Pharmacometric-guided optimisation of early clinical decision-making in oncology:

Identification of drug effect in clinical drug development and prediction of patient response at bedside

Inaugural-Dissertation

to obtain the academic degree

Doctor rerum naturalium (Dr. rer. nat.)

submitted to the

Department of Biology, Chemistry, Pharmacy

of

Freie Universität Berlin

by

Yomna M. Nassar

Berlin 2024

The present thesis was conducted from July 2019 to September 2023 under the supervision of Prof. Dr. Charlotte Kloft at the Institute of Pharmacy, Freie Universität Berlin.

1st reviewer: Prof. Dr. rer. nat. habil. Charlotte Kloft

2nd reviewer: Prof. Dr. med. et rer. nat. Markus Joerger

Date of defense: 29 October 2024

Acknowledgement

I would like to express my thanks and gratitude to:

- First and foremost, ***Prof. Charlotte Kloft***, my supervisor from the Freie Universität Berlin, for allowing me to be part of her working group. Thanks for the interesting scientific and non-scientific discussions, valuable advice and ideas, generous and dedicated input and feedback as well as the support and encouragement to seize opportunities that would further foster my skills and experience.

- ***Prof. Wilhelm Huisinga***, my co-supervisor from the University of Potsdam, for his valuable input, feedback and insights on my projects, publications, and presentations—they were highly appreciated!

- ***The Graduate Research Training Program PharMetrX***, for the multidisciplinary (co-)supervision and exchange platform during the PharMetrX & Research group meetings, the training curriculum laying foundations to both the scientific principles and industry perspectives, the close link to industry and mentorship, the broad networking opportunities, and financial support during my doctoral studies. Extended thanks to ***Cornelia Böhnstedt***, the program administrator, for her support and positive spirit.

- ***The Niklas Werner Memorial Foundation***, for the financial support during the late stage of my doctoral studies.

- ***Boehringer Ingelheim Pharma GmbH & Co. KG*** and ***Dr. Alexander Staab***, for giving me the opportunity to embark on a 3-month internship at the Translational Medicine and Clinical Pharmacology Data Sciences department, to get insights into the real working environment and the daily practical life of a pharmacometrician, foster my research knowledge, expand my network, and gather new experiences. Extended thanks to ***Dr. Valerie Nock*** for making it feasible to have my internship project be part of this doctoral thesis.

- My industry mentors from Boehringer Ingelheim within the PharMetrX program:

  Dr. Alejandro Pérez-Pitarch, for his interest in my project since its early phases; his methodological support, enthusiasm, willingness to help, and most importantly for his invaluable scientific and non-scientific input, advice, and support throughout my doctoral time. Thanks for hosting me in your team during my internship time. It was a pleasure to work with you and under your supervision and I very much appreciate the time you dedicated to our discussions—Thanks for being an outstanding Mentor. I learnt a lot!

  Dr. Karin Haug, for giving me early insights into the working environment in pharmaceutical industry during her short mentorship time.

- My collaboration partners for their valuable input and advice:

  Prof. Markus Jörger from the Department of Medical Oncology and Haematology, Kantonsspital St.Gallen in Switzerland, for sharing the clinical data, conducted by the Central European Society of Anticancer Drug Research (CESAR), with us, believing in the potential of pharmacometrics in bringing better patient care and most importantly for always providing the clinical perspective and sharing his clinical expertise.

  Prof. Stefan Holdenrieder and ***Dr. Kimberly Geiger*** from the Institute of Laboratory Medicine, German Heart Centre in Munich, for providing us with the CEPAC-TDM biomarker data and for the valuable insights into the use of biomarkers in clinical practice.

- ***Prof. Gerd Mikus***, from the Department of Clinical Pharmacology and Pharmacoepidemiology, University Hospital Heidelberg, for the opportunity to work on my first project, his advice when writing my first manuscript, and his support and encouragement thereafter when writing my thesis.

- ***Dr. Sebastian Bossert*** and ***Jonas Schick***, from the Global Biostatistics & Data Sciences at Boehringer Ingelheim Pharma GmbH & Co. KG, for their interest in my internship project and offering their statistical analyses to support our investigation.

- ***Dr. Andrea Henrich*** and ***Dr. Francis Williams Ojara***, former PharMetrX PhD students, for exploring the multi-faceted aspects of the rich and dense CEPAC-TDM data. Your work definitely paved the way for a seamless succession and multiple research ideas.

- ***Dr. Francis Williams Ojara***, my academic peer-mentor, for supporting me in my early doctoral times, and for his dedication thereafter to scientific discussions and feedback.

- ***Dr. Niklas Hartung***, from the University of Potsdam, for sharing his mathematical and statistical knowledge and advice, and ***Dr. Robin Michelet***, from the Freie Universität Berlin, for his support regarding modelling and simulation principles particularly during the early phase of my doctoral studies.

- Current and former ***PharMetrX (Research+ and Training+) students*** and ***members of the working group***, with whom our path crossed, for the scientific and non-scientific discussions and activities beyond modelling and simulations.

- ***Prof. Joseph Standing*** and ***Dr. Frank Kloprogge***, from University College London, for introducing me to the field of pharmacometrics and setting me off on the road of modelling and simulation.

- ***Dr. Joachim Grevel***, ***Dr. Leonid Gibiansky***, ***Prof. Lena Friberg***, ***Dr. Ahmed Suleiman***, and ***Dr. Simon Buatois***, for being so generous to share their expertise.

- Special thanks to ***Alejandro Pérez-Pitarch***, ***Francis Williams Ojara***, ***Franz Weber***, ***Zrinka Duvnjak***, ***Miriam Happ***, ***Fenja Klima***, and ***Jonas Schick*** for sparring the time to proofread and review parts of my thesis, and provide feedback. And to ***Marian Klose*** not only for your incomparable dedicated review and feedback but also for our brainstorming discussions.

- ***Dr. Julten Abdelhalim*** and her family, for helping me settle in Berlin and find my way, and for supporting me when I needed help.
- ***My parents*** for the overseas support and encouragement, ***my sister—Aya***—for her encouragement while writing up the thesis, being my travel companion and lending a listening ear to my end-of-day talks, plannings and stories

  and

  to my ***Grandparents***, who I know are cheering me on from a different place...

Declaration of authorship

I hereby declare that I alone am responsible for the content of my doctoral dissertation and that I have only used the sources or references cited in the dissertation.

Abstract

With an estimated 19 million new cancer cases each year and almost 10 million deaths worldwide, there is a huge need for an optimised development of new therapeutic anticancer compounds as well as prediction of cancer patient response at bedside. Pharmacometric modelling and simulation allow to optimally leverage the rich longitudinal data from clinical studies and the different oncology variables and endpoints as tumour response, biomarker concentrations and survival information to better optimise clinical decision-making in oncology at its different stages: the development of new therapeutic compounds with optimised dosing selection and during therapeutic use to predict patient prognosis for improved treatment decisions at bedside.

Beginning with the development of new therapeutic compounds, identification of the appropriate/optimal dosing for the confirmatory phase III trials that maximises efficacy and minimises toxicity remains the most challenging component of clinical drug development. For this reason, a better understanding of the impact of different dose levels on efficacy and toxicity is needed to characterise the dose-response relationship and offer a more rational derivation of optimal doses/dosing. Compared to traditional pairwise comparisons between different study arms of dose-finding studies, statistical and model-based approaches have been shown to best leverage phase II dose-finding study data and characterise the dose-response relationship. These different approaches have been endorsed by the US Food and Drug Administration (FDA) which has recently initiated "Project Optimus" with the aim to optimise dose finding in oncology. Therefore, *project I* aimed to compare the performance of two model-based approaches within the oncology setting: the recently proposed combined Likelihood Ratio Test (cLRT) which leverages longitudinal phase II data and has shown high power detecting a dose-response relationship—but is computationally expensive, and the Multiple Comparison Procedure (MCP), the earlier and more established approach that has gained the FDA's qualification as "fit-for-purpose" for the design and analysis of phase II studies. A simulation-based framework of a dose-finding phase II study under different study design considerations was established and applied to investigate cLRT and MCP performance. The results showed that, in general, cLRT was associated with higher power (=1−type II error) compared with MCP (89.8% vs 27.0%); however, its type I error (i.e. false positive) was not well controlled (mean: 13%) compared with MCP (<4%). Moreover, cLRT power was less sensitive to the different study design variables (e.g. number of patient with respect to number of dose levels) in contrast with MCP. Therefore, based on these results, before cLRT can be recommended to analyse dose-finding studies in oncology, further investigation of its robustness to different model complexities and study design variables as well as investigation of conditions that would better control type I error are needed to justify its high power at the expense of its computational demands.

Equally important to the development of new therapeutic compounds with optimised dose selection, is the accurate and early prediction of patient prognosis to monitor patient response and improve treatment decisions at bedside. Non-small cell lung cancer (NSCLC), the leading cause of cancer-related death, represents a disease of high burden and poor prognosis. Therefore, *project II* was conducted with the aim to identify early predictors of efficacy for NSCLC patients to spare them the unnecessary exposure to toxicities and contribute to better prediction of treatment outcomes. Clinical data from patients with advanced NSCLC receiving first-line combination therapy with paclitaxel and a platinum-based drug, were leveraged to characterise and quantify the relationships between anticancer drug exposure, tumour dynamics and C-reactive protein (CRP) concentrations—as a measure of the inflammatory level, using pharmacometric modelling. Model-derived variables were then investigated as potential predictors of the most important and commonly adopted efficacy endpoints progression-free survival (PFS) and overall survival (OS) by means of parametric time-to-event modelling, with a special focus on the potential of early longitudinal biomarker information as a potential early prognostic predictor. The results of our modelling framework in which longitudinal CRP concentrations were leveraged for the first time for prognostic investigations, identified the inflammatory level at treatment cycle 3 (CRP_{cycle3}), i.e. day 42 from start of treatment, and the extent of absolute reduction in the inflammatory level between treatment cycles 3 and 2 to be the most significant predictors of PFS. Besides the CRP-related metrics, baseline tumour size and presence/absence of liver lesions were found to be predictors of OS. Nevertheless, CRP_{cycle3} was by far of the highest impact. The identification of CRP at treatment cycle 3 points to the potential and more informative value of longitudinal biomarker data compared to the commonly applied approach in which only baseline (pre-treatment) measurements are investigated and which do not reflect the patient situation and dynamic evolution of the disease. Measuring longitudinal CRP as a routine biomarker allows for the monitoring of inflammatory levels and, along with its reduction across treatment cycles, presents a promising prognostic marker for the timely identification of patients at risk of therapeutic failure, early progression and/or short survival to spare them unnecessary toxicities and provide alternative treatment decisions.

Overall, the two projects presented in this thesis, acknowledged and addressed, by leveraging pharmacometrics methodologies, two critical needs within the scope of oncology drug development and therapeutics that require early and more optimised clinical decisions. First, was the need to optimally identify the drug effect for an accelerated and optimised development of efficacious compounds for oncologic indications. Based on a scientific understanding of the dose-response relationship during early clinical drug development, a more informed and optimised dosing selection can be achieved. Within the investigated approaches, MCP, a robust but less powered approach currently exists. However, for better power, systematic investigations of cLRT are needed to achieve a more robust performance. Through the establishment of a proper understanding of the dose-response relationship, decisions regarding optimal dosing selection can be better informed that would maximise the drug's efficacy, safety and tolerability and would consequently reflect on more successful phase III trials and lower attrition rates. Second, was the need to accurately and timely predict treatment outcomes and prognosis to monitor patient response, and improve treatment decisions at bedside. Our developed modelling framework, successfully identified the minimally invasive and cost-effective biomarker, CRP_{cycle3}, as a significant predictor of PFS and OS. It also proposed that monitoring the inflammatory level was of prognostic value. The application of this modelling framework goes beyond NSCLC and is envisioned to suit other treatment modalities such

as immunotherapies or targeted therapies for prediction of patient response and treatment outcomes in different settings. Finally, with use of proper methodologies as pharmacometric modelling and simulation, different data were successfully leveraged to optimise dosing selection and patient monitoring for a pharmacometric-based optimisation of early clinical decision-making in oncology.

Zusammenfassung

Angesichts von geschätzten 19 Millionen neuen Krebserkrankungen und fast 10 Millionen Todesfällen pro Jahr weltweit besteht ein enormer Bedarf an einer optimierten Entwicklung neuer Antitumortherapeutika sowie an einer verbesserten Vorhersage des Therapieerfolgs im klinischen Alltag. Pharmakometrische Modellierung und Simulation ermöglicht es, umfangreiche longitudinale Daten aus klinischen Studien sowie verschiedene onkologische Messwerte und Endpunkte wie Tumoransprechen, Biomarkerkonzentrationen und Überlebenszeit optimal zu nutzen, um im klinischen Kontext die Entscheidungsfindung in der Onkologie zu verschiedenen Zeitpunkten zu verbessern: sowohl während der Entwicklung neuer Arzneistoffe zur Auswahl einer optimalen Dosierung ebenso wie in der therapeutischen Praxis zur Vorhersage individueller Prognosen zur Verbesserung von Therapieentscheidungen.

In der klinischen Arzneimittelentwicklung ist die größte Herausforderung nach wie vor die Auswahl einer geeigneten/optimalen Dosierung mit maximaler Wirksamkeit und minimaler Toxizität für konformatorische Phase-III-Studien. Um die Dosis-Wirkungs-Beziehung zu charakterisieren und eine rationalere Dosisfindung zu ermöglichen, ist ein verbessertes Verständnis des Einflusses verschiedener Dosierungen auf Wirksamkeit und Toxizität erforderlich. Im Vergleich zum herkömmlichen paarweisen Vergleich verschiedener Arme von Dosisfindungsstudien sind statistische und modellbasierte Ansätze besonders geeignet, um Daten aus Phase-II-Dosisfindungsstudien optimal zu nutzen sowie Dosis-Wirkungs-Beziehung zu beschreiben. Beide Ansätze werden von der US-amerikanischen Food and Drug Administration (FDA) unterstützt, die kürzlich die Initiative "Project Optimus" zur Optimierung der Dosisfindung in der Onkologie angestoßen hat. *Projekt I* hatte daher zum Ziel, zwei modellbasierte Ansätzen in der Onkologie zu vergleichen: der neuere aber rechenintensive *kombinierte Likelihood-Ratio-Test* (cLRT), der longitudinale Phase-II-Daten nutzt und eine hohe Power bei der Erkennung einer Dosis-Wirkungs-Beziehung aufweist, und die etabliertere *Multiple-Comparison-Procedure* (MCP), die von der FDA als zweckmäßig ("fit-for-purpose") für Design und Analyse von Phase-II-Studien eingestuft wurde. Um die Leistungsfähigkeit von cLRT und MCP zu untersuchen, wurde ein Ansatz zur Simulation einer Phase-II-Studie entwickelt und angewandt, der verschiedene Aspekte von klinischen Studiendesigns berücksichtigt. Im Ergebnis zeigte cLRT insgesamt eine höhere Power (=1−Typ-II-Fehler) im Vergleich zu MCP (89.8 % vs 27.0 %); allerdings war der Typ-I-Fehler (d.h. falsch-positiv) weniger gut kontrolliert (Mittelwert: 13 %) als bei MCP (<4 %). Außerdem reagierte die Power von cLRT im Gegensatz zur Power von MCP weniger sensitiv auf verschiedene Variablen des Studiendesigns (z.B. Anzahl der Patienten im Verhältnis zur Anzahl der Dosisstufen). Bevor cLRT daher für die Analyse von Dosisfindungsstudien in der Onkologie empfohlen werden kann, muss er weiterhin auf seine Robustheit gegenüber unterschiedlich komplexen Modellen und verschiedenen Studiendesignvariablen untersucht werden sowie auf Bedingungen, die eine bessere Kontrolle des Typ-I-Fehlers ermöglichen, um die hohe Power auf Kosten des Rechenaufwands zu rechtfertigen.

Ebenso wichtig wie die Auswahl einer optimalen Dosierung bei der Entwicklung neuer Arzneistoffe ist die genaue und frühzeitige Vorhersage von Patientenprognosen in der klinischen Praxis, um das Ansprechen auf eine Therapie zu überwachen und Therapieentscheidungen zu verbessern. Nichtkleinzelliges Lungenkarzinom (NSCLC), die häufigste krebsbedingte Todesursache, ist durch hohe Krankheitslast und schlechte Prognose gekennzeichnet. *Projekt II* hatte daher zum Ziel, frühe Prädiktoren für das Therapieansprechen von NSCLC-Patienten zu identifizieren, um unnötige Toxizitäten zu vermeiden und die Vorhersage des Therapieerfolgs zu verbessern. Auf der Grundlage klinischer Daten von Patienten mit fortgeschrittenem NSCLC, die Paclitaxel und ein platinbasiertes Zytostatikum als Erstlinien-Kombinationstherapie erhielten, wurden die Zusammenhänge zwischen Arzneimittelexposition, Tumordynamik und Konzentration von C-reaktivem Protein (CRP) - als Maß für das Entzündungsniveau - mithilfe pharmakometrischer Modellierung charakterisiert und quantifiziert. Anschließend wurden aus dem Modell abgeleitete Variablen als potenzielle Prädiktoren für die beiden wichtigsten und häufigsten Wirksamkeitsendpunkte *progressionsfreies Überleben* (PFS) und *Gesamtüberleben* (OS) mittels parametrischer Zeit-zu-Ereignis-Modellierung (Time-to-event-Modellierung) untersucht, wobei ein besonderer Schwerpunkt auf dem Potenzial *früher* longitudinaler Biomarkerdaten lag. Unser Modellierungsansatz, bei dem longitudinale CRP-Konzentrationen erstmals für prognostische Untersuchungen genutzt wurden, zeigte, dass die wichtigsten Prädiktoren für PFS das Entzündungsniveau im dritten Behandlungszyklus ($CRP_{Zyklus3}$), d.h. an Tag 42 nach Behandlungsbeginn, und die absolute Verringerung des Entzündungsniveaus zwischen den Behandlungszyklen 3 und 2 waren. Neben diesen CRP-bezogenen Parametern wurden auch die Ausgangsgröße des Tumors und das Vorhandensein bzw. Fehlen von Läsionen in der Leber als Prädiktoren für OS identifiziert. Den bei weitem größten Einfluss zeigte jedoch $CRP_{Zyklus3}$. Die Identifizierung von CRP im dritten Behandlungszyklus als Prädiktor weist auf den potenziellen und aussagekräftigeren Wert von longitudinalen Biomarkerdaten hin, im Vergleich zur in der Klinik üblichen Auswertung von ausschließlich Basislinienmessungen (vor der Behandlung), die die Patientensituation und die dynamische Entwicklung der Krankheit nicht widerspiegeln. Die Messung von CRP im Langzeitverlauf als Routine-Biomarker ermöglicht die Überwachung der Entzündungswerte und ist zusammen mit Verringerung von CRP über Behandlungszyklen hinweg ein vielversprechender prognostischer Marker. Er ermöglicht die frühzeitige Identifizierung von Patienten mit erhöhtem Risiko für Therapieversagen, frühem Progress und/oder einer kurzen Überlebensdauer, und hilft, unnötige Toxizitäten zu vermeiden und alternative Therapieentscheidungen zu treffen.

Insgesamt wurden in den beiden in dieser Arbeit vorgestellten Projekten durch Nutzung pharmakometrischer Methoden zwei kritische Aspekte im Bereich der onkologischen Arzneimittelentwicklung und klinischen Praxis identifiziert und adressiert, die frühzeitige und optimierte Entscheidungsfindung erfordern. Um im ersten Schritt die Arzneimittelentwicklung in der Onkologie zu beschleunigen und zu verbessern, ist es notwendig, die optimale Wirkung eines Arzneimittels zu ermitteln. Auf Basis eines Verständnisses für die Dosis-Wirkungs-Beziehung während der frühen klinischen Arzneimittelentwicklung kann eine fundierte und optimale Dosierungsauswahl getroffen werden. Unter den in dieser Arbeit untersuchten Ansätzen steht mit MCP hierfür eine robuste, aber weniger leistungsfähige Methode zur Verfügung. Für cLRT sind weitere systematische Untersuchungen zur Erhöhung der Power erforderlich, um eine robustere Leistung zu erzielen. Entscheidungen über eine optimale Dosierung mit maximaler Wirksamkeit, Sicherheit und Verträglichkeit können durch ein gutes Verständnis der Dosis-Wirkungs-Beziehung

stark verbessert werden, was sich ebenso in höheren Erfolgsquoten von Phase-III-Studien und niedrigeren Abbruchquoten widerspiegeln würde. Um im zweiten Schritt in der klinischen Praxis das Ansprechen von Patienten zu überwachen und Therapieentscheidungen zu verbessern, ist eine akkurate und frühzeitige Vorhersage von Therapieerfolg und Prognose entscheidend. Mithilfe des entwickelten Modellierungsansatzes konnten wir erfolgreich den minimalinvasiven und kosteneffizienten Biomarker $CRP_{Zyklus3}$ als signifikanten Prädiktor für PFS und OS identifizieren. Zudem zeigte sich, dass Monitoring des Entzündungsniveaus von prognostischem Wert ist. Die Anwendung dieses Modellierungsansatzes geht über NSCLC hinaus und ist prinzipiell auch für andere therapeutische Modalitäten wie Immuntherapien oder zielgerichtete Arzneimittel zur Vorhersage des Ansprechens von Patienten und des Therapieerfolgs in verschiedenen Bereichen geeignet. Insgesamt erlaubte die Anwendung geeigneter Methoden wie der pharmakometrischen Modellierung und Simulation die erfolgreiche Auswertung verschiedener Arten von Daten, um die Dosierungsauswahl sowie das Patientenmonitoring zu verbessern und so die modell-basierte frühzeitige Entscheidungsfindung in der Onkologie zu optimieren.

Relevant publications

Major parts of the presented doctoral thesis have been published or are in the process of being published in the following research articles:

Project II:

- F.W. Ojara[#], **Y.M. Nassar**[#], K. Geiger, M. Joerger, M. van Dyk, S. Holdenrieder, C. Kloft.
  Prognostic impact of serum biomarkers for decision-making in the treatment of advanced non-small cell lung cancer patients: Review of existing approaches for modelling serum biomarker data and potential role for pharmacometrics
  [#]shared first authorship
  In preparation

- **Y.M. Nassar**, F.W. Ojara, A. Pérez-Pitarch, K. Geiger, W. Huisinga, N. Hartung, R. Michelet, S. Holdenrieder, M. Joerger*, C. Kloft*.
  C-reactive protein as an early predictor of efficacy in advanced non-small-cell lung cancer patients: A tumor dynamics-biomarker modeling framework.
  Cancers 15: 5429 (2023).
  *shared senior authorship
  doi: 10.3390/cancers15225429

- F.W. Ojara, A. Henrich, N. Frances, **Y.M. Nassar**, W. Huisinga, N. Hartung, K. Geiger, S. Holdenrieder, M. Joerger, C. Kloft.
  A prognostic baseline serum biomarker and tumour growth kinetics integrated model in paclitaxel/platinum treated advanced non-small cell lung cancer patients.
  CPT Pharmacometrics Syst. Pharmacol. 12: 1714–1725 (2023).
  doi: 10.1002/psp4.12937

Further publications during the doctoral time but outside the topic of this thesis

- **Y.M. Nassar**[#], N. Hohmann[#], R. Michelet, K. Gottwalt, A.D. Meid, J. Burhenne, W. Huisinga, W.E Haefeli, G. Mikus*, C. Kloft*.
  Quantification of the time course of CYP3A inhibition, activation, and induction using a population pharmacokinetic model of microdosed midazolam continuous infusion.
  Clin. Pharmacokinet. 61: 1595–1607 (2022).
  [#]shared first authorship; *shared senior authorship
  doi: 10.1007/s40262-022-01175-6

Additional publications (e.g. conference abstracts) are presented in the **Publications** section.

Table of Contents

List of Tables

List of Figures

List of abbreviations and symbols

Abbreviations

AIC	Akaike information criterion
ALT	Alanine aminotransferase
AST	Aspartate aminotransferase
AUC	Area under plasma concentration-time curve
B1	Baseline method 1
B2	Baseline method 2
B3	Baseline method 3
B4	Baseline method 4
BLTS	Baseline tumour size
BMI	Body mass index
BSA	Body surface area
CA	Cancer antigen
CE-IVD	Conformité Européenne *in vitro* diagnostic
CEA	Carcinoembryonic antigen
CEPAC-TDM	CESAR Study of Paclitaxel Therapeutic Drug Monitoring
CESAR	Central European Society of Anticancer Drug Research
CI	Confidence interval
cLRT-Mod	Combined Likelihood Ratio Test-Modelling
CoxPH	Cox proportional hazards model
CR	Complete response
CRP	C-reactive protein (concentration)
CT	Computed tomography
CV	Coefficient of variation
CWRES	Conditional weighted residuals

CYFRA 21-1	Cytokeratin-Fragment 19
DV	Dependent variable
EBE	Empirical Bayes Estimate
ECOG	Eastern Cooperative Oncology Group
EMA	European Medicines Agency
EOT	End-of-treatment
EVID	Event identifier
FDA	Food and Drug Administration
FO	First-order
FOCE	First-order conditional estimation
FOCE-I	First-order conditional estimation with interaction
G-CSF	Granulocyte-colony stimulating factor
GFR	Glomerular filtration rate
gMCP-Mod	Generalised Multiple Comparison Procedure-Modelling
GOF	Goodness-of-fit
HPC	High-performance computing
HR	Hazard ratio
ICH	International Council for Harmonisation
ID	Individual
IIV	Interindividual variability
IL-6	Interleukin 6
IPRED	Individual prediction
IQR	Interquartile range
KM VPC	Kaplan-Meier visual predictive check
LCC	Large cell carcinoma
LLOQ	Lower limit of quantification
LRT	Likelihood ratio test
MCP-Mod	Multiple Comparison Procedure-Modelling
MDV	Missing dependent variable
MTD	Maximum tolerated dose

NCI-CTCAE	National Cancer Institute Common Terminology Criteria for Adverse Events
NLME	Nonlinear mixed-effects
NONMEM	Nonlinear mixed-effects modelling
NSCLC	Non-small cell lung cancer
NSE	Neuron-specific enolase
OFV	Objective function value
OS	Overall survival
PAGE	Population Approach Group Europe
pCRP	Pentameric C-reactive protein
PD	Pharmacodynamic(s)
PD	Progressive disease
PDF	Probability density function
PET-CT	Positron emission tomography-computed tomography
PFS	Progression-free survival
PK	Pharmacokinetic(s)
PR	Partial response
PRED	Population prediction
PsN	Perl-speaks-NONMEM
Q1	First quartile
Q2	Median
Q3	Third quartile
RECIST	Response Evaluation Criteria in Solid Tumours
RSE	Relative standard error
RUV	Residual unexplained variability
SCC	Squamous cell carcinoma
SCLC	Small cell lung cancer
SCM	Stepwise covariate model
SD	Stable disease
SD	Standard deviation
SE	Standard error

SLD	Sum of longest diameters
TS	Tumour size
TTE	Time-to-event
TTP	Time-to-progression
VPC	Visual predictive check

Symbols

α	Level of significance [statistical context]
α	Shape parameter of Weibull or log-logistic hazard functions [hazard function context]
AUC_{cycle}	Paclitaxel area under plasma concentration-time curve from start to end of a cycle based on a single paclitaxel dose administered on the first day of a 21-day cycle
$\tilde{b}$	Calculated variance associated with $\tilde{\beta}$ [multiple imputation context]
$\hat{b}$	Estimate of b from imputed datasets [multiple imputation context]
β	Parameter in Θ, Ω or Σ [multiple imputation context]
$\tilde{\beta}$	Calculated estimate of β [multiple imputation context]
$\hat{\beta}$	Estimate of β from imputed datasets [multiple imputation context]
β_o	Drug-induced tumour decay rate constant per unit of paclitaxel AUC_{cycle} at time zero
$\beta_1, ..., \beta_n$	Coefficients measuring effect size of each explanatory variable in the Cox proportional hazards model
cov	Covariate
$CovBase$	Baseline of covariate
$CovDiff$	Change from baseline of covariate
Cov_{median}	Median value of covariate
CRP_{cycle1}	C-reactive protein concentration at treatment cycle 1
CRP_{cycle2}	C-reactive protein concentration at treatment cycle 2
CRP_{cycle3}	C-reactive protein concentration at treatment cycle 3
$CRP_{cycle3-2}$	Change in C-reactive protein concentration between treatment cycle 3 and 2
CRP_{ss}	Baseline C-reactive protein steady-state concentration
Δ	Difference/Delta
d_i	Total number of observed events in all groups [log-rank text context]
df	Degree(s) of freedom
ε	Random-effect parameter quantifying residual unexplained variability
ε_{add}	Additive ε
ε_{prop}	Proportional ε
η	Random-effects parameter quantifying interindividual variability
$\eta_{BL,i}$	Interindividual variability on baseline value of i^{th} individual

η_{cov}	Interindividual variability for parameter-covariate relationship
$\eta_{i,k}$	Interindividual variability for k^{th} parameter of i^{th} individual
η_{par_i}	Interindividual variability for individual parameter estimate par_i [multiple imputation context]
$\tilde{\eta}_{par_i}$	Calculated mean individual random effects [multiple imputation context]
$\hat{\eta}_{par_i}$	Estimate of η_{par_i} from imputed datasets [multiple imputation context]
$\tilde{\eta}_{var_i}$	Calculated variance associated with individual tumour dynamics parameters [multiple imputation context]
$\hat{\eta}_{var_i}$	Estimate of individual variance associated with individual random effects [multiple imputation context]
E_0	Baseline drug effect
E_1	Total number of expected events in group 1 [log-rank text context]
ED_{50}	Dose at which drug effect is 50%
E_{max}	Maximum drug effect
f	Function for estimating individual predictions
γ	Shape parameter of Gompertz hazard function
$h(t)$	Hazard function/Modified hazard after inclusion of predictors
$h_o(t)$	Baseline hazard function
$Hill$	Hill (sigmoidicity) coefficient
$H(t)$	Cumulative hazard function
k_{growth}	Linear tumour growth rate constant
K_{in}	Biomarker zero-order production rate constant
$K_{in,CRP}$	C-reactive protein zero-order production rate constant
$K_{in,CRP,basal}$	Basal unperturbed C-reactive protein zero-order production rate constant
$K_{in,CRP,TS}$	Tumour size-dependent C-reactive protein production rate constant
K_{out}	Biomarker first-order degradation rate constant
$K_{out,CRP}$	C-reactive protein first-order degradation rate constant
λ	Rate constant for exponential change in drug effect over time [tumour dynamics context]
λ	Scale parameter of different hazard functions [hazard function context]
λ_o	Baseline scale parameter of different hazard functions

$-2\mathcal{LL}$	-2 · log likelihood
$\mathcal{L}$	Likelihood
μ	Mean parameter of log-normal hazard function
m	Number of replicates γ=1,...,m [multiple imputation context]
N	Normal distribution
N	Number of simulations [project I context]
N/L	Neutrophil-to-lymphocyte ratio
$(N/L)_{cycle1}$	Neutrophil-to-lymphocyte ratio at treatment cycle 1
$(N/L)_{cycle2}$	Neutrophil-to-lymphocyte ratio at treatment cycle 2
n_p	Penalty term accounting for total number (n) of estimated parameters (p) in Akaike information criterion
Ω	Variance-covariance matrix of interindividual variability η
ω^2	Variance of η
ω_k^2	Variance of η of k^{th} parameter
O_1	Total number of observed events in group 1 [log-rank test context]
$OFV_{best-fitting}$	Objective function value of best-fitting dose-response model
OFV_{flat}	Objective function value of flat dose-response model
ϕ	Normal cumulative distribution function
$\phi_{BL,i}$	Individual baseline
ϕ_i	Vector of structural model parameters of i^{th} individual
P	Probability
Par_i	Individual parameter estimate
$q_{\alpha,cLRT}$	Critical value at $\alpha = 0.05$ for cLRT
ρ	Correlation coefficient between two parameters in Ω
r	Pearson correlation coefficient
r_i	Total number of individuals at risk [log-rank text context]
$RSE_{\omega_k^2,CV\%}$	Relative standard error of interindividual variability of k^{th} parameter reported as CV%
RSE_{θ_k}	Relative standard error of fixed-effect parameter θ_k
σ	Standard deviation parameter of log-normal hazard function
σ^2	Variance of elements of random effects ε

Σ	Variance-covariance matrix of elements of random effects ε
$SE_{\omega_k^2}$	Standard error of random-effect parameter ω_k^2
SE_{θ_k}	Standard error of fixed-effect parameter θ_k
slope	Slope of linear function
$S(t)$	Survival function
θ	Fixed-effects parameter i.e. typical population parameter estimate
θ_{BL}	Typical baseline
θ_{BLTS}	Exponent reflecting change in natural log $K_{in,CRP}$ per unit change in baseline tumour size from median value of 8.25 cm
θ_{BLTS}	Parameter quantifying effect of baseline tumour size on hazard [Time-to-event modelling context]
θ_{cov}	Parameter estimate of covariate effect on model parameter for a given functional form
$\theta_{CRP_{cycle3}}$	Parameter quantifying effect of CRP_{cycle3} on hazard
$\theta_{CRP_{cycle3-2}}$	Parameter quantifying effect of $CRP_{cycle3-cycle2}$ on hazard
θ_{IL6}	Fractional change in CRP production rate constant per unit change in baseline IL-6 from median value of 2.57 pg/mL
$\theta_{i,k}$	Individual parameter estimate of k^{th} parameter for i^{th} individual
θ_k	Typical population parameter estimate of k^{th} parameter
$\theta_{k,cov}$	Typical population parameter estimate of k^{th} parameter for a given covariate (*cov*) value
$\theta_{Kin,CRP}$	CRP production rate constant of a non-smoker with disease stage IV, baseline IL-6 of 2.57 pg/mL, and baseline tumour size of 8.25 cm (median values)
θ_{liver}	Parameter quantifying effect of presence of liver lesions on hazard
θ_{RUV}	Residual variability on a standard deviation scale
θ_{SMK2}	Fractional change in $K_{in,CRP}$ in former smokers compared with non-smokers
θ_{SMK3}	Fractional change in $K_{in,CRP}$ in current smokers compared with non-smokers
θ_{stage}	Fractional change in $K_{in,CRP}$ in patients with disease stage IIIB compared with patients with disease stage IV
t	Time
T	Survival time of individual beyond time t
$\mathrm{T}_{c>0.05}$	Time of paclitaxel plasma concentration > 0.05 µmol/L in respective treatment cycle

t_{cLRT}	Test statistic for cLRT
$t_{cLRT_{H0}}$	Test statistic under null hypothesis for cLRT
TS_{week7}	Tumour size at week 7
χ^2	Chi-square distribution
$\bar{x}$	Sample mean
$x_1, ..., x_n$	Explanatory variables in Cox proportional hazards model
x_{ij}	Vector of independent study design variables (e.g. time) for i^{th} individual at j^{th} observation time
Y_{ij}	Measured observation for i^{th} individual at j^{th} observation time
$\hat{Y}_{ij}$	Predicted observation for i^{th} individual at j^{th} observation time

1 | Introduction

Worldwide, there is an estimated 19 million new cancer cases each year, and, with almost 10 million deaths expected, cancer ranks as the second leading cause of deaths globally [1]. With such a disease burden of increasing incidence and mortality, optimised clinical/therapeutic decisions are needed when it comes to the development of new and efficacious medications with optimally selected dosing and the timely monitoring of patients response for the prediction of anticancer treatment success at bedside.

1.1 The need for optimised early clinical decision-making in oncology

Oncology remains a therapeutic area with a huge need for optimised early clinical decision-making at its different stages starting from the process of clinical drug development for better selection of dosing, to the bedside application for patient monitoring and prediction of treatment response.

