)
phecode_805	Fever of unknown origin	0.527 (0.515, 0.539)	0.574 (0.563, 0.584)	0.047 (0.035, 0.059)
phecode_170-19	Neutropenia NOS	0.555 (0.524, 0.582)	0.599 (0.578, 0.631)	0.047 (0.013, 0.068)
phecode_164	Anemia	0.588 (0.58, 0.595)	0.635 (0.628, 0.643)	0.048 (0.041, 0.053)
phecode_596-5	Neuromuscular dysfunction of bladder	0.564 (0.532, 0.611)	0.618 (0.584, 0.654)	0.048 (0.012, 0.089)
phecode_374-55	Puckering of macula	0.66 (0.637, 0.673)	0.706 (0.684, 0.721)	0.048 (0.031, 0.066)
phecode_537-1	Peritonitis	0.508 (0.486, 0.54)	0.557 (0.532, 0.577)	0.048 (0.015, 0.073)
phecode_973	Adverse effect of other drug	0.645 (0.617, 0.675)	0.698 (0.663, 0.726)	0.048 (0.023, 0.08)
phecode_375-11	Open angle glaucoma	0.67 (0.655, 0.685)	0.719 (0.703, 0.734)	0.049 (0.037, 0.062)
phecode_522-12	Ulcerative colitis	0.482 (0.45, 0.512)	0.532 (0.499, 0.562)	0.049 (0.012, 0.085)
phecode_488-6	Wheezing	0.519 (0.512, 0.531)	0.569 (0.56, 0.581)	0.049 (0.039, 0.061)
phecode_504	Periodontal diseases	0.508 (0.487, 0.531)	0.559 (0.539, 0.577)	0.05 (0.023, 0.073)
phecode_809-1	Acute pain	0.494 (0.457, 0.527)	0.545 (0.496, 0.586)	0.051 (0.014, 0.092)
phecode_831	Glycosuria	0.595 (0.582, 0.605)	0.646 (0.632, 0.656)	0.051 (0.041, 0.062)
phecode_684-11	Alopecia Areata	0.624 (0.601, 0.64)	0.676 (0.653, 0.693)	0.051 (0.035, 0.069)
phecode_114	Neuroendocrine tumors	0.579 (0.551, 0.614)	0.632 (0.589, 0.674)	0.052 (0.016, 0.095)
phecode_371	Cataract	0.732 (0.728, 0.736)	0.784 (0.781, 0.789)	0.053 (0.049, 0.055)
phecode_247	Disorders of mineral metabolism and mineral deficiencies	0.563 (0.556, 0.571)	0.616 (0.608, 0.625)	0.053 (0.046, 0.06)
phecode_256-3	Mixed disorder of acid-base balance	0.653 (0.635, 0.672)	0.708 (0.691, 0.727)	0.053 (0.036, 0.068)
phecode_251	Disorders of bilirubin excretion	0.553 (0.52, 0.584)	0.606 (0.569, 0.633)	0.053 (0.033, 0.072)
phecode_824	Other abnormalities of plasma proteins*	0.513 (0.495, 0.527)	0.566 (0.546, 0.585)	0.053 (0.032, 0.077)
phecode_175	Polycythemias	0.618 (0.585, 0.655)	0.673 (0.639, 0.701)	0.053 (0.023, 0.093)
phecode_086	Pediculosis, acariasis and other infestations	0.455 (0.432, 0.488)	0.51 (0.48, 0.541)	0.054 (0.023, 0.087)
phecode_812-2	Angioneurotic edema	0.473 (0.427, 0.5)	0.525 (0.485, 0.561)	0.054 (0.021, 0.101)
phecode_380	Disorders of optic nerve and visual pathways	0.582 (0.554, 0.603)	0.637 (0.613, 0.658)	0.054 (0.032, 0.081)
phecode_281	Substance abuse, dependence, and withdrawal	0.546 (0.537, 0.551)	0.6 (0.594, 0.607)	0.054 (0.049, 0.063)
phecode_351-2	Hypoesthesia of skin*	0.701 (0.688, 0.717)	0.755 (0.741, 0.768)	0.055 (0.041, 0.067)
phecode_098	Carrier or suspected carrier of infectious diseases	0.59 (0.56, 0.622)	0.645 (0.621, 0.681)	0.055 (0.03, 0.082)
phecode_374	Disorders of the retina	0.632 (0.624, 0.637)	0.687 (0.68, 0.693)	0.055 (0.051, 0.06)
phecode_705-5	Rheumatism, unspecified	0.457 (0.422, 0.49)	0.512 (0.473, 0.552)	0.056 (0.005, 0.102)
phecode_251-1	Gilbert syndrome*	0.551 (0.523, 0.59)	0.61 (0.575, 0.639)	0.056 (0.038, 0.071)
phecode_360-5	Noninfectious dermatoses of eyelid	0.595 (0.568, 0.622)	0.652 (0.621, 0.677)	0.056 (0.034, 0.079)
phecode_237	Abnormal weight gain	0.504 (0.477, 0.523)	0.559 (0.541, 0.583)	0.056 (0.03, 0.081)
phecode_679-2	Pallor and flushing	0.516 (0.491, 0.542)	0.573 (0.551, 0.602)	0.057 (0.036, 0.084)
phecode_369-1	Corneal scars and opacities	0.631 (0.587, 0.674)	0.692 (0.655, 0.736)	0.058 (0.023, 0.107)
phecode_343	Disorders of autonomic nervous system	0.52 (0.495, 0.546)	0.579 (0.544, 0.616)	0.059 (0.029, 0.093)
phecode_374-5	Macular degeneration	0.715 (0.707, 0.725)	0.775 (0.769, 0.784)	0.059 (0.052, 0.066)
phecode_232-4	Vitamin D deficiency	0.571 (0.565, 0.578)	0.631 (0.624, 0.64)	0.06 (0.052, 0.067)
phecode_542-2	Fatty liver disease (FLD)	0.525 (0.513, 0.535)	0.585 (0.573, 0.596)	0.06 (0.043, 0.072)
phecode_380-2	Disorders of optic disc	0.551 (0.513, 0.581)	0.611 (0.574, 0.645)	0.06 (0.023, 0.091)
phecode_371-31	Age-related nuclear cataract	0.745 (0.735, 0.751)	0.806 (0.799, 0.812)	0.061 (0.055, 0.069)
phecode_371-3	Nuclear cataract	0.744 (0.736, 0.754)	0.805 (0.799, 0.814)	0.062 (0.054, 0.067)
phecode_232-2	Vitamin B group deficiency	0.568 (0.558, 0.581)	0.63 (0.618, 0.643)	0.062 (0.048, 0.072)
phecode_292	Somatoform disorders	0.487 (0.449, 0.526)	0.546 (0.494, 0.598)	0.062 (0.016, 0.096)
phecode_540-1	Chronic hepatitis	0.535 (0.498, 0.585)	0.598 (0.567, 0.634)	0.063 (0.029, 0.102)
phecode_287	Psychotic disorder	0.531 (0.497, 0.559)	0.592 (0.554, 0.631)	0.064 (0.015, 0.1)
phecode_374-511	Nonexudative (dry) age-related macular degeneration	0.759 (0.736, 0.784)	0.824 (0.806, 0.846)	0.064 (0.045, 0.086)

5 Phenome-wide prediction of disease onset from retinal fundus photographs

Table 23 continued from previous page

Endpoint	Phecode String	Age+Sex	Age+Sex+Retina	Delta C-Index
phecode_281-2	Substance dependence	0.538 (0.53, 0.545)	0.605 (0.596, 0.613)	0.066 (0.058, 0.074)
phecode_135-1	Benign neoplasm of the oral cavity	0.503 (0.466, 0.541)	0.57 (0.533, 0.606)	0.067 (0.015, 0.1)
phecode_542-4	Portal hypertension	0.646 (0.601, 0.681)	0.713 (0.672, 0.749)	0.068 (0.038, 0.098)
phecode_674-1	Hypopigmentation	0.537 (0.498, 0.571)	0.603 (0.562, 0.642)	0.068 (0.021, 0.099)
phecode_160	Nutritional anemias	0.564 (0.556, 0.575)	0.633 (0.625, 0.641)	0.069 (0.061, 0.076)
phecode_280-13	Alcoholic liver disease	0.635 (0.601, 0.66)	0.702 (0.669, 0.733)	0.069 (0.041, 0.095)
phecode_614-54	Abscess or ulceration of vulva	0.418 (0.369, 0.452)	0.484 (0.439, 0.535)	0.069 (0.018, 0.118)
phecode_232	Vitamin deficiencies	0.549 (0.543, 0.556)	0.619 (0.613, 0.624)	0.069 (0.062, 0.077)
phecode_202	Diabetes mellitus	0.593 (0.586, 0.597)	0.662 (0.657, 0.668)	0.07 (0.063, 0.076)
phecode_247-51	Hypocalcemia	0.49 (0.449, 0.531)	0.564 (0.512, 0.597)	0.07 (0.016, 0.113)
phecode_821	Abnormality of red blood cells	0.515 (0.485, 0.542)	0.587 (0.551, 0.611)	0.07 (0.044, 0.099)
phecode_202-2	Type 2 diabetes	0.591 (0.585, 0.599)	0.662 (0.657, 0.667)	0.071 (0.063, 0.077)
phecode_542-1	Fibrosis and cirrhosis of liver	0.636 (0.603, 0.66)	0.705 (0.678, 0.737)	0.071 (0.05, 0.1)
phecode_333-4	Circadian rhythm sleep disorder	0.438 (0.398, 0.481)	0.508 (0.468, 0.55)	0.071 (0.022, 0.119)
phecode_247-7	Disorders of iron metabolism	0.546 (0.538, 0.554)	0.618 (0.61, 0.625)	0.072 (0.063, 0.078)
phecode_542	Chronic liver disease and sequelae	0.517 (0.509, 0.528)	0.589 (0.579, 0.598)	0.072 (0.057, 0.084)
phecode_829-2	Abnormal level of blood mineral*	0.498 (0.481, 0.513)	0.571 (0.552, 0.591)	0.072 (0.055, 0.097)
phecode_236	Overweight and obesity	0.512 (0.506, 0.517)	0.584 (0.58, 0.59)	0.072 (0.066, 0.079)
phecode_097	Drug resistant microorganisms	0.479 (0.441, 0.514)	0.55 (0.517, 0.587)	0.073 (0.039, 0.109)
phecode_236-1	Obesity	0.511 (0.506, 0.516)	0.585 (0.58, 0.589)	0.074 (0.067, 0.08)
phecode_674-11	Vitiligo	0.516 (0.476, 0.558)	0.587 (0.539, 0.638)	0.074 (0.011, 0.13)
phecode_283-4	Patient's noncompliance with medical treatment and regimen	0.497 (0.489, 0.505)	0.572 (0.564, 0.579)	0.074 (0.062, 0.087)
phecode_308-4	Demoralization and apathy	0.548 (0.521, 0.575)	0.622 (0.594, 0.649)	0.074 (0.038, 0.107)
phecode_239-3	Mixed hyperlipidemia	0.48 (0.455, 0.51)	0.56 (0.528, 0.595)	0.075 (0.041, 0.123)
phecode_709-1	Acquired deformities of fingers	0.445 (0.424, 0.476)	0.523 (0.499, 0.549)	0.075 (0.043, 0.109)
phecode_369	Noninflammatory disorders of the cornea	0.633 (0.617, 0.648)	0.709 (0.685, 0.728)	0.076 (0.061, 0.09)
phecode_546-3	Hepatomegaly	0.538 (0.504, 0.585)	0.616 (0.574, 0.67)	0.076 (0.023, 0.134)
phecode_164-1	Microcytic anemia	0.555 (0.544, 0.563)	0.631 (0.62, 0.639)	0.076 (0.068, 0.084)
phecode_160-1	Iron deficiency anemia	0.554 (0.545, 0.565)	0.632 (0.621, 0.639)	0.077 (0.066, 0.085)
phecode_211	Disorders of adrenal glands	0.553 (0.516, 0.592)	0.634 (0.6, 0.673)	0.079 (0.042, 0.123)
phecode_501	Dental caries	0.482 (0.465, 0.498)	0.56 (0.538, 0.579)	0.079 (0.053, 0.104)
phecode_374-42	Diabetic retinopathy	0.63 (0.621, 0.637)	0.708 (0.697, 0.714)	0.079 (0.071, 0.086)
phecode_325-1	Abnormal involuntary movements	0.382 (0.339, 0.427)	0.459 (0.421, 0.503)	0.079 (0.029, 0.133)
phecode_247-72	Iron deficiency	0.545 (0.536, 0.554)	0.624 (0.617, 0.631)	0.079 (0.069, 0.086)
phecode_366-1	Pterygium of eye	0.498 (0.458, 0.536)	0.576 (0.543, 0.631)	0.081 (0.035, 0.136)
phecode_202-1	Type 1 diabetes	0.615 (0.596, 0.638)	0.698 (0.679, 0.719)	0.082 (0.06, 0.11)
phecode_977-72	Long term (current) use of oral hypoglycemic drugs	0.614 (0.606, 0.623)	0.695 (0.688, 0.704)	0.082 (0.073, 0.089)
phecode_367-52	Iridocyclitis	0.523 (0.504, 0.549)	0.608 (0.59, 0.628)	0.082 (0.058, 0.105)
phecode_374-51	Age-related macular degeneration	0.752 (0.738, 0.766)	0.834 (0.819, 0.844)	0.083 (0.071, 0.094)
phecode_498	Asphyxia and hypoxemia	0.616 (0.576, 0.656)	0.7 (0.666, 0.741)	0.083 (0.045, 0.114)
phecode_367-5	Uveitis	0.526 (0.505, 0.551)	0.61 (0.59, 0.636)	0.083 (0.059, 0.11)
phecode_097-1	Methicillin resistant Staphylococcus aureus	0.458 (0.421, 0.492)	0.541 (0.502, 0.583)	0.083 (0.033, 0.125)
phecode_724-51	Calcaneal spur	0.43 (0.389, 0.457)	0.517 (0.486, 0.551)	0.084 (0.047, 0.138)
phecode_168-2	Hyper-coagulability	0.446 (0.4, 0.487)	0.528 (0.485, 0.563)	0.084 (0.034, 0.13)
phecode_374-4	Retinal disorders in diseases classified elsewhere	0.629 (0.622, 0.64)	0.714 (0.705, 0.725)	0.085 (0.074, 0.093)
phecode_202-4	Other specified diabetes*	0.613 (0.607, 0.619)	0.698 (0.693, 0.704)	0.086 (0.08, 0.091)
phecode_977-7	Long term (current) use of insulin or oral hypoglycemic drugs	0.61 (0.601, 0.618)	0.696 (0.687, 0.703)	0.086 (0.078, 0.094)

Supplementary Tables

Table 23 continued from previous page

Endpoint	Phecode String	Age+Sex	Age+Sex+Retina	Delta C-Index
phecode_280-12	Alcohol dependence	0.535 (0.518, 0.553)	0.622 (0.605, 0.637)	0.086 (0.066, 0.105)
phecode_236-11	Morbid obesity	0.577 (0.553, 0.605)	0.666 (0.646, 0.69)	0.088 (0.068, 0.111)
phecode_352-1	Anosmia*	0.455 (0.418, 0.488)	0.544 (0.503, 0.574)	0.088 (0.039, 0.135)
phecode_059	Coronavirus	0.501 (0.465, 0.541)	0.59 (0.536, 0.641)	0.088 (0.041, 0.135)
phecode_385	Abnormal results of function studies of eye	0.587 (0.557, 0.622)	0.678 (0.646, 0.708)	0.093 (0.057, 0.115)
phecode_504-1	Gingivitis	0.422 (0.389, 0.454)	0.515 (0.478, 0.554)	0.093 (0.055, 0.141)
phecode_374-1	Retinal detachments and breaks	0.582 (0.567, 0.599)	0.676 (0.661, 0.688)	0.093 (0.078, 0.111)
phecode_205	Hypoglycemia	0.656 (0.639, 0.68)	0.75 (0.728, 0.769)	0.094 (0.07, 0.119)
phecode_367-6	Episcleritis	0.478 (0.446, 0.509)	0.572 (0.543, 0.61)	0.096 (0.066, 0.13)
phecode_679-21	Pallor	0.462 (0.419, 0.508)	0.558 (0.512, 0.606)	0.097 (0.038, 0.143)
phecode_387-1	Hypermetropia	0.707 (0.671, 0.741)	0.803 (0.777, 0.828)	0.097 (0.063, 0.125)
phecode_374-512	Exudative (wet) age-related macular degeneration	0.751 (0.722, 0.78)	0.85 (0.826, 0.88)	0.101 (0.071, 0.127)
phecode_712-1	Loose body in joint	0.444 (0.412, 0.485)	0.546 (0.511, 0.583)	0.102 (0.058, 0.138)
phecode_676-1	Hypertrophic scar [Keloid scar]	0.507 (0.478, 0.542)	0.613 (0.576, 0.647)	0.103 (0.065, 0.143)
phecode_977	Long term (current) drug therapy	0.489 (0.481, 0.494)	0.593 (0.588, 0.599)	0.105 (0.098, 0.112)
phecode_374-11	Serous retinal detachment	0.59 (0.568, 0.615)	0.696 (0.675, 0.723)	0.105 (0.076, 0.131)
phecode_387	Disorders of refraction and accommodation	0.656 (0.643, 0.666)	0.761 (0.751, 0.772)	0.106 (0.093, 0.119)
phecode_059-1	COVID-19*	0.504 (0.468, 0.542)	0.61 (0.57, 0.651)	0.106 (0.06, 0.17)
phecode_374-8	Retinal edema	0.652 (0.628, 0.682)	0.758 (0.732, 0.785)	0.107 (0.089, 0.126)
phecode_110	Malignant neoplasm of the endocrine glands	0.458 (0.417, 0.492)	0.564 (0.513, 0.611)	0.109 (0.047, 0.158)
phecode_139-52	Lipoma of intrathoracic organs	0.431 (0.383, 0.465)	0.537 (0.495, 0.569)	0.109 (0.061, 0.16)
phecode_337-8	Polyneuropathy in diseases classified elsewhere	0.645 (0.613, 0.681)	0.756 (0.726, 0.8)	0.111 (0.065, 0.155)
phecode_386-1	Amblyopia	0.599 (0.576, 0.622)	0.712 (0.688, 0.738)	0.114 (0.086, 0.136)
phecode_529-6	Halitosis*	0.406 (0.355, 0.442)	0.529 (0.493, 0.567)	0.125 (0.086, 0.168)
phecode_369-5	Hereditary corneal dystrophies	0.64 (0.611, 0.675)	0.766 (0.735, 0.804)	0.126 (0.092, 0.162)
phecode_370	Disorders of iris and ciliary body	0.622 (0.595, 0.648)	0.745 (0.725, 0.774)	0.127 (0.1, 0.152)
phecode_008	Helicobacter [H. pylori]	0.442 (0.421, 0.471)	0.574 (0.548, 0.602)	0.132 (0.1, 0.168)
phecode_096	Contact or exposure to infectious agent	0.429 (0.405, 0.451)	0.563 (0.536, 0.593)	0.135 (0.099, 0.165)
phecode_709-11	Mallet finger	0.396 (0.369, 0.43)	0.535 (0.498, 0.56)	0.138 (0.081, 0.173)
phecode_977-71	Long term (current) use of insulin	0.595 (0.576, 0.616)	0.735 (0.715, 0.755)	0.139 (0.12, 0.161)
phecode_377-5	Vitreous hemorrhage	0.546 (0.507, 0.577)	0.684 (0.645, 0.732)	0.143 (0.094, 0.184)
phecode_139-3	Benign neoplasm of other connective and soft tissue	0.4 (0.352, 0.446)	0.551 (0.502, 0.592)	0.149 (0.096, 0.188)
phecode_005	Mycobacteria	0.485 (0.447, 0.532)	0.651 (0.602, 0.695)	0.169 (0.117, 0.225)
phecode_665-2	Psoriatic arthropathy	0.37 (0.324, 0.405)	0.537 (0.5, 0.572)	0.169 (0.122, 0.214)
phecode_373	Noninflammatory disorders of choroid	0.582 (0.547, 0.613)	0.753 (0.712, 0.783)	0.173 (0.139, 0.203)
phecode_665-3	Other psoriasis	0.373 (0.328, 0.41)	0.547 (0.506, 0.579)	0.175 (0.12, 0.23)
phecode_688-1	Sarcoidosis	0.29 (0.255, 0.319)	0.5 (0.457, 0.543)	0.207 (0.161, 0.263)
phecode_387-2	Myopia	0.585 (0.565, 0.603)	0.818 (0.803, 0.833)	0.232 (0.213, 0.26)
phecode_165	Hemoglobinopathies	0.612 (0.575, 0.655)	0.872 (0.85, 0.905)	0.259 (0.214, 0.312)

5 Phenome-wide prediction of disease onset from retinal fundus photographs

Table 24: Hazard ratios for all endpoints of the retinal risk model.