The need for optimised early clinical decision-making in oncologic clinical drug development stemmed from the selected/approved but often poorly characterised doses for therapeutic compounds. Doses (dosing) are often selected without a thorough understanding of their effect and tolerability leading to unnecessary toxicities with no additional efficacy. This has reflected on the need for post-marketing requirement or post-marketing commitment for further dose optimisation in almost 15.2% of the US Food and Drug Administration (FDA)-approved small molecules and antibody-drug conjugates for oncologic indications between 2019–2021 [2]. Therefore, there is a need for optimised dose selection at the early stages of clinical drug development. A first step to optimise the dose, is the identification of a dose-response relationship indicative of a present drug effect to support a broader understanding of the impact of the different doses on efficacy and toxicity, and identify the best dose for further investigation (Fig. 1.1, left side). This would in return (i) save time, effort, and money; (ii) increase chances of successful trials; (iii) prevent exposure of patients to more toxic or less efficacious doses; and (iv) would reflect on faster approvals of promising therapies with reliably selected doses [3, 4].

Of equal need for optimised early clinical decision-making is the prediction of cancer patients treatment response at bedside. Unfortunately, the majority of cancer patients are still at risk of poor prognosis given their late diagnosis, limited treatment options and/or disease aggressiveness. Therefore, early prediction of patient outcome in response to treatment is crucial. Identification of predictors of treatment efficacy are needed to support timely and individualised treatment decisions, and to optimise therapies for patients with shorter survival. This would then spare them the unnecessary exposure to inefficacious treatments and toxicities and ensure that the best intervention is directed to the patient expected to benefit most from it (Fig. 1.1, right side).

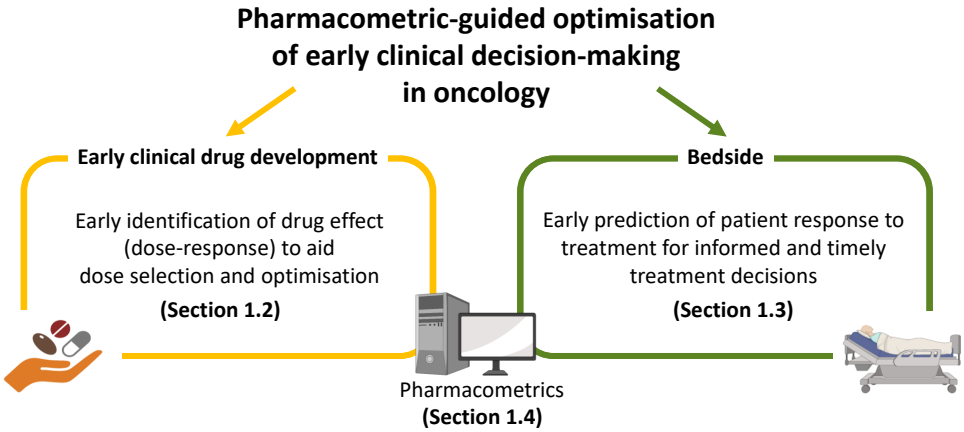

Figure 1.1: Pharmacometric-guided optimisation of early clinical decision-making in oncology in early clinical drug development (left) and at bedside (right) through the use of pharmacometrics (middle)

Therefore, through this doctoral thesis we aimed to address two major critical needs within the scope of oncology: (i) the need for a reliable identification of drug effect for an optimised dose selection during clinical development by supporting early identification of dose-response relationship (Fig. 1.1, left side; section 1.2) and (ii) the need for an early prediction of patient response for timely intervention and better clinical decision-making at bedside by supporting early prediction of prognosis and efficacy endpoints (Fig. 1.1, right side; section 1.3). To best address these needs, the use of pharmacometrics (Fig. 1.1, middle; section 1.4) offered the optimal analysis methodology. Pharmacometrics allowed to analyse and leverage existing oncology data at these different stages and to better answer the challenges in oncological clinical drug development and at bedside, for a pharmacometric-based optimisation of early clinical-decision making in oncology.

1.2 Clinical decision-making in oncology in early clinical drug development

The development of oncologic compounds represents a special area for clinical drug development due to the tolerability and toxicity characterisitcs of anticancer treatments, the therapeutic window and dose-response relationships of the new treatment modalities, and the enrolment of patients, rather than healthy participants, since the early phases of the clinical drug development program. Consequently, dose selection is still suboptimal.

1.2.1 The need for dose optimisation in oncology

The current paradigm for dose selection in oncologic drug development is viewed as non-ideal and suboptimal. The reason for this is that the maximum tolerated dose (MTD) is still conventionally used without further optimisation. The MTD is defined from phase I trials based on escalating drug doses commonly using the 3+3 design [5] until non-tolerated toxicities of predefined rate and grade (i.e. dose-limiting toxicity) are observed. This approach has been historically based on the clinical development of cytotoxic drugs whose higher doses, given their steep dose-response, indicated maximal therapeutic effect at the expense of toxicity. However, this no longer holds true with the introduction

of new treatment modalities as targeted therapies and immunotherapies. These therapies require longer treatment duration and hence, evaluation of their long-term tolerability. They also have a different dose-response relationship and are more target-selective compared to the cytotoxic drugs. Therefore, doses below the MTD may be equally effective and/or the MTD of these compounds may not even be reached [6]. Thus, the continued adoption of the MTD approach does not always result in the optimal dosing and could often result in the inadequate characterisation of doses.

An inappropriately selected high dose is intolerable for a major proportion of cancer patients. It carries the risk of toxicity or adverse events and ultimately requires the need for dose reductions or treatment cessation [6]. Out of the FDA initial approvals for oncology-related indications within 2011–2015 and 2019–2021, approximately one half and one third of the approvals had the MTD as the recommended dose, respectively [2, 7]. Moreover, uncertainties around dose selection have been regarded as the primary and most frequent reason (15.9%) for delayed or rejected marketing applications by the FDA for new molecular entities from 2000 to 2012 [3]. For this reason, in 2021, the FDA initiated "Project Optimus" with the aim to optimise dose finding and selection in oncologic drug development [8].

An optimised dose selection through dose-finding studies, will not only maximise efficacy but also minimise the toxicity of the drug and improve the quality of life of the cancer patients with the aim to bring the right dose to the patients in need.

1.2.2 Dose-finding studies in oncology

Dose-finding studies aim to support a broader understanding of the impact of different doses on efficacy and toxicity. Although these studies are common outside of oncology, nevertheless, within oncologic drug development, they are designed to determine the MTD (i.e. as phase I first-in-human trial). This dose level is then directly investigated in confirmatory trials with no further opportunity for dose optimisation or additional dose-ranging studies [7]. This was reflected in the 57 FDA approvals reported between 2015 and 2017, where 43% of these approvals proceeded directly to confirmatory trials after the first-in-human trials with no further dose optimisation [9]. Moreover, 26% of the 81 anticancer drug combinations approved between 2011 and 2021 lacked a dose-finding study [10]. Hence, an intermediate step for dose-ranging (finding) studies is needed to better understand the drug's efficacy and toxicity across the different doses.

The International Council on Harmonisation (ICH) E4 guideline on dose response information to support drug registration [11] has recommended that dose selection in confirmatory phase III trials should be based on the *observed dose-response relationship* obtained from earlier phase I and II studies. It also recommends that studies investigating a dose-response should focus on characterising the *relationship* rather than pairwise comparisons between dose groups while allowing enough time for the effect to be achieved. Therefore, dose-finding studies should (i) evaluate and compare a range of doses, ideally in a randomised trial, and (ii) select the dose for the confirmatory trials based on the characterised dose-response relationship with the aim of finding the lowest dose beyond which no further benefit is expected and/or the selection of several doses or dose-range for further evaluation, based on pharmacological, tolerability, and safety information [11].

Dose-finding studies are often designed using few dose levels and a narrow dose range and hence, lack the focus on the proper characterisation of the dose-response relationship and the consequent selection of the most appropriate dose [9, 12, 13]. In the recently issued FDA draft guidance on optimising the

dosage of human prescription drugs and biological products for the treatment of oncologic diseases [6], the FDA has recommended that for an optimal dose-finding study, a randomised parallel dose-response study design should be adopted in which multiple doses should be investigated—with the European Medicines Agency (EMA) recommending a minimum of four doses over an at least 10-fold dose range (Fig. 1.2) [13]. In this design, patients are randomly assigned to one of the arms e.g. fixed dose groups, placebo, or active comparator for a time long enough to allow a dose-response comparison and offer an evidence-based way to compare safety and efficacy outcomes (Fig. 1.2) [2]. When it comes to the analysis, the multiple groups in these studies are commonly statistically compared via pairwise comparisons and the best dose is selected for the confirmatory phase III trial [11]. Although this approach is simple; however, it is suboptimal. It requires a large number of individuals in each arm to control false positive results (i.e. type I error), which within a narrow dose range, may reflect on a few number of doses. Moreover, it has limited statistical power, does not efficiently use all the available information, and lacks any assumptions about the underlying dose-response relationship [14]—in contrast to what has been previously recommended by the ICH E4 guideline [11].

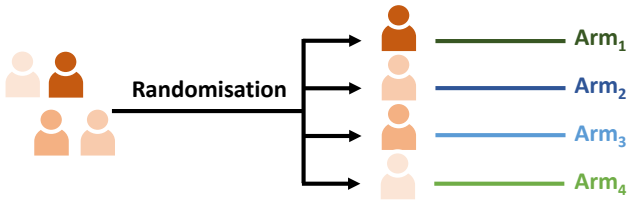

Figure 1.2: General design of a four-arm, randomised, parallel, dose-response study to investigate multiple dose levels. Arms could represent fixed dose groups, placebo, or active comparator.

Therefore, to overcome the challenges associated with the study design and handling of data in typical dose-finding studies, one of the goals of "Project Optimus" has been the focus on strategies and innovative approaches for dose-finding that could better leverage the data. This can be achieved by the use of more methodological and statistical tools (discussed in section 1.4.2.1) which have been shown to significantly reduce the required sample size when it comes to study design and increase accuracy in dose selection when compared with pairwise testing [9] and have additionally been welcomed by the EMA to better design the dose-finding studies [13].

1.3 Clinical decision-making in oncology at bedside

1.3.1 The need for early prediction of response: Lung cancer as case example

Unfortunately, not all cancer types have an equal chance for early diagnosis. Some cancer types such as lung cancer are still less likely to be diagnosed early [15]. Late diagnosis for lung cancer has topped other cancer types such as breast cancer, colorectal cancer, and prostate cancer across Europe [16], Germany [17], Switzerland [18], Australia [19], Canada [20], the United Kingdom [21], and the United States [22]. This late diagnosis reflects on poor long-term survival and consequently develops the need for easily measured predictors to identify early on patients with poor prognosis, for better clinical decision-making when it comes to treatment interventions.

The following sections focus on (i) lung cancer as an example of cancer type with high disease burden due to its late diagnosis and poor prognosis (section 1.3.2) as well as on (ii) the potential prognostic predictors of lung cancer, with a special focus on serum biomarkers (section 1.3.3).

1.3.2 Lung cancer

1.3.2.1 Disease burden: Incidence, mortality, and prognosis

In 2022, lung cancer has been ranked as the most common cancer type and the leading cause of cancer-related death by the World Health organization [23], with a global overall 5-year survival rate of only 13% across Europe [16]—regardless of the disease stage. Amongst the 2.2 million new lung cancer cases each year, almost 1.8 million deaths occur worldwide, which exceed deaths from breast cancer, prostate cancer and colorectal cancer combined [24]. Unfortunately, given that early disease stages are asymptomatic, almost 40% of lung cancer patients are diagnosed with an already advanced metastasised disease, that has spread to other parts of the body (Fig. 1.3, A). Patients with metastatic lung cancer have a very poor life expectancy, besides the limited treatment options. Although the introduction of new treatment modalities as targeted therapy and immunotherapy, in the past 20 years, has resulted in longer survival for these patients especially reflected on the 2-year survival rate [25], the 5-year survival rate still does not exceed 10% for patients diagnosed with advanced disease stage (Fig. 1.3, B) [26].

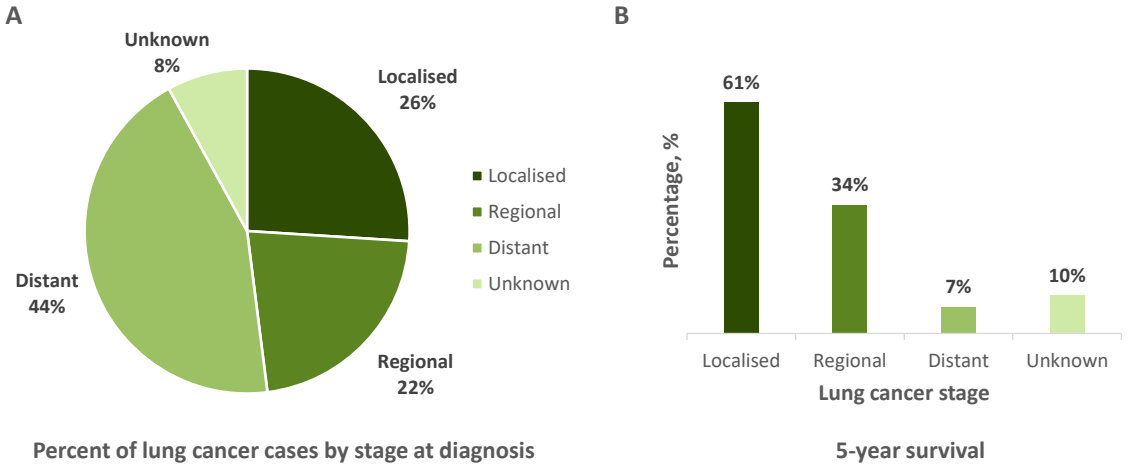

Figure 1.3: Lung cancer **[A]** stage at diagnosis and **[B]** 5-year survival rate.
Lung cancer stages: Localised, cancer confined to primary site; Regional, cancer has spread to regional lymph nodes; Distant, cancer has metastasised; Unknown, unstaged tumour.
Source: American Lung Association: Stage of lung cancer 2022 report [26].

Prognosis of lung cancer patients depends on multiple factors. These include the cancer stage at the time of diagnosis i.e. patients with localised tumour have higher chance of survival compared with patients with metastasised tumour; cancer type; response to treatment; and patient's age and general health [27].

1.3.2.2 Risk factors

The primary risk factor of lung cancer is smoking, accounting for about 80% of lung cancer cases and deaths. However, this is not the only risk factor; exposure to pollutants as asbestos, industrial hazards, air pollution, secondhand smoking and arsenic as well as history of lung disease, precursor lesions, or genetic alterations might also play a role in increasing the risk of lung cancer [25].

1.3.2.3 Types of lung cancer

Traditionally, lung cancers have been divided into non–small cell lung cancer (NSCLC) and small cell lung carcinoma (SCLC). The former accounts for 80% of the lung cancer cases while the latter accounts for the remaining 20%. SCLCs behave aggressively and are treated non-surgically in most cases, whereas NSCLCs are managed by a combination of surgery, radiotherapy and systemic drug treatments. Thus, with around 1.7 million patients diagnosed with NSCLC every year worldwide [24], our focus in this thesis, within *project II*, was on the most common lung cancer type—NSCLC.

Identification of the lung cancer type, and in our case NSCLC, is usually done during diagnosis and through histological examination, to decide on the most appropriate treatment strategy.

1.3.2.4 Diagnosis, staging, and histology

Diagnosis of NSCLC

The diagnosis of NSCLC, and lung cancer in general, usually starts with a chest X-ray where a lung tumour can appear as a white-grey mass. A computed tomography (CT) scan or positron emission tomography-CT (PET-CT) scan usually follows to confirm the suspected diagnosis. Other more invasive procedures can include bronchoscopy for a detailed visualisation of the airways and/or biopsy procedures (e.g. surgical or percutaneous) for histopathological examination of the lung cancer type. Additionally, molecular testing may also be performed to identify patients eligible for targeted therapies. Following the imaging and biopsy procedures, tumour diagnosis with respect to staging and histological subtype can be defined [28].

Staging of NSCLC

Determination of the NSCLC stage aims to identify the extent and severity of the disease, decide on the best treatment strategy, and offer prognostic information. Staging usually follows the TNM system [29] where:

- **T** stands for **T**umour and describes the size and extent of the main tumour. It involves four categories T1–T4.
- **N** stands for **N**odes and describes whether cancer has spread to the lymph nodes or not. It includes three categories N1-N3.
- **M** stands for **M**etastasis and describes whether the cancer has spread to other parts in the body or not. It includes two categories M0-M1.

The combination of the different TNM categories can then be converted into the *stage grouping*, to determine the overall stage of the disease (stage I–stage IV), with the aim that patients in the same stage would have similar treatment strategies and prognosis.

Of relevance in this thesis for *project II*, are the advanced NSCLC stages; specifically stages IIIB and IV. Stage IIIB (TNM categories: Any T with N3 and M0, or T3 or T4 with N2 and M0 [29]) involves tumours that have spread to the lymph nodes but not to distant parts of the body—also known as locally advanced NSCLC. Whereas stage IV (TNM categories: Any T and any N with M1 [29]) involves tumours that have metastasised to the other lung, pleura, pericardium; or have distantly mestastasised to other organs regardless of the main tumour size or involvement of lymph nodes—also known as metastatic advanced NSCLC.

Histology of NSCLC

Lung cancer in general, and within our focus here NSCLC, arises by the transformation of benign bronchial epithelium to neoplastic tissue, triggered by e.g. genetic alterations. Identification of the histological subtype of NSCLC seeks to identify cases eligible for targeted therapies and support therapeutic decision-making for better prognosis. The major NSCLC histological subtypes include adenocarcinoma, squamous cell carcinoma, and large cell carcinoma [30].

Adenocarcinoma

Adenocarcinoma is the most common subtype of NSCLC, accounting for about 40% of lung cancer cases with a steady increase in incidence over the past decade. It is a malignant epithelial neoplasm with glandular differentiation or mucin production that is commonly formed on the periphery. It also shows multiple gene alterations to which approved molecular targeted therapies exist, which has in turn improved its prognosis.

Squamous cell carcinoma

Squamous cell carcinoma (SCC) accounts for about 20% of lung cancer cases. It is associated with smoking and has recently shown a decline in incidence. In contrast to adenocarcinoma, SCC lacks glandular structure or mucin production. It is usually located centrally in the lung along the major pathways near the bronchi and can even form cavities when it becomes large. SCC grows slowly and has a better survival than adenocarcinoma.

Large cell carcinoma

Large cell carcinoma (LCC) is the least common subtype of NSCLC and represents $< 3\%$ of lung cancer cases. LCC is usually peripherally located, bulky and necrotic in appearance. It is devoid of any lineage-specific differentiation, and morphologic and immunohistochemical evidence of adenocarcinoma, SCC, or neuroendocrine carcinoma. Tumours bearing the morphological characteristics of LCC are designated as *not otherwise specified*, as LCC can only be confirmed through a resection specimen with the thoroughly analysed tumour being devoid of lineage-specific differentiation. LCC incidence has declined in recent years due to the reclassification after ancillary tests utilising lineage-specific markers.

1.3.2.5 Treatment strategies

Treatment of NSCLC usually takes into consideration the disease stage, histological subtype, results of molecular testings—which investigate specific gene expressions in the malignant tissue, as well as the general condition of the patient (section 1.3.2.6). Moreover, the choice and intensity of the therapy should focus on the aim; whether it is curative, in case of less advanced disease; or palliative, to reduce disease symptoms and improve the quality of life of the patient. The different treatments options for NSCLC include:

- **Surgery** which involves surgical resection of the tumour and/or part of the lung; however, this is usually reserved for the early lung cancer stages, when the tumour is still localised
- **Radiotherapy** in which ionising radiation targets and kills the malignant cells
- **Targeted therapy** which attacks specific cells and falls into monoclonal antibodies or tyrosine kinase inhibitors
- **Immunotherapy** as immune checkpoint inhibitors
- **Platinum-based chemotherapy** as carboplatin or cisplatin combined with paclitaxel

Locally advanced NSCLC can be treated by a combination of surgery, radiotherapy and systemic drug treatment. However, in case of metastatic advanced NSCLC, which comprises almost 40% of NSCLC patients, systemic treatment is usually offered, most commonly as a combination therapy guided by the patient's general condition, age, and co-morbidities. For patients whose molecular testing has shown a specific gene alternation, targeted therapy is recommended [31].

Of special relevance, is the first-line chemotherapy which was offered to the advanced NSCLC patients enrolled in our clinical study in *project II*. This chemotherapy comprised of taxane (paclitaxel) in combination with a platinum-based compound (cisplatin or carboplatin) [32,33]. Therefore, the focus in the subsequent section would be on the compounds which had been administered to patients in our present dataset:

Paclitaxel

Paclitaxel is a mitotic inhibitor which stabilises the microtubules and consequently hinders the cell's ability to divide [34]. It is approved as a first-line treatment in patients with advanced NSCLC, in combination with a platinum drug. It is also indicated for treatment of ovarian and breast cancers.

Paclitaxel is usually administered as a 3-h infusion, with a 3-week interval at a dose ranging from $100\,\mathrm{mg/m^2}$ to $220\,\mathrm{mg/m^2}$—depending on the indication and combination therapy. Due to the poor aqueous solubility of paclitaxel, it is dissolved in ethanol and Cremophor EL. The latter, causes paclitaxel to exhibit nonlinear pharmacokinetics [35], associated with high interindiviudal variability [36], and holds potential for hypersensitivity reactions that may require pre-medications.

Paclitaxel dose-limiting toxicities include haematological toxicities e.g. neutropaenia, and neurotoxicity e.g. peripheral neuropathy. An association between paclitaxel drug exposure and toxicity was observed. Specifically, time above paclitaxel plasma concentration of 0.05 μmol/L ($T_{>0.05}$) has been shown to be related to the development of peripheral neuropathy and neutropaenia [37]. In fact, $T_{>0.05}$ was a strong predictor of paclitaxel-related neutropaenia [36, 38–44] in which a $T_{>0.05}$ between 24 and 30 h was a discriminator between mild and severe neutropaenia.

Tumour resistance to paclitaxel treatment has been previously reported. The development of resistance was linked to the (i) overexpression of the multi-drug resistance gene i.e. efflux pump; (ii) changes in the target molecule e.g. mutations in the tubulin subunit(s), alterations in their binding regions, or alterations in the stability of the microtubules; (iii) reduced function of apoptopic proteins; or even (iv) alterations in cytokine expression e.g. interleukin-6 [45–47].

Platinum-based compounds

Cisplatin

Cisplatin is a platinum-based drug that is used to treat different types of cancer e.g. testicular cancer, ovarian cancer, cervical carcinoma, bladder carcinoma, SCC of the head and neck, and lung cancer—as monotherapy or in combination with other therapies. Cisplatin binds to the DNA strands by cross-linking, usually through the guanine base, to cause DNA damage and interfere with the mitotic cell division. When DNA repair fails, cell apoptosis is triggered [48].

Cisplatin is administered intravenously (15–120 $\mathrm{mg/m^2}$) every 3 to 4 weeks as a single dose, usually after paclitaxel [49] and requires extensive hydration prior and post treatment administration.

Carboplatin

Carboplatin is a second-generation platinum-based drug used in the treatment of ovarian and lung cancers, alone or in combination with other drugs [50]. Compared with cisplatin, carboplatin has slower DNA binding kinetics; however, it exhibits a similar mechanism of action in which it binds to the DNA strands by attaching an alkyl group to the DNA bases, to interfere with DNA replication and repair. To a lesser extent, carboplatin may additionally interfere with DNA replication by cross-linking the DNA strands.

Carboplatin is administered as a short-term intravenous infusion. To avoid the risk of overdosing and/or renal toxicity, carboplatin dose is adjusted to target a drug exposure corresponding to a specific area under the curve (AUC) of 4–7 mg·min/mL according to the Calvert's formula (Eq. 1.1), where GFR is the glomerular filtration rate [50].

$$\text{Total dose } (mg) = \text{Target AUC } (mg \cdot min/mL) \cdot \big(\text{GFR } (mL/min) + 25\ (mL/min)\big) \tag{1.1}$$

Comparison of cisplatin- and carboplatin-based chemotherapy

In NSCLC, paclitaxel is used with either cisplatin or carboplatin, either of which has been shown to be equally effective with similar response rates [51, 52]. Although previous research has shown a survival benefit of the cisplatin-based combination [53], results were inconclusive and contradictory in other studies [51, 52, 54].

Moreover, the toxicity profiles of the platinum drugs are different, as carboplatin carries the advantage of reduced toxicity compared with cisplatin. While cisplatin has been associated with higher rates of nausea, vomiting, ototoxicity, and nephrotoxocity; carboplatin has been associated with haematological toxicities mainly thrombocytopaenia [51]. Additionally, in case of cisplatin-based chemotherapy, paclitaxel-related neutropaenia was sequence dependent i.e. to reduce the risk of paclitaxel-related neutropaenia, paclitaxel should be administered before cisplatin as the latter reduced paclitaxel clearance [49]. This sequence dependency was not observed with

carboplatin-based chemotherapy. Therefore, the choice of the platinum drug should take into consideration the patient's comorbidities and preferences as well as the expected toxicity profile.

1.3.2.6 Assessments of patient performance and treatment efficacy

Assessment of patient performance

The Eastern Cooperative Oncology Group (ECOG) performance status describes the general condition of the patient with respect to the ability to care for oneself, and perform daily and physical activities. Thus, the ECOG status can assess the progress of the disease and how well the patient can tolerate treatment by evaluating how the disease impacts the patient's life, daily activities, and level of functioning. Furthermore, the ECOG status is commonly used as a selection criteria for patient enrolment in clinical trials, to ensure a consistent and homogeneous patient population. Table 1.1 summarises the different ECOG status grades describing the patient's performance and overall health status [55] which can be classified on a scale from 0 to 5, with ECOG 0 being fully active and ECOG 5 being dead.

Table 1.1: Scale of ECOG performance status

Grade	ECOG performance status
0	Fully active, able to carry on all pre-disease performance without restriction
1	Restricted in physically strenuous activity but ambulatory and able to carry out work of a light or sedentary nature, e.g. light house work, office work
2	Ambulatory and capable of all self-care but unable to carry out any work activities; up and about more than 50% of waking hours
3	Capable of only limited self-care; confined to bed or chair more than 50% of waking hours
4	Completely disabled; cannot carry on any self-care; totally confined to bed or chair
5	Dead

ECOG: Eastern Cooperative Oncology Group.

Assessment of tumour response

The "response evaluation criteria in solid tumours" (RECIST) [56] is a standardised way to measure the tumour response to treatment and is widely used in the clinical evaluation of cancer therapy. A baseline measurement is done using an appropriate radiological imaging technique, usually by CT, and a maximum of five target lesions, representative of all involved organs, are identified. These target lesions should be easily measured, include no more than two per organ, be suitable for accurate repetitive measurements, and with longest diameter ≥ 10 mm (for lymph nodes, short axis diameter should be > 15 mm). The sum of the longest diameters (SLD) of the target lesions (in this thesis referred to as *tumour size* thereafter) is calculated at baseline and serves as the reference for further assessments and later comparisons. Non-target lesions i.e. lesions that cannot be appropriately measured or do not qualify as target lesions, are also identified although their diameters are not necessarily measured. Therefore, they are not considered in the calculation of the SLD but rather followed up to monitor treatment response.

During follow-up, the tumour size is compared to the baseline measurement, non-target lesions are examined for presence/absence, and emergence of new lesions is also monitored. According to this assessment different objective tumour responses can be defined:

- **Complete response (CR)**
  - Disappearance of all target and non-target lesions
  - No new lesions
- **Partial response (PR)**
  - $\geq 30\%$ decrease in tumour size compared to baseline tumour size
  - No progression of non-target lesions
  - No new lesions
- **Progressive disease (PD)**
  - $\geq 20\%$ increase in tumour size compared to the smallest tumour size on the study with an absolute increase of $\geq 5\,\text{mm}$, or
  - Progression of non-target lesions, or
  - Appearance of new lesions
- **Stable disease (SD)**
  - Neither PD nor PR

Assessment of progression and/or survival time

Besides tumour assessment, the time the patient survives with/without progression and/or time taken until the patient progresses are also commonly used to assess treatment efficacy, and can fall under the following endpoints [57]:

- **Overall survival (OS)** defined as the time from the start of treatment until patient death due to any cause. OS is considered the most reliable endpoint to assess clinical benefit, and is easy and precise to measure.
- **Progression-free survival (PFS)** defined as the time from the start of treatment until objective tumour progression or patient death, whichever occurs first. PFS is generally assessed earlier than OS, and since it includes death, it can correlate to OS.
- **Time-to-progression (TTP)** defined as the time from the start of treatment until objective tumour progression and does not include deaths.

1.3.3 Biomarkers as potential prognostic predictors in NSCLC

1.3.3.1 Biomarkers and their roles in clinical setting

A biomarker is defined as a "characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic process or pharmacologic response to a therapeutic intervention" [58]. Biomarkers may include molecular, histologic, radiographic, or physiologic characteristics and can serve valuable roles in both (pre-)clinical drug development and the clinical setting. Of interest in *project II*, is their valuable role in the clinical setting; biomarkers can identify patients at risk for a disease, detect disease presence, or even relate the effect of an intervention to clinical response [58]. Based on the biomarkers role in the clinical setting, they can henceforth be classified into one of the following categories:

- **Diagnostic biomarkers** to identify patients with specific diseases or classify the disease in terms of stage and extent.
- **Predictive biomarkers** to predict and monitor patient response to a therapeutic intervention.
- **Prognostic biomarkers** to predict patient prognosis and the likelihood of a clinical outcome or disease progression.

The latter category is the one of particular interest due to the poor prognosis of NSCLC patients. Therefore, focus in the subsequent sections will only be on the prognostic value of biomarkers with respect to NSCLC.

1.3.3.2 Prognostic (serum) biomarkers in NSCLC

Given the poor prognosis of patients diagnosed with NSCLC, prognostic biomarkers offer a valuable role to identify patients at risk and predict early on a patient's clinical outcome, to support physicians make an informed decision when it comes to treatment choice and long-term outcomes.

Prognostic biomarkers in NSCLC can fall under two major categories: molecular biomarkers and routine biomarkers. Molecular biomarkers include genes, mRNA, proteins, or miRNAs. Protein expressions such as the vascular endothelial growth factor, which plays a role in the formation of new blood vessels for tumour cells to grow, or the Ki-67, which is expressed in actively dividing cells, have been shown to be associated with poor prognosis when overexpressed [59]. However, despite the availability of molecular biomarker data, optimal leveraging of these data for application in the clinical setting is limited due to the lack of reproducible results due to the heterogeneity of the used techniques and/or study designs [59].

On the other hand, routine biomarkers, which have established more presence in the clinical setting, include the tumour TNM staging and patient performance status (section 1.3.2.6) [59–61]. It has been established that patients diagnosed with early stage NSCLC have better prognosis and longer survival compared with patients diagnosed with an advanced disease stage. Taken separately, the extent of lymph node involvement, presence of metastasis, and tumour size also play a role. On the other hand, the ECOG performance status has been useful to guide the choice of therapy and predict patient prognosis [60]. Amongst the other routine biomarkers, which have been investigated for their prognostic potential in NSCLC, are *serum biomarkers*, which present cost-effective, minimally invasive, and easily measured markers. A recent review [62] investigating the prognostic value of carcinoembryonic antigen (CEA), cytokeratin-fragment 19 (CYFRA 21-1), cancer antigen 125 (CA 125), neuron-specific enolase (NSE), SCC antigen, progastrin-releasing-peptide, CA 19-9, and CA 15-3, highlighted CEA, CYFRA 21-1, and NSE to be of high positive prognostic value in advanced NSCLC. To a lesser extent, CA 125, CA 19-9, and SCC antigen were also identified [62]. Additionally, lactate dehydrogenase [63–65], a marker of tissue/cellular damage and the circulating inflammatory markers C-reactive protein (CRP) and interleukin 6 (IL-6) were significantly positively associated with poor prognosis [66–91].

CRP was of particular interest in *project II*. Besides the fact that it represents a serum biomarker with the consequent advantage of being easily measured in a reliable and minimally invasive way—as part of clinical routine, it has been shown to be associated with tumour size, advanced and more aggressive NSCLC stages, as well as poor prognosis [92, 93]. Hence, it holds a strong prognostic value in NSCLC.

1.3.3.3 CRP as a prognostic serum biomarker in NSCLC

Nature, role and characteristics of CRP

CRP is a non-specific acute phase protein that is released in response to inflammation, infection, or tissue injury [94, 95]. In presence of a stimulus, CRP is primarily produced in the hepatocytes in response to increased levels of inflammatory cytokines, especially IL-6 (Fig. 1.4) [94]. Therefore, CRP production can only be impaired in case of liver failure. Apart from that, only the resolution of the pathological trigger either spontaneously or in response to drug effect can reduce CRP concentration [94, 95].

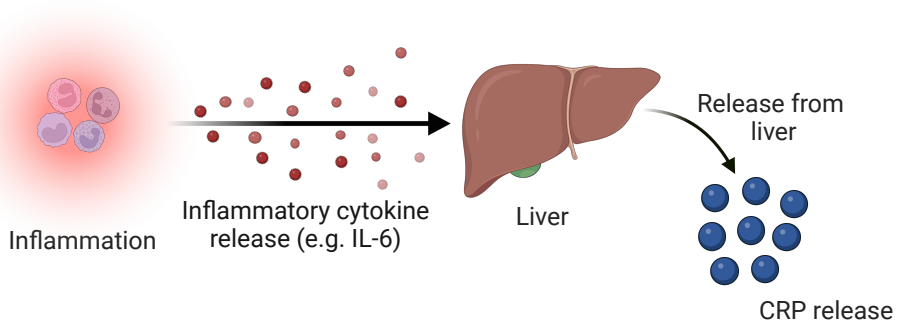

Figure 1.4: Schematic representation of C-reactive protein (CRP) synthesis. In response to inflammation, inflammatory cytokines e.g. interleukin 6 (IL-6) are released which in turn stimulate the hepatocytes to produce CRP. Figure created in BioRender.com.

The quantified serum CRP is a pentameric structure composed of five identical non-glycosylated polypeptide subunits [94]. Median CRP concentration is about 0.8 mg/L [94], although up to 10 mg/L has been observed in healthy adults [94–96]. Following a stimulus, CRP synthesis starts rapidly, its concentration rises within 6 h and peaks around 48 h [94, 95], with a potential of more than 10,000-fold increase in its concentration [94]. CRP has a half-life of 19 h that is unaffected by the (patho)physiological state of the individual [94, 97]. Therefore, CRP synthesis is the sole determinant of its concentration, which consequently reflects the intensity of the pathological process. A sustained elevated concentration of CRP indicates an ongoing pathological condition that has not yet been resolved e.g. chronic inflammation, infections or advancing disease [92, 94]. CRP is generally stable within an individual, with no diurnal variation [98, 99], or seasonal variation—apart from occasions of a pathological stimulus [94]. It is also unaffected by food. On the other hand, it shows a positive association with obesity (as obesity is considered a low-grade systemic inflammatory disease), smoking, and age [94, 95]. The latter being probably because of the associated co-morbid conditions. Moreover, robust, reproducible, and well standardised assays exist for CRP that ease its reliable measurement [92, 94].