Endpoint	PheCode String	unadjusted HR (IQR)	HR adjusted for Age and Sex (IQR)
phecode_701-1	Osteomyelitis	2.86 (0.01, 3.12)	0 (0, 3.03)
phecode_701	Osteomyelitis, periostitis, and other infections involving bone	2.92 (2.61, 3.15)	0.02 (0, 3.06)
phecode_979	Transplated organ	2.9 (0, 3.22)	0.07 (0, 3.66)
phecode_142	Lump or mass in breast or non-specific abnormal breast exam	3.99 (3.55, 4.45)	1.12 (1.07, 1.17)
phecode_520-11	Inguinal hernia	4.06 (3.74, 4.88)	1.13 (1.09, 1.18)
phecode_608	Other disorders of male genital organs	1.21 (1.16, 1.29)	1.14 (1.11, 1.18)
phecode_142-1	Lump or mass in breast	3.53 (3.1, 3.8)	1.17 (1.11, 1.24)
phecode_613-7	Other signs and symptoms in breast	1.2 (1.06, 1.29)	1.18 (1.06, 1.29)
phecode_613	Other nonmalignant breast conditions	1.35 (1.3, 1.5)	1.18 (1.14, 1.28)
phecode_600	Benign prostatic hyperplasia	2.04 (1.94, 2.24)	1.19 (1.15, 1.23)
phecode_709-21	Hallux valgus (Bunion)	3.88 (3.59, 4.43)	1.21 (1.12, 1.3)
phecode_331-6	Migraine	3.23 (2.93, 3.46)	1.22 (1.1, 1.35)
phecode_619	Noninflammatory female genital disorders	1.61 (1.51, 1.78)	1.22 (1.14, 1.28)
phecode_619-4	Noninflammatory disorders of vagina	1.78 (1.64, 1.92)	1.24 (1.17, 1.29)
phecode_603	Disorders and symptoms of testis	1.28 (1.22, 1.35)	1.24 (1.17, 1.32)
phecode_466-1	Acute tonsillitis and adenoiditis	2.91 (2.67, 3.21)	1.25 (1.12, 1.47)
phecode_601-1	Prostatitis	1.58 (1.45, 1.79)	1.25 (1.15, 1.44)
phecode_709-2	Acquired deformities of toe	3.43 (3.1, 3.85)	1.25 (1.16, 1.33)
phecode_601	Inflammatory diseases of prostate	1.57 (1.47, 1.78)	1.26 (1.16, 1.44)
phecode_977-51	Long term (current) use of hormonal contraceptives	4.05 (3.56, 4.54)	1.26 (1.19, 1.37)
phecode_107	Malignant neoplasm of male genitalia	2.24 (2.08, 2.54)	1.27 (1.21, 1.35)
phecode_594-41	Nocturia	3.54 (3.22, 3.87)	1.27 (1.21, 1.35)
phecode_462-1	Acute sinusitis	3.2 (2.9, 3.63)	1.29 (1.08, 1.53)
phecode_290	Reaction to severe stress, and adjustment disorders	3.49 (3.21, 3.85)	1.29 (1.13, 1.42)
phecode_105	Malignant neoplasm of the breast	4.07 (3.7, 5.12)	1.29 (1.2, 1.45)
phecode_627-2	Symptomatic menopause	3.04 (2.71, 3.32)	1.29 (1.22, 1.36)
phecode_608-1	Abnormal findings in semen	4.34 (3.83, 5.02)	1.29 (1.22, 1.37)
phecode_613-5	Mastodynia	1.55 (1.46, 1.74)	1.29 (1.23, 1.41)
phecode_604	Disorders of penis	1.34 (1.16, 1.42)	1.3 (1.13, 1.38)
phecode_107-2	Malignant neoplasm of the prostate	2.25 (2.09, 2.53)	1.3 (1.24, 1.39)
phecode_604-1	Redundant prepuce and phimosis	1.36 (1.19, 1.62)	1.31 (1.12, 1.58)
phecode_603-5	Orchitis and epididymitis	1.34 (1.18, 1.46)	1.31 (1.13, 1.42)
phecode_604-5	Balanoposthitis	1.39 (1.25, 1.55)	1.31 (1.14, 1.44)
phecode_709	Acquired deformities of fingers and toes	3.38 (3.06, 3.71)	1.31 (1.21, 1.38)
phecode_627	Menopausal and postmenopausal disorders	2.53 (2.32, 2.75)	1.31 (1.25, 1.37)
phecode_626	Disorders of menstruation and other abnormal bleeding from female genital tract	2.79 (2.65, 2.96)	1.31 (1.27, 1.38)
phecode_520-1	Hernia of the abdominal wall	3.75 (3.47, 4.21)	1.32 (1.28, 1.38)
phecode_602	Other disorders of prostate	2.03 (1.91, 2.18)	1.32 (1.28, 1.42)
phecode_594-32	Stress incontinence	4.35 (3.69, 5)	1.33 (1.24, 1.47)
phecode_619-5	Noninflammatory disorders of vulva and perineum	1.4 (1.23, 1.65)	1.34 (1.19, 1.59)
phecode_105-1	Malignant neoplasm of the breast, female	1.45 (1.34, 1.64)	1.34 (1.26, 1.51)
phecode_618	Genital prolapse	1.73 (1.51, 1.98)	1.35 (1.22, 1.5)
phecode_139-54	Testicular limpoma	1.82 (1.45, 2.18)	1.36 (0.97, 1.62)
phecode_603-1	Hydrocele	1.52 (1.37, 1.67)	1.36 (1.22, 1.45)
phecode_466	Tonsillitis and adenoiditis	2.85 (2.68, 3.19)	1.36 (1.22, 1.59)
phecode_602-4	Elevated prostate specific antigen [PSA]	2.03 (1.89, 2.17)	1.36 (1.28, 1.48)

Supplementary Tables

Table 24 continued from previous page

Endpoint	PheCode String	unadjusted HR (IQR)	HR adjusted for Age and Sex (IQR)
phecode_444-15	Scrotal varices [Varicocele]	4.42 (3.93, 5.9)	1.37 (1.19, 1.55)
phecode_294	Sexual dysfunction and disorders	4.31 (3.73, 5.02)	1.37 (1.29, 1.44)
phecode_625	Pain and other symptoms associated with female genital organs	2.04 (1.84, 2.2)	1.37 (1.29, 1.47)
phecode_618-2	Uterine/Uterovaginal prolapse	1.84 (1.62, 2)	1.38 (1.27, 1.52)
phecode_977-5	Long term (current) use of agents affecting hormones	4.09 (3.6, 4.39)	1.38 (1.29, 1.48)
phecode_618-11	Cystocele	1.9 (1.72, 2.17)	1.39 (1.28, 1.57)
phecode_626-1	Irregular menstrual cycle/bleeding	3.36 (2.99, 3.72)	1.39 (1.32, 1.45)
phecode_623	Hypertrophy of female genital organs	1.39 (1.21, 1.56)	1.4 (1.22, 1.58)
phecode_146	Benign neoplasm of the genitourinary system	4.23 (3.71, 5.03)	1.4 (1.31, 1.55)
phecode_390-4	Impacted cerumen	2.95 (2.69, 3.15)	1.4 (1.32, 1.48)
phecode_329-41	Attention and concentration deficit	3 (2.59, 3.42)	1.41 (1.2, 1.72)
phecode_618-1	Prolapse of vaginal walls	1.81 (1.56, 2.08)	1.41 (1.26, 1.56)
phecode_626-11	Absent or infrequent menstruation	3.22 (2.94, 3.56)	1.41 (1.29, 1.51)
phecode_592	Cystitis and urethritis	3.87 (3.42, 4.36)	1.41 (1.34, 1.63)
phecode_467-1	Acute laryngitis and tracheitis	3.38 (3.06, 3.79)	1.42 (1.22, 1.59)
phecode_146-2	Benign neoplasm of the prostate	2.22 (2.09, 2.48)	1.42 (1.35, 1.61)
phecode_613-1	Inflammatory disease of breast	2.06 (1.84, 2.72)	1.43 (1.16, 1.97)
phecode_596-1	Bladder neck obstruction	4.09 (3.61, 4.74)	1.43 (1.25, 1.57)
phecode_467	Laryngitis and tracheitis	3.26 (3.03, 3.71)	1.43 (1.26, 1.62)
phecode_603-6	Scrotal pain*	4.86 (4.09, 5.53)	1.43 (1.3, 1.55)
phecode_627-4	Menorrhagia/Excessive and frequent menstruation	3.29 (2.93, 3.55)	1.43 (1.33, 1.49)
phecode_605	Male sexual dysfunction	1.62 (1.54, 1.74)	1.43 (1.36, 1.51)
phecode_605-1	Male erectile dysfunction	1.63 (1.54, 1.74)	1.43 (1.36, 1.52)
phecode_592-1	Cystitis	3.92 (3.5, 4.54)	1.43 (1.37, 1.66)
phecode_597-5	Urethral caruncle	4.06 (3.5, 5.38)	1.44 (1.15, 2)
phecode_628	Ovarian cyst	2.01 (1.8, 2.25)	1.44 (1.29, 1.67)
phecode_140	Benign neoplasm of the breast	4.38 (3.6, 5.36)	1.45 (1.21, 1.86)
phecode_604-3	Peyronie's disease	1.46 (1.26, 1.57)	1.45 (1.27, 1.57)
phecode_627-1	Postmenopausal bleeding	1.65 (1.51, 1.82)	1.46 (1.33, 1.57)
phecode_601-12	Chronic prostatitis	2.16 (1.85, 2.41)	1.47 (1.25, 1.73)
phecode_610	Benign mammary dysplasias	2.21 (2.05, 2.4)	1.47 (1.35, 1.68)
phecode_396	Hearing impairment	3.21 (2.93, 3.49)	1.47 (1.36, 1.59)
phecode_622-1	Polyp of corpus uteri	1.62 (1.52, 2.05)	1.48 (1.37, 1.84)
phecode_628-2	Corpus luteum cyst or hematoma	2.02 (1.83, 2.32)	1.49 (1.36, 1.74)
phecode_593-1	Gross hematuria	3.83 (3.39, 4.29)	1.51 (1.3, 1.78)
phecode_603-2	Spermatocele	1.56 (1.45, 1.69)	1.51 (1.39, 1.63)
phecode_614	Inflammatory diseases of female pelvic organs	1.78 (1.6, 2.04)	1.53 (1.35, 1.76)
phecode_619-3	Noninflammatory disorders of cervix	2.55 (2.32, 2.73)	1.53 (1.38, 1.65)
phecode_394-21	Paroxysmal vertigo	3.33 (2.88, 3.73)	1.53 (1.4, 1.88)
phecode_395-1	Labyrinthitis	4.15 (3.77, 4.73)	1.54 (1.34, 1.8)
phecode_626-14	Irregular menstrual bleeding	3.38 (2.98, 3.67)	1.54 (1.42, 1.65)
phecode_977-52	Hormone replacement therapy (postmenopausal)	3.15 (2.91, 3.5)	1.55 (1.45, 1.71)
phecode_614-5	Inflammatory disease of cervix, vagina, and vulva	1.8 (1.6, 2.12)	1.56 (1.35, 1.85)
phecode_626-13	Irregular menstrual cycle	3.61 (3.19, 4.05)	1.56 (1.47, 1.67)
phecode_144-21	Leiomyoma of uterus	2.61 (2.43, 2.78)	1.56 (1.49, 1.69)
phecode_614-52	Vaginitis and vulvovaginitis	1.78 (1.59, 2.14)	1.57 (1.37, 1.89)
phecode_462	Sinusitis	2.93 (2.68, 3.36)	1.57 (1.43, 1.94)
phecode_144	Gynecological benign neoplasms	2.39 (2.21, 2.57)	1.58 (1.49, 1.75)
phecode_610-1	Cystic mastopathy	2.62 (2.37, 2.89)	1.59 (1.44, 1.89)
phecode_395	Other diseases of inner ear	3.94 (3.54, 4.45)	1.59 (1.45, 1.84)
phecode_682-4	Acne	2.26 (2.11, 2.45)	1.59 (1.46, 1.74)
phecode_615	Endometriosis	2.68 (2.43, 2.92)	1.6 (1.39, 1.81)

5 Phenome-wide prediction of disease onset from retinal fundus photographs

Table 24 continued from previous page

Endpoint	PheCode String	unadjusted HR (IQR)	HR adjusted for Age and Sex (IQR)
phecode_144-2	Benign neoplasms of the uterus	2.42 (2.24, 2.59)	1.6 (1.51, 1.75)
phecode_401-6	Secondary hypertension	3.13 (2.84, 3.55)	1.61 (1.43, 1.75)
phecode_835	Cytology and pathology findings	3.4 (3.19, 3.67)	1.61 (1.55, 1.67)
phecode_675-1	Circumscribed scleroderma	4.04 (3.63, 4.75)	1.62 (1.46, 1.83)
phecode_625-2	Postcoital bleeding	2.98 (2.67, 3.17)	1.62 (1.48, 1.83)
phecode_618-12	Rectocele	1.85 (1.59, 2.11)	1.63 (1.45, 1.81)
phecode_622	Polyp of female genital organs	1.96 (1.8, 2.41)	1.64 (1.49, 1.99)
phecode_619-2	Disorders of uterus, NEC	1.87 (1.73, 2.15)	1.64 (1.52, 2)
phecode_727	Other disorders of bone	3.56 (3.36, 4.05)	1.64 (1.58, 1.73)
phecode_396-21	Sensorineural hearing loss, bilateral	3.21 (2.89, 4.08)	1.65 (1.45, 2.17)
phecode_723-4	Lateral epicondylitis (Tennis elbow)	3.67 (3.26, 4.04)	1.65 (1.48, 1.85)
phecode_675	Atrophic conditions of skin	4.11 (3.63, 4.73)	1.67 (1.5, 1.85)
phecode_377-2	Conjunctival hemorrhage	2.8 (2.53, 3.42)	1.68 (1.49, 2.05)
phecode_394-2	Vertigo	3.3 (2.81, 3.83)	1.71 (1.54, 2.15)
phecode_627-3	Postmenopausal atrophic vaginitis	1.76 (1.54, 1.99)	1.73 (1.52, 1.96)
phecode_394	Disorders of vestibular function	3.23 (2.72, 3.83)	1.73 (1.53, 2.07)
phecode_052-1	Herpes simplex	3.21 (2.97, 3.63)	1.73 (1.58, 2.06)
phecode_390	Disorders of external ear	2.8 (2.66, 3.11)	1.74 (1.64, 1.87)
phecode_664-21	Pityriasis rosea	3.37 (2.79, 4.13)	1.75 (1.43, 2.55)
phecode_200-2	Goiter	3.72 (3.2, 4.6)	1.75 (1.61, 2.12)
phecode_103-3	Carcinoma in situ of skin	3.56 (3.12, 4.33)	1.77 (1.52, 2.09)
phecode_396-2	Sensorineural hearing loss	3.34 (2.99, 3.92)	1.78 (1.67, 2.02)
phecode_592-11	Acute cystitis	4.28 (3.71, 5.1)	1.79 (1.61, 2.03)
phecode_594-11	Urinary hesitancy	3.83 (3.59, 4.56)	1.79 (1.64, 2.14)
phecode_099	Lab findings related to infections	3.45 (2.91, 3.93)	1.81 (1.43, 2.38)
phecode_088	Sexually transmitted disease	3.2 (2.91, 3.76)	1.81 (1.62, 2.28)
phecode_614-53	Cyst or abscess of Bartholin's gland	2.62 (2.22, 3.28)	1.82 (1.5, 2.38)
phecode_620	Dysplasia of female genital organs	2.91 (2.55, 3.11)	1.82 (1.7, 2.05)
phecode_200	Disorders of thyroid gland	3.35 (3.04, 3.81)	1.82 (1.7, 2.12)
phecode_832	Other abnormal findings in urine	3.75 (3.43, 4.21)	1.82 (1.71, 1.98)
phecode_726-1	Osteoporosis	3.76 (3.34, 4.14)	1.82 (1.74, 2.02)
phecode_328-1	Alzheimer's disease	4.05 (3.37, 4.78)	1.83 (1.52, 1.98)
phecode_601-11	Acute prostatitis	1.97 (1.78, 2.2)	1.83 (1.6, 2.02)
phecode_328-8	Dementia in conditions classified elsewhere	3.93 (3.24, 4.66)	1.84 (1.57, 2.07)
phecode_708-11	Primary osteoarthritis of hip, pelvic region and thigh	3.07 (2.7, 3.42)	1.84 (1.63, 2.1)
phecode_625-1	Dyspareunia	2.62 (2.38, 2.98)	1.84 (1.68, 2.08)
phecode_710-32	Genu varum (acquired)	3.48 (2.95, 4.29)	1.85 (1.38, 2.24)
phecode_803	Snoring*	3.57 (3.15, 4.17)	1.85 (1.45, 2.47)
phecode_709-24	Hammer toe	3.25 (2.91, 3.62)	1.85 (1.61, 2.07)
phecode_417-1	Palpitations	3.97 (3.57, 4.36)	1.85 (1.67, 2.05)
phecode_726	Osteoporosis and low bone density	3.73 (3.33, 4.11)	1.86 (1.77, 2.05)
phecode_200-21	Diffuse goiter	3.58 (3.03, 5.24)	1.87 (1.38, 2.51)
phecode_360-4	Blepharitis	3.35 (2.88, 3.67)	1.87 (1.67, 2.1)
phecode_705-3	Polymyalgia rheumatica	3.36 (3.08, 3.7)	1.87 (1.75, 2.13)
phecode_660-13	Pityriasis versicolor	3.01 (2.62, 3.59)	1.88 (1.44, 2.36)
phecode_200-1	Hypothyroidism	3.32 (2.99, 3.78)	1.88 (1.74, 2.22)
phecode_684-11	Alopecia Areata	3.97 (3.48, 4.74)	1.88 (1.8, 2.29)
phecode_624-2	Atrophy of female genital tract	2.25 (1.74, 2.91)	1.89 (1.45, 2.49)
phecode_700-2	Sicca syndrome [Sjogren syndrome]	3.43 (3.12, 3.91)	1.89 (1.62, 2.49)
phecode_614-55	Candidiasis of vulva and vagina	4.4 (3.75, 5.33)	1.91 (1.66, 2.29)
phecode_056-1	Plantar wart	3.56 (2.63, 4.56)	1.92 (1.38, 2.81)
phecode_684-1	Alopecia	3.96 (3.53, 4.32)	1.93 (1.77, 2.12)
phecode_200-14	Hypothyroidism, not specified as secondary	3.33 (3.01, 3.77)	1.94 (1.74, 2.26)

Supplementary Tables

Table 24 continued from previous page

Endpoint	PheCode String	unadjusted HR (IQR)	HR adjusted for Age and Sex (IQR)
phecode_594-4	Frequency of urination and polyuria	3.4 (3.04, 3.67)	1.94 (1.77, 2.15)
phecode_684	Diseases of hair and hair follicles	3.97 (3.55, 4.34)	1.95 (1.81, 2.2)
phecode_520	Hernia	3.28 (2.98, 3.51)	1.95 (1.83, 2.15)
phecode_103-2	Keratinocyte carcinoma	3.29 (3.08, 3.57)	1.96 (1.82, 2.1)
phecode_594-1	Retention of urine	3.62 (3.36, 3.91)	1.97 (1.81, 2.11)
phecode_052-32	Herpes zoster	3.28 (2.88, 3.83)	1.98 (1.79, 2.41)
phecode_103-21	Basal cell carcinoma	3.2 (2.98, 3.57)	1.98 (1.83, 2.16)
phecode_728-1	Chondromalacia	3.53 (3.16, 4.28)	1.99 (1.5, 2.38)
phecode_722-4	Palmar fascial fibromatosis [Dupuytren]	3.07 (2.81, 3.37)	1.99 (1.86, 2.12)
phecode_710-31	Genu valgum (acquired)	3.06 (2.54, 3.82)	2.01 (1.51, 3.38)
phecode_624	Symptoms involving female genital tract	2.29 (1.9, 3.07)	2.01 (1.65, 2.66)
phecode_144-1	Benign neoplasms of external female genital organs and cervix	2.58 (2.22, 2.78)	2.01 (1.71, 2.31)
phecode_144-13	Benign neoplasms of the cervix	2.65 (2.3, 2.89)	2.01 (1.72, 2.36)
phecode_710-3	Acquired deformities of the knee	3.34 (2.88, 3.84)	2.01 (1.73, 2.53)
phecode_106	Gynecological malignant neoplasms	2.32 (1.85, 2.9)	2.02 (1.66, 2.59)
phecode_052-3	Varicella zoster virus	3.28 (2.88, 3.84)	2.02 (1.8, 2.57)
phecode_708	Osteoarthritis	2.79 (2.58, 3.02)	2.02 (1.94, 2.19)
phecode_674-2	Hypopigmentation	3.9 (3.28, 4.72)	2.03 (1.68, 2.42)
phecode_807-1	Chronic fatigue syndrome	3.84 (3.35, 4.41)	2.04 (1.64, 2.29)
phecode_200-22	Uninodular goiter [single thyroid nodule]	4.41 (3.73, 5.84)	2.04 (1.81, 3.02)
phecode_599	Other symptoms/disorders or the urinary system	3.33 (3.07, 3.81)	2.04 (1.83, 2.31)
phecode_721-11	Trigger finger	2.67 (2.39, 3.01)	2.04 (1.84, 2.37)
phecode_626-2	Dysmenorrhea	4.14 (3.56, 4.74)	2.05 (1.78, 2.37)
phecode_417-3	Bradycardia*	3.64 (3.32, 4.14)	2.05 (1.82, 2.34)
phecode_103	Malignant neoplasm of the skin	3.37 (3.08, 3.61)	2.05 (1.88, 2.17)
phecode_622-2	Mucous polyp of cervix	2.66 (2.32, 2.93)	2.06 (1.8, 2.47)
phecode_596-3	Diverticulum of bladder	4.14 (3.76, 5.18)	2.07 (1.79, 2.68)
phecode_112	Malignant neoplasm of other and ill-defined sites	3.52 (3.19, 3.73)	2.08 (1.96, 2.3)
phecode_200-31	Graves' disease [Toxic diffuse goiter]	3.68 (3.01, 4.83)	2.09 (1.41, 3.08)
phecode_288-3	Generalized anxiety disorder	3.75 (3.25, 4.42)	2.09 (1.64, 2.37)
phecode_329-5	Mild cognitive impairment, so stated	3.56 (3.1, 4.08)	2.09 (1.91, 2.38)
phecode_708-9	Heberden's or Bouchard's nodes*	4.07 (3.3, 5.21)	2.1 (1.78, 2.65)
phecode_614-1	Pelvic peritoneal adhesions, female (postoperative) (postinfection)	2.34 (1.95, 2.65)	2.11 (1.74, 2.53)
phecode_396-22	Presbycusis	3.49 (3.04, 4.08)	2.11 (1.82, 2.53)
phecode_444-11	Varicose veins of lower extremities	3.5 (3.15, 4.02)	2.11 (1.92, 2.44)
phecode_621	Endometrial hyperplasia	2.48 (2.15, 2.93)	2.12 (1.73, 2.54)
phecode_721-2	Ganglion cyst	3.31 (2.79, 4)	2.12 (1.76, 2.6)
phecode_394-1	Meniere disease	3 (2.57, 3.35)	2.13 (1.7, 2.43)
phecode_678	Other skin and connective tissue disorders	2.96 (2.71, 3.3)	2.13 (1.97, 2.28)
phecode_720	Spontaneous rupture of synovium and tendon	3.96 (3.19, 4.47)	2.14 (1.66, 2.52)
phecode_594-6	Urinary urgency	3.32 (2.96, 4.04)	2.14 (1.78, 2.54)
phecode_331	Headache	3.3 (2.99, 3.82)	2.14 (1.93, 2.29)
phecode_136-2	Benign neoplasm of stomach	3.24 (2.81, 3.78)	2.15 (1.86, 2.46)
phecode_052	Herpesvirus	3.41 (3.1, 3.98)	2.15 (1.9, 2.65)
phecode_251	Disorders of bilirubin excretion	3.77 (3.04, 5.05)	2.16 (1.58, 2.75)
phecode_703	Chrysal arthropathies	3.45 (3.1, 3.72)	2.16 (2, 2.34)
phecode_103-22	Squamous cell carcinoma of the skin	3.8 (3.51, 4.23)	2.17 (2.01, 2.33)
phecode_673-1	Actinic keratosis	3.03 (2.7, 3.22)	2.17 (2.02, 2.27)

5 Phenome-wide prediction of disease onset from retinal fundus photographs

Table 24 continued from previous page

Endpoint	PheCode String	unadjusted HR (IQR)	HR adjusted for Age and Sex (IQR)
phecode_488-8	Mouth breathing*	3.6 (3.2, 4.07)	2.18 (1.8, 2.51)
phecode_336-55	Lesion of plantar nerve	4.55 (3.83, 5.35)	2.18 (1.91, 2.49)
phecode_832-6	Pyuria*	4.28 (3.81, 4.76)	2.18 (1.98, 2.48)
phecode_594-2	Dysuria	3.59 (3.17, 4.1)	2.19 (1.86, 2.4)
phecode_673	Skin changes due to chronic exposure to nonionizing radiation	3.04 (2.7, 3.22)	2.19 (2.02, 2.29)
phecode_523	Diverticular disease [Bowel diverticulosis]	3.12 (2.81, 3.37)	2.2 (2.05, 2.35)
phecode_130	Cancer (solid tumor, excluding BCC)	3.4 (3.01, 3.7)	2.2 (2.05, 2.49)
phecode_708-1	Primary osteoarthritis	3.1 (2.82, 3.38)	2.21 (2.02, 2.44)
phecode_703-1	Hyperuricemia	3.49 (3.14, 3.8)	2.21 (2.04, 2.39)
phecode_703-11	Gout	3.57 (3.2, 3.9)	2.21 (2.04, 2.41)
phecode_200-23	Multinodular goiter	4.06 (3.34, 5.15)	2.22 (1.98, 2.91)
phecode_200-3	Thyrotoxicosis [hyperthyroidism]	4.24 (3.58, 5.18)	2.23 (1.87, 3.05)
phecode_462-2	Chronic sinusitis	3.48 (3.1, 4.54)	2.23 (1.9, 2.57)
phecode_360	Inflammation of eyelids	4.12 (3.38, 4.9)	2.23 (1.91, 2.51)
phecode_324-11	Parkinson's disease (Primary)	4.26 (3.5, 4.73)	2.23 (1.95, 2.61)
phecode_416-52	Ventricular premature depolarization*	3.52 (3.2, 3.99)	2.24 (1.94, 2.92)
phecode_732	Nonspecific abnormal findings on radiological and other examination of musculoskeletal system	3.23 (2.92, 3.53)	2.24 (2.01, 2.5)
phecode_331-1	Tension headache	3.29 (2.84, 3.91)	2.24 (2.06, 2.75)
phecode_530-3	Rectal prolapse	3.83 (3.52, 4.38)	2.25 (1.8, 2.46)
phecode_624-1	Dystrophy of female genital tract	2.48 (2.18, 3.32)	2.25 (1.93, 3.17)
phecode_132	Sequelae of cancer	4.2 (3.76, 4.68)	2.26 (1.94, 2.63)
phecode_523-2	Diverticula of colon	3.28 (2.99, 3.58)	2.26 (2.07, 2.43)
phecode_363-2	Dry eye syndrome [Tear film insufficiency]	3.36 (3.02, 3.71)	2.26 (2.07, 2.52)
phecode_829	Nonspecific findings on examination of blood	3.54 (3.27, 3.92)	2.26 (2.14, 2.46)
phecode_840-2	Allergy to insects	4.47 (3.4, 5.53)	2.27 (1.7, 2.8)
phecode_200-9	Abnormal thyroid function studies	3.39 (2.91, 3.73)	2.27 (2.07, 2.62)
phecode_331-61	Migraine with aura	4.46 (2.89, 5.96)	2.29 (1.69, 3.26)
phecode_101-4	Malignant neoplasm of the colon and rectum	3.53 (3.07, 3.86)	2.29 (1.98, 2.75)
phecode_336-5	Mononeuritis of lower limb	4.27 (3.58, 5.13)	2.29 (2.02, 2.79)
phecode_367-13	Blepharoconjunctivitis	3.4 (2.96, 3.91)	2.3 (1.84, 2.72)
phecode_660-21	Impetigo	2.64 (2.36, 3.23)	2.3 (1.99, 3.01)
phecode_215	Testicular dysfunction	2.28 (2.03, 2.49)	2.31 (1.91, 2.63)
phecode_515	Heartburn and epigastric pain	3.25 (2.83, 3.65)	2.31 (2.1, 2.65)
phecode_144-3	Benign neoplasms of the ovary	2.33 (1.81, 3.95)	2.32 (1.81, 4.06)
phecode_721	Synoviopathy and bursopathy	2.74 (2.47, 3.31)	2.32 (2.02, 2.62)
phecode_363	Disorders of lacrimal system	3.38 (3.03, 3.74)	2.32 (2.12, 2.59)
phecode_416	Cardiac arrhythmia and conduction disorders	3.46 (3.18, 3.77)	2.32 (2.18, 2.6)
phecode_594-31	Urge incontinence	3.79 (3.4, 4.37)	2.33 (1.96, 2.66)
phecode_324-1	Parkinsonism	4.2 (3.53, 4.65)	2.33 (2, 2.66)
phecode_251-1	Gilbert syndrome*	3.95 (3.12, 5.86)	2.34 (1.69, 3.27)
phecode_708-13	Primary osteoarthritis of the hand	3.5 (2.97, 3.73)	2.34 (2.02, 2.57)
phecode_827-1	Finding of alcohol in blood	3.91 (3.28, 4.37)	2.34 (2.04, 2.57)
phecode_416-43	Right bundle branch block	3.62 (3.33, 4.24)	2.34 (2.08, 2.64)
phecode_717-2	Sacrococcygeal disorders	3.85 (3.45, 4.86)	2.36 (1.83, 2.85)
phecode_108-5	Malignant neoplasm of the bladder	3.57 (3.26, 3.88)	2.36 (2.09, 2.72)
phecode_323	Systemic atrophies primarily affecting the central nervous system	3.79 (3.12, 4.66)	2.37 (1.78, 3.19)
phecode_374-39	Transient retinal arterial occlusion [Amaurosis fugax]	3.66 (3.36, 4.07)	2.37 (2.06, 2.84)
phecode_463-5	Postnasal drip	3.61 (2.84, 4.39)	2.38 (1.97, 2.98)

Supplementary Tables

Table 24 continued from previous page

Endpoint	PheCode String	unadjusted HR (IQR)	HR adjusted for Age and Sex (IQR)
phecode_101-41	Malignant neoplasm of the colon	3.56 (3.01, 3.76)	2.38 (1.98, 2.66)
phecode_431-2	Transient cerebral ischemic attacks	3.33 (2.95, 3.55)	2.39 (2.07, 2.77)
phecode_827	Toxicology findings	3.91 (3.35, 4.25)	2.39 (2.14, 2.62)
phecode_728	Chondropathies	3.63 (3.19, 4.02)	2.39 (2.24, 2.67)
phecode_721-6	Baker's cyst [popliteal cyst]	3.63 (3.19, 4.23)	2.4 (1.93, 2.87)
phecode_101-42	Malignant neoplasm of the rectum	3.58 (3.17, 4.39)	2.4 (2.14, 3.2)
phecode_416-41	Atrioventricular block	3.57 (3.33, 4.08)	2.4 (2.2, 2.59)
phecode_444-5	Venous insufficiency (chronic) (peripheral)	3.22 (2.88, 3.78)	2.4 (2.23, 2.81)
phecode_596-2	Overactive bladder	3.6 (3.19, 4.24)	2.41 (1.99, 2.81)
phecode_390-6	Perichondritis and chondritis of pinna	3.63 (3.22, 4.06)	2.41 (2.06, 2.8)
phecode_288	Anxiety disorders	4.29 (3.9, 5.62)	2.42 (2.07, 3.1)
phecode_597-1	Urethral stricture	3.31 (2.91, 3.47)	2.42 (2.1, 2.67)
phecode_416-21	Atrial fibrillation	3.62 (3.42, 4.12)	2.42 (2.23, 2.77)
phecode_416-71	Sick sinus syndrome*	3.79 (3.31, 4.2)	2.43 (1.99, 2.84)
phecode_416-7	Sinoatrial node dysfunction	3.81 (3.26, 4.12)	2.43 (2, 2.87)
phecode_527	Abdominal pain	3.48 (3.13, 3.99)	2.43 (2.18, 2.82)
phecode_416-5	Premature depolarization [Premature beats]	3.66 (3.17, 4.01)	2.43 (2.29, 3.13)
phecode_416-211	Paroxysmal atrial fibrillation*	3.38 (3.14, 4.03)	2.44 (2.16, 2.88)
phecode_418-1	Abnormal electrocardiogram [ECG] [EKG]	3.44 (3.16, 3.8)	2.44 (2.24, 2.81)
phecode_596	Other disorders of bladder	3.32 (3.01, 3.8)	2.45 (2.18, 2.94)
phecode_728-3	Costochondritis (Tietze's disease)	3.44 (3.03, 3.75)	2.45 (2.21, 2.64)
phecode_416-4	Heart block	3.56 (3.28, 3.97)	2.45 (2.26, 2.7)
phecode_329-6	Transient global amnesia	3.34 (2.8, 3.77)	2.46 (1.85, 2.7)
phecode_200-7	Iodine-deficiency related thyroid disorders*	4.1 (3.39, 5.63)	2.46 (2.03, 3.9)
phecode_329-1	Memory loss	3.68 (3.25, 4.01)	2.46 (2.27, 2.74)
phecode_394-4	Abnormal vestibular function study	3.43 (2.66, 4.36)	2.47 (1.48, 3.55)
phecode_705	Rheumatoid arthritis and other inflammatory polyarthropathies	3.28 (2.87, 3.56)	2.48 (2.22, 2.91)
phecode_465-2	Chronic Pharyngitis	3.17 (2.99, 3.74)	2.48 (2.27, 2.72)
phecode_438-1	Aortic aneurysm and ectasia	4.03 (3.64, 4.7)	2.48 (2.27, 2.78)
phecode_108	Malignant neoplasm of the urinary tract	3.48 (3.17, 3.79)	2.48 (2.29, 2.83)
phecode_510-8	Barrett's esophagus	3.48 (2.99, 4.07)	2.49 (2.1, 3)
phecode_593	Hematuria	3.61 (3.26, 4.03)	2.49 (2.24, 2.82)
phecode_239	Hyperlipidemia	3.14 (2.83, 3.34)	2.49 (2.31, 2.72)
phecode_334-1	Trigeminal nerve disorders [CNS]	4.34 (3.41, 5.11)	2.5 (1.85, 2.91)
phecode_328	Dementias and cerebral degeneration	3.68 (3.24, 4.15)	2.5 (2.25, 2.66)
phecode_403	Angina pectoris	2.99 (2.77, 3.25)	2.51 (2.34, 2.73)
phecode_334-11	Trigeminal neuralgia	4.23 (3.44, 5.46)	2.52 (1.73, 2.85)
phecode_367-1	Conjunctivitis	3.43 (3.12, 3.93)	2.52 (2.29, 2.96)
phecode_416-2	Atrial fibrillation and flutter	3.55 (3.24, 3.86)	2.52 (2.31, 2.82)
phecode_239-1	Hypercholesterolemia	3.21 (2.91, 3.48)	2.52 (2.33, 2.73)
phecode_391-2	Eustachian tube disorders	4.36 (3.68, 5.65)	2.53 (2.16, 3.14)
phecode_239-11	Pure hypercholesterolemia	3.31 (2.91, 3.49)	2.53 (2.32, 2.72)
phecode_106-21	Malignant neoplasm of endometrium	2.84 (2.29, 3.66)	2.55 (2.11, 3.56)
phecode_660-12	Dermatophytosis	3.55 (3.01, 3.94)	2.55 (2.19, 2.72)
phecode_444	Venous insufficiency	3.65 (3.29, 4.25)	2.56 (2.31, 2.9)
phecode_540-3	Viral hepatitis	2.95 (2.44, 3.41)	2.57 (2.07, 3.04)
phecode_054	Hepatovirus	2.94 (2.49, 3.36)	2.57 (2.09, 2.99)
phecode_106-2	Malignant neoplasm of the uterus	2.76 (2.24, 3.54)	2.57 (2.11, 3.43)
phecode_438-11	Abdominal aortic aneurysm	4.31 (3.95, 5.37)	2.57 (2.26, 2.93)

5 Phenome-wide prediction of disease onset from retinal fundus photographs

Table 24 continued from previous page

Endpoint	PheCode String	unadjusted HR (IQR)	HR adjusted for Age and Sex (IQR)
phecode_520-2	Diaphragmatic hernia [Hiatal hernia]	3.2 (2.89, 3.47)	2.57 (2.28, 2.92)
phecode_290-1	Posttraumatic stress disorder	3 (2.77, 3.26)	2.58 (1.75, 3.07)
phecode_550-4	Cholesterolosis of gallbladder	3.63 (2.95, 4.56)	2.58 (2.17, 3.33)
phecode_404-2	Coronary atherosclerosis [Atherosclerotic heart disease]	3.24 (3.01, 3.54)	2.58 (2.42, 2.77)
phecode_354	Dizziness and giddiness	3.29 (2.97, 3.61)	2.59 (2.35, 2.88)
phecode_685-1	Dyshidrosis	2.99 (2.4, 4.01)	2.6 (1.66, 3.48)
phecode_444-1	Varicose veins	3.6 (3.23, 4.26)	2.6 (2.3, 2.92)
phecode_325-23	Unsteadiness on feet*	3.23 (2.84, 3.51)	2.6 (2.34, 2.77)
phecode_715-1	Spondylosis	3.38 (3.1, 3.72)	2.6 (2.46, 2.96)
phecode_593-2	Microscopic hematuria	3.82 (3.4, 5.41)	2.61 (2.14, 3.46)
phecode_660-2	Bacterial infection of the skin	2.79 (2.53, 3.49)	2.61 (2.21, 3.31)
phecode_969	Adverse effects of agents primarily affecting gastrointestinal system	3.44 (3.17, 3.89)	2.61 (2.36, 3.06)
phecode_808	Syncope and collapse	3.48 (3.19, 3.82)	2.61 (2.38, 2.95)
phecode_418	Abnormal results of cardiovascular function studies	3.51 (3.25, 3.8)	2.61 (2.42, 2.91)
phecode_670	Seborrheic keratosis	4.23 (3.12, 4.81)	2.62 (2.06, 3.23)
phecode_733-62	Arthralgia of temporomandibular joint	3.74 (3.21, 4.21)	2.62 (2.21, 3.1)
phecode_436	Atherosclerosis [ASCVD]	3.35 (3.08, 3.68)	2.62 (2.47, 2.8)
phecode_168-4	Abnormal coagulation profile	3.52 (3.13, 4.06)	2.63 (2.22, 3.08)
phecode_416-22	Atrial flutter	3.84 (3.32, 4.46)	2.63 (2.25, 2.99)
phecode_465	Pharyngitis	3.38 (3.16, 3.88)	2.63 (2.41, 2.89)
phecode_465-1	Acute Pharyngitis	3.29 (3.15, 3.92)	2.63 (2.43, 2.87)
phecode_353	Symptoms and signs involving general sensations and perceptions	3.24 (2.91, 3.87)	2.64 (2.38, 3.13)
phecode_708-7	Generalized osteoarthritis	3.21 (2.93, 3.58)	2.64 (2.4, 2.96)
phecode_401	Hypertension	2.88 (2.61, 3.09)	2.64 (2.42, 2.85)
phecode_733	Dentofacial anomalies, including malocclusion	4.01 (3.55, 5.11)	2.65 (2.1, 3.21)
phecode_597	Other disorders of urethra and urinary tract	3.09 (2.59, 3.37)	2.65 (2.26, 3.19)
phecode_328-9	Dementia NOS	3.71 (3.27, 4.1)	2.65 (2.37, 2.83)
phecode_136	Benign neoplasm of the digestive organs	3.24 (2.97, 3.71)	2.65 (2.4, 3.08)
phecode_084	Parasites	2.96 (2.72, 3.38)	2.65 (2.52, 3.18)
phecode_664-2	Pityriasis	3.74 (2.93, 5.58)	2.67 (1.87, 4.71)
phecode_136-41	Benign neoplasm of the colon	3.33 (3.08, 3.75)	2.67 (2.46, 3.06)
phecode_532-1	Intestinal fistula	3.83 (3.24, 4.9)	2.68 (2.1, 3.46)
phecode_733-6	Temporomandibular joint disorders	3.95 (3.5, 5.1)	2.68 (2.21, 3.25)
phecode_401-1	Essential hypertension	2.93 (2.6, 3.08)	2.68 (2.43, 2.84)
phecode_713-3	Pain in joint	2.83 (2.62, 3.42)	2.68 (2.5, 3.07)
phecode_712-5	Disorder of ligament	3.84 (2.56, 5.05)	2.69 (1.75, 4.27)
phecode_361-9	Ectropion of eyelid	3.53 (3.09, 4.18)	2.69 (2.21, 3.23)
phecode_247-52	Hypercalcemia	3.21 (2.86, 3.81)	2.69 (2.52, 3.34)
phecode_463-1	Chronic rhinitis	3.38 (2.97, 4.51)	2.7 (2.32, 4.1)
phecode_708-12	Primary osteoarthritis of knee, lower leg	3.29 (2.96, 4)	2.7 (2.38, 3.22)
phecode_715	Non-inflammatory spondylopathy	3.38 (3.14, 3.78)	2.7 (2.55, 3.12)
phecode_136-4	Benign neoplasm of colon, rectum, anus and anal canal	3.35 (3.07, 3.75)	2.73 (2.49, 3.21)
phecode_404	Ischemic heart disease	3.24 (2.95, 3.48)	2.73 (2.51, 2.94)
phecode_460-1	Acute upper respiratory infection	3.41 (3.09, 3.78)	2.73 (2.55, 3)
phecode_138	Benign neoplasm of the skin	2.94 (2.49, 3.57)	2.74 (2.34, 3.45)
phecode_591	Urinary tract infection	3.71 (3.39, 4.14)	2.74 (2.47, 3.17)
phecode_714	Deforming dorsopathies	4.06 (3.48, 4.61)	2.74 (2.49, 3.24)
phecode_200-4	Thyroiditis	3.34 (2.49, 4.1)	2.75 (1.9, 3.53)
phecode_138-2	Melanocytic nevi*	3.08 (2.49, 3.55)	2.75 (2.25, 3.34)