CRP isoforms and their relation to inflammation and cancer

CRP is released from the liver in the pentameric form (pCRP). Then, it dissociates to the poorly soluble monomeric form (mCRP). Rate of conversion of pCRP to mCRP decreases over time and once mCRP is formed, there is no transition back to pCRP; hence, pCRP concentration in the systemic circulation increases with time (Fig. 1.5) [92, 100]. pCRP has an anti-inflammatory activity and is associated with a continuous, low-level inflammatory response of an unresolved disease. mCRP, which is evident during acute-phase responses, has pro-inflammatory activities and modulates inflammatory

responses by stimulating different cell types at the site of inflammation. Therefore, the reduction of mCRP over time (due to the chronic inflammatory state and the reduced conversion rate of pCRP to mCRP) and the consequent increased levels of pCRP, indicate reduction in the natural host defence response and a persistent tissue damage [92, 100].

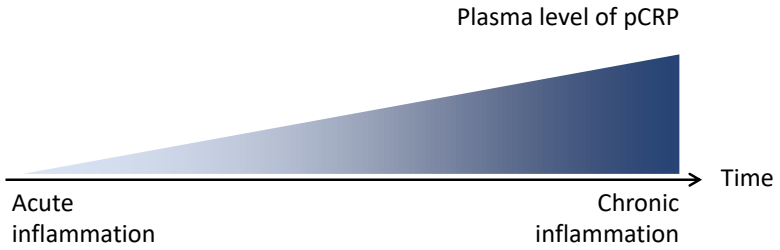

Figure 1.5: Schematic representation of pCRP as a function of inflammation and time. pCRP: pentameric C-reactive protein. Adapted from [92].

The reasons for CRP elevation in cancer patients are not yet fully clear. Although there is no consensus on whether inflammation triggers a carcinogenic effect, tumour cells induce the inflammation, or even if it is an interchange between both postulations [68,92,93,101], there is nevertheless an observed lack of resolved inflammation that is linked to cancer progression, tumour necrosis, and local tissue damage [93]. Given that cancer, in general, and within our focus here NSCLC, in specific, represent a state of chronic inflammation, one of the reasons for the elevated CRP concentration could be attributed to the inflammation caused by the tumour (growth), which promotes the production of inflammatory cytokines and in turn the acute phase protein—CRP (Fig. 1.6). Consequently, pCRP is expected to dominate and reflect the severity of tissue damage associated with tumour growth and progression. Additionally, given its weak anti-inflammatory activity, a unidirectional relationship is assumed, in which the inflammatory signal associated with the tumour is considered to drive CRP concentration (Fig. 1.6). A pro-inflammatory feedback from CRP to trigger tumour growth and metastasis is not expected, given that mCRP only exists in the acute state [92, 93, 101].

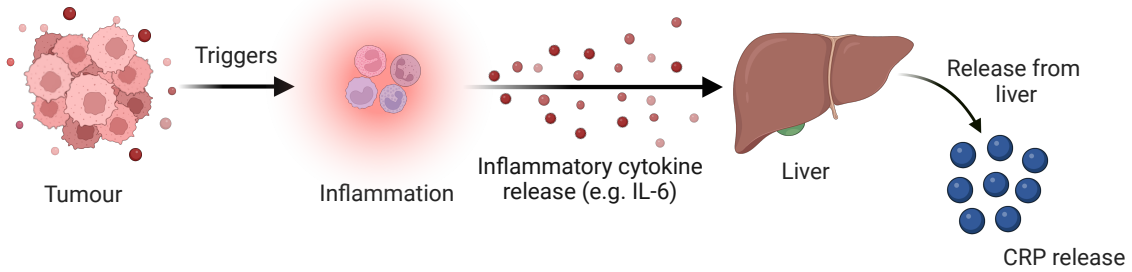

Figure 1.6: Schematic representation of tumour-driven C-reactive protein (CRP) synthesis. IL-6: interleukin 6. Figure created in BioRender.com.

Other reasons that have been postulated for the observed elevated CRP concentration in cancer included an immune response to (i) cancer-related infections, (ii) tumour antigens, or (iii) the production of inflammatory proteins by tumour cells [93, 101]. All of which support that CRP is a response to the malignant state. Additionally, CRP's precursor, IL-6, has been shown to be upregulated in response to paclitaxel and platinum drugs in ovarian and brain cancers, and be

involved in cancer resistance [68]; and in other works was found to be produced by LCC [102] and SCC [103] NSCLC cell lines which may partially explain the connection between elevated CRP concentration and poor prognosis in NSCLC patients.

Leveraging CRP concentration for optimal prognostic value

To understand the prognostic value of CRP, and monitor and predict patient prognosis, optimal leveraging of CRP concentration and the existing clinical data is needed. So far, the majority of research has focused on the baseline measurements to assess the prognostic potential of CRP in NSCLC, neglecting its kinetic change over time [67–74, 74–80, 80–88, 104, 105]. Looking beyond the baseline measurements would have the added benefit of considering the impact of treatment and the disease evolution, and thus better reflect the patient's condition and prognosis during the course of treatment.

For an efficient and better utilisation of existing clinical data, exploiting their informative value, and characterising the kinetics of the biomarker in relation to the response data, the mathematical-based approach—*Pharmacometrics*—presents an optimal methodology to bridge serum biomarker data (i.e. CRP concentration-time course and kinetic profile) to treatment outcomes and efficacy endpoints (i.e. prognostic outcomes).

The following section introduces *Pharmacometrics* and its different approaches. It also highlights its potential in dose optimisation during early clinical drug development, and its role in the early prediction of patient response by leveraging different oncology endpoints including biomarker and efficacy data.

1.4 Pharmacometrics

Pharmacometrics is defined as "the science of developing and applying mathematical and statistical methods to characterise, understand, and predict a drug's pharmacokinetic, pharmacodynamic, and disease-biomarker-outcome behaviour" [106]. Pharmacometrics represents a multidisciplinary field that bridges pharmacy, biology, medicine, mathematics, statistics, and computer programming. Through the development of a simplified representation of the system (i.e. model), pharmacometrics aims to provide a coherent understanding and characterisation of (i) what the body does to the drug (i.e. Pharmacokinetics, PK), (ii) what the drug does to the body (i.e. Pharmacodynamics, PD), and/or (iii) the time-course of drug exposure and its relation to response (PK/PD). Moreover, it can also describe tumour dynamics as well as disease progression and can even infer expected outcomes under different scenarios through further simulations [107].

To date, pharmacometrics has played a crucial role in clinical drug development, and has supported regulatory decision-making and therapy optimisation in the clinical setting.

In clinical drug development, pharmacometrics has optimised clinical study design by aiding dose selection, identifying the optimal dosing, determining sample size, supporting the use of alternative and more informative endpoints, and predicting the risk of toxicity [106, 108]. When it comes to regulatory decision-making, pharmacometrics and further simulations have supported the approval of new drug doses, dosing regimens, and exploration of new treatment indications and dosing in special (sub-)populations without the need to conduct additional clinical trials [106, 108]. Hence, not

only providing an expedited approval to drug applications but also an individual, social, and economic benefit. Between 2000 and 2008, the number of submission to the FDA utilising pharmacometric analysis had a 6-fold increase [109]. Moreover, pharmacometric analysis contributed to drug labelling and dosing decisions resulting in the approval of 64% of the submissions into which it was involved [109]. Although, more recent numbers are lacking, it is estimated that submissions involving pharmacometric analyses have increased by approximately 60% from 2000–2008 to 2012–2021 [110]. In recognition to the pivotal role of pharmacometrics, guidance documents were issued by the FDA and the EMA to support and guide the application and reporting of pharmacometric analyses as an integral part of the drug development and marketing approval processes [111, 112].

Finally, in the clinical setting and at bedside, pharmacometrics helps with individualised patient care through therapeutic drug monitoring based on the patient characteristics and measured drug concentration, and can optimise drug dosing with respect to patient characteristics and expected variability. Moreover, based on the identification of the influential patient/disease characteristics regarding a specific toxicity, it can identify vulnerable populations and patients at risk of developing toxicities and adjust their treatment schedule accordingly.

1.4.1 The population approach

One of the major advancements of pharmacometrics was the development of the population approach for the estimation of parameters in which data regardless of their nature (sparse or dense) could be leveraged. The population approach refers to the quantitative description of the PK and/or PD characteristics, disease progression, biomarker dynamics, or therapeutic outcome of a population of interest, along with their different levels of variability, e.g. interindividual variability, intraindividual variability and residual variability arising from measurement errors as an example [113]. Additionally, the potential sources i.e. factors/covariates impacting the identified variability could also be identified and their impact could be quantified. Population parameters can be estimated by fitting all the individual data together e.g. naïve pooling approach and naïve averaged data approach or by fitting each individual separately and then combining the individual parameter estimates e.g. standard two-stage approach. Each of those approaches has its own limitations. A third approach, the nonlinear mixed-effects (NLME) approach is the gold standard and population approach of choice as it overcomes the limitations of the previous two approaches. These different approaches, along with their limitations, are briefly summarised in this section.

1.4.1.1 Naïve pooling approach

In the naïve pooling approach, population parameter estimates are obtained by fitting and analysing all the individual data together under the assumption that they arose from the same individual. Consequently, variability is summarised under a single term, ignoring its different levels e.g. interindividual, intraindiviudal, measurement errors etc. Despite being an easy, simple, and rapid approach, it does not provide individual parameters estimates and could result in biased estimation of parameters [114]. Hence, it is applicable in situations where interest is only in a rough approximation of the mean profile rather than the variability.

1.4.1.2 Naïve averaged data approach

In the naïve averaged data approach, data from the different individuals are averaged at each time point before data analysis. As with the naïve pooling approach, naïve averaged data approach is simple to implement. Nevertheless, it does not estimate individual parameters, can lead to biased estimates as the variability is masked e.g. fails to account for interindividual variability, and requires identical sampling times across individuals. Thus, it is useful when experimental data with a standardised design and minimal variability are analysed [115].

1.4.1.3 Standard two-stage approach

The standard two-stage approach is a stepwise procedure in which individual parameter estimates are first calculated based on each individual's data and then descriptive statistics of the individual parameter estimates are computed across the individuals. Although this approach focuses on individual parameter estimates, point estimates of the summary statistics e.g. mean can be considered as the population parameter while the distribution e.g. standard deviation or variance, can describe the variability across individuals. However, despite the simplicity of this approach, a key limitation of the two-stage approach is the overestimation of the variability since no distinction exists between the different levels of variability i.e. interindividual variability, intraindividual variability, measurement errors [115]. Moreover, this approach requires dense sampling per individual to reliability estimate the parameters and their variability [116].

1.4.1.4 Nonlinear mixed-effects approach

In the NLME approach, individuals' data are analysed simultaneously (as with the naïve approaches) while taking into account the respective individual data, i.e. which data are from which individual, and the different levels of variability. Moreover, mean population values as well as the different sources of variability are also identified. Compared to the previous approaches, NLME showed less bias and more precise estimates of both mean parameters and different variability [114]. In contrast to the two-stage approach, it can also handle unbalanced data, and both sparse and dense samples [115]. Thus, it overcomes the limitations of the previously described population approaches and best suits the complexity of clinical data.

In this work, NLME was the preferred population analysis approach applied for all population data analyses. Detailed description of the NLME approach methodology is presented in section 2.1.2.

1.4.2 NLME modelling in oncology

Oncology data offer a wide range of variables and endpoints e.g. drug dose and concentrations (exposure metrics), efficacy (response) endpoints as tumour size and survival information, as well as blood cell counts and biomarkers concentrations for monitoring of adverse events and treatment efficacy, respectively. These factors hold a great potential to inform clinical decision-making if properly leveraged, characterised and linked together.

The following sections elaborate on the application of NLME modelling in (i) early clinical drug development where characterisation of the association between the dose and response can help better optimise dose selection; as well as its application (ii) at bedside where characterisation of the potential association among two or more endpoints, e.g. biomarker-tumour-survival, helps to identify key predictors and influencing factors for a better understanding of treatment outcome and improved future predictions.

1.4.2.1 NLME modelling in early clinical drug development

NLME modelling can optimise early clinical drug development by incorporating the PK and PD information to better inform the design of dose-finding studies. Through the explanation and control of the sources of variability in treatment response, NLME modelling can design sufficiently powered studies to detect a drug effect, provide more informative dose selection by evaluating the dose-response relationship, and optimise the benefit-risk ratio of novel therapeutic compounds [11, 13, 117–121].

Modelling of the dose- or exposure-response relationship has been a promising approach and a sought component for optimal dose selection in clinical drug development within oncology [11, 121]. Several model-based approaches have been proposed to better leverage safety and efficacy data from dose-finding studies and allow for more informative characterisation of the dose-response and optimisation of dose selection [117, 122]. Of these model-based approaches that have been recently proposed to identify dose–response relationships and are of interest in this thesis, is the combined Likelihood Ratio Test-Modelling (cLRT-Mod) which is based on the principles of NLME modelling and is commonly compared to the earlier and more established non-NLME model-based approach—the Multiple Comparison Procedure-Modelling (MCP-Mod).

The description of the MCP-Mod and its recent NLME adaptation cLRT-Mod as model-based approaches to optimise dose selection is summarised below.

Model-based approaches to optimise dose selection

Multiple Comparison Procedure-Modelling

MCP-Mod is a statistical methodology which was proposed by Bretz *et al.* [123] for the design and analysis of phase II dose-finding studies. MCP-Mod aims to identify the presence of a significant drug effect (i.e. dose-response), over placebo, using multiple contrast tests and to characterise the expected dose-response relationship that best describes the observed data. Given that multiple models (i.e. relationships) are evaluated at the same time, MCP-Mod provides flexibility in characterising the expected dose-response relationship while accounting for model uncertainty. In the second step, using the best-performing model(s), the dose-response curve is modelled to evaluate the optimal dose, dosing, or dose range that achieves the desired response, for the confirmatory phase III study.

MCP-Mod was initially developed for normally distributed responses, considering a single time point observed in patients enrolled in parallel group study designs. Later, it was extended to what was termed generalised MCP-Mod (gMCP-Mod), to consider non-normally distributed responses (e.g. binary (event/no event), count (number of events), or time-to-event endpoints) and repeated/longitudinal patient observations [13, 124].

Because of its ability to utilise the information from all the active doses and placebo, MCP-Mod is regarded to be more effective than pairwise comparison to generate statistical evidence from phase II trials with regards to dose selection and is associated with a robust control of type I error. Thus, since the method is specified at the study design stage, it can contribute to (i) better design of phase II trials i.e. dose range and dose levels are optimised to better determine the dose-response relationship and ensure a certain power is reached, and (ii) the reduction of phase II failures, and requirements for post-approval dose adjustments [13]. Hence, it has received both an EMA qualification opinion [13] as well as an FDA "fit-for-purpose" designation [125] for the design and analysis of phase II dose-finding studies.

Nevertheless, (g)MCP-Mod has been associated with some limitation such as (i) sensitivity of the method to the choice of the multiple models which may result in a reduced power of the test, and (ii) exploration of a "mean" drug effect per dose group—measured at end of the trial—resulting in reduced information associated with the dose-response analysis, ignored longitudinal information, and power reduction [126]. For these reasons, an extension of MCP-Mod, the cLRT-Mod was proposed.

Combined Likelihood Ratio Test-Modelling

The cLRT-Mod approach was proposed by Buatois *et al.* in 2021 [126] to represent a pharmacometric NLME adaptation of MCP-Mod. Earlier, the use of the likelihood ratio i.e. statistical test to compare how well the model predicts parameters that make observing the data most likely (section 2.1.2.2) has been associated with increased power compared with the contrast tests used in MCP [127]. Hence, cLRT aimed to overcome some of the limitation associated with the use of the multiple contrast tests in MCP by leveraging the likelihood ratio test and principles of NLME (section 2.1.2), to assess dose-response while accounting for model uncertainty. Moreover, in contrast to MCP, cLRT aimed to make use of the full longitudinal data and the repeated measurements over time associated with progressive diseases such as cancer in which the status of the patient evolves with time. In previous clinical trial simulations [126], and compared to MCP-Mod based on an end-of-treatment endpoint, cLRT showed an increase in power in detecting a drug effect for both symptomatic (67% on average) and disease-modifying (11.3% on average) drug effects, with a similar control of type I error.

1.4.2.2 NLME modelling at bedside

At bedside, there is a broad range of oncology-related variables e.g. exposure metrics, tumour sizes, biomarkers concentrations, adverse events, and survival information. NLME modelling allows the characterisation of each of these factors for a quantitative understanding of the different key processes as well as the expected range of outcomes—given the estimated variability. Additionally, the association of different variables and endpoints through a modelling framework e.g. linking drug exposure to tumour size and linking tumour size to survival, offers a broader and more comprehensive understanding of the quantitative relationships between these different variables in relation to the relevant endpoint.

The subsequent sections focus on the basic principles and NLME modelling approaches commonly utilised to leverage oncology-related data at bedside, that were of interest in the work presented in this thesis.

Modelling tumour dynamics

In contrast to the internationally adopted RECIST criteria to monitor tumour response to treatment, in which continuous tumour size data are categorised (section 1.3.2.6) [56], population modelling of tumour size carries the advantage of preserving the continuous nature of tumour size data. Moreover, it can also characterise the dynamic change of tumour size over time, account for the development of resistance, and if treatment information is considered, can explore the impact of different doses and/or treatment schedules other than those investigated. Thus, population modelling of tumour size dynamics offers the valuable opportunity to characterise the time course of tumour size, from the usually sparse tumour size data in clinical trials, for an early assessment of tumour response and treatment efficacy.

The structural complexity of models describing tumour size is highly variable. Models can range from being a simple description of tumour growth, describing the change in tumour size over time informed only by the measured tumour sizes, to being more complex accounting for the influence of treatment, tumour heterogeneity, development of resistance, or even the potential involvement of biological processes e.g. angiogenesis, immune system and targeted biomarkers [128]. Thus, different models exist to describe tumour growth which reflect different hypotheses regarding the tumour environment or treatment effects and commonly include one or more of the following components:

- **Tumour growth**
  - A fundamental component of tumour dynamics models is accounting for tumour growth kinetics. Tumour growth kinetics could represent a non-saturable growth (i.e. unlimited growing pattern) e.g. linear or exponential growth. A linear tumour growth assumes a constant zero-order tumour growth rate (K_{growth}) whereas an exponential tumour growth assumes that the tumour growth rate is proportional to the tumour burden ($K_{growth} \cdot Tumour\ size$) [128]. Alternatively, tumour growth kinetics could represent the less-commonly clinically observed but biologically realistic saturable growth reflective of a maximum tumour growth capacity due to availability of nutrients, oxygen and space e.g. gompertz or logistic growth [128].
- **Impact of treatment**
  - During treatment, a second component of tumour dynamics models is accounting for the impact of treatment on tumour size, which commonly assumes that tumour shrinkage in response to treatment i.e. drug-induced tumour shrinkage, is proportional to tumour burden. Drug-induced tumour shrinkage could also be a function of treatment exposure e.g. drug concentration, drug dose, or AUC. The relationship between treatment exposure/drug dose and tumour size could exhibit either a linear or non-linear relationship (as explored in *project I*). A lag time to account for the delay between drug exposure and treatment effect due to e.g. time needed to reach the site of action or to bind to the target receptor, can also be considered if needed.
- **Resistance**
  - Emergence of resistance to treatment could be represented by a decrease in the efficacy of treatment as proposed by Claret *et al.* [129], in which the efficacy of the drug declines over time, due to emergence of resistance, leading to loss of drug-induced tumour shrinkage. Alternative ways to account for resistance include the assumption of different tumour cell populations with different sensitivities to treatment [128].

Optionally, more complex semi-mechanistic tumour dynamics models could account for tumour heterogeneity assuming that the tumour mass is a mix of proliferative and quiescent cells and/or drug-resistant and sensitive cells. Potential impact of biological factors as angiogenesis dynamics (monitored by respective biomarkers), or the immune system in case of immunotherapy could also be considered [128, 130].

Comprehensive reviews of available models to characterise tumour dynamics in presence/absence of treatment effect and resistance evolution are published [128, 131–133] and provide an overview of the potential modelling frameworks.

Modelling biomarkers

Modelling biomarker data allows the description of their time course as well as their relationship to treatment and/or tumour. Biomarkers are commonly described by the so-called "turnover model" or "indirect response model". The turnover model is a pharmacodynamic model that represents the biomarker concentration as a dynamic process governed by a production process (a zero-order production rate constant, K_{in}) and a degradation process (a first-order degradation rate constant, K_{out}) (Fig. 1.7, A). Alternatively, it is also considered as a one-compartment model with an inflow and outflow represented by K_{in} and K_{out}, respectively.

Such model can describe biomarker concentrations in both steady-state or in response to an external factor e.g. drug effect represented by drug concentration or AUC, or effect of tumour dynamics represented by tumour size (Fig. 1.7, B). In absence of an effect, the biomarker concentration is in steady-state in which the baseline steady-state concentration can be derived from the ratio between K_{in} and K_{out} (Eq. 1.2). In presence of an external factor, its impact on the system can be described in four different ways as either stimulation or inhibition of K_{in} or K_{out} (Fig. 1.7, B) [134–136]. The most appropriate model could be decided based either on prior knowledge of the expected relation between the external factor and the biomarker or through the selection of the best-fitting model [137].

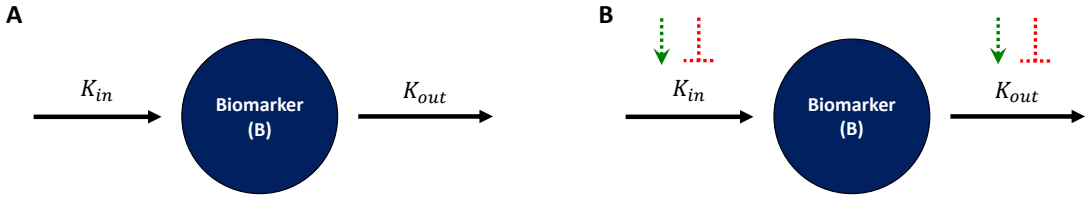

Figure 1.7: Schematic diagram of a biomarker turnover model at **[A]** steady-state and **[B]** in response to external factor showing either stimulation (green arrow) or inhibition (red sign) of K_{in} or K_{out}.
K_{in}: zero-order production rate constant; K_{out}: first-order degradation rate constant.

At steady-state, the baseline biomarker concentration (*baseline*) is given by,

$$Baseline = \frac{K_{in}}{K_{out}} \tag{1.2}$$

and at time (t), the rate of change of the biomarker (B) concentration over time ($\frac{dB(t)}{dt}$) can be described by (Eq. 1.3),

$$\frac{dB(t)}{dt} = K_{in} - K_{out} \cdot B(t) \tag{1.3}$$

Modelling survival

Population modelling of survival data adopts the parametric time-to-event (TTE) modelling approach in which the instantaneous risk of developing an event e.g. death (also known as hazard) is characterised. By calculation of the cumulative risk of having the event, the distribution of these events, over time, are characterised (discussed in details in section 2.2.4).

An advantage of this modelling approach is that it also allows the assessment of different predictors for their quantitative impact on the instantaneous risk of developing the event. These predictors could be (i) observed baseline patient characteristics e.g. ECOG status, disease stage or (ii) model-derived predictors obtained from the characterisation of drug concentration e.g. AUC, the characterisation of tumour size e.g. relative change in tumour size (i.e. tumour shrinkage) at specific time, or the characterisation of biomarker kinetics e.g. change in biomarker from baseline—as investigated in *project II*.

Moreover, a major advantage of the parametric TTE models is that since the shape of the hazard is identified (i.e. constant, increasing, or decreasing risk) along with the impact of the identified significant predictors, simulations could be performed to predict the events under different (i) design settings e.g. different number of dose levels; (ii) predictor levels e.g. extreme values of biomarker concentration or different predictor categories as disease stage; or (iii) dosing schedules e.g. 2- or 3-weekly paclitaxel dosing [137].

1.5 Objectives

The high incidence and mortality rates associated with cancer combined with the high attrition rates during the development of new compounds as well as the high rate of therapeutic failure, rank oncology as a therapeutic area with a huge need for optimised early clinical decision-making. Timely clinical decision-making is valuable at two crucial steps when it comes to (i) the development of new and efficacious medications with optimally selected dosing early on during the initial clinical development phase and (ii) the timely monitoring of patients response for the early prediction of anticancer treatment success/failure at bedside.

Therefore, the objectives of this doctoral thesis aimed to contribute to better leveraging of dose-finding trial data for optimal selection of drug dosing *(project I)* as well as to the early prediction of patient response at bedside for timely decision-making and better informed treatment decisions *(project II)*. To optimally leverage the rich, longitudinal, and heterogeneous oncology data at these different stages, pharmacometric modelling and simulation methodologies were applied to guide early clinical decision-making at these two critical stages represented by the independent objectives of *(project I)* and *(project II)* as follows:

Project I: Simulation-based evaluation of cLRT and MCP approaches in oncology under different clinical trial scenarios

To optimise dose selection in the oncology setting, adequate leveraging of the data of dose-finding studies is essential. The latter can be achieved by model-based approaches that are currently recommended to leverage and analyse phase II dose-finding studies. Therefore, understanding the performance of these approaches within the oncology setting is of high interest and hence, the main objective of *project I* was to assess the performance of the newly proposed cLRT approach *versus* the MCP approach, as two proposed model-based approaches that identify a dose-response relationship to optimise dose selection in dose-finding studies. A simulation-based exploratory assessment was undertaken to evaluate the cLRT and MCP approaches under different study design considerations within an oncology setting with the aim to answer the following questions:

- How does cLRT perform when considering continuous tumour size data in comparison to MCP when using time invariant efficacy endpoint in oncology?
- What is the optimal study design under a fixed total number of patients that would successfully inform early decision-making and achieve high power in detecting a dose-response relationship?

To achieve this, we sought to:

- Compare the performance of cLRT in identifying a dose-response relationship when using continuous tumour size data to MCP when using the best change in tumour size from baseline as endpoint.
- Challenge the performance of cLRT and MCP under two different drug effects (strong *vs* weak) and different dose-response relationships of increasing complexity (linear, log-linear, E_{max}, and sigmoidal E_{max} relationships).
- Evaluate the influence of the different study design variables: total number of patients i.e. total sample size and number of dose levels, on the performance of cLRT and MCP.

Project II: Identification of predictors of efficacy in NSCLC patients: A tumour dynamics - CRP modelling framework

Non-small cell lung cancer (NSCLC) is a disease of high incidence and mortality that is associated with poor prognosis as well as a significant clinical burden and poor quality of life for the patient. Timely and early prediction of outcome is of great benefit to spare the patient the unnecessary toxicity of the treatment, and to better inform early therapeutic decision-making.

Therefore, the main objective of *project II* was to identify early predictors of efficacy in advanced NSCLC patients treated with paclitaxel in combination with either carboplatin or cisplatin chemotherapy. We focused on C-reactive protein (CRP) as a serum biomarker which has been previously shown to be associated with NSCLC prognosis. Further, we sought to explore the informativeness of longitudinal CRP data compared to only baseline data, to evaluate their respective prognostic potential as predictors of progression-free survival (PFS) and overall survival (OS), and to seek answering the main question of whether monitoring specific metrics of CRP, as a minimally invasive routine marker of inflammation, could provide a prognostic potential of disease outcome or not. Through this main objective, we aimed to answer the following questions:

- Can monitoring inflammation, through longitudinal measurement of specific CRP metrics, provide a potential prognostic marker?
- Do longitudinal predictors have a stronger impact compared to baseline predictors?
- Is non-baseline CRP concentration more informative than baseline CRP concentration?
- Going beyond baseline characteristics, which are the most significant predictors of PFS and OS in patients with advanced NSCLC? and what are the characteristics of patient subpopulations at risk of poor prognosis, and those who have good prognosis and better chance of treatment success?
- How does tumour dynamics impact CRP concentration?
- What are the relevant factors impacting CRP production?

To answer these questions within our main objective, the following two sub-objectives were identified:

- Development of a coupled tumour dynamics - CRP model to characterise circulating longitudinal CRP concentration with respect to time. With this coupled model, we aimed to:
  - Link tumour dynamics to CRP production to characterise the relationship between drug exposure, tumour dynamics and CRP concentration; account for the influence of tumour burden on CRP concentration; identify factors impacting CRP production; and characterise the dynamic profile of CRP concentrations.
  - Leverage the developed model to select and derive model-based metrics as potential predictors of efficacy, with a special focus on early longitudinal metrics.
- Characterisation of efficacy endpoints and impact of predictors, with the aim to:
  - Characterise PFS and OS as the most commonly used clinical endpoints.
  - Identify and evaluate the most significant predictors of PFS and OS.
  - Assess the impact of the identified predictors of PFS and OS on median time-to-event.

2 | Data and methods

2.1 Population pharmacometric modelling and simulation

Nonlinear mixed-effects (NLME) modelling and simulation was the main pharmacometric approach used to achieve the objectives of the work presented in this thesis. NLME was adopted to optimally leverage the longitudinal data (*Projects I* and *II*) and allow the quantification of the variability and identification of its sources (*Project II*). For this reason, the principle and description of the applied NLME modelling and simulation methodology as well as the preparation of the data for this type of analysis are presented in the following sections.

2.1.1 Data management and exploratory data analysis

To obtain a dataset that is compatible with our NLME analysis, several steps had to be undertaken (outlined in Fig. 2.1). These included: (i) "data management" to arrange the clinical data in the desired structure and check for its appropriateness and (ii) "exploratory data analysis" to identify trends, variability and outliers within the developed dataset.

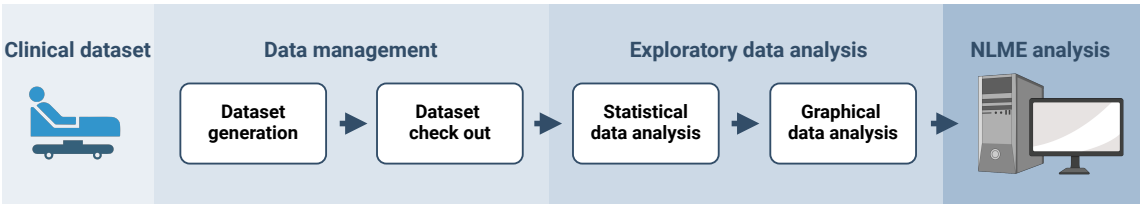

Figure 2.1: Workflow of data management and exploratory data analysis starting with clinical data to create a compatible dataset for nonlinear mixed-effects (NLME) analysis. Figure created with BioRender.com.

2.1.1.1 Dataset generation and check out

Data intended for analysis in a modelling and simulation software require a specific structure. While different software differ in the prerequisites of the data structure, focus here will be on the structure required by the NLME modelling and simulation software, *NON*linear *M*ixed-*E*ffects *M*odelling (NONMEM®), which had been used in the work presented in this thesis.

For a NONMEM®-compatible dataset, all the data were transformed to a numerical format and arranged in a chronological order, per individual, with respect to the occurrence of events. Each record represented a single event (a dosing event, an observation, or other). NONMEM®-specific identifiers were also included to identify the individuals (ID), records with/without the dependent variable (i.e. identifier for missing dependent variable, MDV), and type of event record (i.e. event identifier, EVID).

The dataset(s) used for our modelling purposes included all information ranging from the dependent variable (DV) e.g. drug concentrations or biomarker concentrations to independent variables e.g. time or patient demographics. Additional characteristics, if needed, were also derived during the dataset generation step e.g. deriving body mass index from information on body weight and height. It is worth noting that the necessity of the different information in the dataset depended on the type and purpose of the model e.g. datasets intended for PD models did not necessarily include dosing events, especially if no PK description was included in the model—as in *projects I* and *II*.

Handling of missing data or data below the lower limit of quantification (LLOQ) was performed at this stage of dataset generation. It was important that descriptive information i.e. covariates, were not missing but rather present on every record for a successful characterisation of their impact. As different strategies existed [138, 139], measures to handle these situations are described under the respective project section (*Project I*: not applicable; *Project II*: section 2.6.2.2).

Following the development of the NONMEM®-compatible dataset, a dataset check out was performed to detect inconsistencies, and identify and correct any errors. This involved (i) "cross-column checks" to ensure the correct combinations of the dataset variables e.g. in a correctly built dataset, an observation record should read MDV equals 0, EVID equals 0 and DV does not equal 0; and (ii) "index plots" in which different data variables were plotted against each ID to detect outliers and identify potential clustering.

To ensure reproducibility and traceability, all steps of dataset management (generation and check out, Fig. 2.1) were performed using R and RStudio software (section 2.4).

2.1.1.2 Exploratory data analysis

After the dataset generation and check out, and before model development started, extensive statistical and graphical data analyses were performed to explore the data, detect trends, and identify data points that required further investigation.

Exploratory statistical data analysis

During the statistical data analysis, the distribution and variability of the continuous variables were explored using descriptive summary statistics (section 2.3.1); whereas for categorical variables, the frequencies of the different categories were explored using counts and percentages.

Exploratory graphical data analysis

During the graphical data analysis, histograms and bar plots were used to visualise the frequency distribution for continuous and categorical variables, respectively while box-whisker plots visualised the median, interquartile range, and range of a continuous variable.

Moreover, during the graphical data analysis, the general trend (i.e. typical profile) and variability in the DV (e.g. biomarker concentration in PD analysis) were investigated when plotted against the independent variable (e.g. time). This relationship was visualised as a spaghetti plot in which data points from the same individual were connected.

Additionally, graphical data analyses detected potential trends between the DV and patient characteristics, and identified potential correlations between different covariates to help with covariate screening and analysis (section 2.1.2.1). Correlations between two continuous variables were explored using scatter plots, whereas in presence of a categorical variable, box-whisker plots were used.

Moreover, plots were generated either for the whole patient population or stratified by a specific variable (e.g. covariate) and/or plotted on linear or logarithmic scale for a clearer and more detailed visualisation of the different ranges of values.

2.1.2 Nonlinear mixed-effects modelling and simulation

In NLME modelling, the typical population parameters, that do not vary across individuals (i.e. fixed-effects), were estimated along with the quantification of the variability that contributed to the differences in observations within the population (i.e. random-effects), hence the term "mixed-effects". Additionally, individual parameters were estimated. Besides the ability to deal with unbalanced and sparse data, NLME has the advantage of analysing data of all individuals simultaneously. Therefore, the data of the whole population contributed to the parameter estimation via the maximum likelihood approach, where the most likely set of model parameters' values were determined that maximised the likelihood to observe the data. Approximation methods were used to measure the parameter likelihood, to account for the "nonlinear" relation between the DV, and the model parameters and random-effects.

In general, an NLME model consisted of three major components (Fig. 2.2):

- *the structural model*: to describe the profile of a typical individual in the population,
- *the statistical model*: to quantify and identify the different levels of variability e.g interindividual variability and residual variability (i.e. the variability quantifying the difference between model predictions and observations that could not yet be explained), and
- *the covariate model*: to identify variables contributing to the observed variability.

The structural and statistical models were collectively known as the *base model*.

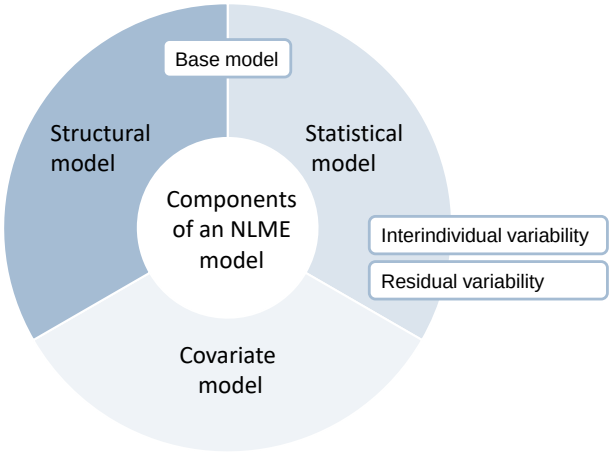

Figure 2.2: Components of a nonlinear mixed-effects (NLME) model

In the following sections, the different NLME model components, parameter estimation methods and means to evaluate an NLME model are described, with respect to the implementation of the NLME framework in the modelling and simulation software NONMEM®.