Supplementary Tables

Table 24 continued from previous page

Endpoint	PheCode String	unadjusted HR (IQR)	HR adjusted for Age and Sex (IQR)
phecode_522-14	Microscopic colitis*	4.37 (3.77, 5.42)	2.75 (2.39, 3.44)
phecode_136-42	Benign neoplasm of rectum and anus	3.55 (3.13, 4.01)	2.76 (2.43, 3.22)
phecode_713	Symptoms related to joints	2.89 (2.56, 3.4)	2.76 (2.57, 3.17)
phecode_471-2	Deviated nasal septum	4.26 (3.38, 4.84)	2.77 (2.33, 3.42)
phecode_660-1	Fungal infection of the skin	3.46 (2.96, 4.05)	2.77 (2.4, 3.14)
phecode_583	Chronic kidney disease	3.4 (3.04, 3.67)	2.77 (2.49, 2.91)
phecode_331-8	Headache NOS	3.5 (3.26, 4.1)	2.77 (2.53, 3.07)
phecode_430-3	Nontraumatic subdural hemorrhage	3.83 (3.42, 4.34)	2.78 (2.37, 3.18)
phecode_823	Abnormal serum enzyme levels	3.36 (2.96, 3.92)	2.78 (2.4, 3.23)
phecode_389	Other disorders of eye	3.1 (2.82, 3.25)	2.78 (2.45, 2.94)
phecode_363-7	Stenosis and insufficiency of lacrimal passages	3.83 (3.02, 4.61)	2.79 (2.22, 3.43)
phecode_367-12	Allergic [atopic] conjunctivitis	3.4 (2.74, 3.95)	2.79 (2.35, 3.31)
phecode_360-51	Eczematous dermatitis of eyelid	5.09 (4.17, 6.94)	2.8 (2.27, 3.76)
phecode_506-5	Disturbances of salivary secretion	3.47 (3.13, 3.86)	2.8 (2.34, 3.13)
phecode_366	Noninflammatory disorders of conjunctiva	2.81 (2.48, 3.21)	2.8 (2.45, 3.12)
phecode_679-1	Rash and other nonspecific skin eruption	3.28 (2.86, 3.86)	2.8 (2.51, 3.1)
phecode_367	Inflammation of the eye	3.3 (3.02, 3.86)	2.8 (2.59, 3.18)
phecode_012	Proteus	2.92 (2.67, 3.38)	2.81 (0.08, 3.31)
phecode_398-1	Abnormal auditory function study	3.31 (2.79, 3.98)	2.81 (2.26, 3.41)
phecode_823-2	Abnormal levels of other serum enzymes	3.39 (2.96, 3.98)	2.81 (2.42, 3.29)
phecode_366-1	Pterygium of eye	2.72 (2.42, 3.2)	2.81 (2.56, 3.33)
phecode_438	Aneurysm or ectasia	3.7 (3.32, 4.19)	2.81 (2.63, 3.09)
phecode_361-1	Entropion and trichiasis of eyelid	3.88 (3.27, 4.37)	2.82 (2.55, 3.59)
phecode_392	Otalgia and effusion of ear	4.26 (3.53, 5.07)	2.83 (2.28, 3.38)
phecode_476	Bronchiectasis	3.68 (3.24, 4.16)	2.83 (2.54, 3.2)
phecode_391-9	Otorrhea	3.06 (2.56, 3.55)	2.84 (2.27, 3.22)
phecode_721-1	Synovitis and tenosynovitis	3 (2.67, 3.36)	2.84 (2.52, 3.19)
phecode_324-4	Tremor	3.45 (3.09, 4.1)	2.84 (2.55, 3.37)
phecode_343	Disorders of autonomic nervous system	2.84 (0, 3.13)	2.85 (0.01, 3.28)
phecode_807-11	Postviral fatigue syndrome*	3.83 (3.34, 4.54)	2.85 (2.29, 3.98)
phecode_704-5	Giant cell arteritis	4.23 (3.6, 5.14)	2.85 (2.31, 3.55)
phecode_529-2	Abdominal distension and flatulence	3.72 (3.35, 4.51)	2.85 (2.59, 3.41)
phecode_580	Glomerular diseases	3.03 (2.76, 3.35)	2.86 (0.16, 3.3)
phecode_349-2	Abnormal results of function studies of peripheral nervous system	3.06 (2.66, 4.02)	2.86 (2.24, 3.76)
phecode_585	Kidney stone disease	3.69 (3.08, 4.14)	2.86 (2.41, 3.06)
phecode_398	Other disorders of ear	3.17 (2.77, 3.51)	2.86 (2.49, 3.29)
phecode_180-3	Paraproteinemias	3.56 (3.29, 4.13)	2.86 (2.53, 3.34)
phecode_101	Malignant neoplasm of the digestive organs	3.56 (3.2, 3.83)	2.87 (2.56, 3.18)
phecode_101-2	Malignant neoplasm of stomach	3.5 (3.3, 4.15)	2.87 (2.56, 3.42)
phecode_377	Hemorrhage of the eye	3.12 (2.82, 3.44)	2.87 (2.58, 3.2)
phecode_374-8	Retinal edema	2.91 (2.71, 3.22)	2.88 (2.69, 3.16)
phecode_707	Other arthropathies	3.43 (3.16, 3.75)	2.88 (2.73, 3.25)
phecode_540	Hepatitis	3.03 (2.61, 3.47)	2.89 (2.45, 3.43)
phecode_841-12	Allergy to antibiotic agent (excluding penicillin)	3.54 (3.29, 3.89)	2.89 (2.73, 3.21)
phecode_106-3	Malignant neoplasm of the ovary	2.98 (2.64, 3.86)	2.9 (2.33, 3.58)
phecode_682	Other follicular disorders	3.41 (2.41, 4.43)	2.9 (2.34, 4.68)
phecode_101-1	Malignant neoplasm of the esophagus	3.97 (3.31, 4.44)	2.9 (2.4, 3.45)
phecode_520-13	Umbilical hernia	3.44 (3.04, 3.83)	2.9 (2.63, 3.17)
phecode_594-3	Urinary incontinence	3.76 (3.47, 4.34)	2.9 (2.69, 3.23)

5 Phenome-wide prediction of disease onset from retinal fundus photographs

Table 24 continued from previous page

Endpoint	PheCode String	unadjusted HR (IQR)	HR adjusted for Age and Sex (IQR)
phecode_419	Presence of cardiac device	3.49 (3.27, 4.08)	2.9 (2.71, 3.19)
phecode_723-51	Achilles tendinitis	3.99 (3.24, 5.37)	2.91 (1.98, 3.96)
phecode_180-31	Monoclonal gammopathy	3.68 (3.25, 4.04)	2.91 (2.56, 3.24)
phecode_446-2	Orthostatic hypotension	3.33 (2.94, 3.57)	2.91 (2.65, 3.19)
phecode_360-5	Noninfectious dermatoses of eyelid	5.24 (4.38, 6.88)	2.92 (2.5, 3.86)
phecode_588-2	Abnormal results of function study of kidney	3.08 (2.76, 3.45)	2.92 (2.68, 3.36)
phecode_666	Urticaria	4.32 (3.88, 5.29)	2.93 (2.53, 3.58)
phecode_584	Renal failure	3.06 (2.68, 3.29)	2.93 (2.54, 3.13)
phecode_086	Pediculosis, acariasis and other infestations	2.89 (2.55, 3.5)	2.93 (2.58, 3.51)
phecode_506	Diseases of salivary glands	3.49 (3.17, 3.87)	2.93 (2.62, 3.4)
phecode_463-21	Seasonal allergic rhinitis	2.94 (2.67, 3.39)	2.93 (2.64, 3.33)
phecode_473	Other diseases of upper respiratory tract	3.76 (2.94, 5.12)	2.94 (2.34, 3.72)
phecode_594	Abnormality of urination	3.66 (3.33, 4.24)	2.94 (2.69, 3.56)
phecode_705-1	Rheumatoid arthritis	3.22 (2.8, 3.71)	2.94 (2.73, 3.52)
phecode_840	Allergy	3.61 (3.19, 4.04)	2.95 (2.54, 3.38)
phecode_416-42	Left bundle branch block	3.84 (3.37, 4.24)	2.95 (2.61, 3.18)
phecode_413-2	Aortic valve disorders	3.58 (3.16, 3.9)	2.95 (2.64, 3.21)
phecode_374-3	Retinal vascular changes and occlusions	3.14 (2.99, 3.69)	2.95 (2.84, 3.36)
phecode_287	Psychotic disorder	3 (2.61, 3.38)	2.96 (2.5, 3.34)
phecode_679	Skin symptoms	3.29 (2.87, 3.76)	2.96 (2.58, 3.35)
phecode_324	Extrapyramidal and movement disorders	3.48 (3.07, 3.95)	2.96 (2.68, 3.6)
phecode_487-1	Epistaxis	3.48 (3.14, 3.77)	2.96 (2.74, 3.39)
phecode_848-2	Nonspecific abnormal findings on radiological and other examination of other intrathoracic organs (echo)	3.52 (3.3, 4.05)	2.96 (2.83, 3.51)
phecode_417	Abnormalities of heart beat	3.59 (3.33, 4.09)	2.97 (2.63, 3.25)
phecode_351-2	Hypoesthesia of skin*	3.12 (2.84, 3.45)	2.97 (2.71, 3.28)
phecode_404-11	Acute myocardial infarction	3.68 (3.19, 3.97)	2.97 (2.74, 3.32)
phecode_328-7	Vascular dementia	3.39 (3.06, 3.88)	2.97 (2.78, 3.4)
phecode_413-21	Aortic stenosis	3.66 (3.32, 3.97)	2.98 (2.68, 3.14)
phecode_588	Disorders and findings from impaired renal function	3 (2.75, 3.43)	2.98 (2.69, 3.38)
phecode_413-6	Heart valve replaced	3.53 (3.06, 4.14)	2.99 (2.68, 3.62)
phecode_424-3	Diastolic heart failure	3.19 (2.84, 4.1)	2.99 (2.73, 3.94)
phecode_487	Hemorrhage from respiratory passages	3.38 (3.04, 3.78)	2.99 (2.79, 3.5)
phecode_404-1	Myocardial infarction [Heart attack]	3.47 (3.02, 3.78)	3 (2.65, 3.2)
phecode_848	Nonspecific abnormal findings of other body structures	3.37 (3.01, 3.58)	3.01 (2.68, 3.26)
phecode_614-54	Abscess or ulceration of vulva	2.9 (1.5, 3.77)	3.02 (1.46, 3.73)
phecode_700	Diffuse diseases of connective tissue	3.48 (3.17, 3.92)	3.02 (2.64, 3.52)
phecode_715-4	Spinal stenosis	3.48 (3.15, 3.99)	3.02 (2.71, 3.43)
phecode_680	Epidermal thickening	3.35 (2.99, 3.78)	3.02 (2.76, 3.4)
phecode_375-1	Glaucoma	3.12 (2.86, 3.45)	3.02 (2.78, 3.38)
phecode_375	Abnormal intraocular pressure	3.01 (2.75, 3.24)	3.02 (2.79, 3.45)
phecode_431	Stroke and transient cerebral ischemic attacks	3.56 (3.28, 3.83)	3.02 (2.8, 3.23)
phecode_352-3	Parageusia*	2.91 (2.46, 3.96)	3.03 (2.4, 3.78)
phecode_540-1	Chronic hepatitis	3.06 (2.64, 3.4)	3.03 (2.6, 3.35)
phecode_070	Candidiasis	3.79 (3.41, 4.45)	3.03 (2.67, 3.48)
phecode_830	Proteinuria	3.36 (3.07, 3.71)	3.03 (2.76, 3.51)
phecode_350-5	Repeated falls*	3.44 (3.17, 3.72)	3.03 (2.82, 3.28)
phecode_714-3	Scoliosis	4.4 (3.73, 4.93)	3.04 (2.65, 3.44)

Supplementary Tables

Table 24 continued from previous page

Endpoint	PheCode String	unadjusted HR (IQR)	HR adjusted for Age and Sex (IQR)
phecode_820	Elevated erythrocyte sedimentation rate and abnormality of plasma viscosity	3.43 (2.93, 4.3)	3.04 (2.67, 3.77)
phecode_524-1	Irritable bowel syndrome	4.15 (3.88, 5.14)	3.04 (2.76, 3.64)
phecode_509	Diseases of tongue	4.04 (3.55, 4.99)	3.04 (2.82, 3.37)
phecode_286-21	Major depressive disorder, recurrent	5.06 (3.59, 8.13)	3.05 (1.7, 5.38)
phecode_367-21	Corneal ulcer	3.97 (2.4, 5.65)	3.05 (1.9, 4.68)
phecode_518	Appendicitis	3.7 (3.07, 5.52)	3.05 (2.53, 5.01)
phecode_149-3	Benign neoplasm of the parathyroid gland	4.82 (4.04, 6.04)	3.06 (2.8, 4.74)
phecode_165	Hemoglobinopathies	2.98 (2.79, 3.21)	3.06 (2.82, 3.31)
phecode_230-21	Abnormal weight loss	3.45 (3.08, 3.82)	3.06 (2.82, 3.45)
phecode_337-8	Polynuropathy in diseases classified elsewhere	3.02 (2.68, 3.26)	3.07 (0.06, 3.33)
phecode_391-21	Eustachian salpingitis	3.7 (2.54, 4.87)	3.07 (2.19, 4.47)
phecode_308-6	Excessive crying of child, adolescent, or adult	4.03 (3.21, 6.28)	3.07 (2.38, 4.56)
phecode_230-2	Abnormal loss of weight and underweight	3.45 (3.04, 3.8)	3.07 (2.78, 3.44)
phecode_708-14	Primary osteoarthritis of the shoulder, upper arm	3.48 (2.94, 3.91)	3.08 (2.66, 3.46)
phecode_841-1	Allergy to other anti-infective agents	3.5 (3.24, 3.99)	3.08 (2.82, 3.6)
phecode_202-1	Type 1 diabetes	3.02 (2.79, 3.32)	3.09 (0.13, 3.38)
phecode_703-2	Chondrocalcinosis	4.39 (3.48, 4.66)	3.09 (2.46, 3.64)
phecode_802	Throat pain	3.81 (3.46, 4.3)	3.09 (2.63, 3.39)
phecode_377-4	Retinal hemorrhage	3.2 (2.89, 3.43)	3.09 (2.75, 3.31)
phecode_375-6	Ocular hypertension	3.02 (2.78, 3.45)	3.09 (2.76, 3.89)
phecode_478	Aspiration pneumonia	3.31 (2.98, 3.61)	3.09 (2.88, 3.37)
phecode_413	Heart valve disorders	3.73 (3.39, 4.01)	3.09 (2.89, 3.43)
phecode_525-3	Disorders of intestinal carbohydrate absorption	3.2 (2.63, 4.07)	3.1 (2.42, 3.95)
phecode_721-5	Bursitis	3.34 (2.84, 4.22)	3.1 (2.63, 3.64)
phecode_715-3	Spondylolisthesis	3.78 (3.43, 4.24)	3.1 (2.82, 3.61)
phecode_583-1	End stage renal disease [CDK, stage 5]	3.14 (2.82, 3.46)	3.11 (0.13, 3.43)
phecode_717	Other and unspecified dorsoopathies	4.75 (4.02, 5.64)	3.11 (2.5, 3.91)
phecode_401-3	Hypertensive chronic kidney disease	3.14 (2.79, 3.76)	3.11 (2.74, 3.64)
phecode_586	Other disorders of the kidney and ureters	3.46 (3.06, 3.94)	3.11 (2.75, 3.63)
phecode_718-2	Cervicalgia	3.54 (3.18, 3.82)	3.11 (2.8, 3.29)
phecode_205	Hypoglycemia	3.11 (2.89, 3.45)	3.12 (0.42, 3.51)
phecode_284-1	Suicidal ideations	4.07 (3.45, 4.9)	3.12 (2.29, 4.17)
phecode_375-12	Angle-Closure Glaucoma	3.37 (3.07, 3.84)	3.12 (2.55, 3.58)
phecode_391	Disorders of the middle ear	3.64 (3.06, 4.69)	3.12 (2.73, 3.77)
phecode_433	Other cerebrovascular disease	3.51 (3.04, 3.84)	3.12 (2.8, 3.44)
phecode_256-4	Hyperkalemia [Hyperpotassemia]	3.18 (2.88, 3.64)	3.12 (2.8, 3.54)
phecode_202	Diabetes mellitus	2.96 (2.73, 3.22)	3.12 (2.87, 3.42)
phecode_232-4	Vitamin D deficiency	3.27 (3.13, 3.65)	3.12 (2.97, 3.42)
phecode_662	Rosacea	3.52 (3, 4.21)	3.13 (2.75, 3.73)
phecode_108-41	Malignant neoplasm of kidney, except pelvis	3.45 (3.13, 3.83)	3.13 (2.81, 3.52)
phecode_202-2	Type 2 diabetes	2.92 (2.73, 3.21)	3.13 (2.89, 3.43)
phecode_977-71	Long term (current) use of insulin	2.98 (2.86, 3.18)	3.14 (0.02, 3.32)
phecode_448-1	Raynaud's syndrome	3.82 (3.3, 4.27)	3.14 (2.69, 3.48)
phecode_116-6	Secondary malignancy of bone	3.79 (3.29, 4.5)	3.14 (2.71, 3.84)
phecode_443	Other specified disorders of arteries and arterioles	3.5 (3.03, 3.85)	3.14 (2.73, 3.31)
phecode_463	Rhinitis and nasal congestion	3.22 (2.71, 3.87)	3.14 (2.73, 3.63)
phecode_494	Voice disturbance	3.82 (3.25, 4.3)	3.14 (2.74, 3.51)

5 Phenome-wide prediction of disease onset from retinal fundus photographs

Table 24 continued from previous page

Endpoint	PheCode String	unadjusted HR (IQR)	HR adjusted for Age and Sex (IQR)
phecode_586-2	Cyst of kidney	3.57 (3.18, 3.9)	3.14 (2.78, 3.55)
phecode_585-1	Renal colic	4.17 (3.89, 4.71)	3.14 (2.82, 3.59)
phecode_530-1	Anal fissure	3.99 (3.24, 5.95)	3.15 (2.21, 4.56)
phecode_308-3	Emotional lability	3.55 (3.31, 4.13)	3.15 (2.37, 3.92)
phecode_719-3	Separation of muscle (nontraumatic)	4.6 (3.76, 5.48)	3.15 (2.46, 3.78)
phecode_840-1	Food allergy	3.33 (3.08, 3.89)	3.15 (2.8, 3.65)
phecode_468-8	Bronchopneumonia	3.36 (3.06, 4.08)	3.15 (2.87, 3.8)
phecode_708-16	Primary osteoarthritis ankle and foot	3.4 (3.07, 3.91)	3.16 (2.63, 3.48)
phecode_230	Malnutrition and underweight	3.45 (3.12, 3.75)	3.16 (2.9, 3.5)
phecode_529	Symptoms involving digestive system	3.68 (3.25, 4.06)	3.16 (2.91, 3.59)
phecode_337	Polyneuropathies	3.17 (2.75, 3.51)	3.17 (2.76, 3.55)
phecode_686-2	Non-pressure chronic ulcer	3.23 (2.96, 3.68)	3.17 (2.84, 3.56)
phecode_433-2	Occlusion and stenosis of pre-cerebral arteries	3.61 (3.18, 4.1)	3.17 (2.89, 3.65)
phecode_377-5	Vitreous hemorrhage	3.08 (2.87, 3.32)	3.18 (0.02, 3.46)
phecode_683	Nail disorders	3.3 (2.59, 3.93)	3.18 (2.44, 3.87)
phecode_363-5	Epiphora	3.72 (3.1, 4.18)	3.18 (2.45, 3.66)
phecode_709-22	Hallux rigidus	3.89 (2.86, 5.02)	3.18 (2.58, 3.94)
phecode_724-52	Osteophyte*	3.44 (2.81, 4.37)	3.18 (2.66, 4.03)
phecode_366-42	Conjunctival hyperemia	3.2 (2.77, 3.72)	3.18 (2.79, 3.69)
phecode_413-11	Mitral valve insufficiency	3.83 (3.49, 4.08)	3.18 (2.85, 3.46)
phecode_388	Blindness and low vision	3.38 (3.14, 3.67)	3.18 (2.96, 3.42)
phecode_674	Disorders of pigmentation	3.91 (3.41, 4.9)	3.19 (2.71, 3.72)
phecode_164-6	Anemia secondary to chronic diseases and conditions	3.21 (2.83, 3.51)	3.19 (2.82, 3.64)
phecode_369-4	Corneal degenerations	3.59 (3.16, 3.91)	3.19 (2.83, 3.65)
phecode_009	Pseudomonas	3.25 (2.9, 3.78)	3.19 (2.87, 3.77)
phecode_433-21	Carotid artery stenosis	3.63 (3.17, 4.01)	3.19 (2.89, 3.59)
phecode_108-4	Malignant neoplasm of the kidney	3.52 (3.25, 3.93)	3.2 (2.87, 3.58)
phecode_416-12	Ventricular tachycardia	3.69 (3.37, 4.12)	3.2 (2.9, 3.52)
phecode_329	Symptoms and signs involving cognitive functions and awareness	3.68 (3.42, 4.04)	3.2 (2.93, 3.55)
phecode_386-9	Visual distortions and subjective visual disturbances	4.29 (3.58, 6.25)	3.21 (2.78, 4.51)
phecode_510-5	Dyskinesia of esophagus	3.54 (3, 4.04)	3.21 (2.81, 3.84)
phecode_410-2	Endocarditis	3.69 (3.18, 4.22)	3.21 (2.82, 3.73)
phecode_841-13	Allergy to sulfonamides	3.75 (3.07, 4.4)	3.21 (2.82, 4.26)
phecode_448-9	Peripheral vascular disease NOS [includes PAD]	3.43 (3.05, 3.71)	3.21 (2.83, 3.4)
phecode_204-1	Impaired fasting glucose	3.44 (3.23, 4.19)	3.21 (2.98, 3.89)
phecode_707-8	Polyarthritits	4.22 (3.63, 4.72)	3.22 (2.73, 3.76)
phecode_437	Vascular insufficiency of intestine	3.33 (2.93, 3.87)	3.22 (2.84, 3.79)
phecode_374-42	Diabetic retinopathy	3.16 (2.81, 3.42)	3.22 (2.87, 3.47)
phecode_204	Elevated blood glucose level	3.42 (3.08, 3.83)	3.22 (2.89, 3.61)
phecode_248	Disorders of plasma-protein metabolism, NEC	3.23 (2.95, 3.49)	3.22 (2.94, 3.55)
phecode_281-1	Substance abuse	4.01 (3.57, 4.35)	3.22 (2.99, 3.55)
phecode_440-2	Arterial embolism and thrombosis	3.36 (2.89, 3.72)	3.23 (2.72, 3.54)
phecode_723-2	Rotator cuff tear or rupture	2.94 (2.73, 3.74)	3.24 (2.81, 3.98)
phecode_524	Functional intestinal disorder	4.13 (3.66, 4.92)	3.24 (2.92, 3.84)
phecode_681	Localized swelling, mass and lump of skin and subcutaneous tissue	3.45 (3.02, 3.97)	3.24 (2.95, 3.63)
phecode_400-2	Chronic rheumatic heart diseases	3.65 (3.28, 4.03)	3.24 (2.96, 3.54)
phecode_446	Hypotension	3.45 (3.16, 3.79)	3.24 (2.97, 3.6)
phecode_523-4	Diverticulitits	3.54 (3.17, 4.45)	3.24 (3.06, 3.9)
phecode_557-2	Blood in stool	3.42 (2.91, 3.78)	3.25 (2.8, 3.58)

Supplementary Tables

Table 24 continued from previous page

Endpoint	PheCode String	unadjusted HR (IQR)	HR adjusted for Age and Sex (IQR)
phecode_440-11	Deep vein thrombosis [DVT]	3.69 (3.08, 4.25)	3.25 (2.8, 3.83)
phecode_230-22	Underweight	3.58 (3.13, 5.31)	3.25 (2.87, 4.15)
phecode_679-4	Pruritus	3.52 (3.22, 4.07)	3.25 (2.89, 3.62)
phecode_977-7	Long term (current) use of insulin or oral hypoglycemic drugs	3.08 (2.8, 3.27)	3.25 (2.91, 3.41)
phecode_325-2	Abnormality of gait and mobility	3.43 (3.13, 3.72)	3.25 (2.97, 3.62)
OMOP_4306655	All-Cause Death	3.68 (3.34, 3.97)	3.25 (3, 3.53)
phecode_840-9	Anaphylactic reaction	3.31 (2.79, 7.54)	3.26 (2.86, 6.81)
phecode_361	Disorders of eyelid function	3.64 (3.23, 4)	3.26 (2.91, 3.66)
phecode_447	Nonspecific low blood-pressure reading	3.31 (3.07, 3.74)	3.26 (2.94, 3.62)
phecode_008	Helicobacter [H. pylori]	3.25 (3.05, 3.54)	3.26 (3.04, 3.58)
phecode_095	Sequela of infection	3.27 (2.63, 4.33)	3.27 (2.07, 6.02)
phecode_424-1	Left heart failure	3.61 (3.13, 4.04)	3.27 (2.92, 3.72)
phecode_400	Rheumatic fever and chronic rheumatic heart diseases	3.63 (3.28, 4.08)	3.27 (2.96, 3.58)
phecode_256-6	Fluid overload	3.36 (3.07, 3.83)	3.27 (2.99, 3.75)
phecode_529-5	Constipation	3.68 (3.37, 4.01)	3.27 (3.02, 3.61)
phecode_723-5	Tendinitis	3.78 (3.19, 6.85)	3.28 (2.77, 6.25)
phecode_381	Strabismus	3.43 (3, 3.73)	3.28 (2.82, 3.48)
phecode_200-13	Postprocedural hypothyroidism	4.59 (3.88, 5.28)	3.29 (2.26, 3.88)
phecode_704	Systemic vasculitis	4.14 (3.48, 4.63)	3.29 (2.85, 3.77)
phecode_479-3	Respiratory failure	3.33 (2.88, 3.72)	3.29 (2.89, 3.75)
phecode_391-7	Perforation of tympanic membrane	2.95 (2.58, 3.41)	3.29 (2.89, 3.94)
phecode_424	Heart failure	3.67 (3.26, 4.06)	3.29 (2.91, 3.67)
phecode_416-11	Supraventricular tachycardia	3.89 (3.37, 4.88)	3.29 (2.98, 4.15)
phecode_280-11	Alcohol abuse	4.05 (3.61, 4.4)	3.29 (3.06, 3.62)
phecode_256-31	Acidosis	3.27 (2.79, 3.57)	3.3 (0.22, 3.63)
phecode_471-3	Hypertrophy of nasal turbinates	3.22 (2.88, 3.91)	3.3 (2.51, 4.15)
phecode_841	Drug and medical agent allergy	3.72 (3.42, 4.17)	3.3 (2.99, 3.89)
phecode_463-2	Allergic rhinitis	3.34 (3.01, 3.8)	3.3 (3, 3.77)
phecode_325	Symptoms and signs related to movement disorders	3.51 (3.17, 3.84)	3.3 (3.01, 3.74)
phecode_686	Chronic ulcer of skin	3.43 (3.17, 3.85)	3.3 (3.03, 3.74)
phecode_374-38	Retinal vein occlusions	3.44 (2.94, 3.76)	3.31 (2.81, 3.49)
phecode_507-11	Recurrent oral aphthae [Recurrent aphthous stomatitis]	3.78 (3.31, 4.71)	3.31 (2.84, 4.21)
phecode_510	Diseases of esophagus	3.42 (3.03, 3.67)	3.31 (2.92, 3.62)
phecode_349-13	Abnormal findings on diagnostic imaging of skull and head	3.37 (3.02, 4.06)	3.31 (2.96, 3.93)
phecode_002-1	Staphylococcus aureus	3.18 (2.9, 3.57)	3.32 (0.05, 3.75)
phecode_672	Other acute skin changes due to ultraviolet radiation	4.24 (3.07, 5.21)	3.32 (2.32, 4.71)
phecode_477	Inhalation lung injury	4.01 (3.71, 5.1)	3.32 (2.94, 4.36)
phecode_514-1	Esophageal obstruction (Stricture and stenosis of esophagus)	3.9 (3.32, 4.41)	3.32 (2.98, 3.81)
phecode_840-8	Allergies related to other diseases/symptoms	3.43 (3.15, 3.85)	3.32 (3, 3.69)
phecode_232	Vitamin deficiencies	3.39 (3.17, 3.74)	3.32 (3.11, 3.68)
phecode_526-2	Viral enteritis	3.32 (2.85, 3.6)	3.33 (0.01, 3.72)
phecode_517	Gastrointestinal angiodysplasia	3.29 (2.91, 3.83)	3.33 (2.75, 3.95)
phecode_513-3	Duodenal ulcer	3.5 (3.21, 4.08)	3.33 (2.78, 3.93)
phecode_713-2	Effusion of joint	3.37 (2.93, 3.71)	3.33 (2.93, 3.59)
phecode_308	Signs and symptoms involving emotional state	3.91 (3.56, 4.85)	3.33 (2.93, 3.92)
phecode_977-72	Long term (current) use of oral hypoglycemic drugs	3.19 (2.86, 3.4)	3.33 (2.95, 3.57)
phecode_413-1	Mitral valve disorders	3.85 (3.52, 4.15)	3.33 (3.06, 3.63)
phecode_426	Other heart disorders in diseases NEC	3.62 (3.3, 4.11)	3.33 (3.08, 3.93)
phecode_286	Mood [affective] disorders	4.23 (3.64, 5)	3.34 (2.71, 3.81)
phecode_324-41	Essential tremor*	3.94 (3.39, 4.54)	3.34 (2.72, 3.77)

5 Phenome-wide prediction of disease onset from retinal fundus photographs

Table 24 continued from previous page

Endpoint	PheCode String	unadjusted HR (IQR)	HR adjusted for Age and Sex (IQR)
phecode_374-4	Retinal disorders in diseases classified elsewhere	3.23 (2.83, 3.43)	3.34 (2.92, 3.53)
phecode_420	Cardiac arrest	3.47 (3.13, 3.66)	3.34 (2.98, 3.52)
phecode_431-11	Cerebral infarction [Ischemic stroke]	3.72 (3.26, 4.12)	3.34 (3.02, 3.67)
phecode_385	Abnormal results of function studies of eye	3.26 (3.03, 3.59)	3.34 (3.03, 3.83)
phecode_431-1	Stroke	3.67 (3.31, 4.04)	3.34 (3.12, 3.67)
phecode_581-3	Obstructive and reflux uropathy	3.92 (3.25, 4.23)	3.35 (2.84, 3.6)
phecode_483	Pleural effusion	3.5 (3.11, 3.78)	3.35 (2.99, 3.68)
phecode_812	Edema	3.52 (3.14, 3.73)	3.35 (3.11, 3.73)
phecode_809	Pain	3.12 (2.77, 3.94)	3.35 (3.13, 3.86)
phecode_329-9	Delirium	3.62 (3.36, 4.09)	3.35 (3.16, 3.82)
phecode_288-4	Phobic disorders	4.38 (3.9, 5.34)	3.36 (2.69, 4.12)
phecode_509-11	Glossodynia	4.04 (3.08, 5.38)	3.36 (2.71, 4.07)
phecode_416-212	Persistent atrial fibrillation*	4.38 (3.81, 4.96)	3.36 (3.01, 3.94)
phecode_256-3	Mixed disorder of acid-base balance	3.32 (2.86, 3.61)	3.37 (2.86, 3.65)
phecode_387-3	Astigmatism	3.71 (3.25, 4.35)	3.37 (2.93, 3.98)
phecode_488-1	Dyspnea [Shortness of breath]	3.4 (3.06, 3.66)	3.37 (3.03, 3.68)
phecode_676-2	Scar conditions and fibrosis of skin	3.49 (3.22, 4.25)	3.37 (3.15, 4.12)
phecode_002	Staphylococcus	3.27 (2.97, 3.53)	3.38 (0, 3.65)
phecode_718-4	Low back pain	3.45 (3.03, 4.02)	3.38 (2.96, 3.83)
phecode_480	Pulmonary edema	3.53 (3.1, 3.93)	3.38 (2.98, 3.73)
phecode_821	Abnormality of red blood cells	3.33 (2.96, 4.02)	3.38 (2.99, 4.05)
phecode_280	Substance related disorders	3.61 (3.21, 3.94)	3.38 (3.04, 3.67)
phecode_256-7	Volume depletion	3.49 (3.16, 3.73)	3.38 (3.05, 3.63)
phecode_676-1	Hypertrophic scar [Keloid scar]	3.32 (3.09, 3.61)	3.38 (3.14, 3.64)
phecode_500-41	Impacted teeth*	4.47 (3.74, 8.95)	3.39 (2.27, 10.76)
phecode_116-4	Secondary malignant neoplasm of liver	3.59 (3.14, 3.92)	3.39 (2.93, 3.79)
phecode_005	Mycobacteria	3.43 (3.05, 3.93)	3.39 (3.03, 3.88)
phecode_247-4	Disorders of magnesium metabolism	3.4 (3.07, 3.91)	3.39 (3.04, 4)
phecode_327	Other degenerative diseases of nervous system	3.51 (3.19, 3.77)	3.39 (3.06, 3.72)
phecode_841-11	Penicillin allergy	3.51 (3.17, 4.06)	3.39 (3.08, 4.13)
phecode_280-1	Alcohol use disorders	3.62 (3.22, 3.9)	3.4 (3.04, 3.68)
phecode_726-2	Pathologic fracture	3.93 (3.55, 4.91)	3.4 (3.05, 3.94)
phecode_139-52	Lipoma of intrathoracic organs	3.41 (3.09, 4.47)	3.4 (3.09, 4.84)
phecode_721-12	Radial styloid tenosynovitis [de Quervain]	3.98 (3.3, 4.51)	3.41 (2.77, 3.97)
phecode_683-2	Nail dystrophy*	3.32 (2.83, 3.95)	3.41 (2.93, 4.4)
phecode_801	Cough	2.91 (2.59, 3.38)	3.41 (2.94, 3.89)
phecode_232-29	Folate deficiency [Vitamin B9]	3.52 (3.06, 4.16)	3.41 (2.98, 4.02)
phecode_582	Acute kidney failure	3.52 (3.13, 3.79)	3.41 (3.01, 3.67)
phecode_507-1	Stomatitis	3.81 (3.31, 4.48)	3.41 (3.04, 4)
phecode_286-2	Major depressive disorder	4.32 (3.71, 4.93)	3.42 (2.81, 3.88)
phecode_116	Secondary malignant neoplasm	3.79 (3.38, 4.16)	3.42 (3.03, 3.81)
phecode_491	Pleurisy	3.83 (3.45, 4.28)	3.42 (3.04, 3.98)
phecode_374-55	Puckering of macula	3.54 (3.17, 4.04)	3.42 (3.12, 3.91)
phecode_514-2	Intestinal obstruction	3.63 (3.16, 4.1)	3.43 (3.03, 4.03)
phecode_514	Gastrointestinal obstruction	3.7 (3.25, 4.41)	3.43 (3.04, 4.18)
phecode_479	Pulmonary insufficiency and acute respiratory distress syndrome	3.49 (3.15, 3.73)	3.43 (3.12, 3.72)
phecode_374	Disorders of the retina	3.34 (3.05, 3.56)	3.43 (3.16, 3.68)
phecode_256-1	Hyperosmolality and/or hypernatremia	3.43 (3.01, 3.96)	3.44 (0.3, 3.96)
phecode_204-2	Impaired glucose tolerance (oral)	3.67 (3.18, 3.92)	3.44 (2.94, 3.66)
phecode_841-3	Allergy to narcotic agent	3.88 (3.35, 4.41)	3.44 (2.98, 3.81)
phecode_168-1	Hypo-coagulability	3.35 (2.99, 3.89)	3.44 (3.04, 4.04)
phecode_425	Cardiomegaly	3.58 (3.33, 3.93)	3.44 (3.13, 3.76)

Supplementary Tables

Table 24 continued from previous page

Endpoint	PheCode String	unadjusted HR (IQR)	HR adjusted for Age and Sex (IQR)
phecode_514-3	Ileus	3.48 (3.13, 3.85)	3.45 (0.27, 3.81)
phecode_379-21	Infection of the eye, herpes	3.03 (2.56, 3.46)	3.45 (2.58, 4.47)
phecode_386-2	Diplopia	3.71 (3.34, 5.14)	3.45 (2.69, 4.48)
phecode_247-5	Disorders of calcium metabolism	3.4 (3.08, 3.9)	3.45 (3.09, 3.97)
phecode_413-3	Tricuspid valve disorders	3.77 (3.47, 4.26)	3.45 (3.15, 3.78)
phecode_514-21	Impaction of intestine	3.49 (2.82, 4.88)	3.46 (0.06, 5.18)
phecode_500	Disorders of tooth development	4.68 (3.47, 7.61)	3.46 (2.14, 6.77)
phecode_308-1	Irritability	3.24 (2.76, 4.05)	3.46 (2.41, 4.65)
phecode_362	Other disorders of the eyelids	3.78 (3.32, 4.81)	3.46 (2.69, 4.13)
phecode_379	Infection of the eye	3.08 (2.46, 3.41)	3.46 (2.89, 4.27)
phecode_371	Cataract	3.83 (3.37, 4.11)	3.46 (3.01, 3.7)
phecode_376-21	Crystalline deposits in vitreous body	3.6 (3.06, 4.2)	3.46 (3.02, 4.01)
phecode_718-5	Sciatica	3.76 (3.43, 4.39)	3.46 (3.03, 3.89)
phecode_443-1	Stricture of artery [Arterial stenosis]	3.66 (3.23, 4.41)	3.46 (3.03, 4.11)
phecode_826	Other abnormal immunological findings in serum	3.52 (3.14, 3.92)	3.46 (3.04, 3.91)
phecode_529-3	Fecal incontinence	3.59 (3.22, 4.01)	3.46 (3.05, 3.85)
phecode_686-1	Pressure ulcer	3.51 (3.12, 4.03)	3.46 (3.05, 3.9)
phecode_810	Shock	3.57 (3.14, 4.32)	3.46 (3.05, 4.22)
phecode_202-4	Other specified diabetes*	3.38 (3.02, 3.63)	3.46 (3.06, 3.68)
phecode_120-21	Mature B-cell	3.93 (3.44, 4.71)	3.46 (3.08, 4.7)
phecode_440	Embolism and thrombosis	3.6 (3.2, 3.89)	3.46 (3.1, 3.8)
phecode_826-3	Raised antibody titer*	3.49 (3.17, 3.79)	3.46 (3.21, 3.88)
phecode_831	Glycosuria	3.22 (2.87, 3.5)	3.47 (3.05, 3.76)
phecode_433-1	Occlusion and stenosis of cerebral arteries	3.62 (3.36, 4.01)	3.47 (3.13, 3.75)
phecode_513-2	Gastric ulcer	3.31 (3.05, 3.84)	3.47 (3.13, 3.98)
phecode_665	Psoriasis	2.77 (2.29, 3.1)	3.48 (2.91, 4.28)
phecode_092	Bacteremia, Sepsis, and SIRS	3.51 (3.07, 3.85)	3.48 (3, 3.82)
phecode_092-2	Sepsis	3.5 (3.08, 3.91)	3.48 (3.01, 3.89)
phecode_495-1	Solitary pulmonary nodule	3.25 (2.85, 3.66)	3.48 (3.02, 4.18)
phecode_416-1	Paroxysmal tachycardia	3.83 (3.28, 4.3)	3.48 (3.03, 4.05)
phecode_680-3	Xerosis cutis*	3.7 (3.36, 4.2)	3.48 (3.04, 3.9)
phecode_373	Noninflammatory disorders of choroid	3.42 (3.15, 3.71)	3.48 (3.15, 3.75)
phecode_809-3	Pain in limb	3.28 (2.87, 3.88)	3.48 (3.17, 3.76)
phecode_481	Interstitial pulmonary diseases	3.8 (3.42, 4.22)	3.48 (3.17, 3.87)
phecode_371-3	Nuclear cataract	3.8 (3.27, 4.21)	3.49 (3.02, 3.8)
phecode_430	Nontraumatic Intracranial hemorrhage	3.63 (3.31, 4.19)	3.49 (3.09, 4.15)
phecode_448	Peripheral vascular disease	3.57 (3.22, 3.86)	3.49 (3.1, 3.82)
phecode_719-7	Muscle weakness (generalized)	3.59 (3.19, 4.45)	3.49 (3.24, 4.47)
phecode_537-1	Peritonitis	3.45 (3.13, 4.25)	3.5 (0.02, 4.47)
phecode_414-2	Dilated cardiomyopathy*	3.63 (3.24, 4.14)	3.5 (2.99, 3.93)
phecode_815	Symptoms and signs concerning food and fluid intake	3.47 (3.13, 4.53)	3.5 (3.12, 4.52)
phecode_487-3	Hemoptysis	3.56 (3.08, 4.09)	3.5 (3.16, 3.93)
phecode_554-2	Cyst and pseudocyst of pancreas	3.61 (3.1, 4.56)	3.51 (2.89, 4.38)
phecode_509-1	Glossitis	4.17 (3.25, 6.54)	3.51 (2.92, 4.71)
phecode_371-31	Age-related nuclear cataract	3.82 (3.26, 4.19)	3.51 (3.01, 3.78)
phecode_718	Back pain	3.41 (3.04, 4.41)	3.51 (3.06, 4.22)
phecode_580-3	Nephrotic syndrome	3.51 (3.15, 4.61)	3.51 (3.11, 4.85)
phecode_256	Disorders of fluid, electrolyte and acid-base balance	3.56 (3.21, 4.04)	3.51 (3.14, 4.06)
phecode_468-9	Lobar pneumonia*	3.63 (3.32, 3.85)	3.51 (3.2, 3.76)
phecode_004	Streptococcus	3.34 (2.87, 4.04)	3.52 (0.04, 4.46)
phecode_708-8	Secondary osteoarthritis	3.48 (3.09, 4.4)	3.52 (2.8, 5.03)
phecode_356-2	Aphasia and dysphasia	3.89 (3.24, 4.51)	3.52 (2.83, 4.28)
phecode_089-2	Viral infections	4.08 (3.48, 4.73)	3.52 (3.11, 4.47)
phecode_479-6	Pulmonary collapse [Atelectasis]	3.57 (3.24, 3.85)	3.52 (3.18, 3.8)
phecode_124	Myeloproliferative disorder	3.75 (3.31, 4.37)	3.52 (3.18, 4.24)
phecode_718-3	Mid back pain	4.35 (3.86, 6.58)	3.53 (2.97, 5.04)