2.1.2.1 Model components

Structural model

The structural model described the typical profile of the DV (e.g. biomarker concentration) as a function of the independent variable (e.g. time)—in absence of covariates. It also estimated the typical (mean) value of the population parameters e.g. population value of the biomarker production rate constant.

It is therefore considered a mathematical representation of the relevant underlying process and could consequently be described as a mathematical function, f, including study design variables, x, and structural model parameters, ϕ. Therefore, for a measured observation, j, of the i^{th} individual, a general mathematical representation of the structural model can be represented as follows (Eq. 2.1),

$$\hat{Y}_{ij} = f(\phi_i, x_{ij}) \tag{2.1}$$

where $\hat{Y}_{ij}$ is the predicted observation for the i^{th} individual at the j^{th} observation time, f is the function for predicting the observation, j, of the i^{th} individual and is dependent on ϕ_i, the vector of structural model parameters of the i^{th} individual and x_{ij}, the vector of independent study design variables (e.g. time).

In this thesis, the structural model was leveraged to describe (i) biomarker concentration (as implemented in *project II*) and (ii) tumour dynamics (as implemented in *projects I* and *II*).

Statistical model

The statistical model aimed to quantify the different levels of variability within the population under investigation. In the presented work, the interindividual variability and residual variability were of relevance and are described below. However, it is worth noting that additional variability components such as interoccasion or interstudy variability could also be investigated, when applicable.

Interindividual variability

The interindividual variability (IIV) was quantified to allow the typical value of the parameter, estimated from the structural model, to vary across individuals and estimate the parameter value for each individual—so called Empirical Bayes Estimates (EBEs). Hence, IIV quantified the deviation of the individual parameter value from the typical population parameter value (i.e. population mean) for the different model parameters. Since model parameters are typically assumed to follow a log-normal distribution, IIV was implemented on model parameters using an exponential relationship (Eq. 2.2). This ensured that individual parameters were always positive (e.g. biomarker production rate constant) and were not estimated to have physiologically implausible values [140, 141].

$$\theta_{i,k} = \theta_k \cdot e^{\eta_{i,k}} \qquad \eta_{i,k} \sim N\left(0, {\omega_k}^2\right) \tag{2.2}$$

In Eq. 2.2, $\theta_{i,k}$ is the i^{th} individual parameter value for the k^{th} parameter, θ_k is the typical population parameter value for the k^{th} parameter and $\eta_{i,k}$ is the deviation of the i^{th} individual value from the typical population parameter value, θ_k. $\eta_{i,k}$ was assumed to be normally distributed with a mean of zero and a variance ${\omega_k}^2$. Therefore, when reporting IIV, the variance ${\omega_k}^2$, which was on a log-scale, was converted to the original scale as coefficient of variation (CV%) to ease interpretation according to Eq. 2.3. For small variances (e.g. < 0.4), a simpler approximate was used (Eq. 2.4) [140, 141].

$$CV\% = \sqrt{e^{{\omega_k}^2} - 1} \cdot 100\% \tag{2.3}$$

$$CV\% = \sqrt{{\omega_k}^2} \cdot 100\% \tag{2.4}$$

The variances of all IIV parameters, ${\omega_k}^2$, were combined in a variance-covariance matrix called Ω. To account for the potential correlation between the different ${\omega_k}^2$, a non-zero off-diagonal variance-covariance matrix (Ω) could be generated, in which the diagonal elements represented the variance (${\omega_k}^2$) and the off-diagonal elements represented the covariance terms of the IIV parameters. Thus, for k parameters 1 and 2, Ω can be expressed as in Eq. 2.5,

$$\Omega = \begin{bmatrix} (\omega_{1,1})^2 & \\ (\omega_{1,2})^2 & (\omega_{2,2})^2 \end{bmatrix} \tag{2.5}$$

where $(\omega_{1,1})^2$ and $(\omega_{2,2})^2$ are the variances for parameters 1 and 2, respectively and $(\omega_{1,2})^2$ is the covariance term between parameters 1 and 2. The correlation coefficient (ρ) between parameters 1 and 2 ($\rho_{1,2}$) could then be calculated as follows (Eq. 2.6),

$$\rho_{1,2} = \frac{(\omega_{1,2})^2}{\sqrt{(\omega_{1,1})^2 \cdot (\omega_{2,2})^2}} \tag{2.6}$$

It is worth noting that unless explicitly estimated, the off-diagonal elements were commonly assumed to be zero. This was the case in *projects I* and *II*, ${\omega_k}^2$ were assumed to be independent of each other.

Residual variability

The residual unexplained variability (RUV) represented the deviation between the observed DV and the model-predicted observations which could not be explained by the other variability components as IIV. RUV could be attributed to model misspecification, imprecision in the assay method of the DV, or errors in the documentation of dosing information or sampling collection times.

RUV could be implemented with respect to the DV using different models e.g. additive, proportional, or combined additive and proportional models, represented by Eqs. 2.7, 2.8, and 2.9, respectively.

Knowing that the model prediction, $f(\phi_i, x_{ij})$, equalled $\hat{Y}_{ij}$ (Eq. 2.1), then

$$Y_{ij} = \hat{Y_{ij}} + \varepsilon_{ij,add} \qquad \varepsilon_{ij,add} \sim N(0, {\sigma_{ij,add}}^2) \tag{2.7}$$

$$Y_{ij} = \hat{Y_{ij}} \cdot (1 + \varepsilon_{ij,prop}) \qquad \varepsilon_{ij,prop} \sim N(0, {\sigma_{ij,prop}}^2) \tag{2.8}$$

$$Y_{ij} = \hat{Y_{ij}} \cdot (1 + \varepsilon_{ij,prop}) + \varepsilon_{ij,add} \qquad \varepsilon_{ij,add} \sim N(0, {\sigma_{ij,add}}^2) \tag{2.9}$$

$$\varepsilon_{ij,prop} \sim N(0, {\sigma_{ij,prop}}^2)$$

where Y_{ij} is the measured observation (j) of the i^{th} individual, $\hat{Y_{ij}}$ is the predicted observation (j) of the i^{th} individual, and ε_{ij} is the random-effect parameter quantifying the discrepancy between the measured and predicted observation and is assumed to be independent and normally distributed, with a mean of zero and variance σ^2.

For DV spanning a wide range of magnitude as it was the case in our data, modelling them in the log scale was recommended for increased model stability and improved precision (as applied in *projects I* and *II*). In this case, a log-transformed both sides approach for the RUV was useful in which both the measurements and model predictions were log-transformed and the residual variability was implemented as an additive component on the log-scale (Eq. 2.10)—corresponding to the exponential component on the normal scale (Eq. 2.11). Besides numerical stability, this approach also had the advantage of constraining predictions in the positive domain [140].

$$ln(Y_{ij}) = ln(\hat{Y_{ij}}) + \varepsilon_{ij,add} \qquad \varepsilon_{ij,add} \sim N(0, {\sigma_{ij,add}}^2) \tag{2.10}$$

$$Y_{ij} = \hat{Y_{ij}} \cdot e^{\varepsilon_{ij,add}} \qquad \varepsilon_{ij,add} \sim N(0, {\sigma_{ij,add}}^2) \tag{2.11}$$

Alternatively, the additive RUV could be estimated as a fixed-effect parameter (θ_{RUV}) representing the standard deviation (SD) of the variability component (as applied in *project II*), then the random-effect parameter (ε) was fixed to 1 and the variance (σ^2) equalled ${\theta_{RUV}}^2$ (Eq. 2.12) [142].

$$Y_{ij} = \hat{Y_{ij}} + \theta_{RUV} \cdot \varepsilon_{ij,add} \qquad \varepsilon_{ij,add} \sim N(0,1) \tag{2.12}$$

When RUV was reported, the additive RUV component was reported as SD with the same unit as the DV, whereas the proportional RUV component was reported as CV%, similar to the IIV (Eq. 2.13).

$$CV\% = \sqrt{\sigma^2} \cdot 100 \tag{2.13}$$

Depending on the nature of the dataset (e.g. different DV types in case of joint modelling, different patient populations, or parent-metabolite modelling), more than one residual component, specific to each term, could be estimated. In this case, variances of all RUV parameters σ^2 were combined in a variance-covariance matrix called Σ in which elements of the diagonal were the σ^2 of each term and the off-diagonal elements were estimates of the correlation between elements of the diagonal of the matrix. This correlation was assumed to be zero unless otherwise specified.

Covariate model

The covariate model followed the development of the base model (i.e. structural+statistical model) and aimed to increase the model's predictive performance by (i) identifying variables that explained the reasons behind the observed IIV and (ii) quantifying their impact. These variables, known as *covariates*, were factors related to the individual, drug, or disease.

Covariates could be categorised into two types: non-time-varying or time-varying. Non-time varying covariates were, as the name implied, not changing with time (i.e. one value per individual throughout the observation period) whereas time-varying covariates were longitudinally measured and changed in magnitude with time (i.e. multiple values per individual throughout the observation period).

Furthermore, covariates could be classified as continuous or categorical. Continuous covariates could take on any possible value e.g. weight, tumour shrinkage whereas categorical covariates were discrete in nature e.g. sex, disease stage.

Dependent on the type of covariate, different mathematical functional forms existed to implement and describe the parameter-covariate relationship in the model.

Continuous covariates

The effect of continuous covariates on a model parameter was described by a linear, power or exponential functional form. A linear parameter-covariate relationship described a linear change (increase or decrease) in the model parameter with respect to the change in the covariate value over the full range of the observed covariate values. Whereas, nonlinear relationships e.g. power and exponential parameter-covariate relationships were more flexible given the magnitude of the exponent value. Hence, they were biologically more justifiable and could appropriately explain different relationships compared with the linear relationship. In addition to the functional form, covariates could be centred relative to the median covariate value, to allow the parameter estimates to be defined within the observed data range for an improved precision.

Hence, for a continuous covariate (Cov) with median value (Cov_{median}), the typical population parameter (θ_k) for the given covariate ($\theta_{k,cov}$) was described using one of the following functional forms:

- Linear function (Eq. 2.14),

$$\theta_{k,cov} = \theta_k \cdot \left(1 + \theta_{cov} \cdot (Cov - Cov_{median})\right) \tag{2.14}$$

  where the covariate effect, θ_{cov}, is the fractional change in the typical population parameter estimate θ_k per unit change of Cov from Cov_{median}, and $\theta_{k,cov}$ is the population value of parameter k at the covariate level Cov.

- Power function (Eq. 2.15),

$$\theta_{k,cov} = \theta_k \cdot (\frac{Cov}{Cov_{median}})^{\theta_{cov}} \tag{2.15}$$

  where θ_{cov} is the exponent representing the change in the natural log of the parameter k per unit change in the natural log of Cov relative to Cov_{median}. The variation in the exponent value granted the power function great flexibility to characterise the parameter-covariate relationship.

- Exponential function (Eq. 2.16),

$$\theta_{k,cov} = \theta_k \cdot exp^{(Cov - Cov_{median}) \cdot \theta_{cov}} \tag{2.16}$$

  where θ_{cov} is the exponent estimate reflecting the change in the natural log of the parameter k per unit change of Cov from Cov_{median}.

Besides the above mentioned functional forms, additional approaches were considered when handling *time-varying covariates* to account for the variation of the covariate over time and the extra information they contained compared to non-time-varying covariates [143, 144]. These approaches included:

- Interpolation: this further included (i) backward interpolation between the covariate measurements (the default procedure in NONMEM®) i.e. covariate measurement at time t_2 was carried backward and assumed to be constant until the previous measurement at time t_1; (ii) forward interpolation in which the last covariate measurement was carried forward until the next measurement; or (iii) linear interpolation in which the slope between the covariate measurements was calculated.
- Accounting for IIV in the parameter-covariate relationship: this included addition of a variability component (η_{cov}) to the parameter-covariate relationship to account for the variability in the covariate relationship across individuals. Hence, Eq. 2.14 can be rewritten as follows (Eq. 2.17),

$$\theta_{k,cov} = \theta_k \cdot \left(1 + \theta_{cov} \cdot (Cov - Cov_{median}) \cdot e^{\eta_{cov}}\right) \tag{2.17}$$

- Accounting for intraindividual variability in the parameter-covariate relationship: this included the quantification of the additive impact of both the baseline covariate ($CovBase$) and the change from the baseline covariate ($CovDiff$). Hence, Eq. 2.14 can be rewritten as follows (Eq. 2.18),

$$\theta_{k,cov} = \theta_k \cdot \left(1 + \theta_{cov,Base} \cdot (CovBase - CovBase_{median}) + \theta_{cov,Diff} \cdot (CovDiff)\right) \tag{2.18}$$

Categorical covariates

Depending on the nature of the discrete categories, categorical covariates could be further divided into:

- *Nominal data* in which there was no hierarchy or intrinsic ordering to the categories e.g. sex: males vs females or NSCLC histology: adenocarcinoma, bronchioalveolar carcinoma, carcinoma not otherwise specified, or squamous-cell carcinoma
- *Ordinal data* in which there was a specific order or hierarchy to the categories e.g. smoking status: non-smokers, former smokers, current smokers or ECOG performance status: 0, 1, 2

In contrast to continuous covariates whose parameter-covariate relationship was explained by different functional forms, the relation between a categorical covariate and a parameter was described by a proportional model in which the covariate effect on a parameter was described as a fractional (proportional) change relative to a reference category. This reference category was usually set as the group with the highest number of observations in the dataset (i.e. most common category). However, when there was an inherent ordering (i.e. ordinal categorical data), it was indeed useful to consider it when deciding for the reference category to ease of interpretation e.g. setting non-smokers as the reference group for the different smoking status.

For a dichotomous categorical covariate with two potential categories (COV = 1, 2), a fractional change model is represented in Eq. 2.19 which could be extended for each additional category in case of polychotomous covariates (i.e. one extra line for each extra category).

$$\theta_{k,cov} = \theta_k \cdot \theta_{cov} \qquad \text{where } \theta_{cov} = \begin{cases} 1, & \text{If COV} = 1 \text{ (reference)} \\ 1 + \theta_{cov2}, & \text{If COV} = 2 \end{cases} \tag{2.19}$$

Where θ_k represents the typical population parameter estimate in the reference covariate category and θ_{cov2} represents the fractional (proportional) change in θ_k encountered in presence of the other covariate category (i.e. COV=2).

Covariate model building

To build a covariate model, the process started with the pre-selection of potential covariates. This pre-selection was usually guided by different criteria: biological plausibility, prior knowledge of the underlying mechanism, previously reported correlations in the literature, graphical exploration of $\eta_{i,k}$ of the base model versus potential covariates (section 2.1.1.2), and data quality. Moreover, for a successful detection of a covariate effect, the pre-selected covariate of interest had to be present in the dataset across a wide range/distribution in case of continuous covariates or be appropriately balanced with sufficient number of patients across the different categories in case of categorical covariates.

As a second step, the impact of the pre-selected covariates was examined on the relevant model parameters using a pre-selected functional form or by investigation of different functional forms within one of the existing covariate model building strategies. Amongst these covariate model building strategies are the stepwise covariate model (SCM) [145, 146] and its recent upgrade SCM+ [147], full covariate model [148–151], lasso [152], and machine learning methods [153, 154], to name a few. The choice of the method depended on the nature of the dataset and the intended use of the model. The SCM approach, the most common approach for covariate analysis, was adopted in *project II* as the covariate model building strategy and is therefore described in details in the next section.

Stepwise covariate model building

The SCM was a 2-step covariate model building approach which based covariate selection on statistical significance. Since the inclusion of covariates resulted in nested models, this allowed the covariate effect to be assessed through the extent of numerical model improvement i.e. change in -2 log likelihood, also known as objective function value (OFV), in relation to the number of added coefficients (i.e. degrees of freedom, *df*). The change in OFV also translated to a corresponding *p*-value for ease of statistical assessment (more details on numerical assessment in section 2.1.2.4).

The 2-step process of the SCM included forward inclusion and backward deletion (Fig. 2.3). In the forward inclusion process (Fig. 2.3, left), parameter-covariate relationships were added one at a time to the base model, and the parameter-covariate relationships which resulted in the most significant reduction in OFV, according to a pre-defined significance level as threshold, was retained in the model. In the subsequent round, all the remaining covariates were re-tested and the process was repeated until no added parameter-covariate relationship was associated with a significant reduction in OFV, resulting in the *full covariate model*.

Next, the full covariate model was subjected to backward deletion (Fig. 2.3, right) in which the model was tested for the removal of parameter-covariate relationship one at a time using a pre-defined stricter criteria to account for multiple testing. The eliminated parameter-covariate relationship, which did not produce a significant increase in OFV i.e. was not associated with a significant worsening of model performance, was removed from the model and the remaining covariates were re-tested for elimination, one at a time, until the removal of any of the remaining parameter-covariate relationships was associated with a significant increase in OFV (i.e. only parameter-covariate relationships that resulted in a significant increase in OFV were retained). This resulted in the *final covariate model*.

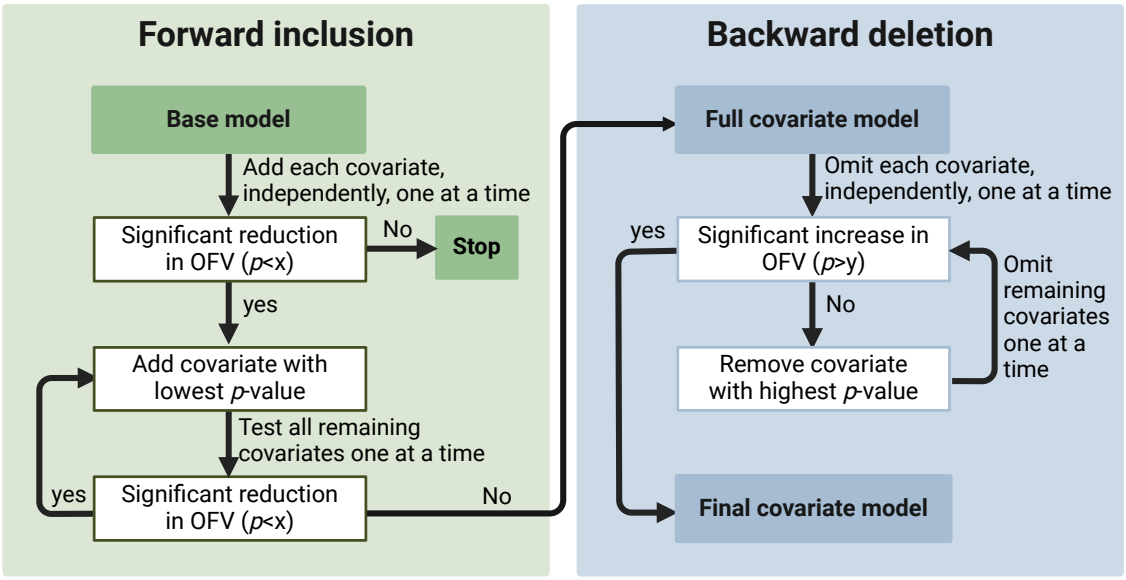

Figure 2.3: Overview of stepwise covariate model building. OFV: objective function value; x: significance level α for forward inclusion; y: significance level α for backward deletion. Figure created with BioRender.com.

Because covariate selection was mainly governed by the change in OFV (i.e. statistical significance), further refinement of the selected covariates based on precision of the quantified effect, clinical relevance, or scientific plausibility always followed.

Whenever the SCM approach was used in *project II*, a p-value of < 0.05 per $1\,df$ (corresponding to $\Delta\,\mathrm{OFV} > 3.84$) was set as the threshold for inclusion of a parameter-covariate relationship in the forward inclusion steps and a stricter criteria of p-value < 0.01 per $1\,df$ (corresponding to $\Delta\,\mathrm{OFV} > 6.64$) was set as threshold for exclusion of a parameter-covariate relationship in the backward deletion steps to retain covariates of highest impact in the final covariate model.

2.1.2.2 Parameter estimation and estimation methods

NLME modelling determined the parameters' values that best described the observed data. This was performed through the maximum likelihood approach, in which the aim was to determine the set of parameters (θ, Ω, Σ), given the developed model, that maximised the likelihood of observing the data.

The likelihood ($\mathcal{L}$) for a series of observations in a number of individuals was the product of the likelihood of individual observations. Thus, for n number of observations, $\mathcal{L}$ can be described as follows (Eq. 2.20),

$$\mathcal{L} = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-\frac{1}{2\sigma_i^2}(Y_i - \hat{Y}_i)^2} \tag{2.20}$$

where Y_i is the measured observation, $\hat{Y}_i$ is the predicted observation, and σ^2 is the variance of the model. By log transforming both sides and multiplying by -2, to simplify Eq. 2.20, the latter can be rewritten as follows (Eq. 2.21),

$$-2 \cdot log(\mathcal{L}) = n \log(2\pi) + \sum_{i=1}^{n} \left(log(\sigma_i^2) + \frac{(Y_i - \hat{Y}_i)^2}{\sigma_i^2} \right) \tag{2.21}$$

Therefore, based on Eq. 2.21, in order to maximise the likelihood, NONMEM® minimised the -2 $\cdot$ log likelihood (-2$\mathcal{LL}$, also known as the objective function value, OFV) such that the best set of parameters for a developed model i.e. maximum likelihood, translated to the lowest -2$\mathcal{LL}$ (OFV) [140, 155]. In practice, this was an iterative procedure, in which, as a first step, NONMEM® evaluated the likelihood given the set of parameters that the user had specified. Then in the second iteration, the parameter values were updated in the direction where the -2$\mathcal{LL}$ decreased. This process continued until the convergence criteria was reached in which the best set of parameter estimates, that no further minimised the -2$\mathcal{LL}$, had been found or alternatively the maximum number of iterations had been reached (set to 9999 iterations) which might not then—in this case—represent the best parameter set.

Because of the nonlinear dependency of the parameters on the random-effects (η), there was no analytical solution for Eq. 2.21 and the -2$\mathcal{LL}$ had to be approximated numerically for which different estimation methods existed. Amongst which were the first-order (FO) method and the conditional estimation methods: first-order conditional estimation (FOCE) method and Laplace method, which are summarised briefly below.

The FO method

The FO method was the first method to be introduced to NONMEM®. It approximated the solution of the likelihood function by linearisation of the model around $\eta_i = 0$ through the use of the first-order Taylor series expansion. This linear approximation was then used to only estimate the population parameters of the NLME model. Thus, to derive individual-specific parameter estimates, the final population model parameter estimates were later used as priors for the calculation of the EBEs (i.e. individual parameters) in a *post-hoc* step [115, 156].

The FOCE method

The FOCE method used the first-order Taylor series expansion to approximate the NLME model around the conditional estimates of η_i rather than zero. In contrast to the FO method, the EBEs were estimated after each iteration step, hence, influencing the final population parameter estimates.

The FOCE method with the interaction option (**FOCE-I**) further improved the accuracy of the estimation as it accounted for the interaction between IIV (η) and RUV (ϵ) [115, 156].

The Laplace method

The Laplace method was similar to the FOCE method except for that it used the second-order Taylor series expansion for approximation [115, 156].

In comparison, despite the simplicity of the FO method and the need for a short computation time, this method was only useful for sparse data and cases of small IIV e.g. non-repeated time-to-events with no IIV. Also, the assumption of $\eta_i = 0$, compromised the method's accuracy and could produce biased parameter estimates [157]. Whereas, the FOCE method, despite its longer computation time, was considered to be more accurate and more suitable for dense data and cases of high IIV, with the exception of the FOCE-I method that became more useful in cases of small RUV (ϵ). The Laplace method required longer computation time compared with the FOCE method; however, it was particularly useful in cases of categorical data and time-to-event modelling when IIV was considered e.g. repeated events, or for the analysis of continuous type data when LAPLACE INTER was defined.

In this thesis, the Laplace method was used for the (re-)estimations in *project I* to account for the heterogeneous endpoints and modelling of dropouts. In *project II*, the FOCE-I method was used for the development of the tumour dynamics model, the CRP model and the coupled tumour size - CRP model for more accurate parameter estimates, whereas the FO method was used in the time-to-event modelling framework for a shorter computation time due to the presence of a single event per patient (i.e. there was no need to incorporate IIV). Details on the estimation methods for *project I* are provided in section 2.5.2 and for *project II* are provided in sections 2.6.3.1 and 2.6.3.2.

2.1.2.3 Modelling approaches to handle baseline information

An inherent characteristic of our PD measurements was the availability of baseline observations. Baseline observations usually carried information that—if correctly utilised—better informed the model and avoided potential biases [158]. To handle PD baseline measurements, different modelling approaches had been previously reported which could be summarised to the following four methods: *B1 method* where the typical value and IIV of the baseline were estimated; *B2 method* where the observed baseline was included as a covariate while accounting for RUV; *B3 method* where both the observed baseline and typical baseline values were utilised and weighted by RUV and IIV, respectively; and *B4 method* where observations were normalised by the baseline value [158]. In this thesis, only the B1 and B2 methods were used due to their minimal bias and imprecision, and proven superiority [158] and hence, are further described below:

B1 method

In the B1 method, the typical value of the baseline and its IIV were estimated. Individual baselines were assumed to follow a log-normal distribution with a random variability having a mean of zero and a variance of ω^2. RUV in baseline data was treated as with non-baseline data. This is represented as follows (Eq. 2.22),

$$\phi_{BL,i} = \theta_{BL} \cdot e^{\eta_{BL,i}} \qquad \eta_{BL,i} \sim N\left(0, \omega^2\right) \tag{2.22}$$

where $\phi_{BL,i}$ is the individual baseline value, θ_{BL} is the typical baseline value and $\eta_{BL,i}$ is the IIV with a mean of zero and variance ω^2.

The B1 method was associated with the smallest bias and imprecision and therefore is regarded as the gold standard [158]. It has been applied in *project I* to characterise baseline tumour size information (section 2.5.1.2).

B2 method

In the B2 method, the observed baseline was included as a covariate and was assumed to be measured with the same variability as the other non-baseline measurements. Individual baselines were estimated to deviate from the observed individual baseline by a random variable which had the same magnitude as the RUV and was derived from non-baseline data. Therefore, both baseline and non-baseline data had the same variability. This is represented as follows (Eq. 2.23),

$$\phi_{BL,i} = BL_{i,o} \cdot e^{\eta_{BL,i} \cdot \theta_{RUV}} \qquad \eta_{BL,i} \sim N(0,1) \tag{2.23}$$

where $\phi_{BL,i}$ is the individual baseline value; $BL_{i,o}$ is the observed individual baseline value; θ_{RUV} is the RUV on a SD scale; and $(\eta_{BL,i} \cdot \theta_{RUV})$, in which the variance of $\eta_{BL,i}$ is fixed to 1 and scaled by θ_{RUV}, is the random component constrained to have the same variance as the RUV (σ^2) and defining the deviation of $\phi_{BL,i}$ from $BL_{i,o}$.

The B2 method was not recommended when RUV was high and/or number of observations per individual was low. Nevertheless, it was useful when the baseline distribution was difficult to capture as it did not involve any assumptions about the shape of the distribution of the baseline values in the population [158]. The B2 method has been applied in *project II* to characterise baseline tumour size information (section 2.6.3.1).

2.1.2.4 Model selection and evaluation

During the model development process, different structural models, RUV functional forms, and parameter-covariate relationships were usually explored. Therefore, criteria were needed for (i) comparison between competing models to select the best model in terms of precision in describing the data, robustness, and predictive performance; and (ii) evaluation of the final model to make sure it suited its intended use and described the data well.

According to the parsimony principle, if two models of different complexities described the data equally well, then the less complex model was usually considered. Consequently, model selection and evaluation did not only seek to find the "best" model describing the data but also balanced the decision with regards to parameter estimates' plausibility and precision, intended purpose of the model, quality of its predictions, and the parsimony principle.

Thus, to aid the comparison and selection of the most appropriate model, several numerical and graphical methods were applied. These methods could be further classified into internal and external methods (Fig. 2.4). Internal methods described evaluation methods that used the same dataset as that which was used for model development, whereas the external methods described methods which used external datasets having similar characteristics to that on which model development was based. As it was not usually feasible to obtain additional data with similar characteristics from the clinical setting, only internal evaluation methods were applied in the work presented in this thesis and are therefore discussed in details in the following sections.

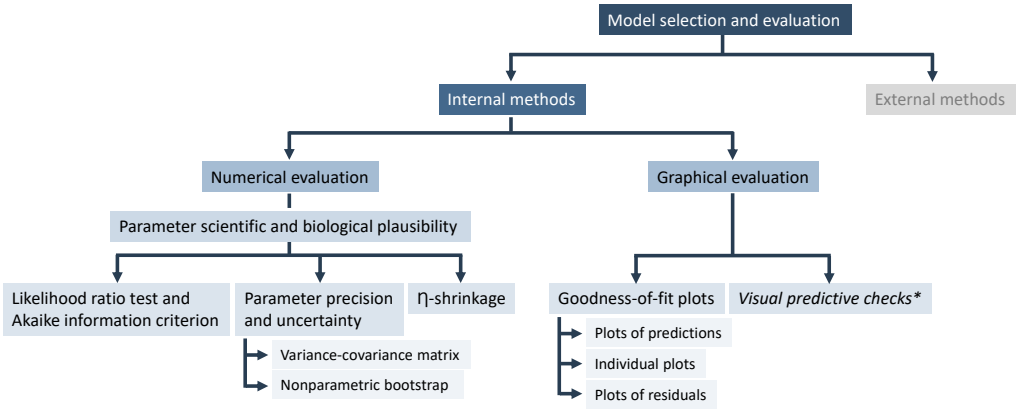

Figure 2.4: Overview of the different model selection and evaluation techniques
* simulation-based

Numerical evaluation criteria

As a first evaluation step, the parameter estimates were assessed for scientific and biological plausibility. Subsequently, model selection and assessment of the precision of the estimated parameters followed, using the different numerical evaluation methods.

Likelihood ratio test and Akaike information criterion

The likelihood ratio test (LRT) and the Akaike information criterion (AIC) both represented tools to assess model performance, and support model comparison and selection based on the OFV [140, 155, 159].

The LRT was a test for statistical significance [155] and assessed a model's superiority and better fit to the data by assessment of the significance of model improvement/worsening with respect to the added/removed model parameter(s) compared to a reference model, based on the ratio of their likelihoods. However, the LRT was only applicable to nested models. Models were said to be nested if one model was the subset of the other [140] i.e. setting parameters of the more complex model to the null hypothesis values collapses the model to the simpler one (e.g. covariate model vs base model). Since the OFV was defined as twice the negative log-likelihood ($-2LL$, Eq. 2.24) (section 2.1.2.2), the LRT between two models could be also represented as the difference between OFV or referred to as Δ OFV (Eqs. 2.24 and 2.25).

$$OFV = -2LL \tag{2.24}$$

$$\begin{aligned}\Delta OFV(LRT) &= OFV_{complex} - OFV_{simple} \\ &= -2 \cdot (LL_{simple} - LL_{complex}) = \frac{L_{simple}}{L_{complex}}\end{aligned} \tag{2.25}$$

The distribution of the likelihood ratio is approximately χ^2 distributed, with n df equal to the difference in the number of estimated parameters between the two models. Therefore, the threshold for significance was usually defined relative to the number of df and significance level, α. If the Δ OFV was larger than the defined threshold, then the model with additional parameter has provided a significantly improved description of the data (e.g. at $\alpha = 0.05$ and $df = 1$, a model was deemed superior if the drop in OFV was > 3.84 per addition of one parameter) [159].

Alternatively, for non-nested models, the AIC was computed (Eq. 2.26), in which a penalty term (n_p) accounting for the total number of estimated parameters was included.

$$AIC = OFV + 2 \cdot n_p \tag{2.26}$$

When comparing two models, the lower the AIC value, the better the model. A difference of ≥ 2 points in AIC was considered as a threshold for model comparison, and selection over another [140].

Parameter precision and uncertainty

The precision of the parameter estimate was commonly evaluated as relative standard error (RSE) or confidence intervals (CI). Summarised below are the different methods which derived these metrics and which differed in computation time, complexity and the underlying assumptions:

Standard errors from the variance-covariance matrix

Precision of each parameter estimate could be obtained from NONMEM® through the variance-covariance matrix of the parameter estimates. The square root of the diagonal elements of this matrix represented the standard errors (SE) of the model parameters which, for easier interpretation, were presented in relation to the parameter estimate, as RSE (Eq. 2.27). For the random-effects parameters, the RSE was transformed to the SD scale as expressed in Eq. 2.28, since these parameters were usually reported as CV%. Therefore, for a fixed-effects parameter θ_k and a random-effects parameter ω_k^2, RSE can be represented as follows [142],

$$RSE_{\theta_k} = \frac{SE_{\theta_k}}{\theta_k} \cdot 100 \tag{2.27}$$

$$RSE_{\omega_k^2,CV\%} = \frac{SE_{\omega_k^2}}{2 \cdot \omega_k^2} \cdot 100 \tag{2.28}$$

An RSE of approximately $< 30\%$ for the fixed-effect parameters and approximately $< 50\%$ for the random-effects parameters were considered acceptable. Yet, different limits could be specified provided that they were appropriately justified e.g. smaller sample sizes or low information content in the dataset [140].

Additionally, the SE could be used to calculate the CIs (Eq. 2.29). However, this was based on the assumption that the number of individuals used in the estimation was large and that SE were asymptotically normally distributed and consequently that CIs were symmetric. Unfortunately, the number of individuals was usually not large enough to fulfil this assumption, which would have lead to the underestimation of the parameter uncertainty [140]. Moreover, NONMEM® could sometimes fail to generate the variance-covariance matrix especially for complex models. Therefore, alternative methods which did not assume a normal distribution of the parameter were preferred to derive CIs e.g. the nonparametric bootstrap.

$$95\%CI = \theta \pm 1.96 \cdot SE \tag{2.29}$$

Nonparametric bootstrap

The nonparametric bootstrap has been often viewed as the gold standard to assess model robustness and obtain information on parameter precision and accuracy. In contrast to the variance-covariance matrix which assumed a normal distribution of uncertainty, the nonparametric bootstrap lacked such a parametric assumption [140].

The process involved generation of a large number of replicate datasets by repeated random sampling from the original dataset, with replacement, to result in multiple datasets of the same size (i.e. number of individuals) as the original dataset. Then, the model was fitted to each of these replicate datasets and parameters were estimated [160].

Based on the models with successful minimisation, different descriptive statistics were calculated for each parameter estimate across the bootstrap replicates e.g. mean, median, and SE. The 90% CIs could also be obtained from the 5^{th} and 95^{th} percentiles of the parameter estimates across the bootstrap replicates. The SE could be used along with the parameter estimate to calculate the RSE whereas the median and CI could then be compared to the original parameter estimates to identify the parameter's accuracy and precision, respectively. Additionally, the convergence rate (i.e. percentage of successful runs to the total number of runs) was reported as a measure of model robustness [160].

The number of bootstrap samples depended on the intended purpose e.g. for calculation of CI, $n_{samples}$ should be $\approx$1000 [140]. An increase in the number of samples was associated with better precision, more representative parameter estimates to the true value, and narrower CI. Yet, performing a nonparametric bootstrap was computationally intensive [161].

η-shrinkage

η-shrinkage was a measure of how informative the data were and consequently the reliability of the diagnostic plots. In the presence of sparse data, individual data would be of little influence when it comes to the estimation of individual parameters and consequently, they would tend to 'shrink' towards the population values. As a result, individual predictions would shrink towards the corresponding observations and the individual IIV distribution (η_i) would shrink towards zero [162].

For parameter k, calculation of η-shrinkage can be described as in Eq. 2.30,

$$\eta_k - shrinkage[\%] = 1 - \frac{SD(EBE_{\eta_{k,i}})}{\sqrt{{\omega_k}^2}} \cdot 100 \tag{2.30}$$

where SD is the standard deviation of the individual values of the Empirical Bayesian Estimates of $\eta_{k,i}$ ($EBE_{\eta_{k,i}}$), and ${\omega_k}^2$ is the variance.

The significance of η-shrinkage would arise when its value is high ($>30\%$) as it affected both model interpretation and evaluation. On the parameter level, shrinkage might lead to biased individual parameter estimates and altered (masked or induced) correlations between parameters. Similarly, on the covariate level, the relation between EBEs or $\eta_{i,k}$ and covariates might be masked, induced, or distorted. Moreover, diagnostic plots involving individual predictions i.e. EBE-based diagnostics and conditional weighted residuals would be often misleading.

Therefore, if η-shrinkage was high, model evaluation was primarily based on OFV and the predicted population parameters. Otherwise, it was reported as a measure of the reliability of the parameter estimates, the diagnostic plots, and the correlations between parameters and covariates [140, 162].

Graphical evaluation criteria

Graphical evaluation aimed to assess model performance and predictivity to aid model selection and assessment. The different methods utilised in this thesis are described below:

Goodness-of-fit plots

Goodness-of-fit (GOF) plots represented a group of commonly used graphical analyses that evaluated the model predictions compared to the observed data and aided the selection of the appropriate model by observing the adequacy of data description. GOF plots could be generated for the whole population, or stratified per patient group or a specific explanatory variable e.g. covariate. Thus, they represented a powerful tool to diagnose potential model misspecifications and biases throughout the whole model development process [140, 142]. The most commonly used plots are highlighted below:

Plots of observations versus predictions

A scatterplot of observations against the population predictions (PRED) or individual predictions (IPRED) was plotted. Ideally, data points were expected to show an even, random, and uniform distribution on either side of the line of identity (intercept = 0, slope = 1) and across the range of observations. Observed systematic deviations or biases could be suggestive of model misspecifications/misfit. Generally, observations versus PRED showed more spread of data points around the line of identity compared with the observations versus IPRED in which data points were more clustered and fell closer to the line of identity as a result of accounting for individual-specific random-effects [142].

Plots of individual observations and predictions versus time

These plots overlaid the observations, PRED, and IPRED against the independent variable (e.g. time) per individual. Ideally, IPRED had to be as close as possible to the observations whereas PRED had to capture the central tendency of the data, representing a typical individual [142]. Besides investigating the appropriateness of model predictions relative to observations over time, these plots also allowed to identify poorly characterised individuals, hence, guiding further model optimisation by considering inclusion of covariates, as an example.

Plots of residuals

Residuals were calculated as the difference between observations and predictions. They are commonly weighted to account for the magnitude of data and ease interpretation. Conditional weighted residuals (CWRES), an indicator of model performance, were weighted residuals calculated based on the FOCE approximation method (section 2.1.2.2) [163]. They followed a normal distribution with a mean of 0 and a variance of 1. Commonly, CWRES versus time or PRED were plotted. In these residual plots, data points had to be evenly and randomly distributed around the line of identity (intercept = 0, slope = 0), within a range of approximately ± 3 SD of CWRES, across the entire range. Data points beyond approximately ± 3 SD of CWRES were considered to be less well-predicted.

Systematic trends in the plot of CWRES versus time could identify structural model misspecifications, while in the CWRES versus PRED plot, inappropriately characterised RUV could be identified [140, 142].

Visual-predictive checks

The visual predictive check (VPC) was a simulation-based graphical diagnostic that aimed to test the predictive performance of the model based on the assumption that data simulated from the model would have similar characteristics to the original observed data upon which the model was developed.

The procedure involved simulating a large number of datasets (in this thesis, $n_{replicates} = 250$ [164]) according to the same study design, and using the model of interest and the final parameter estimates while maintaining the same variability as that described by the model (i.e. stochastic simulation, section 2.1.2.5). The 5^{th}, 50^{th} and 95^{th} percentiles (corresponding to median and 90% prediction interval) of the observed and simulated data were then computed. The percentiles of the simulated data along with their 95% CI (typically shown as a shaded area) were then overlaid against the corresponding percentiles of the original observed data and plotted relative to the independent variable, usually time, for graphical comparison. Ideally, percentiles of the observed data had to overlay the respective percentiles of the simulated data. Nevertheless, sufficient prediction was achieved if the percentiles of the observed data fell within the 95% CI of the simulated percentiles [165].

As individuals varied in their actual sampling times, during the VPC procedure observations/simulations were grouped into time windows (bins) for which percentiles and prediction interval were then generated based on these bins for a less erratic-looking profile. A smooth time course then facilitated an easier interpretation of the model performance and identification of potential misspecifications. Number and boundaries of bins could be manually set to include equal distribution of observations, equal time intervals, or alternatively allowed to be automatically computed [160]. In this thesis, the number of bins was set manually to ensure a balanced number of observations within each bin.

2.1.2.5 Simulations

Once a model has been developed and evaluated, simulations could be performed to explore the model's performance under different scenarios e.g. explore the impact of different dose-response relationships on a virtual population as in *project I* or explore the impact of specific covariate levels or different combination of covariates on model parameters as in *project II*. Based on presence/absence of random-effects, simulation could be either *deterministic* or *stochastic*.

Deterministic simulations

In deterministic simulations, the typical profile of an individual was simulated given the design variables e.g. covariates and time but without considering any sources of variability i.e. only the fixed-effects parameters were considered. Thus, allowing to explore the general time course and assess the impact of inclusion of e.g. covariates on the typical profile. In *project II*, deterministic simulations were performed to explore the impact of predictors on the median time-to-events e.g. impact of various levels of CRP concentrations on the typical survival profile (section 3.2.2.3).

Stochastic simulations

In stochastic simulations, fixed- and random-effects were included to represent the virtual patient population. Stochastic simulations were not only used to judge the model's predictive performance (section 2.1.2.4), but also explore the *range* of the expected outcome given all the model parameters and under different scenarios e.g. clinical trial design, specific patient population, different dosing regimens, or as explored in *project I*: specific study design and/or dose-response relationships. The simulated individuals were generated by randomly sampling individual parameters from the respective parameter distribution using a random number generator i.e. for each parameter with IIV an η-value was randomly sampled from the elements of the omega matrix (Ω) and for every observation, an ϵ-value was randomly sampled from the elements of the sigma matrix (Σ) [166].

Although the higher the number of simulations, the better; we were constrained by the computation time due to model complexity, data size and computational capacity.

2.2 Survival analysis

In survival analysis, the main focus was the time taken for an event of interest to occur. When performing a survival analysis, we needed to define two time points; the start of observation and the time of occurrence of the event of interest. Although this type of analysis, in which time of occurrence of the event is of interest, is traditionally called "survival analysis", it is rather a general analysis concept for any event (i.e. not necessary death) and in a broader context could be a non-survival-related event e.g. time-to-occurrence of adverse event or discharge from hospital. In the context of this thesis, the event of interest was related to the commonly used efficacy endpoints in oncology: time-to-progression (TTP), progression-free survival (PFS), and overall survival (OS)—which have been previously defined in section 1.3.2.6.

Survival analysis was governed by two main—and related—functions: the survival function and the hazard function.

Survival function

The survival function ($S(t)$) defined the probability (P) of the survival time (T) of an individual (i.e. not having the event of interest) beyond time t (Eq. 2.31).

$$S(t) = P(T > t) \tag{2.31}$$

Hazard function

The hazard function ($h(t)$) described the instantaneous risk of an event occurring at time t, conditional on survival up to that time. The integral of the hazard with respect to time i.e. from time zero (e.g. start of observation) to time (t) is known as the cumulative hazard function ($H(t)$, Eq. 2.32) and described the cumulative risk of an event occurring up to time t.

$$H(t) = \int_0^t h(t)dt \tag{2.32}$$

Exponentiation of the negative of the $H(t)$ related $S(t)$ to $h(t)$ (Eq. 2.33), whereas the product of $S(t)$ and $h(t)$ yielded the probability density function (PDF) which described the likelihood of observing an event at a particular time t (Eq. 2.34).

$$S(t) = e^{-H(t)} \tag{2.33}$$

$$PDF(t) = h(t) \cdot S(t) \tag{2.34}$$

Different techniques exist for survival analysis and could be broadly classified into: nonparametric analysis (Kaplan-Meier analysis) which is purely descriptive of the observed events; parametric time-to-event analysis which characterises the shape of the hazard underlying the distribution of the observed events (i.e. increasing risk, decreasing risk, constant risk); and semi-parametric analysis (Cox proportional hazards model) which is considered an intermediate between nonparametric and parametric analyses. Nonetheless, censoring, a key characteristic feature of survival data, had to be primarily accounted for to avoid bias [167] (section 2.2.1).

2.2.1 Censoring

An important feature to be considered when performing a survival analysis was censoring. Censoring was said to exist if the exact time to occurrence of the event was unknown. Thus, an individual was censored if the exact time-to-event could not be precisely identified. Different types of censoring exist depending on when the event was likely to occur [168]; however, *right censoring* in which the event did not occur until the end of the observation time represented the most common and relevant type of censoring in oncology clinical studies and our here presented work. Right censoring could occur due to patient not developing the event until the end of observation time, being lost to follow-up, or dropping out before study termination.

2.2.2 Nonparametric analysis: Kaplan-Meier analysis

The Kaplan-Meier analysis was a nonparametric statistical method for the estimation of the probability of an event occurring as a function of time while accounting for the presence of censored individuals [169, 170].

The Kaplan-Meier analysis was based on three main assumptions [170]: (i) censored individuals had the same probability of having the event of interest in the future as uncensored individuals, (ii) probability of having the event was the same regardless of the start of the observation time, and (iii) the exact time of the occurrence of the event of interest was known [170].

For an event of interest being e.g. death, at each time interval (t_i), the survival function ($S(t_i)$) i.e. the probability of being alive until the time interval (t_i), was computed as follows (Eq. 2.35),

$$\begin{aligned} S(t_i) &= \frac{\text{number of surviving individuals at } t_i}{\text{number of individuals at risk at } t_i} \\ &= \frac{\text{number of individuals at risk at } t_i - \text{number of individuals who died at } t_i}{\text{number of individuals at risk at } t_i} \end{aligned} \tag{2.35}$$

Censored individuals, if present, were not considered at risk and were thus excluded from the number of individuals at risk. The total survival probability until any given time e.g. survival until time interval k (i.e. $S(t_k)$), was then calculated as the product of all survival probabilities at all the preceding time intervals $(1, ..., k)$ (Eq. 2.36) [169, 170].

$$S(t_k) = S(t_1) \times S(t_2) \times ... \times S(t_k) \tag{2.36}$$

Results of the Kaplan-Meier analysis were visualised graphically (Fig. 2.5, A). Probability of an event (i.e. $S(t)$) was plotted as a step function versus the observation time, which was from a defined point e.g. time of treatment start/patient enrolment until follow-up/study termination, and censored individuals were marked with e.g. a cross or vertical line. The median survival time (Fig. 2.5, dashed lines), a parameter of interest commonly derived from the Kaplan-Meier plots, was defined as the time at which probability of survival was 50% [170].

Assessment of the potential influence of different covariates on survival could only be explored by stratification of the Kaplan-Meier plots with respect to the different categories of the covariate of interest (Fig. 2.5, B). Strata were then compared for presence of a statistical significant difference using the log-rank statistical test (section 2.3.2) [170]. Thus, a disadvantage of Kaplan-Meier analysis was that it could look into the impact of only one potential covariate without taking other variables into account and that this was handled only through stratification which often lead to loss of power. Stratification worked best in case of categorical covariates. However, it became challenging to stratify continuous covariates, e.g. biomarker concentrations, which were then dichotomised based on a descriptive statistic e.g. median (Fig. 2.5, B). This however came at the expense of losing information which was inherent to the nature of continuous data.

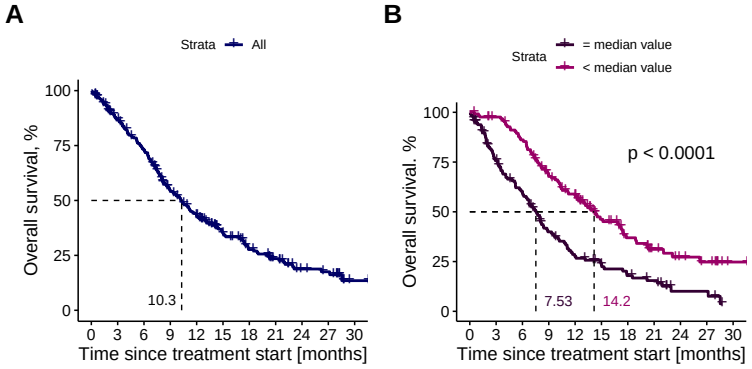

Figure 2.5: Kaplan-Meier curve of overall survival for **[A]** all patients i.e. unstratified and for **[B]** patients stratified by median baseline C-reactive protein (CRP) concentration based on the dataset of *project II*. Dashed lines: median overall survival; p-value: result of the log-rank test (section 2.3.2) to assess if survival is significantly different between the subpopulation with baseline CRP concentration below (pink) and above (purple) median baseline CRP concentration; Cross: censored individuals corresponding to the time of the patient's last participation in the study.

Moreover, as a nonparametric approach, the analysis was mainly descriptive with no knowledge on the underlying distribution of the survival data i.e. no assumption on the baseline hazard was made. Thus, it was difficult to extrapolate the results beyond the observed events or perform further simulations. Nevertheless, Kaplan-Meier analyses are regularly performed to statistically compare treatment outcome or efficacy between different study arms or patient groups and was therefore applied in *project II* to graphically explore our observed data.

2.2.3 Semi-parametric analysis: Cox proportional hazards model

In contrast to the Kaplan-Meier analysis, the Cox proportional hazards model (CoxPH) had the advantage of handling both continuous and categorical covariates as well as quantifying the impact of multiple explanatory variables, i.e. covariates, such as disease stage and baseline CRP concentration, simultaneously. In a CoxPH model, the association between the response variable, represented by the hazard, $h(t)$, and the predictor variables, represented by the different covariates (x), was investigated and described as follows (Eq. 2.37),

$$h(t) = h_o(t) \cdot e^{\beta_1 \cdot x_1 + \beta_2 \cdot x_2 + ... + \beta_n \cdot x_n} \tag{2.37}$$

where $h_o(t)$ is the baseline hazard at time t when all the explanatory variables $(x_1, ..., x_n)$ i.e. covariates are zero, and β_1,...,β_n are the coefficients of the effect size of each respective covariate.

For two groups i and ii that differed in their covariate value (x) i.e. x_i and x_{ii}, respectively, exponentiation of the coefficient, β, resulted in the hazard ratio (HR) which described the effect size of the covariate x and was defined as the ratio between the hazard of two groups if the associated covariate changed by one unit, given that all other covariates remained constant (Eq. 2.38).

$$HR = \frac{h_i(t)}{h_{ii}(t)} = \frac{h_o(t) \cdot e^{\beta \cdot x_i}}{h_o(t) \cdot e^{\beta \cdot x_{ii}}} = e^{\beta(x_i - x_{ii})} \tag{2.38}$$

A HR > 1 indicated a higher risk with the increase in the covariate value i.e. the covariate positively correlated with an increase in hazard (bad prognostic factor) while a HR < 1 indicated lower risk with increase in the covariate value (good prognostic factor) [171, 172].

For a valid CoxPH result, specific assumptions had to be initially checked: (i) hazards in the different groups were proportional i.e. survival curves did not cross (Fig. 2.5, B), (ii) effect of covariate did not change with time i.e. constant HR over time, and (iii) linear association existed between log hazard and covariate [173, 174]. The overall significance of the model was then assessed by the LRT in which the difference between the log likelihood of the reduced model (without the covariate) and the log likelihood of the full model (with the covariates) was computed. The likelihood-ratio was approximately χ^2 distributed with *df* equal to the number of additional covariates. A p-value < 0.05 indicated a significant impact of the covariate on the hazard [171].

Although the CoxPH quantified the magnitude of the covariate effect on the hazard, the estimates were always normalised to a reference arm i.e. estimates were relative rather than absolute (Eq. 2.38). Moreover, there was no assumption about the shape of the baseline hazard function. Therefore, similar to the Kaplan-Meier analysis, it was not suited for simulations and was applied in *project II* to only explore our observed data.

2.2.4 Parametric time-to-event analysis

Parametric time-to-event analysis (TTE) represented the optimal way to characterise survival data. Besides accounting for censoring, TTE analysis defined the underlying baseline hazard function and allowed to quantitatively estimate the impact of predictors (i.e. covariates) on the risk of occurrence of the event of interest. Consequently, the developed TTE models were used for simulations under different conditions in which the endpoint was the TTE. To characterise the distribution of the event of interest, different hazard functions were explored. These included [175]:

Constant hazard function

This hazard function represented the simplest form, where the hazard (baseline) was constant over time (λ), i.e. time independent (Fig. 2.6, A). The constant hazard function and the survival function are represented in Eq. 2.39 and Eq. 2.40, respectively,

$$h(t) = \lambda \tag{2.39}$$

$$S(t) = e^{-\lambda t} \tag{2.40}$$

where λ is the scale parameter (i.e. baseline hazard). Under the constant hazard function, the PDF (Eq. 2.34) showed an exponential distribution, hence the term *exponential* model.

Weibull hazard function

The Weibull hazard function allowed for the change in hazard over time. It was parameterised by both a scale parameter (λ) and a shape parameter (α). When $\alpha > 1$, the hazard would be increasing over time whereas at $\alpha < 1$, the hazard would be decreasing over time (Fig. 2.6, B). Eq. 2.41 and Eq. 2.42 represent the Weibull hazard and survival functions, respectively.

$$h(t) = \lambda\alpha(\lambda t)^{\alpha-1} \tag{2.41}$$

$$S(t) = e^{-(\lambda t)^{\alpha}} \tag{2.42}$$

Gompertz hazard function

As with the Weibull hazard function, the Gompertz hazard function also allowed for the change in hazard over time. It was parameterised by both a scale parameter (λ) following the Gompertz distribution and a shape parameter (γ). When $\gamma > 0$, the hazard would be increasing over time whereas at $\gamma < 0$, the hazard would be decreasing over time (Fig. 2.6, C). Eq. 2.43 and Eq. 2.44 represent the Gompertz hazard and survival functions, respectively.

$$h(t) = \lambda e^{\gamma t} \tag{2.43}$$

$$S(t) = e^{-\frac{\lambda}{\gamma}(e^{\gamma t-1})} \tag{2.44}$$

Log-normal hazard function

In the log-normal hazard function, the hazard initially increased to a maximum, then decreased resulting in a log-normal distribution of the event time (Fig. 2.6, D) [175]. It was parameterised by the mean (μ) and standard deviation (σ) as well as the standard normal cumulative distribution function (ϕ). Eq. 2.45 and Eq. 2.46 represent the log-normal hazard and survival functions, respectively.

$$h(t) = \frac{1}{t \cdot \sqrt{2\pi\sigma^2}} \cdot \frac{e^{-\frac{1}{2}\cdot(\frac{log(t)-\mu}{\sigma})^2}}{1-\phi(\frac{log(t)-\mu}{\sigma})} \tag{2.45}$$

$$S(t) = 1 - \phi(\frac{log(t)-\mu}{\sigma}) \tag{2.46}$$

Log-logistic hazard function

Similar to the Weibull hazard function, the log-logistic hazard function allowed for the change in hazard over time. It was parameterised by both a scale parameter (λ) and a shape parameter (α). When $\alpha > 1$, the hazard would be increasing over time to a maximum whereas at $\alpha < 1$, the hazard would be decreasing over time (Fig. 2.6, E). Eq. 2.47 and Eq. 2.48 represent the log-logistic hazard and survival functions, respectively.

$$h(t) = \frac{\lambda\alpha t^{\alpha-1}}{1+(\lambda t)^{\alpha}} \tag{2.47}$$

$$S(t) = \frac{1}{1+(\lambda t)^{\alpha}} \tag{2.48}$$

The shape of the hazard over time for the previously described hazard functions is depicted in Fig. 2.6, showing the time independency for the constant hazard function (Fig. 2.6, A) and the variation in hazard over time for the Weibull, Gompertz, log-normal and log-logistic hazard functions (Fig. 2.6, B–E).

To explore the predictive value of the different covariates on the event of interest, covariates were included on the hazard function to represent a change in hazard. Eq. 2.49 and 2.50 show the different ways of covariate inclusion, exemplified by the Weibull hazard function in which the covariate (COV) could either be added on the hazard function, as a whole (Eq. 2.49), or on specific parameter(s) e.g. scale parameter (λ) of the hazard function (Eq. 2.50).

$$h(t) = \lambda\alpha(\lambda t)^{\alpha-1} \cdot COV \tag{2.49}$$

$$\lambda_{cov} = \lambda_o \cdot COV \tag{2.50}$$
$$h(t) = \lambda_{cov}\alpha(\lambda_{cov} t)^{\alpha-1}$$

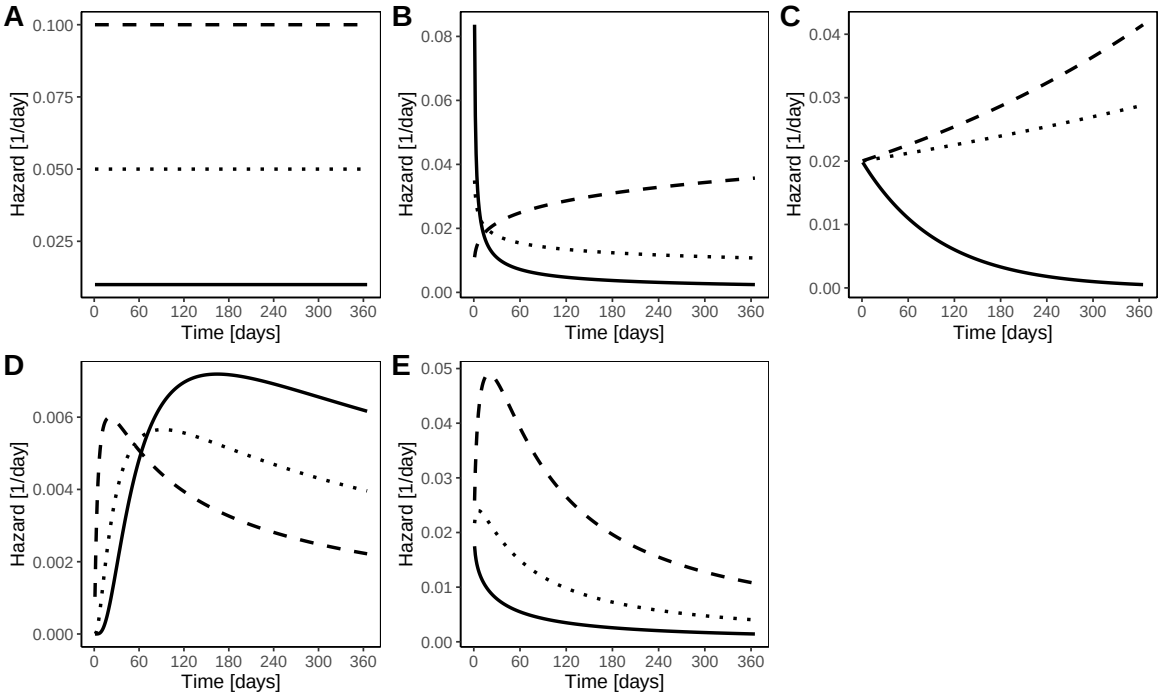

Figure 2.6: Profiles of hazard over time for the different hazard functions: **[A]** Constant, **[B]** Weibull, **[C]** Gompertz, **[D]** Log-normal, and **[E]** Log-logistic.
λ: scale parameter of the different hazard functions; α: shape parameter of the Weibull and log-logistic hazard functions; γ: shape parameter of the Gompertz hazard function, μ: mean parameter for the log-normal function, σ: standard deviation parameter of the log-normal hazard function.
[A] For the constant hazard model: solid line ($\lambda = 0.01$), dotted line ($\lambda = 0.05$), dashed line ($\lambda = 0.1$).
[B] For the Weibull hazard model: $\lambda = 0.02$; solid line ($\alpha = 0.4$), dotted line ($\alpha = 0.8$), dashed line ($\alpha = 1.2$).
[C] For the Gompertz hazard model: $\lambda = 0.02$; solid line (γ = -0.01), dotted line ($\gamma = 0.001$), dashed line ($\gamma = 0.002$).
[D] For the log-normal hazard model: μ=5; solid line ($\sigma = 0.75$), dotted line ($\sigma = 1$), dashed line ($\sigma = 1.5$).
[E] For the log-logistic hazard model: $\lambda = 0.02$; solid line ($\alpha = 0.9$), dotted line ($\alpha = 1.1$), dashed line ($\alpha = 1.3$).

To account for the impact of post-baseline covariate values, a special type of survival analysis known as *Landmark analysis* was performed. In the landmark analysis, a time point was selected *after* the start of the observation (known as the landmark time) at which the post-baseline covariate was available and from which follow-up was calculated and survival analysis was conducted. Therefore, only patients who have not yet had an event i.e. still at risk/surviving or had not been right censored until this landmark time were included in the analysis [176, 177]. Thus, with this approach the post-baseline covariate of interest would be already known/available/measured for the patients-at-risk by the landmark time.

Although, evaluation and selection of TTE models could be based on the OFV or AIC (section 2.1.2.4), assessments of their appropriateness and predictive performance were mainly based on simulations in which a VPC using Kaplan-Meier representation was investigated. In the Kaplan-Meier VPC (KM VPC), which had the advantage of handling censoring, simulations are performed in which the time of the occurrence of the event, for each individual, was predicted using the final model parameters and the same trial design as the dataset. For a realistic scenario that would allow simulated events to occur at any possible observation time, simulations were performed at small time-steps. Kaplan-Meier survival probabilities were then calculated, over time, for the observed and simulated datasets and plotted on the same axis. The Kaplan-Meier profile of the observed data was then compared against the median profile of the simulated data and its associated CI, to evaluate model predictiveness.

In *project II*, nonparametric and semi-parametric analyses were applied to explore the impact of baseline CRP on the efficacy endpoints (section 2.6.2.4) and parametric TTE analysis was used to describe PFS and OS, identify and quantify potential prognostic predictors, and perform further simulations (section 2.6.3.2). Landmark analysis was applied to explore the impact of early non-baseline longitudinal covariates (section 2.6.3.2).

2.3 General statistics

2.3.1 Descriptive statistics

Descriptive statistics were used to summarise and describe properties of continuous data by measurements of central tendency, dispersion, and percentiles.

2.3.1.1 Measures of central tendency

These were single values that identified the central position of the data distribution. The measures of central tendencies included:

Mean

The mean was used in case of non-skewed distribution and in absence of outliers to describe the average value of the data. As described in Eq. 2.51, the mean was calculated as the sum of all observations in the data divided by the total number of observations.

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{2.51}$$

Median

The median was less affected by outliers. Therefore, it was preferred, over the mean, as a descriptor of central tendency in case of skewed data distribution and/or in presence of outliers. As described in Eq. 2.52, the median was defined as the value that divided the data into two equal halves when arranged according to magnitude. The median of an odd number of observations was given by the middle number while for an even number of observations, the median was given by the mean of the two middle observations.

$$Median = \begin{cases} x\left(\frac{n+1}{2}\right) & \text{if } n \text{ is odd} \\ \frac{x\left(\frac{n}{2}\right) + x\left(\frac{n}{2}+1\right)}{2} & \text{if } n \text{ is even} \end{cases} \tag{2.52}$$

2.3.1.2 Measures of dispersion

Measures of dispersion described the spread and variability of the data around the central value. They included:

Variance

The variance described how far data points deviated from the mean and was represented as the mean of the squared deviations from the mean (Eq. 2.53).

$$variance = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 \tag{2.53}$$

Standard deviation

The SD was defined as the square root of the variance (Eq. 2.54). Therefore, it was easier to interpret compared with the variance since it had the same unit as the observed values.

$$SD = \sqrt{variance} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2} \tag{2.54}$$

Coefficient of variation

The CV described the relative dispersion. It was defined as the ratio between the SD and the mean, expressed as a percentage (Eq. 2.55). Since it was unitless, it permitted the comparison between data with different scales of measurement.

$$CV\% = \frac{SD}{\bar{x}} \cdot 100 \tag{2.55}$$

Range

The range was defined as the interval between the minimum and maximum value in the data (Eq. 2.56). However, it could be misleading in presence of extreme outliers.

$$Range = [x_{minimum} - x_{maximum}] \tag{2.56}$$

Quartiles and interquartile range

Quartiles were values in the dataset that divided the data into four equal parts, denoted by Q1 (first quartile), Q2 (median), and Q3 (third quartile). The interquartile range (IQR) represented the difference between Q3 and Q1, in which the middle 50% of the distribution lay. The IQR had the advantage of not being influenced by outliers.

2.3.1.3 Percentiles

An x^{th} percentile was the value below which x percentage of the data fell i.e. if the value of the 15^{th} percentile was 10, it meant that 15% of the data were ≤ 10.

2.3.2 Inferential statistics

Inferential statistics were used to infer conclusions about the population of interest from the observed data. The inferential statistical test, the log-rank test, described in the next section, complemented the NLME modelling and simulation statistical principles which were applied to *Projects I* and *II* to test for statistical significance and which have been previously described in their respective sections 2.1.2.2 and 2.1.2.4.

Log-rank test

The log-rank test was an inferential statistical method which was used to test the null hypothesis that there was no difference in the survival probability between two groups at any time point.

The log-rank test is represented by Eqs. 2.57 to 2.59, in which the analysis was based on calculating the total numbers of observed events (e.g. death) in groups 1 and 2 (O_1 and O_2, respectively), and the total number of expected events in groups 1 and 2 (E_1 and E_2, respectively, Eq. 2.57).

$$\textit{Log-rank test statistic} = \frac{(O_1 - E_1)^2}{E_1} + \frac{(O_2 - E_2)^2}{E_2} \tag{2.57}$$

For group 2, the *total* number of expected events (E_2) was calculated as the sum of the expected number of events calculated at each time of event (Eq. 2.58). This expected number of events at time of event i was calculated as the risk of event at that time multiplied by the number of individuals at risk in that group. This is represented in Eq. 2.58,

$$\text{For a total of } n \text{ distinct time of events, } E_2 = \sum_{i=1}^{n} \frac{d_i}{r_i} \cdot r_{2i} \tag{2.58}$$

where d_i is the total number of observed events in both groups, r_i is the total number of individuals at risk in both groups, $\frac{d_i}{r_i}$ is the risk of event at that time and r_{2i} is the number of individuals at risk in group 2 i.e. alive or not having the event.

Subsequently, the total number of expected events in group 1 (E_1) was calculated as the difference between the total observed events in both groups and the total number of expected events in group 2 (E_2) (Eq. 2.59) [170, 172, 178].

$$E_1 = (O_1 + O_2) - E_2 \tag{2.59}$$

The Log-rank test statistic result was then compared with the χ^2 distribution at $df = 1$. A significant p-value meant rejecting the null hypothesis and thus the presence of a statistically significant difference in the survival probability between both groups. The log-rank test was used to only test for significance as it could not provide information on the magnitude of difference between groups [178].

2.4 Software

Modelling and simulation tasks were performed using NONMEM® in combination with Perl-speaks-NONMEM® (PsN) and aided by Piraña as a graphical user interface to maintain an overview of the modelling activities. Statistical analyses; dataset preparation, exploration, and visualisation during the exploratory data analysis; and post-processing and evaluation of results after model development were performed in Microsoft Excel and R aided with RStudio for a more friendly user interface. A summary of the different software used in the execution of the presented work is listed in Table 2.1. Moreover, the different functionalities and packages that were utilised in the respective software listed in Table 2.1 are summarised in Table 2.2 along with their intended roles.

Table 2.1: List of software used in the analyses presented in this thesis

Software	Version	Projects	Reference
NONMEM®	7.4.3	I, II	Icon Development Solutions, Ellicott City, MD, USA (https://www.iconplc.com/solutions/technologies/nonmem/)
PsN	4.8.1	I, II	Uppsala University, Uppsala, Sweden (uupharmacometrics.github.io/PsN)
Piraña	2.9.4	II	Certara Inc., Princeton, New Jersey, USA (www.certara.com/software/pirana-modeling-workbench)
R	3.5.3	I, II	The project for statistical computing. Vienna, Austria (www.CRAN.R-project.org)
R Studio Server Pro	1.4.1717-3	I	RStudio: Integrated Development for R, Boston, MA, USA. (www.rstudio.com)
R Studio	2021.09.2	II	RStudio: Integrated Development for R, Boston, MA, USA. (www.rstudio.com)
Microsoft Excel	2021	II	Microsoft Corporation. (https://office.microsoft.com/excel)

NONMEM: nonlinear mixed-effects modelling; PsN: Perl-speaks-NONMEM

Table 2.2: List of functionalities and packages used in specific software

Software	Role	Functionality/package	Version	Projects	Reference
R	Data modification and wrangling	dplyr	1.0.7	I	[179]
			1.1.2	II	
		plyr	1.8.4	II	[180]
		zoo	1.8.6	II	[181]
		reshape2	1.4.3	II	[182]
		tibble	3.2.1	II	[183]
		tidyr	1.1.4	I	[184]
			1.3.0	II	
		stringr	1.4.0	I	[185]
		magrittr	2.0.1	II	[186]
	Data visualisation	ggplot2	3.3.5	I	[187]
			3.4.2	II	
		xpose4	4.7.1	I	[188]
			4.6.1	II	
		ggpubr	0.4.0	I	[189]
			0.2.5	II	
		ggforce	0.3.1	II	[190]
		vctrs	0.6.2	II	[191]
		corrplot	0.84	II	[192]
		vpc	1.2.2	II	[193]
		grid	3.5.3	II	[194]
		ggtext	0.1.1	II	[195]
	Manage NONMEM® output files	GlimmeRcore	2.1.1	I	*
	MCP analysis	DoseFinding	1.0-5	I	[196]
	Survival analysis	survminer	0.4.7	II	[197]
		survival	3.2.12	II	[198, 199]
PsN	Covariate testing	scm functionality	—	II	[146]
	Parameter precision and uncertainty	bootstrap functionality	—	II	[200]
	Stochastic simulation and estimation	sse functionality	—	I	[201]
	Model predictivity	vpc functionality	—	II	[202]

*In-house package, not publicly available
MCP: Multiple Comparison Procedure; NONMEM: nonlinear mixed-effects modelling; PsN: Perl-speaks-NONMEM; scm: stepwise covariate model; sse: stochastic simulation and estimation; vpc: visual predictive check

2.5 Project I: Simulation-based evaluation of cLRT and MCP approaches in oncology under different clinical trial scenarios

To evaluate the performance of cLRT and MCP approaches within the oncology setting, data were simulated using a previously developed model (later referred to as the data-generating model) [203] and a traditional design of a phase II dose-finding study which considered a placebo arm. The simulated data considered the probability of dropping out, disease progression, and end of follow-up to produce a realistic scenario, and characterised the time course of tumour size as a function of drug dose, given a pre-specified dose-response relationship in the data-generating model. Hence, given the pre-defined study design, strength of drug effect, and the different dose-response relationships of the data-generating models, different datasets were simulated to reflect different scenarios. The following sections describe the simulation study (section 2.5.1), the principles of cLRT (section 2.5.2) and MCP (section 2.5.3) and their implementation, and the evaluation criteria used to assess and compare their performance (section 2.5.4).

2.5.1 Simulation study

The simulation study comprised the study design and the pharmacometric models. The different simulation scenarios were generated based on the different study design variables, drug strengths, and the dose-response relationships of the data-generating models. The details of these components are highlighted below.