5 Phenome-wide prediction of disease onset from retinal fundus photographs

Table 24 continued from previous page

Endpoint	PheCode String	unadjusted HR (IQR)	HR adjusted for Age and Sex (IQR)
phecode_410	Inflammation of the heart [Carditis]	3.85 (3.28, 4.35)	3.53 (3.02, 4.13)
phecode_510-2	Esophagitis	3.43 (2.99, 3.7)	3.53 (3.05, 3.76)
phecode_124-5	Essential thrombocythemia	3.52 (3.11, 4.19)	3.53 (3.08, 4.26)
phecode_208-21	Primary hyperparathyroidism	4.34 (3.62, 5.16)	3.53 (3.13, 4.75)
phecode_706	Other inflammatory spondylopathies	3.64 (3.36, 4.4)	3.53 (3.23, 4.3)
phecode_281	Substance abuse, dependence, and withdrawal	4.07 (3.72, 4.54)	3.53 (3.25, 3.92)
phecode_282-1	Current tobacco use and nicotine dependence	3.91 (3.52, 4.36)	3.53 (3.26, 3.97)
phecode_500-4	Disturbances in tooth eruption	4.5 (3.36, 8.34)	3.54 (2.27, 8.11)
phecode_708-15	Primary osteoarthritis of the wrist, forearm	3.89 (3.2, 4.76)	3.54 (2.77, 4.93)
phecode_688-3	Pyogenic granuloma of skin and subcutaneous tissue	3.11 (2.65, 3.9)	3.54 (2.83, 4.6)
phecode_528-2	Vomiting	3.77 (3.31, 4.24)	3.54 (3.14, 3.85)
phecode_528	Nausea and vomiting	3.74 (3.31, 4.26)	3.54 (3.14, 3.86)
phecode_528-1	Nausea	3.78 (3.33, 4.24)	3.54 (3.17, 3.84)
phecode_556-8	Nonspecific abnormal findings in stool contents	3.51 (3.18, 3.98)	3.54 (3.19, 3.96)
phecode_386	Visual disturbances	3.25 (2.97, 3.6)	3.54 (3.2, 3.76)
phecode_688-1	Sarcoidosis	3.4 (3.13, 3.93)	3.54 (3.27, 4.31)
phecode_679-7	Abnormal granulation tissue, NOS	4.03 (3.4, 4.61)	3.55 (3.01, 4.19)
phecode_550-2	Cholecystitis	3.51 (3.03, 4.09)	3.55 (3.03, 4.28)
phecode_423	Abnormal cardiac sounds	4.02 (3.62, 4.5)	3.55 (3.11, 3.91)
phecode_724-1	Myalgia	4.01 (3.53, 4.52)	3.55 (3.17, 3.84)
phecode_449	Other disorders of the circulatory system	3.68 (3.26, 4.23)	3.55 (3.19, 4.28)
phecode_719	Disorders of muscle	3.4 (3.04, 3.76)	3.55 (3.28, 4.03)
phecode_247-3	Disorder of phosphorus metabolism	3.62 (3.01, 4.19)	3.56 (0.55, 4.38)
phecode_440-3	Pulmonary embolism	3.58 (3.18, 3.96)	3.56 (3.05, 3.87)
phecode_103-1	Melanomas of skin	4.2 (3.78, 5.06)	3.56 (3.15, 4.23)
phecode_414	Cardiomyopathy	3.76 (3.57, 4.28)	3.56 (3.28, 4.03)
phecode_208	Disorders of parathyroid gland	3.6 (3.07, 4.24)	3.57 (3.03, 4.21)
phecode_355-2	Alteration of consciousness	3.5 (3.11, 4.16)	3.57 (3.1, 4.32)
phecode_413-32	Tricuspid valve insufficiency*	3.69 (3.28, 4.48)	3.57 (3.2, 4.19)
phecode_341-2	Hemiplegia and hemiparesis	3.66 (3.31, 4.19)	3.57 (3.21, 4.03)
phecode_375-14	Low-tension glaucoma (Normal-tension glaucoma)	3.72 (3.41, 4.09)	3.57 (3.33, 3.9)
phecode_376-2	Vitreous opacities	3.81 (3.06, 4.19)	3.58 (3.01, 3.92)
phecode_120-1	Myeloid	3.91 (3.48, 4.2)	3.58 (3.1, 4.03)
phecode_440-13	Phlebitis and thrombophlebitis	3.48 (3.03, 4.07)	3.58 (3.11, 4.41)
phecode_350	Other symptoms involving nervous system	3.93 (3.44, 4.23)	3.58 (3.12, 3.95)
phecode_351	Disturbances of skin sensation	3.27 (3.05, 3.62)	3.58 (3.19, 3.89)
phecode_520-12	Femoral hernia	3.96 (3.51, 5.5)	3.59 (2.89, 4.8)
phecode_702	Infective and reactive arthropathies	3.63 (2.98, 4.57)	3.59 (2.93, 4.63)
phecode_438-12	Thoracic aneurysm	3.94 (3.37, 4.87)	3.59 (2.93, 4.67)
phecode_376	Disorders of vitreous body	3.81 (3.08, 4.19)	3.59 (3.01, 3.94)
phecode_375-11	Open angle glaucoma	3.59 (3.32, 3.93)	3.59 (3.27, 3.96)
phecode_439	Hemorrhoids	3.52 (2.68, 4.56)	3.6 (2.61, 4.39)
phecode_168-19	Spontaneous ecchymoses	3.24 (2.74, 3.94)	3.6 (3, 4.84)
phecode_973	Adverse effect of other drug	3.52 (3.07, 3.69)	3.6 (3.09, 3.95)
phecode_817	Motion sickness	5.23 (3.57, 6.71)	3.61 (2.62, 5.1)
phecode_511	Gastro-esophageal reflux disease	3.62 (3.06, 4.16)	3.61 (3.07, 4.32)
phecode_355	Coma and other alteration of consciousness	3.63 (3.16, 4.1)	3.61 (3.09, 4.24)
phecode_411-2	Pericardial effusion (noninflammatory)*	3.48 (3.09, 3.8)	3.61 (3.14, 4.14)
phecode_468-1	Viral pneumonia	3.61 (3.2, 3.95)	3.61 (3.18, 4.07)

Supplementary Tables

Table 24 continued from previous page

Endpoint	PheCode String	unadjusted HR (IQR)	HR adjusted for Age and Sex (IQR)
phecode_351-1	Anesthesia of skin*	3.67 (3.14, 4.41)	3.61 (3.21, 4.39)
phecode_555	Ascites	3.56 (3.22, 3.99)	3.61 (3.25, 4.18)
phecode_468	Pneumonia	3.73 (3.44, 4.01)	3.61 (3.32, 3.89)
phecode_366-4	Vascular abnormalities of conjunctiva	3.14 (2.57, 3.66)	3.62 (2.81, 4.42)
phecode_135-5	Benign neoplasm of the paranasal sinus and nasal cavity	3.82 (3.35, 5.03)	3.62 (2.83, 4.87)
phecode_381-1	Paralytic strabismus [Neurogenic strabismus]	3.95 (3.3, 4.37)	3.62 (3.07, 4.12)
phecode_680-1	Corns and callosities	3.43 (2.82, 4.04)	3.62 (3.09, 4.15)
phecode_349-1	Abnormal findings on diagnostic test of central nervous system	3.61 (3.24, 4.03)	3.62 (3.2, 4.03)
phecode_710	Acquired deformities of limbs	3.62 (3.27, 4.46)	3.62 (3.22, 4.36)
phecode_481-4	Pulmonary fibrosis	4.12 (3.64, 5.39)	3.62 (3.24, 4.84)
phecode_660-6	Cellulitis and abscess	3.5 (3.11, 3.78)	3.62 (3.25, 4.05)
phecode_495	Abnormal findings on diagnostic imaging of lung	3.55 (3.27, 4.01)	3.62 (3.37, 4.22)
phecode_554	Diseases of the pancreas	3.55 (3.15, 4.07)	3.63 (3.19, 4.25)
phecode_374-5	Macular degeneration	3.81 (3.46, 4.27)	3.63 (3.29, 4.07)
phecode_474	Chronic obstructive pulmonary disease [COPD]	3.67 (3.23, 3.88)	3.63 (3.3, 3.96)
phecode_529-1	Diarrhea	3.77 (3.22, 4.52)	3.64 (3.07, 4.3)
phecode_101-6	Malignant neoplasm of the liver and intrahepatic bile ducts	3.88 (3.3, 4.48)	3.64 (3.11, 4.27)
phecode_406	Chronic pulmonary heart disease	3.8 (3.38, 4.28)	3.64 (3.21, 4.06)
phecode_411	Other diseases of pericardium	3.56 (3.14, 3.83)	3.64 (3.25, 4.05)
phecode_413-22	Aortic insufficiency	4.07 (3.4, 4.53)	3.65 (3.01, 4.19)
phecode_239-12	Familial hypercholesterolemia*	4.44 (3.69, 5.39)	3.65 (3.14, 4.89)
phecode_232-2	Vitamin B group deficiency	3.51 (3.15, 3.68)	3.65 (3.17, 3.92)
phecode_440-1	Venous thromboembolism	3.48 (3.07, 4.12)	3.65 (3.21, 4.4)
phecode_391-1	Otitis media	4.08 (3.69, 4.89)	3.65 (3.3, 4.48)
phecode_668-5	Lichen simplex chronicus	3.72 (2.9, 5)	3.66 (2.79, 4.71)
phecode_504	Periodontal diseases	3.72 (3.18, 4.59)	3.66 (3.15, 4.71)
phecode_244	Disorders of lipoprotein metabolism and other lipidemias	3.58 (3.08, 3.88)	3.66 (3.16, 4.04)
phecode_139-53	Lipoma of other skin subcutaneous tissue	4.03 (2.98, 5.33)	3.67 (2.74, 4.54)
phecode_466-4	Hypertrophy of tonsils and adenoids	3.45 (3.06, 4.49)	3.67 (2.75, 4.99)
phecode_430-2	Nontraumatic intracerebral hemorrhage	3.74 (3.32, 4.37)	3.67 (3.14, 4.32)
phecode_169-1	Thrombocytopenia	3.93 (3.38, 4.36)	3.67 (3.17, 4.09)
phecode_406-1	Pulmonary hypertension	3.86 (3.43, 4.26)	3.67 (3.24, 4.03)
phecode_807	Malaise and fatigue	4.25 (3.96, 4.94)	3.67 (3.27, 4.29)
phecode_513	Peptic ulcer	3.28 (3.06, 3.69)	3.67 (3.32, 4.03)
phecode_341	Cerebral palsy and other paralytic syndromes	3.75 (3.38, 4.15)	3.67 (3.33, 4.09)
phecode_522-8	Duodenitis	3.97 (3.46, 4.43)	3.67 (3.38, 4.19)
phecode_164-1	Microcytic anemia	3.7 (3.51, 3.97)	3.67 (3.49, 4.05)
phecode_015-2	Clostridium difficile	3.61 (3, 4.15)	3.68 (0.29, 4.24)
phecode_537	Abnormality of the peritoneum	3.51 (3.03, 4.1)	3.68 (2.61, 4.46)
phecode_516	Other diseases of stomach and duodenum	3.69 (3.21, 4.2)	3.68 (3.1, 4.42)
phecode_468-2	Bacterial pneumonia	3.59 (3.17, 4.21)	3.68 (3.18, 4.47)
phecode_668	Dermatitis [eczema]	4.03 (3.52, 4.57)	3.68 (3.18, 4.52)
phecode_180	Other disorders involving the immune mechanism	3.73 (3.37, 4.27)	3.68 (3.22, 4.3)
phecode_160-1	Iron deficiency anemia	3.7 (3.52, 4)	3.68 (3.5, 4.07)
phecode_706-1	Sacroiliitis NEC	3.78 (2.91, 5.03)	3.69 (2.63, 4.78)
phecode_208-2	Hyperparathyroidism	3.65 (2.99, 4.14)	3.69 (2.95, 4.14)
phecode_507	Lesions of mouth	4.2 (3.49, 4.99)	3.69 (2.98, 4.2)
phecode_664-1	Lichen planus, nitidus, or striatus	4.03 (3.24, 4.68)	3.69 (2.99, 4.22)
phecode_003	Escherichia coli	3.69 (3.19, 4.12)	3.69 (3.12, 4.33)
phecode_374-52	Macular cyst, hole, or pseudohole	3.72 (3.31, 4.14)	3.69 (3.17, 4.06)

5 Phenome-wide prediction of disease onset from retinal fundus photographs

Table 24 continued from previous page

Endpoint	PheCode String	unadjusted HR (IQR)	HR adjusted for Age and Sex (IQR)
phecode_722-1	Plantar fascial fibromatosis [Plantar fasciitis]	4.08 (3.66, 5.46)	3.69 (3.23, 4.62)
phecode_660	Infection of the skin	3.5 (2.92, 4.02)	3.69 (3.24, 4.27)
phecode_230-3	Anorexia	3.66 (3.29, 4.16)	3.69 (3.31, 4.25)
phecode_413-42	Pulmonary valve insufficiency*	4.49 (3.54, 5.14)	3.7 (3.12, 4.43)
phecode_169	Platelet defects	3.95 (3.37, 4.31)	3.7 (3.18, 4.08)
phecode_361-4	Blepharochalasis	3.46 (3.1, 4.12)	3.7 (3.22, 4.56)
phecode_556	Other symptoms involving the digestive system and abdomen	3.62 (3.22, 3.89)	3.7 (3.25, 3.97)
phecode_719-1	Cramp and spasm	3.49 (3.1, 3.87)	3.7 (3.26, 4.1)
phecode_089-1	Bacterial infections	3.64 (3.28, 4.23)	3.7 (3.27, 4.29)
phecode_177-4	Lymphedema	4.08 (3.58, 4.71)	3.7 (3.29, 4.6)
phecode_356	Speech disturbance	3.73 (3.38, 4.76)	3.7 (3.29, 4.79)
phecode_664	Papulosquamous disorders	3.85 (3.27, 4.42)	3.71 (3.12, 4.45)
phecode_386-4	Visual field defects	3.57 (3.13, 4.49)	3.71 (3.24, 4.58)
phecode_532-4	Volvulus	3.9 (3.59, 4.6)	3.71 (3.36, 4.49)
phecode_247-72	Iron deficiency	3.71 (3.49, 4)	3.71 (3.48, 4.05)
phecode_123	Multiple myeloma and malignant plasma cell neoplasms	4.13 (3.68, 4.84)	3.72 (3.21, 4.52)
phecode_168	Coagulation defects, purpura and other hemorrhagic conditions	3.74 (3.34, 4.61)	3.72 (3.31, 4.76)
phecode_424-2	Systolic heart failure	4.29 (3.79, 4.71)	3.72 (3.41, 4.16)
phecode_284	Suicide ideation and attempt or self harm	4.36 (3.83, 5.07)	3.73 (3.02, 4.53)
phecode_402	Elevated blood pressure reading without diagnosis of hypertension	3.28 (2.84, 3.84)	3.73 (3.14, 4.24)
phecode_423-1	Cardiac murmurs	4.13 (3.68, 4.68)	3.73 (3.26, 4.07)
phecode_059-1	COVID-19*	3.79 (3.29, 4.42)	3.73 (3.27, 4.35)
phecode_349	Disorder of nervous system	3.8 (3.39, 4.34)	3.73 (3.32, 4.24)
phecode_211	Disorders of adrenal glands	3.42 (2.97, 4.15)	3.74 (0.01, 4.41)
phecode_463-4	Nasal congestion*	4.21 (3.33, 6.62)	3.74 (2.92, 5.81)
phecode_724-3	Nontraumatic hematoma of soft tissue	3.82 (3.24, 5.32)	3.74 (2.99, 5.27)
phecode_526-12	Intestinal infection due to C. difficile	3.63 (3.08, 4.19)	3.74 (3.1, 4.28)
phecode_542-4	Portal hypertension	3.46 (3.12, 4.79)	3.74 (3.25, 5.29)
phecode_360-2	Chalazion	4.09 (3.39, 9.25)	3.75 (2.94, 10)
phecode_148-2	Benign neoplasm of meninges (Meningioma)	4.67 (4.01, 7.38)	3.75 (3.04, 5.12)
phecode_148	Benign neoplasm of the eye, brain and other parts of central nervous system	4.82 (4.23, 7.15)	3.75 (3.08, 5.31)
phecode_102-1	Malignant neoplasm of the of bronchus and lung	3.82 (3.3, 4.34)	3.75 (3.31, 4.33)
phecode_674-1	Hypopigmentation	3.68 (3.33, 4.15)	3.75 (3.33, 4.42)
phecode_374-51	Age-related macular degeneration	3.92 (3.51, 4.42)	3.75 (3.35, 4.19)
phecode_247-7	Disorders of iron metabolism	3.81 (3.55, 4.01)	3.75 (3.54, 4.04)
phecode_841-5	Allergy to serum and vaccine	4.1 (3.32, 5.32)	3.76 (2.55, 5.34)
phecode_367-2	Keratitis	3.52 (2.98, 4.62)	3.76 (2.88, 4.65)
phecode_666-2	Idiopathic urticaria	3.88 (3.08, 4.84)	3.76 (3.02, 4.73)
phecode_464-1	Acute nasopharyngitis	3.5 (3.08, 4.36)	3.76 (3.07, 4.63)
phecode_688	Granulomatous disorder of the skin	3.27 (2.62, 4.65)	3.76 (3.07, 5.34)
phecode_102	Malignant neoplasm of the thoracic and respiratory organs	3.86 (3.35, 4.25)	3.76 (3.25, 4.18)
phecode_160	Nutritional anemias	3.81 (3.58, 4.04)	3.76 (3.53, 4.08)
phecode_557-1	Hematemesis	3.69 (3.13, 4.05)	3.77 (0.04, 4.24)
phecode_581-11	Acute pyelonephritis	4.3 (3.64, 5.67)	3.77 (2.84, 4.99)
phecode_120	Hemo onc - by cell of origin	4.23 (3.62, 4.8)	3.77 (3.01, 4.49)
phecode_120-2	Lymphoid	4.36 (3.69, 5.09)	3.77 (3.26, 5.16)
phecode_247	Disorders of mineral metabolism and mineral deficiencies	3.82 (3.54, 4.09)	3.77 (3.48, 4.07)

Supplementary Tables

Table 24 continued from previous page

Endpoint	PheCode String	unadjusted HR (IQR)	HR adjusted for Age and Sex (IQR)
phecode_668-1	Atopic dermatitis	3.89 (2.94, 4.45)	3.78 (2.85, 4.47)
phecode_256-5	Hypokalemia [Hypopotassemia]	3.7 (3.07, 4.02)	3.78 (3.04, 4.24)
phecode_552	Other diseases of biliary tract	3.67 (3.09, 4.16)	3.78 (3.06, 4.51)
phecode_464	Nasopharyngitis	3.48 (3.08, 4.45)	3.78 (3.06, 4.76)
phecode_122-2	Non-Hodgkin lymphoma	4.09 (3.5, 5.34)	3.78 (3.33, 5.27)
phecode_502	Other diseases of teeth and supporting structures	3.6 (3.37, 4.54)	3.78 (3.4, 4.58)
phecode_375-113	Primary open angle glaucoma	3.79 (3.56, 4.36)	3.78 (3.44, 4.47)
phecode_841-4	Allergy to analgesic agent	3.8 (3.36, 4.33)	3.78 (3.46, 4.38)
phecode_501	Dental caries	3.62 (3.15, 4.37)	3.79 (3.28, 4.52)
phecode_374-511	Nonexudative (dry) age-related macular degeneration	4.02 (3.49, 4.79)	3.79 (3.29, 4.47)
phecode_369	Noninflammatory disorders of the cornea	3.64 (3.13, 3.94)	3.8 (3.16, 4.29)
phecode_488	Abnormalities of breathing	3.64 (3.17, 3.9)	3.8 (3.3, 4.23)
phecode_101-8	Malignant neoplasm of the pancreas	4.09 (3.65, 4.81)	3.8 (3.36, 4.69)
phecode_557-8	Hemorrhage of rectum and anus	3.62 (2.92, 4.17)	3.81 (3.01, 4.67)
phecode_581-31	Hydronephrosis	4.04 (3.63, 4.52)	3.81 (3.3, 4.3)
phecode_374-512	Exudative (wet) age-related macular degeneration	3.98 (3.48, 4.61)	3.81 (3.3, 4.38)
phecode_336	Mononeuropathies	4.13 (3.26, 5.15)	3.82 (2.99, 4.7)
phecode_809-1	Acute pain	4.06 (3.22, 5.65)	3.82 (3.04, 5.41)
phecode_123-1	Multiple myeloma	4.12 (3.73, 4.85)	3.82 (3.34, 4.64)
phecode_542-3	Hepatic failure	3.61 (3.22, 4.04)	3.82 (3.4, 4.32)
phecode_164	Anemia	3.84 (3.5, 4.17)	3.82 (3.44, 4.13)
phecode_679-22	Flushing	4.78 (2.88, 6.15)	3.83 (1.86, 5.26)
phecode_376-1	Vitreous degeneration	3.82 (3.11, 4.36)	3.83 (3.27, 4.3)
phecode_059	Coronavirus	3.89 (3.37, 4.41)	3.83 (3.35, 4.36)
phecode_256-2	Hyposmolality and/or hyponatremia	3.93 (3.49, 4.7)	3.83 (3.41, 4.32)
phecode_545	Nonspecific abnormal results of function study of liver	3.61 (3.18, 4.29)	3.83 (3.42, 4.61)
phecode_471-5	Nasal polyps	3.71 (3.4, 5.13)	3.84 (2.94, 5.17)
phecode_526	Intestinal infection	3.84 (3.24, 4.65)	3.84 (3.21, 4.8)
phecode_724	Other symptoms and disorders of the soft tissue	3.92 (3.48, 4.6)	3.84 (3.31, 4.39)
phecode_162	Aplastic anemia	3.87 (3.55, 4.7)	3.84 (3.4, 4.67)
phecode_380	Disorders of optic nerve and visual pathways	3.55 (3.39, 3.96)	3.84 (3.6, 4.27)
phecode_660-4	Carbuncle and furuncle	3.92 (3.19, 4.66)	3.85 (3.1, 4.57)
phecode_814	Jaundice (not of newborn)	3.86 (3.28, 4.65)	3.85 (3.16, 4.62)
phecode_015	Clostridium	3.66 (3.1, 3.99)	3.85 (3.18, 4.22)
phecode_522-9	Gastritis	3.61 (3.13, 4.02)	3.85 (3.35, 4.5)
phecode_121	Leukemia	4.14 (3.54, 4.78)	3.85 (3.39, 4.65)
phecode_504-3	Periodontitis	3.95 (3.29, 4.86)	3.86 (3, 4.61)
phecode_503	Diseases of pulp and periapical tissues	3.93 (3.32, 5.05)	3.86 (3.23, 5.01)
phecode_724-5	Exostosis	3.71 (3.1, 4.53)	3.86 (3.26, 4.87)
phecode_503-5	Periapical abscess	3.95 (3.25, 5.11)	3.86 (3.32, 4.99)
phecode_506-3	Sialoadenitis	3.88 (3.18, 4.62)	3.86 (3.33, 4.65)
phecode_116-3	Secondary malignant neoplasm of digestive systems	3.82 (3.27, 4.4)	3.86 (3.39, 4.51)
phecode_169-11	Immune thrombocytopenic purpura [ITP]	3.77 (3.38, 4.42)	3.86 (3.42, 4.39)
phecode_164-2	Macrocytic anemia	3.79 (3.47, 4.4)	3.86 (3.44, 4.74)
phecode_829-2	Abnormal level of blood mineral*	3.88 (3.5, 5.15)	3.86 (3.52, 5.13)
phecode_174	Diseases of spleen	3.64 (3.38, 4.58)	3.87 (0.02, 5)
phecode_581	Renal tubulo-interstitial diseases	3.69 (3.22, 4.31)	3.87 (3.18, 4.67)
phecode_413-4	Pulmonary valve disorders	4.68 (3.8, 5.54)	3.87 (3.27, 4.92)
phecode_535	Intestinal or peritoneal adhesions	3.76 (3.37, 4.63)	3.87 (3.29, 4.73)
phecode_550-1	Gallstones [Cholelithiasis]	3.9 (3.43, 4.52)	3.87 (3.54, 4.81)
phecode_727-1	Osteonecrosis	3.53 (2.99, 4.54)	3.88 (3.21, 5.84)

5 Phenome-wide prediction of disease onset from retinal fundus photographs

Table 24 continued from previous page

Endpoint	PheCode String	unadjusted HR (IQR)	HR adjusted for Age and Sex (IQR)
phecode_281-2	Substance dependence	4.22 (3.84, 4.69)	3.88 (3.53, 4.24)
phecode_104	Malignant sarcoma-related cancers	3.36 (2.78, 4.48)	3.89 (3.02, 4.83)
phecode_280-12	Alcohol dependence	3.8 (3.31, 4.19)	3.89 (3.4, 4.39)
phecode_353-1	Hallucinations	3.71 (3.3, 4.73)	3.9 (3.34, 5.11)
phecode_542-1	Fibrosis and cirrhosis of liver	3.68 (3.25, 4.36)	3.9 (3.38, 4.92)
phecode_679-3	Changes in skin texture	3.92 (3.27, 4.73)	3.9 (3.39, 4.56)
phecode_387	Disorders of refraction and accommodation	3.62 (3.26, 3.97)	3.9 (3.48, 4.19)
phecode_581-1	Pyelonephritis	3.97 (3.49, 5.71)	3.91 (2.96, 5.14)
phecode_668-2	Seborrheic dermatitis	3.33 (2.78, 3.89)	3.91 (3.26, 4.81)
phecode_170-19	Neutropenia NOS	3.86 (3.47, 4.44)	3.91 (3.32, 4.54)
phecode_096	Contact or exposure to infectious agent	3.89 (3.33, 4.29)	3.91 (3.36, 4.34)
phecode_683-1	Ingrowing nail	4.14 (3.5, 5.33)	3.92 (3.03, 4.92)
phecode_525	Intestinal malabsorption	3.02 (2.52, 3.39)	3.92 (3.18, 4.55)
phecode_513-1	Esophageal ulcer	3.53 (3.07, 4.02)	3.92 (3.41, 4.66)
phecode_474-1	Emphysema	4 (3.54, 4.47)	3.92 (3.45, 4.35)
phecode_283-8	Other problems related to lifestyle	2.74 (2.49, 3.09)	3.93 (3.34, 4.72)
phecode_160-2	Megaloblastic anemia	3.8 (3.39, 4.41)	3.93 (3.4, 4.79)
phecode_122	Lymphoma	4.31 (3.74, 5.26)	3.93 (3.4, 5.09)
phecode_121-21	Chronic lymphoid leukemia	4.64 (3.68, 5.51)	3.94 (3.07, 4.99)
phecode_171	Increased white blood cell count	3.92 (3.28, 5.02)	3.94 (3.2, 5.2)
phecode_668-4	Dermatitis due to substances taken internally	3.75 (3.29, 4.2)	3.95 (0.01, 4.58)
phecode_552-1	Cholangitis	3.46 (3.05, 4.01)	3.95 (3.16, 4.29)
phecode_089-3	Fungal infections	4.12 (3.46, 4.67)	3.95 (3.25, 4.25)
phecode_361-3	Ptosis of eyelid	3.95 (3.49, 4.22)	3.95 (3.35, 4.1)
phecode_723-6	Impingement syndrome of shoulder*	3.42 (2.98, 4.21)	3.95 (3.48, 4.89)
phecode_819	General symptoms and other findings	4.15 (3.84, 4.44)	3.95 (3.59, 4.17)
phecode_718-1	Radiculopathy	3.85 (3.1, 5.28)	3.96 (3.22, 5.5)
phecode_324-8	Restless legs syndrome	4.29 (3.58, 5.21)	3.96 (3.33, 4.91)
phecode_387-2	Myopia	3.83 (3.45, 4.15)	3.96 (3.52, 4.28)
phecode_550	Disorders of the gallbladder	3.98 (3.49, 4.57)	3.96 (3.59, 5.04)
phecode_509-3	Hypertrophy of tongue papillae	3.96 (2.82, 5.5)	3.97 (2.88, 5.32)
phecode_346	Brain damage and brain death	3.94 (3.28, 4.72)	3.97 (3.2, 4.82)
phecode_431-12	Hemorrhagic stroke	3.71 (3.25, 4.38)	3.97 (3.25, 4.75)
phecode_800	Chest pain	3.5 (3, 3.98)	3.97 (3.44, 4.38)
phecode_805	Fever of unknown origin	3.78 (3.37, 4.08)	3.97 (3.47, 4.56)
phecode_496	Abnormal results of pulmonary function studies	3.6 (3.16, 4.05)	3.97 (3.5, 4.8)
phecode_666-1	Allergic urticaria	4.49 (3.75, 5.53)	3.98 (3.16, 4.72)
phecode_336-1	Carpal tunnel syndrome	4.02 (3.42, 4.91)	3.98 (3.51, 4.93)
phecode_061	Influenza virus	3.95 (3.28, 5.24)	3.99 (3.25, 5.35)
phecode_417-2	Tachycardia	3.74 (3.18, 4.36)	3.99 (3.35, 5.06)
phecode_379-2	Eye infection, viral	3.57 (2.53, 3.86)	4 (3.02, 5.09)
phecode_523-1	Diverticula of small intestine	4.51 (3.68, 5.38)	4 (3.21, 5.15)
phecode_416-51	Atrial premature depolarization [Supraventricular premature beats]	3.94 (3.5, 5.46)	4 (3.33, 5.38)
phecode_526-1	Bacterial enteritis	4.09 (3.42, 5.48)	4 (3.39, 5.24)
phecode_100	Malignant neoplasm of the head and neck	3.97 (3.59, 4.67)	4 (3.43, 4.74)
phecode_558	Abnormal findings on diagnostic imaging of the digestive tract	3.81 (3.38, 4.36)	4 (3.5, 4.83)
phecode_367-6	Episcleritis	4.2 (3.34, 5.14)	4.01 (3.23, 5.02)
phecode_508	Diseases of lips	4.34 (3.64, 5.23)	4.01 (3.45, 4.74)
phecode_512	Dysphagia	3.76 (3.41, 4.15)	4.01 (3.48, 4.47)
phecode_722	Fasciopathy	3.79 (3.26, 4.4)	4.01 (3.48, 4.5)
phecode_355-1	Coma	4.21 (3.41, 4.82)	4.02 (3.23, 4.62)

Supplementary Tables

Table 24 continued from previous page

Endpoint	PheCode String	unadjusted HR (IQR)	HR adjusted for Age and Sex (IQR)
phecode_336-2	Lesion of median, ulnar, radial nerve	3.96 (3.54, 4.92)	4.02 (3.44, 4.85)
phecode_804	Other symptoms and signs involving the circulatory and respiratory system	3.67 (3.28, 4.13)	4.02 (3.53, 4.45)
phecode_283	Other behavioral problems	3.18 (2.78, 3.72)	4.02 (3.54, 5.02)
phecode_532	Other disorders of the intestines	3.87 (3.59, 4.4)	4.02 (3.63, 4.83)
phecode_239-2	Hyperglyceridemia	3.87 (3.5, 4.84)	4.02 (3.63, 5.04)
phecode_122-22	Diffuse large B-cell lymphoma*	3.99 (3.28, 5.04)	4.03 (3.14, 5.45)
phecode_546	Other disorders of liver	3.78 (3.26, 4.29)	4.03 (3.39, 4.67)
phecode_011	Klebsiella	3.89 (3, 4.28)	4.04 (3.01, 4.57)
phecode_674-11	Vitiligo	3.93 (3.38, 4.53)	4.04 (3.48, 5.17)
phecode_098	Carrier or suspected carrier of infectious diseases	3.97 (3.43, 4.46)	4.04 (3.5, 4.64)
phecode_172	Other disorders of white blood cells	4.45 (3.79, 5.02)	4.04 (3.63, 4.73)
phecode_148-1	Benign neoplasm of eye	3.86 (3.19, 4.47)	4.05 (3.3, 4.76)
phecode_369-1	Corneal scars and opacities	3.9 (3.38, 4.45)	4.05 (3.41, 4.68)
phecode_324-3	Dystonia	3.87 (3.22, 4.83)	4.05 (3.45, 5.09)
phecode_824	Other abnormalities of plasma proteins*	4.13 (3.31, 5.25)	4.05 (3.46, 5.29)
phecode_333-1	Sleep apnea	3.85 (3.34, 4.09)	4.05 (3.54, 4.42)
phecode_204-4	Prediabetes*	3.13 (2.81, 3.75)	4.06 (3.05, 5.15)
phecode_116-1	Secondary malignancy of lymph nodes	4.01 (3.17, 5.12)	4.07 (3.12, 5.11)
phecode_530-2	Anorectal abscess	5.58 (4.35, 8.02)	4.07 (3.22, 5.51)
phecode_530	Disease of anus and rectum	4.52 (3.79, 5.57)	4.07 (3.37, 5.22)
phecode_283-4	Patient's noncompliance with medical treatment and regimen	3.98 (3.59, 4.62)	4.07 (3.67, 4.77)
phecode_181	Autoimmune disease	3.57 (3.1, 4.42)	4.08 (3.51, 4.96)
phecode_557	Gastrointestinal hemorrhage	3.95 (3.38, 4.45)	4.09 (3.41, 4.59)
phecode_352-1	Anosmia*	4.01 (3.37, 5)	4.09 (3.44, 5.27)
phecode_472	Diseases of vocal cords and larynx, not elsewhere classified	3.36 (2.84, 4.33)	4.1 (3.13, 5.3)
phecode_352	Disturbances of sensation of smell and taste	3.93 (3.26, 5.42)	4.1 (3.29, 5.78)
phecode_174-2	Splenomegaly	3.82 (3.27, 4.64)	4.11 (3.35, 5.04)
phecode_387-1	Hypermetropia	3.73 (3.38, 4.43)	4.11 (3.48, 4.8)
phecode_723	Enthesopathy/Enthesitis/Tendinopathy	5.5 (4.85, 6.79)	4.11 (3.68, 4.74)
phecode_148-16	Benign neoplasm of choroid	3.87 (3.27, 4.69)	4.12 (3.36, 5.17)
phecode_383	Irregular eye movements	3.87 (3.25, 4.86)	4.12 (3.47, 5.03)
phecode_522	Gastrointestinal inflammation	3.72 (3.22, 4.08)	4.12 (3.66, 4.62)
phecode_135	Benign neoplasm of the head and neck	3.52 (3.09, 4.22)	4.13 (3.13, 5)
phecode_380-2	Disorders of optic disc	3.75 (3.35, 4.16)	4.13 (3.67, 4.74)
phecode_460	Acute respiratory infection	4.21 (3.66, 5.06)	4.14 (3.5, 5)
phecode_504-32	Chronic periodontitis	4.53 (3.7, 5.83)	4.15 (3.23, 5.66)
phecode_504-1	Gingivitis	3.85 (3.14, 4.56)	4.15 (3.31, 4.84)
phecode_711-1	Derangement of meniscus	4.6 (3.3, 5.76)	4.16 (2.86, 5.19)
phecode_121-2	Chronic leukemia	4.47 (3.6, 5.35)	4.16 (3.12, 5.09)
phecode_308-7	Restlessness and agitation*	4.07 (3.57, 4.96)	4.16 (3.56, 5.7)
phecode_812-2	Angioneurotic edema	4.22 (3.64, 5.57)	4.16 (3.59, 6.29)
phecode_232-27	Vitamin B12 deficiency	3.7 (3.45, 4.4)	4.16 (3.76, 5.63)
phecode_280-13	Alcoholic liver disease	4.02 (3.32, 4.63)	4.17 (3.39, 5.18)
phecode_723-1	Adhesive capsulitis of shoulder	4.61 (3.65, 5.43)	4.18 (3.25, 4.82)
phecode_452	Hemorrhage, NOS	3.76 (3.06, 4.33)	4.2 (3.07, 4.86)
phecode_374-11	Serous retinal detachment	3.63 (3.38, 4)	4.2 (3.92, 4.66)
phecode_430-1	Nontraumatic subarachnoid hemorrhage	3.88 (3.26, 5.06)	4.21 (3.23, 5.79)
phecode_360-11	Hordeolum externum	4.69 (3.99, 7)	4.21 (3.53, 6.32)
phecode_176	Other diseases of blood and blood-forming organs	3.74 (3.01, 4.46)	4.22 (3.23, 6.08)
phecode_498	Asphyxia and hypoxemia	3.92 (3.44, 5.04)	4.22 (3.56, 5.24)

5 Phenome-wide prediction of disease onset from retinal fundus photographs

Table 24 continued from previous page

Endpoint	PheCode String	unadjusted HR (IQR)	HR adjusted for Age and Sex (IQR)
phecode_288-2	Panic disorder [episodic paroxysmal anxiety]	5.05 (4.17, 6.41)	4.23 (3.3, 5.67)
phecode_682-12	Pilar and trichodermal cyst	2.82 (1.87, 4.57)	4.24 (2.88, 7.97)
phecode_581-33	Stricture or kinking of ureter	4.47 (3.61, 6.3)	4.24 (3.46, 6.59)
phecode_554-1	Pancreatitis	3.99 (3.35, 4.81)	4.25 (3.29, 5.3)
phecode_325-3	Lack of coordination	4.12 (3.54, 5.25)	4.25 (3.32, 5.57)
phecode_679-21	Pallor	3.91 (3.29, 5.68)	4.25 (3.49, 5.86)
phecode_331-3	Headache syndromes, non migraine	4.03 (3.05, 5.62)	4.26 (3.26, 5.6)
phecode_334-2	Facial nerve disorders and weakness	3.93 (3.39, 4.57)	4.27 (3.66, 4.71)
phecode_460-2	Acute lower respiratory infection	3.9 (3.47, 4.34)	4.27 (3.88, 4.87)
phecode_109	Malignant neoplasm of the eye, brain and other parts of central nervous system	3.86 (3.24, 4.6)	4.28 (3.25, 5.69)
phecode_360-1	Hordeolum	4.71 (3.93, 7.2)	4.28 (3.46, 6.31)
phecode_177	Abnormality of the lymph nodes	4.24 (3.75, 4.91)	4.28 (3.76, 4.93)
phecode_330-3	Convulsions	4.03 (3.67, 4.96)	4.28 (3.77, 5.65)
phecode_554-11	Acute pancreatitis	4.01 (3.18, 4.83)	4.29 (3.26, 5.86)
phecode_593-3	Recurrent and persistent hematuria*	4.02 (3.08, 5.73)	4.3 (3.32, 7.08)
phecode_484	Pneumothorax and air leak	4.19 (3.53, 4.78)	4.3 (3.71, 5.03)
phecode_710-4	Acquired deformities of the ankle and foot	3.85 (3.43, 4.62)	4.3 (3.78, 4.94)
phecode_374-1	Retinal detachments and breaks	3.75 (3.34, 4.08)	4.3 (3.91, 4.74)
phecode_702-3	Enteropathic arthropathies	4.15 (3.13, 5.23)	4.32 (3.13, 5.67)
phecode_667	Erythematous conditions	3.85 (3.23, 5.46)	4.32 (3.88, 6.24)
phecode_711	Disorder of patella	4.54 (3.43, 5.96)	4.33 (2.99, 5.38)
phecode_324-34	Torticollis	3.98 (3.27, 4.98)	4.33 (3.66, 5.11)
phecode_592-12	Chronic cystitis	4.48 (3.41, 5.42)	4.34 (3.28, 5.04)
phecode_723-3	Medial epicondylitis (Golfer's elbow)	6.93 (5.14, 10.26)	4.34 (3.32, 6.21)
phecode_177-2	Enlargement of lymph nodes [Lymphadenopathy]	4.29 (3.55, 5.03)	4.34 (3.53, 4.93)
phecode_308-5	Nervousness	4.83 (3.69, 6.5)	4.35 (3.12, 5.99)
phecode_109-3	Malignant neoplasm of brain	4.04 (3.53, 4.66)	4.35 (3.78, 5.42)
phecode_734-9	Jaw pain	5.29 (4.39, 7.04)	4.36 (3.46, 5.67)
phecode_520-15	Incisional hernia	3.81 (3.13, 4.18)	4.36 (3.5, 5.31)
phecode_522-7	Ulceration of the lower GI tract	3.81 (3.26, 4.74)	4.36 (3.51, 5.2)
phecode_800-2	Precordial pain	3.65 (3.05, 4.14)	4.38 (3.55, 5.07)
phecode_542	Chronic liver disease and sequelae	4.11 (3.58, 4.59)	4.38 (3.87, 5.01)
phecode_384	Anomalies of pupillary function	4.01 (3.5, 5.82)	4.39 (3.55, 6.68)
phecode_116-2	Secondary malignancy of respiratory organs	3.81 (3.37, 4.5)	4.39 (3.71, 5.44)
phecode_734	Diseases of the jaws	5.27 (4.52, 7.31)	4.39 (3.71, 5.91)
phecode_367-52	Iridocyclitis	3.95 (3.29, 4.68)	4.39 (3.76, 5.32)
phecode_236-1	Obesity	3.72 (3.38, 4.21)	4.39 (3.89, 4.91)
phecode_342-4	Monoplegia	4.05 (3.53, 5.7)	4.41 (3.7, 6.92)
phecode_329-4	Other specified cognitive deficit	4.63 (2.92, 6.2)	4.42 (2.99, 6.07)
phecode_668-3	Contact dermatitis	4.43 (3.58, 5.09)	4.42 (3.56, 5.34)
phecode_367-5	Uveitis	3.92 (3.31, 4.78)	4.42 (3.83, 5.42)
phecode_236	Overweight and obesity	3.71 (3.38, 4.22)	4.42 (3.89, 4.88)
phecode_526-11	Intestinal e.coli	3.84 (3.32, 4.5)	4.43 (3.52, 5.82)
phecode_333-2	Insomnia	4.72 (4.06, 5.61)	4.43 (3.72, 5.15)
phecode_486-5	Abnormal sputum	3.91 (3.26, 4.37)	4.44 (3.72, 4.96)
phecode_522-1	Inflammatory bowel disease	4.57 (3.65, 6.45)	4.45 (3.29, 6.38)
phecode_716-3	Spinal disc displacement (herniation)	4.34 (3.34, 4.7)	4.46 (3.49, 5.57)
phecode_529-6	Halitosis*	3.96 (3.03, 4.91)	4.47 (3.27, 5.69)
phecode_334	Disorders of other cranial nerves	4.11 (3.27, 5.23)	4.48 (3.78, 5.58)
phecode_179	Immunodeficiencies	4.51 (3.02, 6.91)	4.49 (3.01, 7.92)
phecode_292	Somatoform disorders	4.66 (3.7, 6.66)	4.52 (3.45, 6.43)
phecode_712	Other specific joint derangements	4.94 (3.75, 7.38)	4.52 (3.46, 6.15)