2.5.1.1 Study design

A randomised, placebo-controlled, parallel group, phase II study design was simulated based on the different dose-response relationships of the data-generating models representing the true data-generating model. The study was designed with parallel dose groups including a placebo arm and active dose arms, with a balanced allocation in the number of patients per arm. Observation times were scheduled at baseline and every 7 days for 4 months (i.e. 112 days) (Fig. 2.7). Different sample sizes (i.e. total number of patients), number of dose groups, and strengths of drug effect were considered and defined the different scenarios to be investigated under the assumption of the different true data-generating dose-response models (section 2.5.1.3).

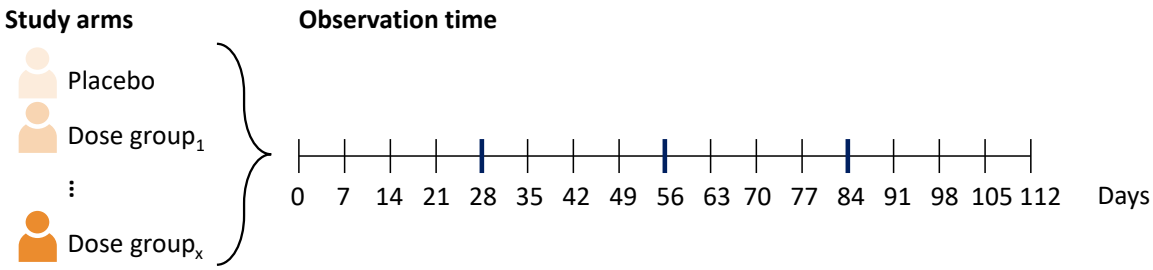

Figure 2.7: General design of the simulated placebo-controlled, parallel group, phase II studies to explore performance of combined likelihood ratio test (cLRT) and multiple comparison procedure (MCP)

2.5.1.2 Pharmacometric models

Data-generating models

The data-generating models (i.e. true models) utilised a previously developed longitudinal continuous tumour response model based on clinical data from phase III trials [203]. The model aimed to describe the sub-endpoints namely: the change in sum of longest diameters of the target lesions over time (i.e. tumour dynamics model), the probability of appearance of new lesions, the probability of progression of non-target lesions, and the death events prior to progression, as determinants of PFS (section 1.3.2.6), besides accounting for dropouts due to other reasons rather than progression.

A tumour dynamics model with an exponential tumour growth and a resistance component described the change in tumour size over time (Eq. 2.60),

$$\frac{dTS}{dt} = K_{growth} \cdot TS - \text{drug effect} \cdot TS \cdot e^{-\lambda \cdot t} \tag{2.60}$$

where K_{growth} was the first-order tumour growth rate constant, λ was the rate constant for appearance of resistance, TS was the tumour size, and drug effect was defined by the different dose-response relationships. The baseline tumour size was described by the B1 method (section 2.1.2.3) in which the typical value of the baseline and its IIV were estimated. Besides the baseline tumour size, IIV was accounted for on the tumour growth rate constant (K_{growth}), the rate constant for appearance of resistance (λ), and the drug effect component i.e. *slope* or E_{max} (Table 2.3).

A logistic regression model was used to describe the probability of the appearance of new lesions and the progression of non-target lesions, while TTE models (section 2.2.4) with an exponential (Eq. 2.39) and Weibull hazard functions (Eq. 2.41) described the death and dropout events, respectively [203].

Candidate models

A suitable set of plausible models likely to represent the dose-response relationships were identified to be tested against the simulated trial data and termed "candidate models". These candidate models were selected to cover the diverse set of plausible dose-response relationships—hence differing in the relationship of the dose-response—and represented a simplified version of the data-generating model. The simplified candidate models retained the description of the change in the sum of longest diameters and accounted for events occurring in the other sub-endpoints due to the progression of target lesion, the appearance of new lesions or death, as reasons for dropout from the study. These events were described by a parametric TTE model with an exponential hazard function (section 2.2.4, Eq. 2.39).

Dose-response relationships of data-generating and candidate models

Different relationships of the dose-response were considered for the data-generating true models and the candidate models. In the data-generating true models, to account for the drug effect (i.e. impact of the drug dose on tumour size), five different dose-response relationships were investigated on the tumour dynamics model, relating the drug dose to the drug effect (Table 2.3), for impact on the tumour size (Eq. 2.60). The different relationships included: flat (no dose-response), linear, log-linear, E_{max} and sigmoidal functions (Table 2.3). However, for the candidate models only the non-flat dose-response relationships were considered in which the flat relationship was regarded as the reference (i.e. $n_{candidate\ models} = 4$).

Table 2.3: Different dose-response relationships and their functional forms

Dose-response relationship	Model function (drug effect)
Flat	E_0
Linear	$E_0 + slope \cdot dose$
Log-linear	$E_o + slope \cdot ln(dose + 1)$
E_{max}	$E_0 + \frac{E_{max}}{ED_{50}+dose} \cdot$dose
Sigmoidal	$E_0 + \frac{E_{max}}{ED_{50}^{Hill}+dose^{Hill}} \cdot dose^{Hill}$

E_0 was fixed to zero i.e. no baseline
E_0: baseline drug effect; E_{max}: maximum drug effect; ED_{50}: dose at which drug effect is 50%; *slope*: slope of linear dose-response function; Hill: Hill coefficient.

Examples of the data-generating (true) model and the candidate model, based on a linear dose-response relationship are available in Appendix C, sections C.1 and C.2, respectively.

2.5.1.3 Simulation scenarios

The combinations of the different study design variables as well as the dose-response relationships and strength of drug effect defined the different simulation scenarios to be explored. These variables are depicted in Fig. 2.8 and are summarised in the following section.

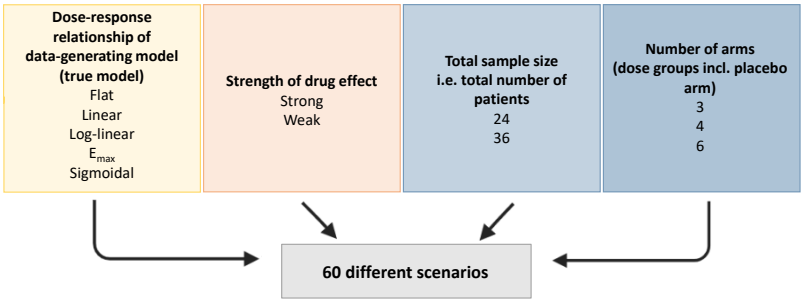

Figure 2.8: Variables defining the different simulation scenarios to explore performance of combined likelihood ratio test (cLRT) and multiple comparison procedure (MCP). Blue shade: study design-related variables.

Dose-response relationships of data-generating models

Different relationships of the dose-response in the data-generating true models were investigated as previously described in section 2.5.1.2.

Strength of drug effect

A total of two different strengths of drug effect were investigated. A drug effect, defined by the parameter estimates, that would result in ≈80% power per each dose-response relationship in the dose-response model was identified as a strong drug effect. Consequently, a weak drug effect was calculated as half the value of the parameter estimates defining the strong drug effect (i.e. slope parameter for the linear and log-linear relationships and E_{max} parameter for the E_{max} and sigmoidal relationships). The parameter estimates used for the simulations under the strong and weak drug effects, for each dose-response relationship of the data-generating true model, are presented in Table 2.4.

Table 2.4: Model specifications for the data-generating dose-response true models used in the simulations under a strong or weak drug effect

Parameter	Strong drug effect				Weak drug effect			
	Linear	Log-linear	E_{max}	Sigmoidal	Linear	Log-linear	E_{max}	Sigmoidal
Slope	1.8	25	—	—	0.9	12.5	—	—
E_{max}	—	—	100	95	—	—	50	47.5
ED_{50}	—	—	10	10	—	—	10	10
Hill	—	—	—	2	—	—	—	2

Note: values are arbitrary and reflect the conventional units i.e. slope $[effect \cdot dose^{-1}]$; E_{max} $[effect]$; ED_{50} $[dose]$; Hill [unitless].
E_{max}: maximum drug effect; ED_{50}: dose at which drug effect is 50%; Hill: Hill coefficient

Total sample size and number of arms

A total sample size of either 24 or 36 patients was considered to mimic a phase II study design. Based on a balanced allocation of patients per arm and according to the number of arms (3, 4 or 6), a unique sample size per arm was calculated per each group of scenarios differing only in the dose-response relationship of the data-generating true model and strength of drug effect (Table 2.5). The number of dose levels was selected to ensure an even distribution between a dose range of 0 mg (i.e. placebo) and 60 mg, defined according to the number of arms (Appendix Fig. B.1.1).

Table 2.5: Investigated scenarios based on the different dose-response relationships of the data-generating true models, drug strengths, and study design variables (n = 60 scenarios)

Dose-response relationship of true model	Number of arms	Dose levels	Strong drug effect Total n = 24	Strong drug effect Total n = 36	Weak drug effect Total n = 24	Weak drug effect Total n = 36
Flat Linear Log-linear E_{max} Sigmoidal	3	0, 30, 60	$n_{arm} = 8$	$n_{arm} = 12$	$n_{arm} = 8$	$n_{arm} = 12$
Flat Linear Log-linear E_{max} Sigmoidal	4	0, 20, 40, 60	$n_{arm} = 6$	$n_{arm} = 9$	$n_{arm} = 6$	$n_{arm} = 9$
Flat Linear Log-linear E_{max} Sigmoidal	6	0, 12, 24, 36, 48, 60	$n_{arm} = 4$	$n_{arm} = 6$	$n_{arm} = 4$	$n_{arm} = 6$

n: total sample size; n_{arm}: sample size per arm.

Consequently, based on the different dose-response relationships of the data-generating models (i.e. true models) (Table 2.3), drug strengths, and the different study design variables, a total of 60 different scenarios were investigated (Fig. 2.8, Table 2.5).

2.5.2 cLRT-Mod approach: Principle, procedure and implementation

The cLRT-Mod included two steps [126]. The first step, the "cLRT" step, involved fitting each candidate model to the dataset and then the investigation of the presence of a drug effect by comparing the best-fitting candidate model to that of the flat dose-response model. In the second step, the "mod" step, the responses after various doses were predicted from either the best-fitting candidate model (i.e. model selection) or a weighted mixture of candidate models (i.e. model averaging). In our work presented in this thesis, we only focused on the investigation of the cLRT step.

cLRT procedure and implementation within the simulation framework

The cLRT approach [126] relied on the principles of NLME modelling in which the model describing the null hypothesis of a flat-dose response was assumed to be nested in each candidate model and parameters were estimated by the maximum likelihood (section 2.1.2.2). In this work, the likelihood was computed by the Laplace estimation method (section 2.1.2.2).

The overall framework of cLRT in relation to a simulation scenario is depicted in Fig. 2.9. Starting with a specific scenario, defined by the dose-response relationship of the data-generating true model, strength of drug effect, and variables of the study design (e.g. sample size, number of arms), N datasets were simulated (Fig. 2.9, section 1). Afterwards, the flat model along with the four candidate models (differing in the dose-response relationship) were fitted to each of the simulated datasets and their OFV and AIC were derived (Fig. 2.9, section 2.1). As a next step, the best-fitting candidate model was identified. The best-fitting candidate model was defined as the one with the lowest AIC value amongst the different candidate models, after fitting each candidate model to the dataset (Fig. 2.9, section 2.2). A cLRT test was then performed between the LRT i.e. OFV, of the best-fitting candidate model ($OFV_{best-fitting}$) and that of the flat dose-response model (OFV_{flat}) to obtain the respective test statistic, t_{cLRT} (Eq. 2.61; Fig. 2.9, section 2.3).

$$t_{cLRT} = OFV_{best-fitting} - OFV_{flat} \tag{2.61}$$

Next, t_{cLRT} was compared against a critical value, $q_{\alpha,cLRT}$, derived under the assumption of a null hypothesis (Fig. 2.9, section 3). Based on the union-intersection principle, one could reject the global null hypothesis of no drug effect if at least the null hypothesis of one of the candidate models was rejected. Therefore, if t_{cLRT} was smaller than the calculated $q_{\alpha,cLRT}$, where α was the significance level and nominal type I error, then the global null hypothesis could be rejected (Fig. 2.9, section 4) and the cLRT test was considered significant i.e. a significant dose-response (drug effect) could be identified (Fig. 2.9, section 5).

Calculation of the critical value $q_{\alpha,cLRT}$ (Fig. 2.9, section 3)

To derive the critical value, data had to be generated under the null hypothesis. Therefore, per each N simulated dataset , M datasets were simulated using the flat dose-response model and the parameter estimates obtained by fitting the flat dose-response model to the placebo arm data of the respective N dataset (i.e. no drug effect, Fig. 2.9, sections 3.1 and 3.2). The flat and candidate models were then fitted to each M simulated dataset, per each N simulated dataset i.e. $N{\cdot}M$ times (Fig. 2.9, section 3.3), the best-fitting candidate model was selected (Fig. 2.9, section 3.4) and the test statistic, $t_{cLRT_{Ho}}$ was computed for each best-fitting candidate model (Fig. 2.9, section 3.5). At $\alpha = 0.05$, the α

quantile of the $t_{cLRT_{Ho}}$ distribution (Fig. 2.9, section 3.6) was then used as the cut-off value to define the critical value, $q_{\alpha,cLRT}$. Therefore, the value of $q_{\alpha,cLRT}$ was derived from the α quantile of the $t_{cLRT_{Ho}}$ distribution obtained from the same procedures as outlined in Fig. 2.9, sections 2.1 to 2.3, but using data generated under the null hypothesis.

2.5.3 MCP-Mod approach: Principle, procedure and implementation

MCP-Mod included two steps. The first step, the "MCP" step, applied multiple contrast tests to establish evidence of drug effect—in contrast to cLRT which relied on maximum likelihood—whereas the second step, the "Mod" step, involved modelling of the dose-response relationship to evaluate the optimal dose/dose range to achieve the desired response. As with cLRT-Mod, our presented work in this thesis only focused on the MCP step for comparison with cLRT performance.

MCP procedure and implementation within the simulation framework

For a standardised basis of comparison between the performance of cLRT and MCP, the N simulated datasets, based on each specific scenario (Fig. 2.9, section 1), were utilised for the MCP analysis (Fig. 2.9, orange label). Since MCP relied on time-invariant responses as endpoint, the best change in tumour size from baseline was selected as the endpoint of interest (i.e. response) to closely mimic the longitudinal change in tumour size upon which cLRT relied.

The MCP analysis started with a pre-selected set of dose-response candidate models which were informed by the investigated dose-response relationships in the cLRT step i.e. $n_{candidate\ models} = 4 +$ flat as reference. These models were first defined for a dose effect by standardised parameters for each of the different relationships. The standardised parameter values of each dose-response relationship of the model were based on prior knowledge (as applied here) from similar studies/treatments or the true underlying dose-response profile, to calculate the mean responses for each dose-response relationship of the model based on the observed study data. Then, each candidate model was tested using contrast tests. These are hypothesis tests that aimed to establish evidence of a drug effect, based on the observed study data, while accounting for model uncertainty. The multiple contrast tests approach relied on the principle of the union-intersection test. Therefore, if one dose-response model was significant (i.e. if a dose-response relationship was identified), the null hypothesis could be rejected and a significant drug-effect was identified.

2.5.4 Evaluation of cLRT and MCP performance

Before evaluating cLRT and MCP, exploratory data check was performed to identify the robustness of cLRT and stability of the simulations. The percentage of successful M simulations-estimations within each N simulation (Fig. 2.9, sections 3.2–3.4) was explored and N simulations which failed to show $\geq 90\%$ successful executions after model-fitting and selection, under the null hypothesis, were excluded. Moreover, test statistics (t_{cLRT}, $t_{cLRT_{Ho}}$) showing implausible values e.g. better performance of the flat model over the candidate model, were excluded. Lastly, extreme outlying values in the $t_{cLRT_{Ho}}$ distribution ($t_{cLRT_{Ho}} < -60$) indicating failure to reach the best-set of model parameters (i.e. local minima) were identified and excluded from the $t_{cLRT_{Ho}}$ distribution of the respective N simulated dataset.

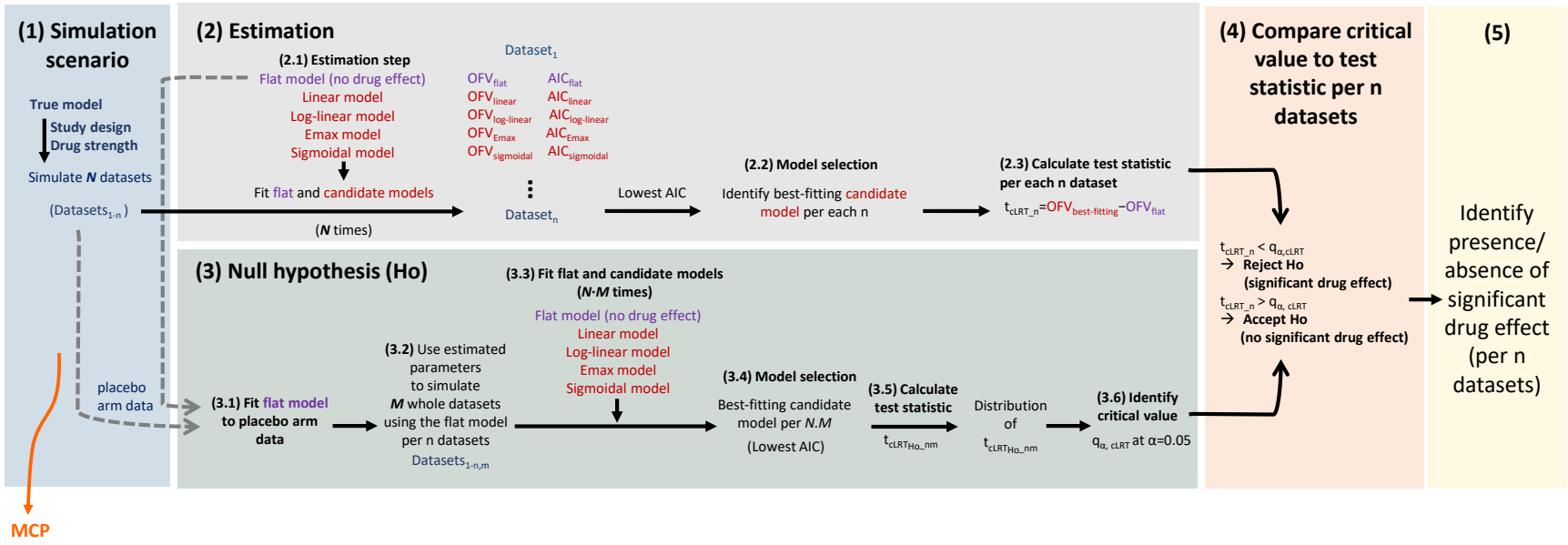

Figure 2.9: cLRT as implemented within the simulation framework showing the common datasets used for both cLRT and MCP analyses in this thesis. Starting with a specific simulation scenario, defined by the dose-response relationship of the data-generating true model, drug strength and variables of the study design, N datasets were simulated (section 1). Afterwards, the flat model (purple) along with the four candidate models (red) were fitted to each of the simulated datasets and their OFV and AIC were derived (section 2.1). Next, the best-fitting candidate model (candidate model with the lowest AIC value) was selected (section 2.2) and a cLRT test was performed to derive the test statistic, t_{cLRT} (section 2.3). The t_{cLRT} was then compared against a critical value derived under the assumption of a null hypothesis i.e. no drug effect (section 3). To derive the critical value, per each N simulated datasets, M datasets were simulated using the flat model and the parameter estimates obtained by fitting the flat dose-response model to the respective N placebo arm data (sections 3.1 and 3.2). The flat and candidate models were then fitted to each M simulated datasets per N simulated datasets (section 3.3) and the best-fitting candidate model was selected (section 3.4) and the test statistic, $t_{cLRT_{Ho}}$, was computed (section 3.5). At $\alpha = 0.05$, the critical value was identified at the α quantile of the $t_{cLRT_{Ho}}$ distribution (section 3.6). Finally, if the test statistic was smaller than the identified critical value, the global null hypothesis was rejected (section 4) and a significant drug effect was identified (section 5). Orange arrow indicates that the same N simulated datasets were utilised for the MCP analysis. AIC: Akaike information criterion; OFV: objective function value.

Following the exploratory data check, different criteria were selected to evaluate and compare the performance of cLRT and MCP:

2.5.4.1 Type I error calculation

Type I error referred to the rejection of the null hypothesis when it was true (i.e. a significant drug effect was detected when the true dose-response model was flat). Type I error was evaluated, per each group of scenarios sharing the same study design variables and differing only in the dose-response relationship of the true model and strength of drug effect since under "no effect" there was no dose-response relationship or different strengths of drug effect. Type I error was compared between cLRT and MCP.

2.5.4.2 Power calculation

Power referred to the probability of detecting a significant dose-response signal (i.e. drug effect/non-flat dose-response) when the drug effect was present in the data-generating model, regardless of correctly identifying the *true* dose-response relationship of the underlying dose-response model. Power is also referred to as $1-\beta$, where β is the type II error representing false negative results. The power of cLRT and MCP was compared, per scenario.

2.5.4.3 Selection of best-fitting candidate model

The probability of being selected as the best model was investigated for each of the different candidate models, per scenario.

2.5.4.4 Identification of true underlying dose-response model

The probability of correctly identifying the true dose-response relationship of the data-generating dose-response model from the different candidate models was computed, per scenario.

2.5.5 Execution and software

To implement cLRT, we developed an automated framework in R using RStudio Server Pro as a graphical user interface. Within this framework, NONMEM® and PsN were used for simulations and (re-)estimations, and different R packages were used for dataset optimisation, post-processing, and graphical representations (further details on the software versions and R packages are mentioned in section 2.4).

For each scenario, $N = 100$ datasets were simulated and for each simulated dataset, $q_{\alpha,cLRT}$ was approximated through $M = 100$ simulations from data generated under the null hypothesis.

The simulation framework was executed on a high-performance computing (HPC) cluster using 4 cpus-per-task while setting the number of tasks to 4. Hence, the total computation time to simulate and re-estimate the N datasets for all the 60 scenarios was ≈ 12 days.

The MCP analysis step of MCP-Mod was performed by members in the Global Biostatistics & Data Sciences at Boehringer Ingelheim Pharma GmbH & Co.KG using the DoseFinding package in R (section 2.4).

2.6 Project II: Identification of predictors of efficacy in NSCLC patients: A tumour dynamics - CRP modelling framework

Project II aimed to identify early prognostic predictors of efficacy (i.e. progression and/or survival) in advanced NSCLC patients, with a special focus on the potential of longitudinal CRP concentrations. To achieve this objective, data from a previously conducted clinical study in advanced NSCLC patients were retrospectively leveraged by means of population pharmacodynamic (PD) and parametric TTE models. The following sections elaborate on the source of the clinical data and the modelling framework which was adopted to achieve our objective. *Project II* has been published in Cancers 15: 5429 (2023) [204].

2.6.1 Clinical study data: The CEPAC-TDM study

The CESAR Study of Paclitaxel Therapeutic Drug Monitoring (CEPAC-TDM study) was a prospective, open-label, randomised, two-arm, phase III trial conducted in 10 centres in Germany and Switzerland between March 2011 and April 2014 (ClinicalTrials.gov Identifier: NCT01326767, EudraCT No: 2010-023688-16) [33]. The CEPAC-TDM study offered a rich database of drug exposure, clinical chemistry including serum biomarkers as well as efficacy and survival data. Thus, it was well suited to support our proposed modelling framework in which the relationship between anticancer drug exposure, tumour dynamics, and CRP concentrations could be characterised and subsequently leveraged to identify predictors of efficacy endpoints.

The study was conducted by the Central European Society of Anticancer Drug Research-EWIV (CESAR), approved by the respective institutional review boards (Ostschweiz, St. Gallen in Switzerland; Eberhard-Karls-Universität, Tübingen in Germany), and conducted in accordance with the ICH Harmonised Tripartite Guidelines for Good Clinical Practice 1996, Directive 91/507/EEC, Declaration of Helsinki, Directive 2001/20/EC, and local legislation. All patients provided written informed consent before study initiation.

2.6.1.1 Study population

Patients diagnosed with histologically confirmed advanced NSCLC (stage IIIB–IV) and considered for first-line combination chemotherapy with paclitaxel and either carboplatin or cisplatin were enrolled in the study.

Eligible patients included both men and women, aged 18–75 years with an ECOG performance status ≤ 2, and at least one measurable tumour according to the RECIST criteria v.1.1 [56] (section 1.3.2.6). Patients should have received no prior systemic treatment for advanced NSCLC with paclitaxel, carboplatin, or cisplatin nor have had pre-existing neuropathy $>$ grade I according to the National Cancer Institute Common Terminology Criteria for Adverse Events (NCI-CTCAE), version 4 [205]. Patients with asymptomatic brain metastasis were eligible in case corticosteroids were not indicated and irradiation completed >4 weeks prior to the first treatment cycle. The full list of the inclusion and exclusion criteria is available under `https://clinicaltrials.gov/ct2/show/NCT01326767`.

2.6.1.2 Study design

All patients enrolled in the CEPAC-TDM study were randomised at a ratio of 1:1 to receive up to six cycles of 3-weekly paclitaxel at either the conventional standard dose (treatment arm A, Fig. 2.10, upper panel) or an individualised dose (treatment arm B, Fig. 2.10, lower panel). Paclitaxel was administered in combination with a platinum-based drug (Fig. 2.10, upper and lower panels: arms A and B, label: Dose) for which randomisation was also stratified i.e. balanced allocation was maintained between cisplatin and carboplatin. The reason for this design was the original objective of whether individualised and PK-guided paclitaxel dosing would significantly reduce paclitaxel-related grade 4 neutropaenia compared with the conventional body surface area (BSA)-guided paclitaxel dosing in patients with advanced NSCLC without compromising efficacy, represented by PFS and OS [33].

Chemotherapy was given for up to six cycles, until disease progression or intolerable toxicity, whichever occurred first. The patient was then evaluated in an end-of-treatment (EOT) visit which occurred 27–35 days after the last dose. Post-treatment, patients were followed up until death, loss to follow-up, or end-of-study. Patients who stopped treatment prior to disease progression were also monitored for disease progression through tumour assessments during the follow-up phase. An overview of this study design is depicted in Fig. 2.10.

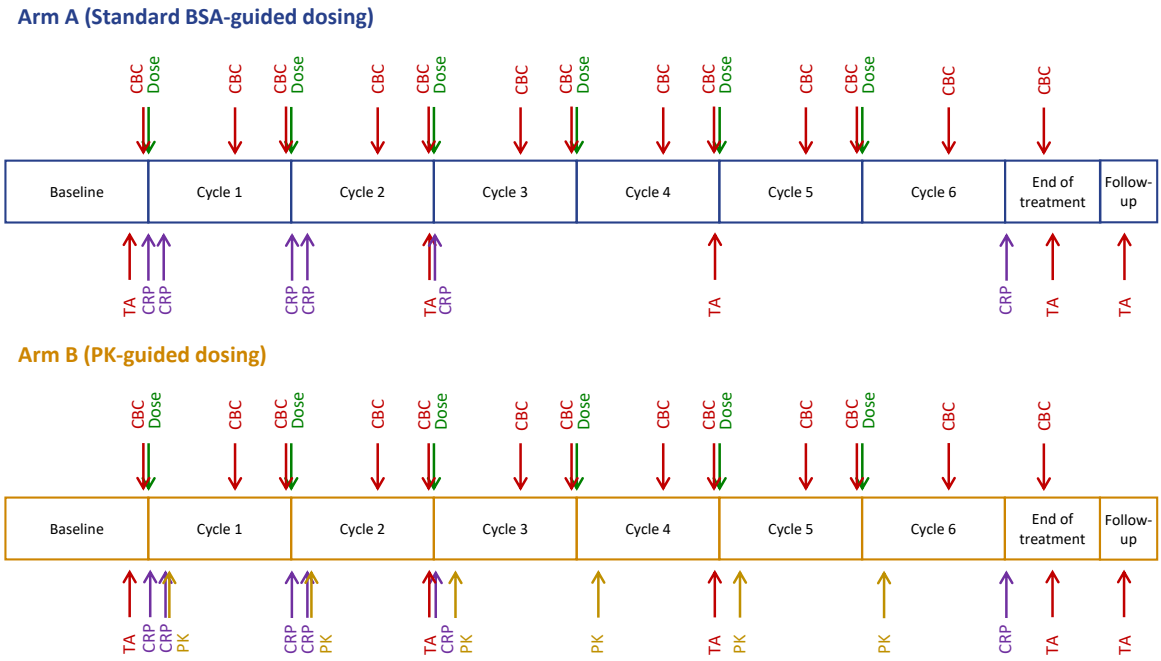

Figure 2.10: Design of the CEPAC-TDM study showing the two randomised study arms. Patients were randomised to either the conventional standard BSA-guided dosing arm (arm A, upper panel) or the individualised PK-guided dosing arm (arm B, lower panel). Arrows indicate the approximate time points of the different assessments within the study duration. Adapted from Henrich [206].
BSA: body surface area; CBC: complete blood count; CRP: C-reactive protein; PK: pharmacokinetic (sample); TA: tumour assessment.

Before study initiation, the patients' disease status was evaluated (diagnosis, tumour histology, prior cancer treatment, tumour size, medical history, and concomitant medications). Their demographic characteristics (age, body weight, body height, and sex), complete blood count and ECOG performance status were also documented. Additionally, assessments of body weight and

ECOG performance status were performed at the start of each cycle and at the EOT; complete blood count with white blood cell differentiation was evaluated at baseline, day 1 and day 15±2 of each cycle and at the EOT (Fig. 2.10, upper and lower panels: arms A and B, label: CBC); and clinical chemistry and coagulation panel were performed at baseline and at the EOT.

2.6.1.3 Study treatment and dosing

Pre-medication

To avoid paclitaxel-related hypersensitivity reactions (section 1.3.2.5), all patients received either dexamethasone or an antihistaminic drug prior to paclitaxel administration.

Paclitaxel treatment

Paclitaxel was administered at the beginning of each cycle (i.e. day 1) as a 3-h infusion every three weeks (Fig. 2.10, upper and lower panels: arms A and B, label: Dose). Patients who were randomised to the conventional treatment arm (arm A) received the standard BSA-based paclitaxel dosing of 200 mg/m^2, according to the drug label. While patients who were randomised to the investigational treatment arm (arm B) received an individualised paclitaxel dose according to a previously developed and published dosing algorithm [207]. In this dosing algorithm (Fig. 2.11), the first dose of paclitaxel was based on the patient's age, BSA, and sex. For the subsequent treatment cycles, paclitaxel dose was adjusted according to the patient's paclitaxel exposure (i.e. estimated individual $T_{>0.05}$ which was calculated as described in section 2.6.1.4) and the grade of experienced neutropaenia in the previous cycle. These dose adjustments aimed to target $T_{>0.05}$ of 26–31 h (section 1.3.2.5). Maximum paclitaxel dose increase was limited to 320 mg/m^2.

Platinum treatment

Paclitaxel administration was followed either by carboplatin or cisplatin (Fig. 2.10, upper and lower panels: arms A and B, label: Dose). Cisplatin was administered as an intravenous 2-h infusion either as one dose of 80 mg/m^2 on day 1 or as split doses of 40 mg/m^2 on days 1 and 2. Carboplatin was administered as a 30-min infusion (target AUC: 6 mg·min/L).

Treatment adaptations

In general, dose adjustments for patients in arm A were based on the recommendations in the paclitaxel drug label. Whereas for patients in arm B, dose adjustments were performed according to the dosing algorithm (Fig. 2.11). In case toxicities were observed, additional safety measures were adopted in the subsequent cycle for all three drugs. These measures are outlined in Appendix A.1, Table A.1.1.

Discontinuation of chemotherapy

Patients had to discontinue chemotherapy in case of experiencing intolerable adverse events e.g. grade 3 or 4 irreversible polyneuropathy, prolonged cytopaenia (absolute neutrophil count $< 1{\cdot}10^9$ cells/L, thrombocytes $< 100{\cdot}10^9$ cells/L) for more than three weeks, or irreversible renal dysfunction (GFR < 30 mL/min), despite adequate dose reductions.

Cycle 1

Age [years]	Dose women [mg/m²]	Dose men [mg/m²]
< 46	185	200
46–50	180	195
51–55	175	190
56–60	170	185
61–65	160	175
> 65	150	165

Cycles 2–6

Previous-cycle neutropaenia grade 0-2		Previous-cycle neutropaenia grade 3		Previous-cycle neutropaenia grade 4	
$T_{C>0.05}$ [h]	Dose adjustment	$T_{C>0.05}$ [h]	Dose adjustment	$T_{C>0.05}$ [h]	Dose adjustment
> 50	-30%	> 50	-30%	> 50	-40%
41–50	-25%	41–50	-25%	41–50	-30%
31–40.9	-20%	31–40.9	-20%	31–40.9	-25%
26–30.9	No change	< 31	No change	< 31	-20%
20–25.9	+10%				
10–19.9	+20%				
< 10	+30%				

Figure 2.11: Paclitaxel dosing algorithm for patients in treatment arm B of the CEPAC-TDM study [207]. $T_{>0.05}$: time of paclitaxel plasma concentration $> 0.05\,\mu$mol/L.

2.6.1.4 Blood sampling for paclitaxel analysis

To determine paclitaxel plasma concentration and exposure for the dosing algorithm, only patients in the PK-guided dosing arm (i.e. arm B) underwent blood sample collection, once per cycle, 24 h [16–30 h] after the start of paclitaxel infusion (Fig. 2.10, lower panel: arm B, label: PK). Samples were centrally analysed at the West German Cancer Center and University Hospital Essen, Institute of Analytical Pharmacology using a previously validated method, according to the EMA guideline. The method utilised high-performance liquid chromatography with ultraviolet detection and an LLOQ of 0.015 mg/L [208]. The cycle-specific plasma sample was then used to calculate the individual and cycle-specific $T_{>0.05}$. $T_{>0.05}$ was derived from the Bayesian estimates (*post-hoc*, section 2.1.2.1) using the previously developed paclitaxel PK model [207], to inform paclitaxel dose adaptations according to the developed dosing algorithm, in treatment cycles 2–6 for patients in arm B (section 2.6.1.3).

2.6.1.5 Blood sampling for CRP analysis

CRP was measured in a subset of patients who were enrolled in the CEPAC-TDM study i.e. biomarker substudy of the CEPAC-TDM study. Blood sampling for CRP analysis was performed along with the routine blood sampling or with the PK sample, for the PK-guided dosing arm patients, at the following time points (Fig. 2.10, upper and lower panels: arms A and B, label: CRP):

- Day 1 of treatment cycles 1 (i.e. baseline before the start of treatment), 2 and 3
- Day 2 of treatment cycles 1 and 2 (24 h after paclitaxel administration)
- At the EOT

CRP quantification was performed at the Institute of Laboratory Medicine, Munich Biomarker Research Centre, German Heart Centre of the Free State of Bavaria, Technical University Munich, Germany, using a validated commercial Conformité Européenne *in vitro* diagnostic (CE-IVD) assay, used in routine diagnostics. The CE-IVD assay was a latex-enhanced turbidimetry method on the Cobas C501 analyzer (Roche Diagnostics, Mannheim, Germany) with an LLOQ of 0.3 mg/L and an upper limit of quantification of 350 mg/L. All measurements were done according to the requirements of the guidelines of the German Chamber of Physicians (Rili-BäK) [209].

2.6.1.6 Assessment of efficacy

The efficacy of the therapy was evaluated during the treatment and the follow-up phase by assessment of tumour size and survival.