Supplementary Tables

Table 24 continued from previous page

Endpoint	PheCode String	unadjusted HR (IQR)	HR adjusted for Age and Sex (IQR)
phecode_709-1	Acquired deformities of fingers	4.04 (2.64, 4.89)	4.53 (3.42, 5.72)
phecode_333-11	Obstructive sleep apnea	4.77 (4.3, 5.23)	4.53 (3.68, 5.24)
phecode_716	Intervertebral disc disorder	4.48 (3.67, 4.95)	4.54 (3.81, 5.13)
phecode_348-2	Myelopathies	4.38 (3.46, 4.96)	4.55 (3.69, 5.82)
phecode_386-1	Amblyopia	3.6 (3.11, 3.95)	4.56 (3.9, 5.27)
phecode_520-14	Ventral hernia	3.92 (3.24, 4.38)	4.57 (3.62, 4.92)
phecode_236-11	Morbid obesity	4.31 (3.8, 5.17)	4.57 (3.77, 5.77)
phecode_676	Hypertrophic conditions of skin	4.1 (3.2, 4.91)	4.58 (3.67, 5.46)
phecode_280-8	Other psychoactive substance related disorders	4.73 (4.1, 5.83)	4.58 (4.07, 6.01)
phecode_713-4	Stiffness of joint	4.48 (3.56, 6.65)	4.6 (3.35, 6.4)
phecode_469-2	Chronic bronchitis	4.27 (3.73, 4.91)	4.6 (4.08, 5.34)
phecode_139-5	Lipoma	4.44 (3.27, 6.33)	4.61 (3.3, 6.99)
phecode_800-11	Pleurodynia*	4.28 (3.58, 5.27)	4.61 (3.71, 5.82)
phecode_369-5	Hereditary corneal dystrophies	3.94 (3.2, 4.38)	4.61 (3.71, 5.27)
phecode_351-3	Paresthesia of skin*	4.35 (3.7, 5.15)	4.61 (4.02, 5.86)
phecode_685	Disorders of sweat glands	5.23 (4.26, 5.99)	4.62 (3.58, 5.19)
phecode_552-2	Obstruction of bile duct	4.21 (3.65, 5.05)	4.62 (3.64, 5.21)
phecode_977	Long term (current) drug therapy	4.27 (3.79, 4.92)	4.62 (3.91, 5.13)
phecode_679-2	Pallor and flushing	4.39 (3.89, 5.66)	4.63 (4.09, 6.48)
phecode_237	Abnormal weight gain	4.38 (3.63, 5.04)	4.64 (3.83, 5.38)
phecode_348	Other diseases of spinal cord	4.35 (3.67, 5.77)	4.65 (3.78, 6.27)
phecode_682-1	Cutaneous cyst	2.81 (2.11, 4.54)	4.66 (3.2, 7.99)
phecode_391-12	Chronic otitis media	4.32 (3.11, 7.48)	4.66 (3.32, 7.27)
phecode_116-5	Secondary malignancy of brain/spine	3.95 (3.33, 4.87)	4.67 (3.95, 5.81)
phecode_089	Infections	4.87 (3.96, 5.27)	4.68 (3.83, 5.4)
phecode_170-1	Neutropenia	4.34 (3.9, 4.96)	4.71 (4.07, 5.58)
phecode_356-1	Dysarthria	4.48 (3.74, 5.36)	4.73 (3.98, 5.59)
phecode_362-5	Cysts of eyelid	4.35 (3.36, 6.26)	4.74 (3.58, 5.89)
phecode_234	Other nutritional deficiencies	4.39 (3.49, 5.47)	4.74 (3.59, 5.52)
phecode_724-51	Calcaneal spur	4.19 (3.23, 5.81)	4.75 (3.72, 6.91)
phecode_170	Decreased white blood cell count	4.52 (4, 5.26)	4.77 (4.01, 5.81)
phecode_342	Plegia and unspecified paralysis	4.69 (3.98, 7.28)	4.79 (3.89, 8.06)
phecode_097-1	Methicillin resistant Staphylococcus aureus	4.23 (3.69, 5.07)	4.79 (3.95, 6.04)
phecode_370	Disorders of iris and ciliary body	3.98 (3.65, 4.67)	4.8 (4.37, 5.51)
phecode_832-5	Acetonuria	4.14 (3.62, 4.74)	4.81 (3.9, 5.69)
phecode_721-4	Calcium deposits in tendon and bursa	5.73 (4.25, 11.62)	4.82 (3.7, 10.72)
phecode_280-82	Other psychoactive substance dependence	5.27 (4.24, 7.14)	4.82 (3.79, 6.63)
phecode_334-21	Bell's palsy	4.38 (3.61, 5.29)	4.84 (3.97, 5.62)
phecode_333	Sleep disorders	4.18 (3.46, 5.07)	4.85 (3.91, 5.66)
phecode_522-12	Ulcerative colitis	3.31 (2.42, 4.71)	4.86 (3.37, 6.18)
phecode_347	Other disorders of the brain and CNS	4.17 (3.45, 4.59)	4.86 (3.69, 6.02)
phecode_556-3	Abdominal or pelvic swelling, mass, or lump	4.39 (3.51, 5.39)	4.86 (3.76, 6.28)
phecode_335	Nerve root and plexus disorders	4.6 (3.72, 5.88)	4.87 (4.02, 6.19)
phecode_471	Other disorders of nose and nasal sinuses	3.81 (3.01, 5.2)	4.89 (3.67, 6.19)
phecode_391-11	Acute otitis media	5.06 (3.97, 6.64)	4.9 (3.95, 6.49)
phecode_660-11	Candidiasis of skin and nails	4.21 (3.69, 4.99)	4.91 (4.07, 6.22)
phecode_284-29	Intentional self-harm*	5.16 (3.99, 6.65)	4.93 (3.91, 6.44)
phecode_344	Disorders of the circulation of the cerebrospinal fluid	4.23 (3.7, 5.81)	4.94 (4.04, 6.59)
phecode_712-6	Instability of joint	5.67 (3.67, 10.55)	4.95 (3.6, 9.32)
phecode_299	Mental disorder, not otherwise specified	4.63 (3.95, 5.72)	4.95 (4.15, 6.25)
phecode_097	Drug resistant microorganisms	4.22 (3.57, 4.96)	4.96 (3.95, 5.73)
phecode_284-2	Suicide and self-inflicted harm	5.16 (4.04, 6.11)	4.97 (3.93, 5.99)
phecode_329-42	Cognitive communication deficit	5.67 (4.13, 7.29)	4.98 (4.08, 7.41)
phecode_330-11	Generalized epilepsy	4.42 (3.67, 5.6)	5 (3.76, 6.31)

5 Phenome-wide prediction of disease onset from retinal fundus photographs

Table 24 continued from previous page

Endpoint	PheCode String	unadjusted HR (IQR)	HR adjusted for Age and Sex (IQR)
phecode_800-1	Chest pain on breathing	4.31 (3.35, 5.29)	5.01 (4.01, 5.89)
phecode_685-8	Hyperhidrosis	5.48 (4.17, 6.34)	5.04 (3.86, 5.78)
phecode_709-11	Mallet finger	4.46 (3.11, 6.1)	5.05 (3.94, 6.57)
phecode_469-1	Acute bronchitis	4.27 (3.76, 5.26)	5.05 (3.94, 7.08)
phecode_149	Benign neoplasm of the endocrine glands	6.04 (4.69, 7.14)	5.06 (4.04, 6.35)
phecode_239-3	Mixed hyperlipidemia	3.96 (3.14, 5.68)	5.06 (4.07, 6.93)
phecode_542-2	Fatty liver disease (FLD)	4.72 (4.08, 5.92)	5.08 (4.4, 6.74)
phecode_344-1	Hydrocephalus	3.99 (3.6, 5.07)	5.1 (4.17, 6.36)
phecode_685-82	Generalized hyperhidrosis	5.61 (4.13, 6.4)	5.11 (3.82, 5.7)
phecode_390-1	Otitis externa	3.99 (3.04, 4.45)	5.17 (4.07, 5.79)
phecode_668-6	Prurigo	3.09 (2.66, 3.57)	5.19 (3.72, 6.98)
phecode_247-51	Hypocalcemia	5.14 (4.02, 6.4)	5.25 (4.08, 6.76)
phecode_469	Bronchitis	4.63 (3.96, 5.98)	5.25 (4.31, 6.9)
phecode_168-2	Hyper-coagulability	5.29 (4.27, 7.53)	5.29 (4.35, 7.69)
phecode_685-4	Prickly heat and miliaria	5.65 (4.09, 9.25)	5.31 (3.98, 7.56)
phecode_171-1	Lymphocytosis (symptomatic)	5.59 (3.78, 6.68)	5.32 (3.63, 6.85)
phecode_486	Other respiratory disorders	4.55 (3.94, 5.28)	5.35 (4.6, 6.2)
phecode_525-1	Celiac disease	3.44 (2.89, 4.42)	5.37 (4.2, 6.41)
phecode_394-22	Vestibular neuronitis	3.89 (2.83, 5.05)	5.38 (3.12, 6.65)
phecode_665-2	Psoriatic arthropathy	4.58 (3.19, 5.62)	5.43 (3.27, 7.96)
phecode_330	Epilepsy, recurrent seizures, convulsions	5.36 (4.79, 6.96)	5.43 (4.75, 7.26)
phecode_486-21	Bronchospasm	5.03 (3.92, 6.99)	5.44 (4.54, 7.02)
phecode_138-1	Nevus, non-neoplastic	4.95 (3.58, 7.53)	5.49 (3.87, 8.52)
phecode_135-1	Benign neoplasm of the oral cavity	4.51 (3.73, 5.49)	5.57 (4.59, 7.92)
phecode_056	Human papillomavirus	5.1 (3.67, 7.5)	5.58 (3.86, 8.06)
phecode_030	Campylobacter	5.6 (4.58, 7.43)	5.59 (4.29, 6.89)
phecode_716-2	Degenerative disc disease	4.25 (3.18, 4.96)	5.61 (3.82, 6.82)
phecode_139	Benign sarcoma-related cancers	5.46 (4.01, 8.01)	5.61 (4.59, 8.31)
phecode_175	Polycythemias	4.09 (3.54, 4.76)	5.63 (4.57, 6.74)
phecode_114	Neuroendocrine tumors	4.82 (4.22, 5.78)	5.66 (4.78, 7.04)
phecode_175-2	Secondary polycythemia	4.32 (3.78, 4.86)	5.68 (4.74, 6.72)
phecode_333-4	Circadian rhythm sleep disorder	3.85 (3.05, 8.19)	5.7 (3.7, 13.08)
phecode_325-1	Abnormal involuntary movements	5.66 (4.49, 10)	5.72 (4.73, 10.14)
phecode_731	Symptoms involving musculoskeletal systems	5.42 (4.25, 8.78)	5.73 (4.55, 10.37)
phecode_730	Other disorders and symptoms of the musculoskeletal system	5.08 (3.73, 6.67)	5.74 (4.13, 7)
phecode_397-1	Tinnitus	4.97 (3.24, 8.49)	5.74 (4.55, 9.51)
phecode_596-5	Neuromuscular dysfunction of bladder	5.38 (4.35, 7.61)	5.78 (4.24, 7.93)
phecode_336-52	Meralgia paresthetica	5.62 (3.54, 7.85)	5.84 (3.68, 7.82)
phecode_209	Disorders of the pituitary gland and its hypothalamic control	5.25 (4.72, 6.02)	5.87 (4.78, 7.19)
phecode_139-3	Benign neoplasm of other connective and soft tissue	5.02 (4.1, 6.69)	5.96 (4.93, 7.62)
phecode_110	Malignant neoplasm of the endocrine glands	4.53 (3.94, 6.16)	5.97 (4.7, 8.87)
phecode_330-1	Epilepsy	5.63 (4.67, 8.69)	6.01 (4.68, 9.53)
phecode_209-1	Pituitary hyperfunction	5.43 (4.45, 6.14)	6.13 (4.59, 7.61)
phecode_397	Other hearing abnormality	4.93 (3.14, 7.59)	6.16 (4.86, 9.38)
phecode_488-6	Wheezing	5.17 (3.72, 6.52)	6.24 (5.12, 7.73)
phecode_389-1	Ocular pain	4.52 (3.62, 5.47)	6.32 (4.73, 8.05)
phecode_710-41	Flat foot [pes planus]	5.08 (3.96, 8.5)	6.33 (4.05, 9.86)
phecode_559	Other disease of digestive system, NOS	4.41 (3.66, 5.54)	6.33 (4.55, 8.09)
phecode_257	Polydipsia	5.87 (3.62, 12.35)	6.38 (4.66, 15.76)
phecode_546-3	Hepatomegaly	4.8 (4.34, 5.88)	6.4 (5.43, 8.95)
phecode_007	Hemophilus infection	5.57 (4.69, 6.81)	6.42 (5.61, 8.78)
phecode_286-1	Bipolar disorder	5.99 (4.72, 7.21)	6.51 (4.64, 8.49)
phecode_139-6	Hemangioma and lymphangioma	4.63 (3.34, 7)	6.52 (4.06, 8.89)

Supplementary Tables

Table 24 continued from previous page

Endpoint	PheCode String	unadjusted HR (IQR)	HR adjusted for Age and Sex (IQR)
phecode_522-11	Crohn's disease	6.37 (4.79, 8.81)	6.52 (4.77, 9.52)
phecode_682-11	Sebaceous cyst [Epidermal cyst]	4.79 (2.86, 13.22)	6.62 (3.53, 13.9)
phecode_486-2	Other diseases of bronchus	5.54 (4.15, 6.84)	6.66 (5.3, 7.78)
phecode_360-12	Hordeolum internum	4.58 (3.32, 5.5)	6.75 (4.64, 8.84)
phecode_007-1	Hemophilus influenzae	5.56 (4.62, 7.14)	6.75 (5.47, 9.02)
phecode_977-4	Long term (current) use of steroids	5.52 (4.15, 7.69)	6.8 (5.86, 10.31)
phecode_139-61	Hemangioma	4.69 (3.36, 7.82)	6.91 (3.97, 9.88)
phecode_977-41	Long term (current) use of inhaled steroids*	5.49 (4.33, 7.64)	6.93 (5.71, 10.07)
phecode_330-12	Partial epilepsy	6.51 (5.06, 11.03)	6.94 (5.07, 12.37)
phecode_396-1	Conductive hearing loss	6.06 (4.61, 9.33)	7.13 (5.21, 12.18)
phecode_475	Asthma	7.08 (5.69, 10.81)	7.13 (5.58, 11.12)
phecode_179-9	Immunodeficiency NOS	8.16 (5.69, 10.34)	7.14 (5.15, 9.89)
phecode_705-5	Rheumatism, unspecified	6.89 (4.4, 14.08)	7.21 (4.85, 15.39)
phecode_665-3	Other psoriasis	6.29 (4.91, 10.39)	7.34 (5.44, 11.06)
phecode_413-12	Mitral valve prolapse*	9.64 (7.23, 15.5)	7.78 (5.5, 12.08)
phecode_475-5	Exercise induced bronchospasm	6.18 (4.19, 8.07)	7.89 (5.34, 13.42)
phecode_719-11	Spasm of muscle	6.08 (3.95, 9.28)	7.95 (5.49, 12.21)
phecode_712-1	Loose body in joint	5.91 (3.6, 10.35)	8.46 (4.76, 21.97)
phecode_308-4	Demoralization and apathy	6.85 (5.54, 8.8)	9.1 (6.72, 16.11)

5 Phenome-wide prediction of disease onset from retinal fundus photographs

Table 25: Discriminative performances of retinal state addition over cardiovascular predictor sets for the cardiovascular endpoints.

PheCode String	Comparison	Median Delta C-index	Lower limit 95% CI	Upper limit 95% CI
All-Cause Death	Age+Sex vs. Age+Sex+Retina	0.01	0.005	0.015
All-Cause Death	SCORE2 vs. SCORE2+Retina	0.002	-0.002	0.006
All-Cause Death	ASCVD vs. ASCVD+Retina	0.001	-0.004	0.005
All-Cause Death	QRISK3 vs. QRISK3+Retina	-0.002	-0.006	0.002
All-Cause Death	SCORE2 vs. Age+Sex+Retina	-0.011	-0.016	-0.006
All-Cause Death	ASCVD vs. Age+Sex+Retina	-0.011	-0.017	-0.006
All-Cause Death	QRISK3 vs. Age+Sex+Retina	-0.024	-0.03	-0.02
Ischemic stroke	Age+Sex vs. Age+Sex+Retina	0.006	0	0.016
Ischemic stroke	SCORE2 vs. SCORE2+Retina	0.001	-0.005	0.01
Ischemic stroke	ASCVD vs. ASCVD+Retina	-0.004	-0.008	0.005
Ischemic stroke	QRISK3 vs. QRISK3+Retina	-0.006	-0.011	0.003
Ischemic stroke	SCORE2 vs. Age+Sex+Retina	-0.006	-0.012	0.004
Ischemic stroke	ASCVD vs. Age+Sex+Retina	-0.011	-0.016	-0.002
Ischemic stroke	QRISK3 vs. Age+Sex+Retina	-0.016	-0.022	-0.005
Heart failure	Age+Sex vs. Age+Sex+Retina	0.01	0.004	0.015
Heart failure	SCORE2 vs. SCORE2+Retina	0.003	-0.001	0.007
Heart failure	ASCVD vs. ASCVD+Retina	-0.001	-0.005	0.003
Heart failure	QRISK3 vs. QRISK3+Retina	-0.004	-0.008	-0.001
Heart failure	SCORE2 vs. Age+Sex+Retina	-0.009	-0.014	-0.004
Heart failure	ASCVD vs. Age+Sex+Retina	-0.015	-0.019	-0.01
Heart failure	QRISK3 vs. Age+Sex+Retina	-0.034	-0.039	-0.028
Ischemic heart disease	Age+Sex vs. Age+Sex+Retina	0.005	0.001	0.008
Ischemic heart disease	SCORE2 vs. SCORE2+Retina	0	-0.002	0.003
Ischemic heart disease	ASCVD vs. ASCVD+Retina	-0.001	-0.004	0.001
Ischemic heart disease	QRISK3 vs. QRISK3+Retina	-0.003	-0.005	-0.001
Ischemic heart disease	SCORE2 vs. Age+Sex+Retina	-0.015	-0.019	-0.011
Ischemic heart disease	ASCVD vs. Age+Sex+Retina	-0.022	-0.027	-0.018
Ischemic heart disease	QRISK3 vs. Age+Sex+Retina	-0.04	-0.043	-0.035
Myocardial infarction	Age+Sex vs. Age+Sex+Retina	0.007	0.001	0.011
Myocardial infarction	SCORE2 vs. SCORE2+Retina	0.001	-0.003	0.005
Myocardial infarction	ASCVD vs. ASCVD+Retina	-0.001	-0.005	0.002
Myocardial infarction	QRISK3 vs. QRISK3+Retina	-0.003	-0.006	0
Myocardial infarction	SCORE2 vs. Age+Sex+Retina	-0.023	-0.029	-0.018
Myocardial infarction	ASCVD vs. Age+Sex+Retina	-0.029	-0.036	-0.024
Myocardial infarction	QRISK3 vs. Age+Sex+Retina	-0.042	-0.05	-0.035