Assessment of tumour size

Tumour size was evaluated according to the RECIST criteria v.1.1 [56] (section 1.3.2.6). It was radiologically assessed at baseline, within five days prior to the start of treatment cycles 3 and 5 (i.e. every six weeks), at the EOT, and every eight weeks during follow-up—if the patient had not progressed during the time of treatment—until progression was observed (Fig. 2.10, upper and lower panels, label: TA).

Assessment of survival status

Survival of patients was assessed every three months during the follow-up phase.

2.6.1.7 Assessment of safety

Toxicities and adverse events were monitored and evaluated by laboratory tests and physical examinations according to the NCI-CTCAE criteria during the whole treatment duration and later on every three months during the follow-up phase.

To assess neutropaenia and other haematologic toxicities, a complete blood count with white blood cell differentiation was performed at baseline, day 1 and day 15 ± 2 of each treatment cycle, and at the EOT (Fig. 2.10). Primary prophylaxis with granulocyte-colony stimulating factor (G-CSF) was not allowed. Nevertheless, therapeutic G-CSF was allowed in patients with febrile neutropaenia, prolonged grade 4 neutropaenia (≥ 7 days), or grade 4 neutropaenia persisting after at least one prior dose reduction of paclitaxel (secondary prophylaxis).

2.6.2 Data management and exploratory data analysis

2.6.2.1 Dataset building

The original dataset, which was developed from the CEPAC-TDM study and previously utilised in the works of Henrich [206, 210] and Ojara [67, 211, 212], was updated with the longitudinal CRP information. Since CRP concentrations were reported with respect to treatment cycle days rather than date/time, therefore, for each respective treatment cycle, CRP concentrations were assigned to the respective treatment cycle date and time as follows:

- Day 1 CRP samples were aligned with the time of paclitaxel administration (i.e. time zero) of the respective cycle
- Day 2 CRP samples for patients in the PK-guided dosing arm (i.e. arm B) were aligned with the time recorded for the PK sampling which corresponded to ≈ 24 h after paclitaxel administration of the respective cycle. For patients in the BSA-guided dosing arm (i.e. arm A), day 2 CRP samples were aligned with the time which corresponded to 24 h post paclitaxel administration of the respective cycle
- EOT CRP samples were aligned with the first record of the EOT visit

Next, different NONMEM®-compatible datasets were generated to suit the different modelling objectives, in which the dependent variable (e.g. CRP concentration, time-to-event) and potential covariates (e.g. disease characteristics, patient demographics) were specified. A summary of the created columns in the different datasets, their interpretation, and any modifications or imputations performed is available in Appendix A.2, Tables A.2.1 and A.2.2. All datasets were created using R software to ensure traceability and reproducibility of the process (section 2.4).

2.6.2.2 Handling of missing values and measurements below limit of quantification

Missing DV e.g. CRP concentrations were assumed to be missing completely at random and were not imputed. However, concentrations at specific time points which were of interest e.g. to be later used as potential predictors in the TTE model, were utilised from the model-predicted CRP concentrations.

For missing independent variables e.g. missing covariate information, it was necessary to not have any missing values for a successful covariate analysis to be performed within NONMEM®. Hence, missing values of continuous covariates were imputed with the median value of the whole patient population (e.g. imputation of missing and zero baseline IL-6 concentration with median baseline IL-6 concentration). For times between covariate measurements, the last observation was carried forward until the subsequent measurement was performed (e.g. alanine aminotranferase (ALT) and aspartate aminotransferase (AST) measurements). Other methods, as linear interpolation were not required in the current datasets. Our datasets also did not have any missing categorical covariates. For a detailed description on how missing values were handled for each covariate see Appendix A.2, Tables A.2.1 and A.2.2.

For the below LLOQ samples, the choice of the means to handle them was dependent on their proportion in the dataset [139, 213]. In our case, for a low percentage of below LLOQ samples (i.e. $< 5\%$), the decision to exclude these samples from the analysis has been shown to perform equally well compared to methods where values of below LLOQ samples were imputed or more complex methods where the likelihood of a value being below the LLOQ was estimated [213].

2.6.2.3 Handling of implausible values and outliers

The dataset was evaluated for implausible values and outliers in each column by graphical evaluation and physiological plausibility e.g. samples outside the upper limit of quantification. In general, implausible values were identified, flagged, and temporarily excluded during model development to avoid distortion of potential relations, especially during covariate analysis. However, they were later returned to the datasets of the final/key models after each stage to assess their impact on model performance and parameter estimates.

2.6.2.4 Exploratory data analysis

To explore the data, descriptive statistics of the patient population and CRP concentrations were summarised, with the latter being further explored regarding sampling frequency. In addition, exploratory plots of the efficacy metrics in the patient population were generated i.e. plots of tumour size over time and Kaplan-Meier plots of TTP, PFS, and OS. To identify the general trend in the concentration-time profile of CRP and the extent of CRP variability, and to detect outliers, CRP concentrations and fold change in CRP concentrations relative to baseline were plotted versus time. Different stratifications (e.g. by treatment cycle) were applied for comparison and, if needed, variables were plotted on a log-scale to account for potential skewness in the data distribution and have a better view of the lower values. Additionally, the relationship of CRP concentration versus tumour size was explored to inform the modelling strategy and support the modelling assumptions. For a preliminary assessment of the prognostic value of CRP concentrations, association between baseline CRP concentration and efficacy endpoints was investigated using nonparametric Kaplan-Meier and semi-parametric CoxPH analyses, before investigating the prognostic value of longitudinal CRP concentrations in a fully parametric TTE model. Finally, to guide the covariate analysis process, pre-selected covariates were further explored for potential correlations to identify strongly correlated pairs of covariates and avoid their simultaneous inclusion (for details on the used software and packages see section 2.4).

2.6.3 Modelling framework to identify predictors of efficacy in NSCLC patients

Based on the clinical data from the patients with advanced NSCLC enrolled in the CEPAC-TDM study (section 2.6.1) and to identify predictors of efficacy endpoints, a modelling framework was developed which comprised two stages (Fig. 2.12):

- In the first stage (Fig. 2.12, A), the quantitative and structural relationship between anticancer drug exposure, tumour dynamics, and CRP concentrations was established. To achieve this, a tumour dynamics model that was developed to characterise the change in tumour size over time while accounting for the impact of drug exposure was leveraged as a source of individual tumour size data [67] (Fig. 2.12, A, section 1). As a next step, model-predicted individual longitudinal tumour size (as a metric of tumour dynamics) was coupled to CRP model to characterise circulating CRP concentrations using a pharmacodynamic NLME model (Fig. 2.12, A, sections 2 and 3). The coupled model was later used to derive potential early longitudinal CRP- and tumour size-related predictors to inform the second stage (Fig. 2.12, A, sections 5 and 6).

- In the second stage (Fig. 2.12, B), efficacy endpoints (PFS and OS) were characterised (Fig. 2.12, B, sections 7 and 8). CRP- and tumour size-related metrics (derived from the first stage) (Fig. 2.12, A, sections 5 and 6) along with other patient and disease characteristics (Fig. 2.12, A, section 4) were explored as potential predictors of PFS and OS by means of parametric TTE models, to explore their impact on TTE occurrence.

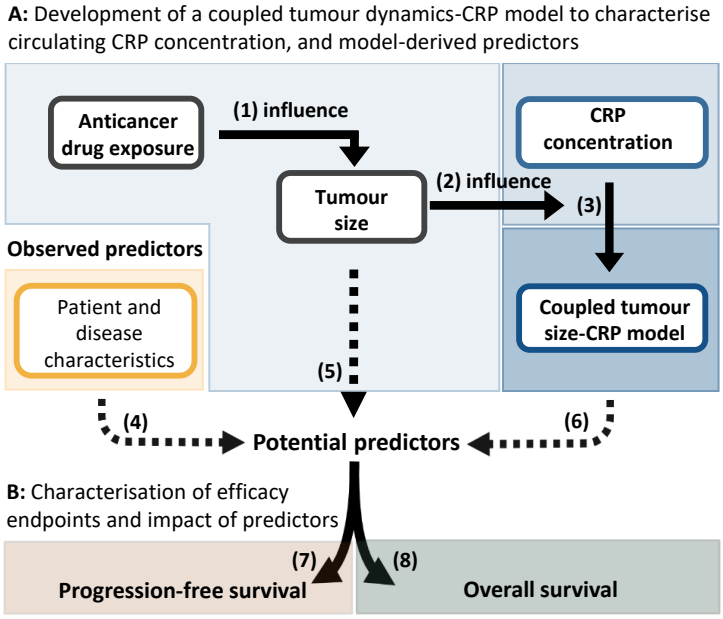

Figure 2.12: Schematic overview of the modelling framework highlighting the different stages undertaken to (**A**) characterise circulating CRP concentration and derive potential CRP-, tumour size-, patient-, and disease-related predictors; and (**B**) identify significant predictors of the prognostic outcomes progression-free survival and overall survival. CRP: C-reactive protein.

These two stages are explained in further details in the following sections.

2.6.3.1 Development of a coupled tumour dynamics - CRP model to characterise circulating CRP concentration

Model development to characterise tumour dynamics

To explore the influence of tumour dynamics on CRP (Fig. 2.12, A, sections 1–3), a previously developed anticancer drug-driven tumour dynamics model was leveraged to predict dense individual longitudinal tumour sizes corresponding to more frequent time points than those originally observed. The tumour dynamics model which characterised the time-course of tumour size, was developed based on the whole patient population of the CEPAC-TDM study (i.e. not only the patients in the biomarker substudy). It considered data only from the first 30 weeks after treatment start since after which data were too sparse due to patient death and biased due to the high dropout rate. Details of the different strategies which were investigated during the development of the tumour dynamics model were described in [67, 211].

Starting with data from only the PK-guided dosing arm due to the availability of paclitaxel PK information, the change in tumour size over time was described considering two processes represented by a zero-order linear tumour growth rate constant and a first-order drug-induced tumour decay. The drug-induced tumour decay described the efficacy of the administered drug on tumour cells as a first-order decay of the drug effect with respect to tumour size. Drug effect was in turn described as a function of drug exposure represented by paclitaxel AUC between the start and end of a treatment cycle (AUC_{cycle}). AUC_{cycle} was derived from the individual paclitaxel PK, estimated using the previously developed in house paclitaxel PK model which was based on the patients in the PK-guided dosing arm [206]. Furthermore, drug effect was allowed to change over time to account for a potential change in drug efficacy due to emergence of resistance. Therefore, drug effect exponentially declined, from the start of treatment, over time.

Thus, the tumour dynamics model was described as follows (Eq. 2.62),

$$\frac{dTS(t)}{dt} = K_{growth}(t) - \beta_o \cdot e^{-\lambda \cdot t} \cdot AUC_{cycle} \cdot TS(t) \tag{2.62}$$

where $\frac{dTS(t)}{dt}$ represents the change in tumour size (TS) over time; k_{growth} is the zero-order linear tumour growth rate constant; β_o is the drug-induced tumour decay rate constant per unit of paclitaxel AUC_{cycle} at time zero (i.e. cycle 1); λ is the rate constant for the exponential change in drug effect over time; and AUC_{cycle} is the paclitaxel area under plasma concentration-time curve from start to end of a cycle based on a single paclitaxel dose administered on the first day of a 21-day cycle.

Baseline tumour size was described by the B2 method which utilised the individual measured baseline tumour sizes while accounting for RUV (section 2.1.2.3). To account for the diverse individual tumour size profiles, IIV was added to the tumour growth parameter (K_{growth}) and the drug effect parameters (β_o and λ). No covariates were found to be significant on the model parameters representing the tumour growth rate or the drug-induced tumour decay rate constant. For the estimation of the tumour dynamics model parameters, the FOCE-I method was used (section 2.1.2.2).

As the tumour dynamics model was initially developed starting with the PK-guided dosing arm data, expansion of the model to include both study arms (BSA-guided and PK-guided dosing arms) was achieved by the use of the multiple imputation approach [214–216] to estimate the tumour dynamics model parameters for the whole study population. Within the multiple imputation framework (Fig. 2.13), the paclitaxel PK model that was developed based on the subset of patients with PK information (i.e. patients in the PK-guided dosing arm, treatment arm B) [206, 210] was leveraged to stochastically simulate the PK information for all individuals, based on their dosing information and individual characteristics in the dataset with missing PK information (Fig. 2.13, lower panel, facet 1). Stochastic simulations were performed to generate m replicates (here $m = 50$) of complete datasets in which missing PK information was simulated (Fig. 2.13, lower panel, facet 2) and consequently, individual paclitaxel exposure (i.e. AUC_{cycle}) was derived (Fig. 2.13, lower panel, facet 3). As a next step, now with complete PK information, the tumour dynamics model was fitted to each dataset replicate (Fig. 2.13, lower panel, facet 4) to estimate the model parameters for both study arms (Fig. 2.13, lower panel, facet 5). The results, now available as m replicates, were summarised—as explained in details in the following section—as described by Rubin [217] (Fig. 2.13, lower panel, facet 6) to derive population and individual tumour dynamics model parameters that would later inform CRP concentrations (Fig. 2.13, lower panel, facet 7).

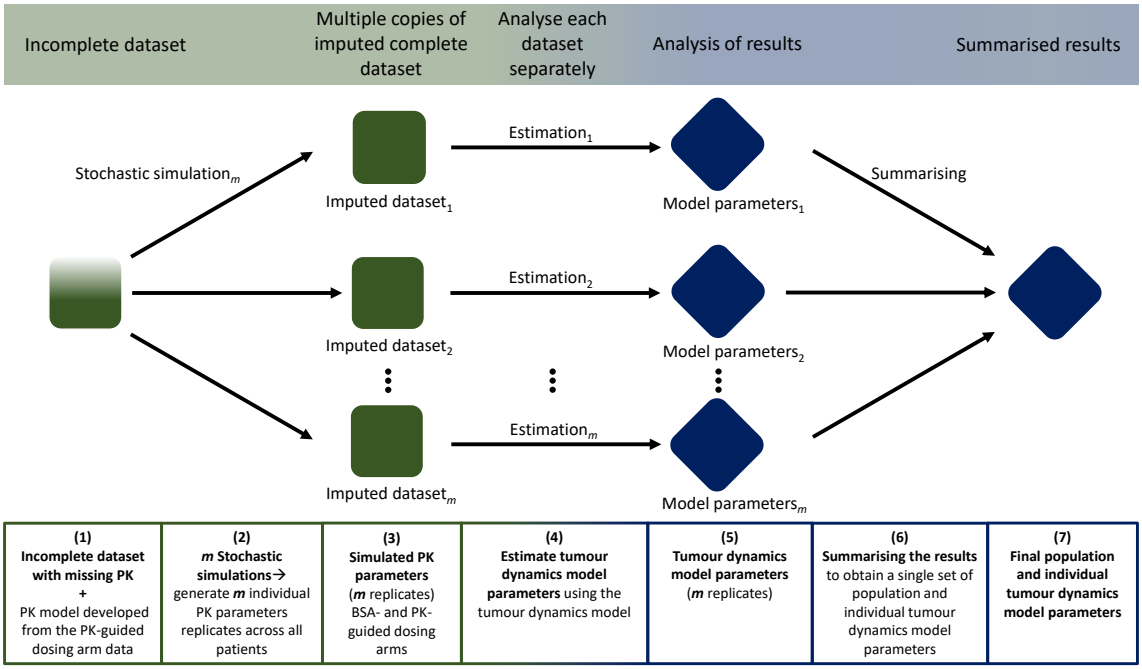

Figure 2.13: Schematic diagram of the multiple imputation approach applied to the tumour dynamics model development. *Upper and middle panels* show the general outline of the multiple imputation approach. Starting with an incomplete dataset (left side) and using the appropriate model to describe these data, stochastic simulations are performed to impute the missing information and generate m multiple copies of complete datasets. Next, each dataset is analysed separately using the model under development to generate m copies of model parameters which are then summarised to obtain a single set of the respective model parameters (end in the right side). *Lower panel, facets 1–7* show the detailed application of the multiple imputation steps with respect to the development of the tumour dynamics model. Here $m = 50$.
BSA: body surface area; PK: pharmacokinetic.

Derivation of population and individual tumour dynamics model parameters

As mentioned in the previous section, the tumour dynamics model, now fitted to each dataset replicate, resulted in m replicates of the tumour dynamics model parameters. To summarise the results, and in contrast to only taking the arithmetic mean of the parameter estimates across the m replicates of the multiple imputation, the equations described by Rubin (Eq. 2.63, 2.64) [217] ensured that the variances associated with the parameter estimates were also computed. Hence, we took into consideration the variability and uncertainty associated with the multiple imputation approach. Derivation of population and individual tumour dynamics model parameters are outlined in steps 1 and 2, respectively.

Population tumour dynamics model parameters (step 1)

Population tumour dynamics parameter estimates (i.e. K_{growth}, λ, β_o) from the 50 replicates of the multiple imputation were pooled according to Rubin's rule (Eq. 2.63, 2.64).

$$\tilde{\beta} = \frac{1}{m}\sum_{\gamma=1}^{m}\hat{\beta}^{(\gamma)} \tag{2.63}$$

$$\tilde{b} = \frac{1}{m}\sum_{\gamma=1}^{m}\hat{b}^{(\gamma)} + \frac{m+1}{m(m-1)}\sum_{\gamma=1}^{m}(\hat{\beta}^{(\gamma)} - \tilde{\beta})^2 \tag{2.64}$$

In Eq. 2.63 and 2.64, m is the number of replicates (here $m = 50$), β is a parameter in the Θ, Ω or Σ matrices i.e. a fixed-effects or a random-effects parameter, $\hat{\beta}^{(\gamma)}$ is the estimate of β from the imputed dataset γ ($\gamma = 1, ..., m$), $\tilde{\beta}$ is the calculated estimate of β, $\tilde{b}$ is the calculated variance (b) associated with β, and $\hat{b}^{(\gamma)}$ is the estimate of b from the imputed data set γ ($\gamma = 1, ..., m$).

Application of Rubin's rule to summarise the results, resulted in a single population parameter set in which the mean of the population parameter estimate for each of the fixed- and random-effects parameters across the 50 replicates was computed (Eq. 2.63) along with its associated variance and standard error (Eq. 2.64).

Individual tumour dynamics model parameters (step 2)

Following the same principle described in step 1 and to obtain individual tumour dynamics parameter estimates (i.e. $K_{growth,i}$, λ_i, $\beta_{o,i}$) from the 50 replicates of the multiple imputation, mean individual parameter estimates, $\tilde{\eta}_{par_i}$, of the interindiviudal random-effects ($\eta_{k_{growth}}$, η_λ, η_β) were computed at an individual level to obtain a single parameter set per individual according to Eq. 2.63 (now re-written as in Eq. 2.65).

$$\tilde{\eta}_{par_i} = \frac{1}{m} \sum_{\gamma=1}^{m} \hat{\eta}_{par_i}{}^{(\gamma)} \tag{2.65}$$

The computed mean individual random-effects ($\tilde{\eta}_{par_i}$) were then transformed to the individual parameter estimate as parameterised in the tumour dynamics model and according to Eq. 2.66,

$$Par_i = \tilde{\beta} \cdot e^{\tilde{\eta}_{par_i}} \tag{2.66}$$

where Par_i is the individual parameter estimate, $\tilde{\beta}$ is the typical (mean) population parameter estimate obtained from Eq. 2.63 in step 1 (i.e. the arithmetic mean of the population parameters across the m replicates) and $\tilde{\eta}_{par_i}$ is the computed mean individual random-effects from Eq. 2.65.

The calculated variance, $\tilde{\eta}_{var_i}$ (Eq. 2.64, now re-written as in Eq. 2.67), associated with the individual tumour dynamics parameter was therefore the sum of the mean of the individual variances, $\hat{\eta}_{var_i}$, associated with the individual random-effects, η_{par_i}, from the imputed dataset γ ($\gamma = 1, ..., m$) i.e. $\frac{1}{m} \sum_{\gamma=1}^{m} \hat{\eta}_{var_i}{}^{(\gamma)}$ (derived from the NONMEM output *.phi file); and the second term $\frac{m+1}{m(m-1)} \sum_{\gamma=1}^{m} (\hat{\eta}_{par_i}{}^{(\gamma)} - \tilde{\eta}_{par_i})^2$ in which $\hat{\eta}_{par_i}$ is the estimate of the individual random-effects, η_{par_i}, from the imputed data set γ ($\gamma = 1, ..., m$) and $\tilde{\eta}_{par_i}$ is the computed mean individual estimate of the random-effects from Eq. 2.65. The calculated variance was afterwards used as the source to calculate the uncertainty (SE) associated with the individual tumour dynamics parameters.

$$\tilde{\eta}_{var_i} = \frac{1}{m} \sum_{\gamma=1}^{m} \hat{\eta}_{var_i}{}^{(\gamma)} + \frac{m+1}{m(m-1)} \sum_{\gamma=1}^{m} (\hat{\eta}_{par_i}{}^{(\gamma)} - \tilde{\eta}_{par_i})^2 \tag{2.67}$$

Derivation of individual tumour size over time

The development of the tumour dynamics model under the multiple imputation framework hindered a direct utilisation of the model-predicted individual tumour sizes due to the presence of multiple replicates of model-predicted time-dependent tumour size estimate per individual. Moreover, simply averaging the m model-predicted individual tumour sizes would have:

- discarded the uncertainty associated with each replicate within the multiple imputation (Rubin's rule was not applicable since tumour size was a derived parameter with no associated random-effect and/or variance)
- restricted estimates to a definite time grid (i.e. averaged tumour sizes would have been obtained for only the time records that were available in the tumour dynamics model dataset and which did not necessarily coincide with the CRP measurements)

Therefore, the derived individual tumour dynamics parameter estimates (i.e. $K_{growth,i}$, λ_i, $\beta_{o,i}$) and their associated uncertainty (SE), obtained from step 2, were leveraged to estimate dense individual tumour size over time using a sequential modelling approach [218]. In this approach, individual tumour dynamics parameter estimates along with their uncertainty informed longitudinal tumour size estimation using the model parameterisation described in Eq. 2.62 and under a denser time grid (i.e. use of the ordinary differential equations of the tumour dynamics model under the $DES part in the NONMEM model code). This approach was used for:

- its higher precision and less bias compared to only using the population tumour dynamics parameter estimates for individual tumour size estimation [218],
- to account for the variability and uncertainty encountered from the multiple imputation and
- obtain denser, i.e. more frequent, individual tumour size estimates over time to better inform CRP concentrations.

Model development to characterise CRP concentration

To characterise CRP concentration-time profiles and identify influential covariates on CRP production, a CRP model was developed as follows:

Development of structural model of CRP

Circulating CRP concentrations were described by a structural turnover model governed by two processes represented by a zero-order production rate constant ($K_{in,CRP}$) and a first-order degradation rate constant ($K_{out,CRP}$) (Section 1.4.2.2; Fig. 1.7). During the initial phase of model development, no influence of the tumour was considered—at this stage—on CRP production and/or degradation. Therefore, a steady state was *initially* assumed and the baseline CRP steady-state concentration (CRP_{ss}) corresponded to the ratio between $K_{in,CRP}$ and $K_{out,CRP}$ as represented in Eq. 2.68,

$$CRP_{ss} = \frac{K_{in,CRP}}{K_{out,CRP}} \tag{2.68}$$

such that at time (t), the change in CRP concentration over time ($\frac{dCRP(t)}{dt}$) could be described as in Eq. 2.69,

$$\frac{dCRP(t)}{dt} = K_{in,CRP} - K_{out,CRP} \cdot CRP_{ss} \tag{2.69}$$

Since the system was governed by three parameters in which—at steady-state—identification of two parameters was sufficient to derive the third, the model was parameterised in terms of $K_{in,CRP}$ and $K_{out,CRP}$, and baseline was derived according to Eq. 2.68. To grant stability to the model and avoid the problem of unidentifiability (i.e. inability of the model to reliably estimate the true value of a parameter), $K_{out,CRP}$ was fixed to result in a CRP plasma half-life of 19 h which has been previously reported to be unchanged regardless of the health or disease status of the individual [97].

CRP concentrations were modelled in the log domain to account for the skewness in their distribution and the broad range of their reported values. Since the distribution of baseline CRP concentrations was known to follow a log-normal distribution, baseline CRP concentrations were handled according to the B1 method in which a typical baseline value was estimated (section 2.1.2.3).

Development of statistical model of CRP

Since $K_{out,CRP}$ was assumed not to vary across individuals, and CRP baseline was a derived parameter, IIV was considered only on the estimable parameter, $K_{in,CRP}$; as it was the only determinant of CRP plasma concentration. IIV was implemented as an exponential relationship assuming a log-normal distribution of $K_{in,CRP}$ (Eq. 2.2). For the RUV, because CRP concentrations were log-transformed, a log-transformed both sides approach was adopted (Eq. 2.10) in which RUV was estimated as a fixed-effect parameter (Eq. 2.12) and implemented as an additive component, corresponding to the exponential relation on the linear scale.

Development of covariate model of CRP

Covariates were pre-selected based on (i) clinical relevance e.g. covariates that reflect disease aggressiveness, inflammatory status or correlate to the biochemical synthesis of CRP; and (ii) graphical exploration e.g. covariate values showing potential trends against individual variability of CRP parameter(s). Potential correlations between the pre-selected covariates were then investigated to identify correlated covariates and avoid their simultaneous inclusion. Pre-selected covariates were then investigated for their statistically significant impact only on $K_{in,CRP}$, as the precursor and only determinant of CRP production, using the SCM approach, as describe earlier in section 2.1.2.1. For the continuous covariates, each parameter-covariate relationship was tested leveraging different functional forms (linear, exponential, and power) whereas for the categorical covariates, a fractional change model was explored.

Of note, to avoid a biased estimation of the parameter-covariate relationship, the developed base model had the outlying concentrations of both the dependent and independent variables *temporarily* excluded during the covariate model development steps but were later re-introduced to the final covariate model.

The impact of the identified significant covariates on $K_{in,CRP}$ was assessed at specific covariate values using the final covariate model. For the categorical covariates, comparison was performed between the different categories against the reference category while for the continuous covariates, comparison was between the lower (5^{th}) and upper (95^{th}) percentiles of the respective covariate distribution.

Parameter estimation, model evaluation and selection

To characterise CRP concentrations, parameters were estimated using the FOCE-I method (section 2.1.2.2). Since the CRP turnover model before considering the tumour influence, was not expected to mimic the observed CRP concentrations i.e. it represented a steady-state, no GOF plots or VPC were performed. Instead, model evaluation was based on the plausibility and precision of parameter estimates. The latter was derived from a nonparametric bootstrap with n = 1000 samples (section 2.1.2.4).

Linking tumour dynamics to CRP

Development of structural model of coupled tumour dynamics and CRP

To simultaneously associate the estimated individual tumour size to CRP production, and establish the quantitative and structural relationship between tumour dynamics and CRP concentration (Fig. 2.12, A, sections 2 and 3), different models were explored to relate tumour size-metric to $K_{in,CRP}$. These included (i) models assuming a direct influence, representing a direct impact of tumour load/burden on CRP production, through e.g. linear, exponential, power, or fractional change models; as well as (ii) models investigating the potential for a delayed effect between tumour load and CRP production, representing the IL-6-mediated activation of hepatocytes to synthesise and release CRP, through e.g. inclusion of an effect compartment or considering transit compartment(s) (with n = 1, 2, 3, 4 compartments) [140, 219]. Moreover, different tumour size-related metrics were explored and included: the tumour size, the change (difference) in tumour size from baseline tumour size, the fold change in tumour size from baseline tumour size i.e. ratio of tumour size to baseline tumour size, and the tumour size change over time i.e. first derivative over time. Additionally, a combined influence of the baseline tumour size with change in tumour size from baseline, or relative change in tumour size to baseline was also investigated.

Development of statistical model of coupled tumour dynamics and CRP

In a subsequent step, to account for the heterogeneity in the patients' CRP concentration-time profiles and the diverse individual tumour sizes, the impact of accounting for IIV on the parameters of the function describing the relation between tumour size and CRP production was explored. IIV was implemented as an exponential relationship assuming a log-normal distribution (Eq. 2.2).

Parameter estimation, model evaluation and selection

Parameters were estimated using the FOCE-I method (section 2.1.2.2). Evaluation and selection of the best model to describe the relation between tumour size-metric and CRP concentration were based on parameter precision, numerical improvement of model performance demonstrated by a drop in the AIC of the model under investigation compared with the base model and/or the other investigated structural models, graphical diagnostics through the GOF plots, and predictive performance demonstrated by the VPC for models that successfully converged. Model robustness and parameter uncertainty were derived from a nonparametric bootstrap test with n = 1000 samples (section 2.1.2.4). The NONMEM model codes for each of the developed models are available in Appendix C, sections C.3 and C.4.

2.6.3.2 Characterisation of efficacy endpoints and impact of predictors

Efficacy endpoints, namely PFS and OS, were characterised by means of parametric TTE models (section 2.2.4) to evaluate the prognostic value of different baseline and early longitudinal covariates i.e. predictors (Fig. 2.12, B, sections 7 and 8). Besides the ability to quantify the magnitude of impact of the identified significant predictors, parametric TTE modelling also allowed to perform simulations under different scenarios e.g. to explore the impact of different values of the identified predictors on median TTE.

Since the primary objective was to identify early predictors of PFS and OS, we focused on potential predictors derived from only the first three treatment cycles e.g. CRP concentration at the beginning of each of the first three treatment cycles. A landmark time (section 2.2.4) was chosen at the beginning of treatment cycle 3 (i.e. day 42 from the start of treatment) such that in the context of TTE modelling, PFS would be defined as the time from the start of treatment cycle 3 until objective tumour progression or death, whichever occurred first and OS would be defined as the time from the start of treatment cycle 3 until patient death.

The model development processes for PFS and OS followed the same framework and are henceforth, jointly described in the next section.

Development of base model of efficacy endpoints

An exponential, Weibull, Gompertz, log-normal, and log-logistic hazard functions (section 2.2.4) were explored to characterise the observed PFS events over time. For OS, exponential, Weibull, and Gompertz hazard functions were explored. No IIV was considered, as each individual contributed to only one observation i.e. occurrence of either progression or death event for PFS, occurrence of death event in case of OS. The identified base model was then adopted for the subsequent covariate analysis.

Development of covariate model of efficacy endpoints

Covariates were pre-selected based on what has been previously reported in the literature, clinical relevance, and patient- and disease-related factors that would relate to disease aggressiveness and overall health status of the patient. However, a key feature in the selection of these covariates was the focus that they can provide an *early* predictive potential. Hence, the pre-selected longitudinal covariates and their associated metrics were derived from their respective longitudinal data of only the first three treatment cycles.

These pre-selected covariates included:

- Baseline patient- and disease-related characteristics :
  - Baseline ECOG status, smoking status, disease stage, presence/absence of liver lesions, presence/absence of brain lesions, tumour histology, number of target lesions, number of non-target lesions, and sum of target and non-target lesions (Fig. 2.12, A, section 4)
- Markers of inflammation:
  - Neutrophil-to-lymphocyte ratio-related metrics derived from the first two treatment cycles, and CRP-related metrics derived from the first three treatment cycles (Fig. 2.12, A, section 6)
- Tumour size-related metrics derived from the first three treatment cycles (Fig. 2.12, A, section 5)

CRP- and tumour size-related metrics were respectively obtained from the individual CRP and tumour size predictions of the coupled tumour dynamics - CRP model developed in section 2.6.3.1 (Fig. 2.12, A, sections 5 and 6), based on their respective longitudinal data from only the first three treatment cycles—as previously mentioned. The full list of the derived metrics for the markers of inflammation and tumour size is provided in Table 2.6.

The pre-selected covariates were first explored for potential correlations to identify strongly correlated pairs of covariates and guide the covariate analysis process. Afterwards, covariates were tested on the hazard function that best described the distribution of the respective events using an SCM approach as described earlier in section 2.1.2.1, to identify significant predictors of PFS and OS. Besides the previously mentioned statistical criteria (section 2.1.2.1), the precision of parameter and covariate estimates (RSE $< 50\%$) and the non-correlation to a previously included covariate were additional criteria for acceptance of a parameter-covariate relation during the covariate model selection process.

Parameter estimation, model evaluation and selection

Parameters were estimated with the FO method (section 2.1.2.2) since no IIV was considered. Evaluation and selection of the TTE base and covariate models were based on the precision of parameter estimates, numerical improvement in model performance assessed by the AIC, and the model's predictive performance evaluated through 250 simulations followed by generating KM VPC (section 2.2.4). To enable simulations of events at different time points when no observations were made, extra records were added to the dataset every 7 days until the last observation time of 28.8 months for PFS and 32.6 months for OS. The covariate information in these additional records was linearly interpolated between the original records. For the final TTE base and covariate models, a nonparametric bootstrap with 1000 samples was performed (section 2.1.2.4) from which model robustness (i.e. percent minimisation) and parameter uncertainty (i.e. 95% CI) were derived.

The NONMEM model codes for each of the developed models are provided in Appendix C, sections C.5 and C.6.

Assessment of the impact of the identified predictors of efficacy

The impact of the identified significant predictors on the median TTE, in comparison to the observed median TTE of PFS and OS, was assessed through simulations ($n = 250$) of the final covariate model at specific covariate values. For the categorical covariates, comparison was performed between the different categories against the reference category while for the continuous covariates, comparison was between the lower (5^{th}) and upper (95^{th}) percentiles of the respective covariate distribution. Additional interim values e.g. at the 25^{th}, 50^{th}, and 75^{th} percentiles of the continuous covariate were also explored—if deemed useful.

Table 2.6: List of the derived metrics of the covariates tested in the parametric time-to-event models as potential early predictors of efficacy endpoints

Covariate	Derivation
Markers of inflammation	
CRP-related metrics	
Observed baseline CRP	—
Model-estimated CRP concentration:	
cycle 1 day 1	CRP_{cycle1}
cycle 2 day 1	CRP_{cycle2}
cycle 3 day 1	CRP_{cycle3}
Difference in CRP concentration:	
cycle 2 from cycle 1	$CRP_{cycle2} - CRP_{cycle1}$
cycle 3 from cycle 1	$CRP_{cycle3} - CRP_{cycle1}$
cycle 3 from cycle 2	$CRP_{cycle3} - CRP_{cycle2}$
Relative change in CRP concentration:	
cycle 2 from cycle 1	$\frac{CRP_{cycle2}-CRP_{cycle1}}{CRP_{cycle1}}$
cycle 3 from cycle 1	$\frac{CRP_{cycle3}-CRP_{cycle1}}{CRP_{cycle1}}$
cycle 3 from cycle 2	$\frac{CRP_{cycle3}-CRP_{cycle2}}{CRP_{cycle2}}$
Fold change in CRP concentration:	
cycle 2 from cycle 1	$\frac{CRP_{cycle2}}{CRP_{cycle1}}$
cycle 3 from cycle 1	$\frac{CRP_{cycle3}}{CRP_{cycle1}}$
cycle 3 from cycle 2	$\frac{CRP_{cycle3}}{CRP_{cycle2}}$
Neutrophil-to-lymphocyte ratio-related metrics	
Observed cycle 1 day 1	$(N/L)_{cycle1}$
Observed cycle 2 day 1	$(N/L)_{cycle2}$
Difference in neutrophil-to-lymphocyte ratio:	
cycle 2 from cycle 1	$(N/L)_{cycle2} - (N/L)_{cycle1}$
Relative change in neutrophil-to-lymphocyte ratio:	
cycle 2 from cycle 1	$\frac{(N/L)_{cycle2}-(N/L)_{cycel1}}{(N/L)_{cycle1}}$
Fold change in neutrophil-to-lymphocyte ratio:	
cycle 2 from cycle 1	$\frac{(N/L)_{cycle2}}{(N/L)_{cycle1}}$
Tumour size-related metrics	
Observed baseline tumour size	—
Model-estimated tumour growth rate	—
Model-estimated tumour size at week 7 relative to baseline	$(\frac{TS_{week7}}{BLTS}) \cdot 100$

CRP: C-reactive protein; TS: tumour size; BLTS: baseline tumour size; TS_{week7}: model-estimated tumour size at week 7; N/L: neutrophil-to-lymphocyte ratio

3 | Results

3.1 Project I: Simulation-based evaluation of cLRT and MCP approaches in oncology under different clinical trial scenarios

In *project I*, following an exploratory data check to identify the robustness of cLRT and the stability of the simulations, evaluation of cLRT and MCP performances was based on the previously selected criteria: type I error, power, selection of the best-fitting candidate model and identification of the true underlying dose-response model—with respect to the different simulation scenarios. The type I error and power were also directly compared between cLRT and MCP.

3.1.1 Evaluation of cLRT and MCP performance

Across the different scenarios, the majority (88%–95%) of the N simulations under the null hypothesis had 100% successful M simulations-estimations i.e. 100% successful runs after model-fitting and selection under the null hypothesis (Fig. 2.9, sections 3.2–3.4). Only 1%–4% N simulations showed $< 90\%$ successful runs with an observed minimum successful runs of 69% (Appendix Table A.3.1) and thus these respective N simulations were excluded.

Implausible positive t_{cLRT} values were mainly observed with scenarios investigating a weak drug effect and ranged from 0% to 6%, given the study design variables and dose-response relationship of the data-generating true model. Only one positive t_{cLRT} value was observed for the scenario investigating a strong drug effect based on a true E_{max} relationship and three dose levels for a total number of patients of $n = 24$ (Appendix Table A.3.2). Under the null hypothesis, implausible positive $t_{cLRT_{Ho}}$ values were more prominent and were observed in the $t_{cLRT_{Ho}}$ distribution of all N simulations. Across the different simulation scenarios, irrespective of the true dose-response relationship, the proportion of implausible positive $t_{cLRT_{Ho}}$ values ranged from 11%–45% (Appendix Table A.3.3). These implausible values which indicated an unrealistic better performance of a no dose-response, i.e flat dose-response relationship, over the candidate model were excluded.

Lastly, extreme outlying values in the $t_{cLRT_{Ho}}$ distribution ($t_{cLRT_{Ho}} < -60$), indicative of the failure to reach the best-set of model parameters (i.e. local minima), were observed in only 9%–16% of the N simulations under the null hypothesis. Amongst which, these outlying values, which were excluded, comprised 1.2%–38.9% of the $t_{cLRT_{Ho}}$ values in the respective distribution of each of the N simulations. Further details on these outliers, stratified per N simulations and study design are provided in Appendix Table A.3.4 and Appendix Fig. B.2.1.

3.1.1.1 Type I error calculation

Type I error referred to the percentage of detecting a significant drug effect when the true dose-response model was flat. Type I error, for both cLRT and MCP, is shown in Fig. 3.1 per each group of scenarios sharing the same study variables i.e. number of dose levels and total number of patients, and differing only in the dose-response relationship of the true model and/or drug strength—since type I error was independent of the dose-response relationship of the true model and drug strength.

For all scenarios, cLRT type I error was inflated and fell within 11.1%–15.1% (Fig. 3.1 left column). On the other hand, MCP type I error was more controlled and reached a maximum of only 4% (Fig. 3.1, right column). An MCP type I error of 4% meant that, given $N = 100$ simulations and after selecting the best-fitting candidate model, 4 models out of 100 were expected to be *falsely* identified to have a significant drug effect when no drug effect actually existed.

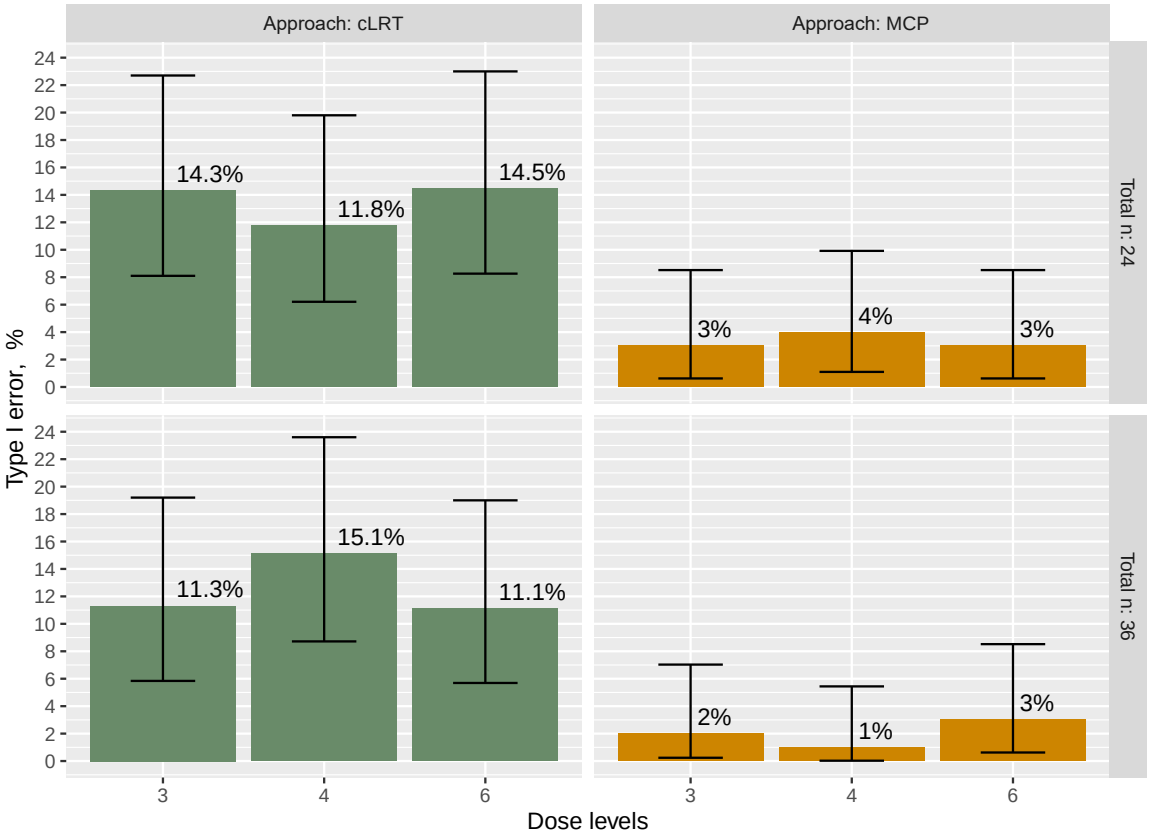

Figure 3.1: Type I error of cLRT and MCP approaches per each group of scenarios sharing the same study variables (number of dose levels, total number of patients) and differing only in drug strength and the dose-response relationship of the true model, stratified per approach (cLRT approach: left column or MCP approach: right column) and total number of patients (n = 24, top panel or n = 36, bottom panel).
Bar plots: type I error (i.e. percentage of detecting a significant drug effect when the true dose-response model was flat) per each group of scenarios sharing the same study variables i.e. number of dose levels and total number of patients, and differing only in the dose-response relationship of the true model and/or drug strength.
Error bars: 95% confidence interval of a binomial distribution with a probability of success of x% on 100 trial replicates, where x is the respective value of the type I error (i.e. number shown on the respective bar plot).
cLRT: combined likelihood ratio test; MCP: multiple comparison procedure; n: total number of patients.

3.1.1.2 Power calculation

The power of cLRT versus MCP is shown in Fig. 3.2; or in other words: Fig. 3.2 shows the percentage of simulated studies per scenario identified to have a significant dose-response relationship (i.e. drug effect/non-flat dose-response/rejected null hypothesis) for cLRT and MCP when a drug effect was present in the data-generating dose-response true model. As expected, the power was highly dependent on (i) the analysis approach, (ii) strength of the drug effect, and (iii) study design.

Scenarios based on a strong drug effect had higher power compared with the respective scenarios based on a weak drug effect. Moreover, cLRT was associated with higher power compared to MCP across all the different investigated scenarios.

In scenarios exploring a strong drug effect, 96%–100% of studies were detected to have a significant dose-response, regardless of the true dose-response relationship—if analysed using cLRT, whereas only 17%–52% of studies would have been detected—if analysed by MCP. Similarly, in scenarios exploring a weak drug effect, 66.3%–90.8% of studies were detected to have a significant dose-response—if analysed using cLRT in contrast to the detection of only 10%–24% if analysed by MCP.

Additionally, there was a consistent trend of higher power associated with the larger sample size (i.e. greater number of patients). Scenarios with total number of patients of $n = 36$ had a higher power compared with scenarios with total number of patients of $n = 24$, regardless of the analysis approach. Nevertheless, this was more prominent with MCP in comparison with cLRT.

Finally, with MCP, a clear trend of increasing power with fewer dose level and larger sample size per arm across the same total number of patients was observed. cLRT lacked such a discriminating power when it came to study design variables and showed almost a plateau across the different number of dose levels. Details of the power associated with each scenario is provided in Appendix Table A.4.

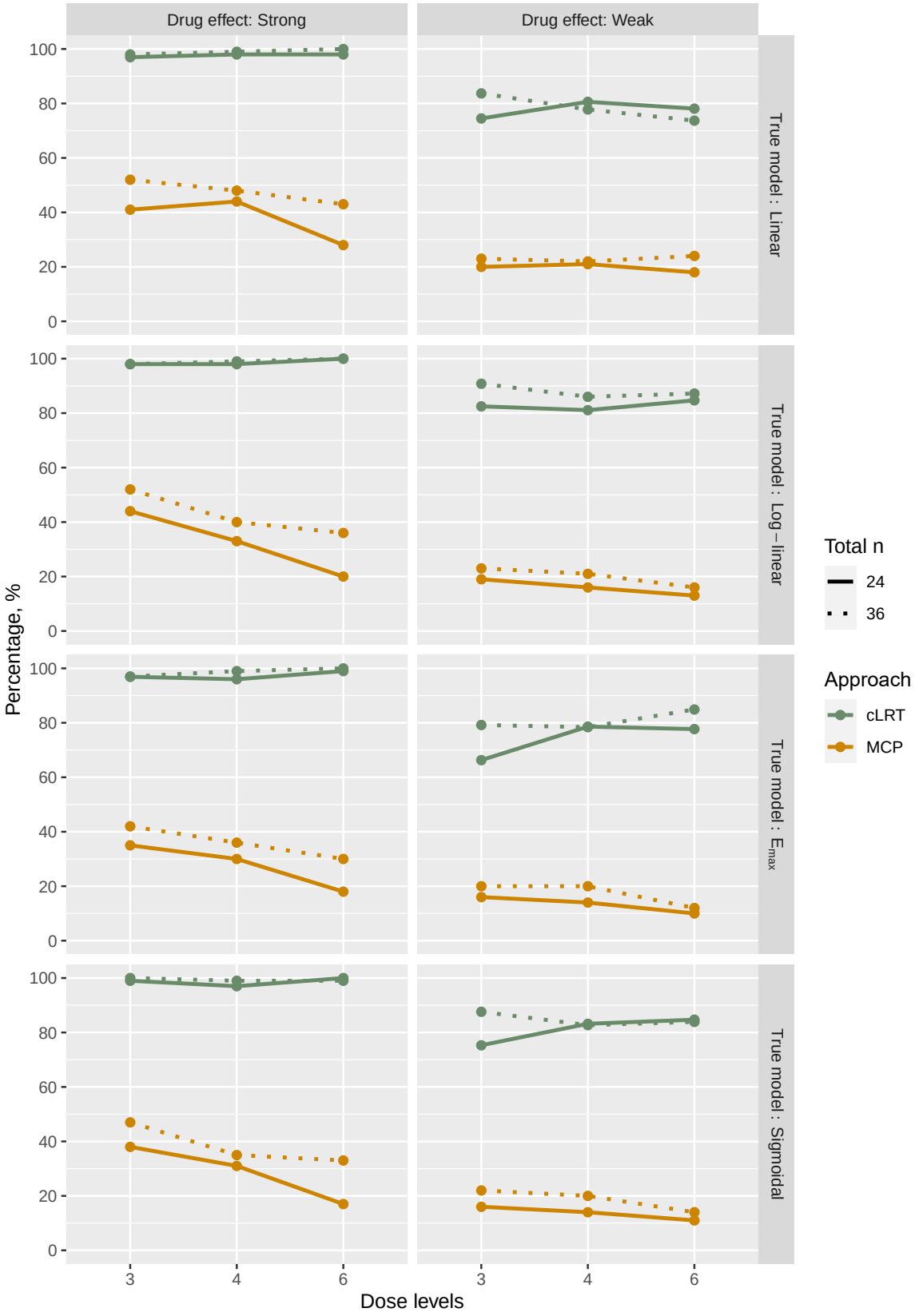

Figure 3.2: Power of cLRT and MCP approaches across the different study design variables stratified by strength of drug effect and dose-response relationship of the data-generating true model.
Colour coding: different analysis approaches (cLRT: green; MCP: orange), line type: total number of patients (dotted: n = 36; solid: n = 24). cLRT: combined likelihood ratio test; MCP: multiple comparison procedure.

3.1.1.3 Selection of best fitting candidate model and identification of true underlying dose-response model

The proportion of the selected best-fitting candidate dose-response models are depicted in Figs. 3.3 and 3.4 for the strong and weak drug effect, respectively, per each underlying dose-response relationship of the true model and for each group of scenarios sharing the same study variables (total number of patients, number of arms), and differing only in the dose-response relationship of the true model. Regardless of the study design variables and the dose-response relationship the true model, there was a clear tendency towards selection of the "simplest" dose-response relationship (i.e. models with lowest number of estimated parameters). The linear relationship was predominately selected if the true underlying dose-response relationship was also linear, whereas if the true underlying dose-response relationship was otherwise (i.e. log-linear, E_{max}, or sigmoidal), then selection favoured both the linear and log-linear relationships. For scenarios under the strong drug effect (Fig. 3.3), the selection of each dose-response relationship was as follows: linear relationship (30.3%–74.2%) followed by the log-linear relationship (11.3%–43.4%), and then alternating between the more complex E_{max} (6.1%–28.3%) and sigmoidal (3.1%–21.1%) relationships. For scenarios under the weak drug effect (Fig. 3.4), a similar pattern was observed. The selection of each dose-response relationship was as follows: linear relationship (34.1%–80.5%) followed by the log-linear relationship (10.7%–45.1%), and then alternating between the more complex E_{max} (1.2%–16.0%) and sigmoidal (1.4%–13.3%) relationships.

Identification of the true underlying dose-response relationship of the data-generating model was independent of the study design variables (bold barplots in Figs. 3.3 and 3.4). A total number of patients of $n = 36$ did not necessarily translate to a better chance of detecting the true dose-response relationship compared with a total number of patients of $n = 24$. Similarly with the number of dose levels. More complex models (E_{max} and sigmoidal) had a weaker chance of being successfully detected compared with the simpler models (linear and log-linear)—although they were generally better detected under the scenarios of the strong drug effect (Fig. 3.3) compared with the scenarios of the weak drug effect (Fig. 3.4). The best performance was observed only when the true dose-response relationship was linear.

Results for MCP are provided in Appendix Figs. B.3.1 and B.3.2.

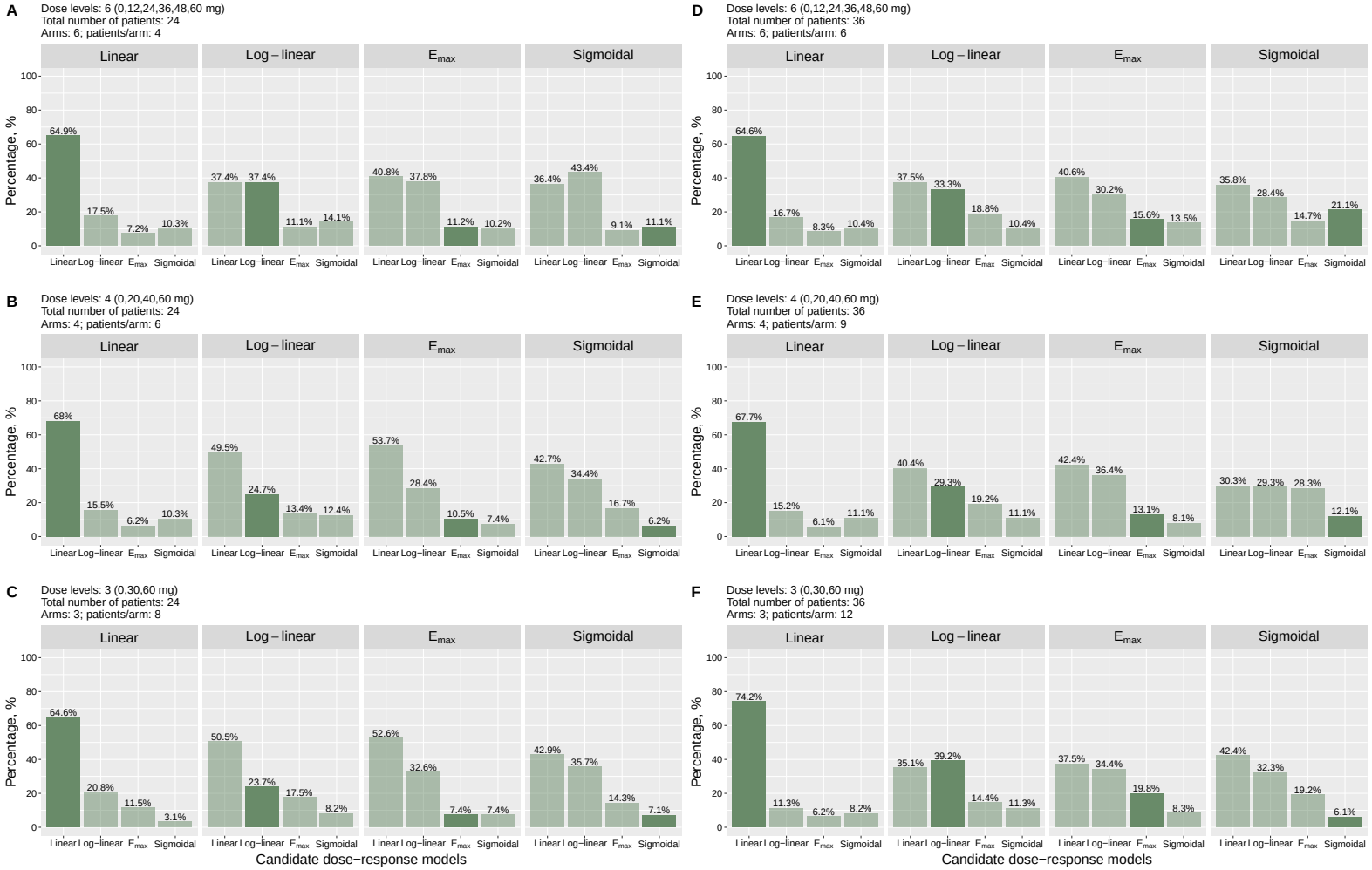

Figure 3.3: Proportion of the selected best-fitting candidate dose-response models under the ***strong*** drug effect highlighting (darker bars) the proportion of the identified true underlying relationship of the dose-response model for combined likelihood ratio test (cLRT), per each group of scenarios (panels **A**–**F**) sharing the same set of study design variables (i.e. number of dose levels, total number of patients, number of arms) and split (**facets**) by the underlying dose-response relationship of the true data-generating model. Number of dose levels: top panel (**A, D**) 6, middle panel (**B, E**) 4, lower panel (**C, F**) 3; Total number of patients: left panel (**A**–**C**): 24, right panel (**D**–**F**): 36.

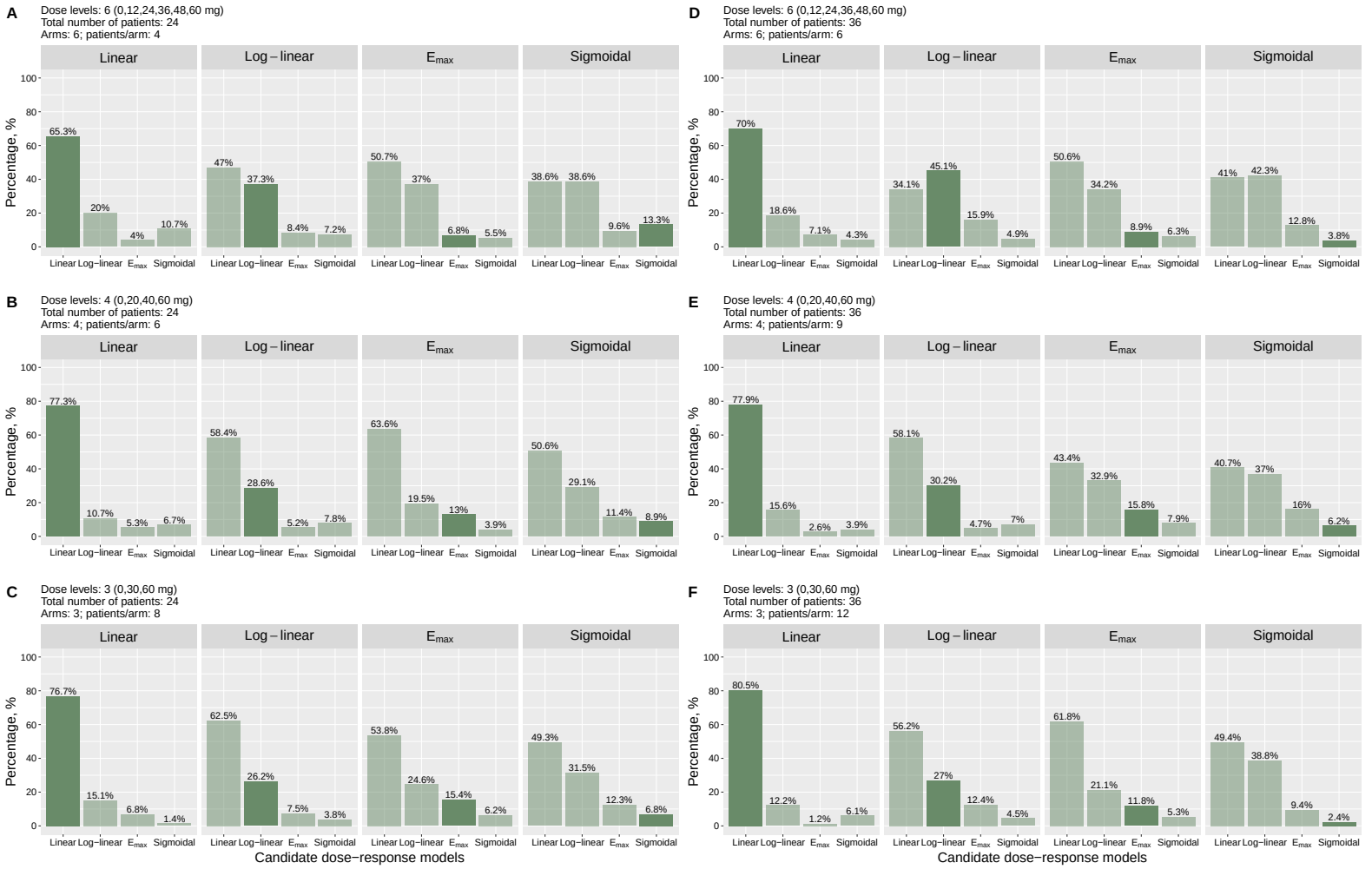

Figure 3.4: Proportion of the selected best-fitting candidate dose-response models under the ***weak*** drug effect highlighting (darker bars) the proportion of the identified true underlying relationship of the dose-response model for combined likelihood ratio test (cLRT), per each group of scenarios (panels **A**–**F**) sharing the same set of study design variables (i.e. number of dose levels, total number of patients, number of arms) and split (**facets**) by the underlying dose-response relationship of the true data-generating model. Number of dose levels: top panel (**A, D**) 6, middle panel (**B, E**) 4, lower panel (**C, F**) 3; Total number of patients: left panel (**A**–**C**): 24, right panel (**D**–**F**): 36.

3.2 Project II: Identification of predictors of efficacy in NSCLC patients: A tumour dynamics - CRP modelling framework

In *project II*, following an exploratory analysis of the clinical data, a framework combining anticancer drug exposure, tumour dynamics and CRP concentrations was developed to identify and quantify the significance of early longitudinal predictors of PFS and OS in patients with advanced NSCLC receiving first-line combination chemotherapy with paclitaxel and either carboplatin or cisplatin.

3.2.1 Clinical study data: The CEPAC-TDM study

The CEPAC-TDM study included a total of 365 patients. A detailed description of this population has been extensively reported in Joerger *et al.* [33] and Henrich [206]. However, in this project, the focus was on the subcohort of patients with CRP concentrations (n = 258), for which the extensive data exploration and model development were performed and for which the term "*patient population*" will refer to thereafter—unless otherwise stated. Yet, for the development of the tumour dynamics model under the multiple imputation (section 3.2.2.1), the whole population (n = 365) was used to obtain the population and individual tumour dynamics parameters. However, only the individual tumour dynamics parameters of our "*patient population*" were leveraged for the subsequent model development.

3.2.1.1 Patient demographics and clinical characteristics

Our clinical dataset included 258 patients with advanced NSCLC, balanced between both treatment arms. The patient population could be described as middle aged to elderly adults (median age: 64 years) with the majority being males (64.7%), and approximately half (51.2%) being former smokers. Patients varied from being underweight to extremely obese (body mass index (BMI) range: 16.8–41.7 kg/m^2). Markers of liver function (ALT, AST) were mildly elevated in few patients—as depicted by their respective range (5.00 U/L–125 U/L and 9.00 U/L–212 U/L)—although the majority of patients were with physiological values as reflected by the median ALT and AST of 23 U/L and 21 U/L, respectively. Adenocarcinoma was the most frequent histological NSCLC type (64%). The majority of patients had advanced metastatic disease i.e. stage IV (84.1%) and a systemic inflammation as depicted by the elevated baseline neutrophil-to-lymphocyte ratio (median > 2). Nevertheless, > 90% of patients had an ECOG performance status ≤ 1, and only < 20% of patients had brain and/or liver lesions. Detailed baseline patient demographics and clinical characteristics are shown in Table 3.1 and a stratification based on the two treatment arms is provided in Appendix Table A.5.1.

Table 3.1: Baseline patient demographics and clinical characteristics

Patient characteristic	All patients (n = 258)
Age [years]	
median	64.0
[range]	[41.0–78.0]
Sex, n (%)	
Female	91 (35.3)
Male	167 (64.7)
Treatment arm, n (%)	
BSA-guided paclitaxel dosing arm	126 (48.8)
PK-guided paclitaxel dosing arm	132 (51.2)
Body weight [kg]	
median	74.0
[range]	[42.0–135]
Body height [cm]	
median	171
[range]	[146–194]
Body mass index [kg/m^2]	
median	24.9
[range]	[16.8–41.7]
Body surface area [m^2]	
median	1.86
[range]	[1.34–2.49]
Neutrophil-to-lymphocyte ratio	
median	4.28
[range]	[0.798–94.0]
Alanine aminotransferase activity [U/L]	
median	23.0
[range]	[5.00–125]
Aspartate aminotransferase activity [U/L]	
median	21.0
[range]	[9.00–212]
Smoking status, n (%)	
Non-smokers	28 (10.9)
Former smokers	132 (51.2)
Current smokers	98 (38.0)
Disease stage, n (%)	
IIIB	41 (15.9)
IV	217 (84.1)
NSCLC histology, n (%)	
Adenocarcinoma	165 (64.0)
Bronchioalveolar carcinoma	1 (0.388)
Carcinoma, not otherwise specified	32 (12.4)
Squamous-cell carcinoma	60 (23.3)
Brain lesions, n (%)	
No	224 (86.8)
Yes	34 (13.2)
Liver lesions, n (%)	
No	207 (80.2)
Yes	51 (19.8)
Baseline ECOG performance status, n (%)	
0	135 (52.3)
1	105 (40.7)
2	18 (6.98)

BSA: body surface area; ECOG: Eastern Cooperative Oncology Group; NSCLC: non-small cell lung cancer; PK: pharmacokinetic

3.2.1.2 Assessment of efficacy: Tumour size, time-to-progression, progression-free survival and overall survival

A total of 995 tumour size assessments (RECIST v1.1 [56]), over a time span of 108 weeks, were available from the 258 patients (median: 5.53 cm; range: 0.6 cm–38.3 cm) with a median of four measurements per patient (range: 1–17 measurements). The distribution of tumour size over time is shown in Fig. 3.5. The median tumour size initially declined within the first 6 weeks, after the start of treatment. It then continued to gradually decline slowly until ≈19 weeks, after which a plateau was observed until the last observation at ≈108 weeks, possibly indicative of resistance development over time and/or drop out of patients with recorded progression. The tumour regrowth that could be seen in some of the individual profiles, after initial response, could be attributed to resistance development or EOT.

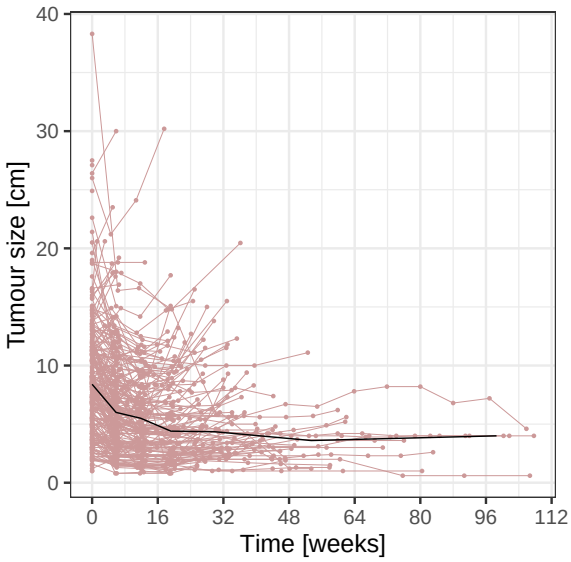

Figure 3.5: Individual tumour size versus time until 108 weeks (n = 995 measurements from 258 patients). Data points: tumour size data evaluated using the Response Evaluation Criteria in Solid Tumors (RECIST, version 1.1) [56], connected to show individual profiles; solid black line: median of observed tumour size data. Note the reduction in the data points with time due to recorded progression.

A total of 151 patients (58.5%) had documented times of tumour progression whereas 175 patients (67.8%) were recorded to have died, during the follow-up time. Patients with unknown time of event i.e. their progression or survival times were not recorded, were assumed to be alive, for OS information, or alive without progression, for PFS information, until their last follow-up visit and were consequently right censored at their time of last observation. The maximum recorded observation time for progression (i.e. time-to-progression and progression-free survival) was 28.7 months (123 weeks) whereas the maximum recorded observation time for survival was 32.5 months (140 weeks), with a median follow-up time of 8.92 months.

Not all patients completed the six treatment cycles, rather a median of four cycles were completed. The reason for not completing all the six treatment cycles (n = 168, 65.1%) did not necessarily imply that patients dropped out of the study, but it could have been due to exclusion from receiving treatment as a result of reported progressive disease, following tumour assessments (17.8%), death during the course of the study (8.93%), or other/unknown reasons (73.2%) e.g. intolerable toxicities or deterioration of health status. Nevertheless, the majority of these patients continued to be followed-up.

Hence, for the 168 patients who did not complete all the treatment cycles, progression and survival information were available for 71.4% (n = 120) and 58.3% (n = 98) of patients, respectively. Therefore, not completing all the six treatment cycles did not consequently imply a loss to follow-up nor was it a reason for right censoring i.e. patients who exited the study before its termination still contributed to the overall event information, if their TTE was available, and were alternatively right censored if their TTE was missing, and no special consideration (e.g. independent modelling of drop outs) was adopted to handle their missing information.

Fig. 3.6 shows the Kaplan-Meier plots of the efficacy endpoints: TTP, PFS and OS. The proportion of patients reporting a progressive disease gradually increased until 18 months with a median TTP of 7.26 months. After 18 months, no further progression events were recorded, denoted by the horizontal line with censored events at the end (Fig. 3.6, A). The proportion of patients who neither progressed nor died (Fig. 3.6, B) or were still alive (Fig. 3.6, C) gradually declined over time, from the time when treatment started, with a median PFS and OS of 5.97 months and 10.3 months, respectively. The PFS and OS profiles in this subset of patients (Fig. 3.6, B and C) were similar and representative of the profiles of the entire population from the CEPAC-TDM study [33] (no information on TTP was available in [33] for comparison): 25% PFS was ≈ 9 months vs ≈ 8 months, 50% PFS was ≈ 6 months vs ≈ 5 months, and 75% PFS was ≈ 3 months vs ≈ 2.5 months, respectively. Whereas 25% OS was ≈ 20 months vs ≈ 20 months, 50% OS was ≈ 10 months vs ≈ 10 months, and 75% OS was ≈ 6 months vs ≈ 5.5 months, respectively.

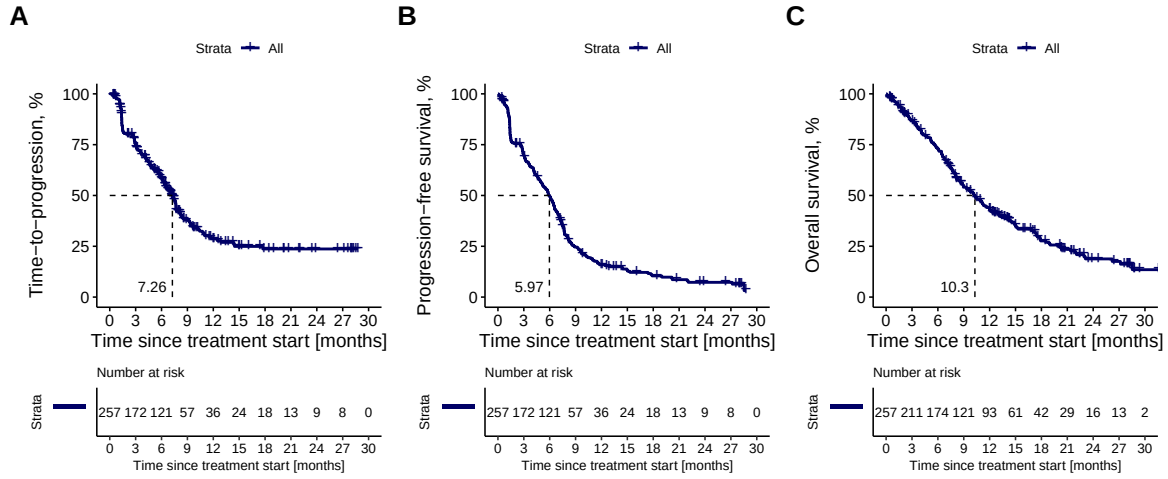

Figure 3.6: Kaplan-Meier plots of **[A]** time-to-progression, **[B]** progression-free survival and **[C]** overall survival. Cross: censored events; dashed line: line indicating median time-to-event.
Time-to-progression: time from the start of treatment until objective tumour progression—does not include deaths; progression-free-survival: time from the start of treatment until objective tumour progression or death, whichever occurs first; overall survival: time from the start of treatment until patient death of any cause [57]. n = 257 based on the final analysis dataset (section 3.2.1.3)

3.2.1.3 Exploratory analyses of CEPAC-TDM CRP data

In total, 945 CRP observations over the time span of 177 days were available from 258 individuals. Six samples (0.635%) were below the LLOQ, whereas only one sample (0.106%) was above the upper limit of quantification and was validated for dilution integrity. Because of the low percentage of below LLOQ samples, they were excluded from the analysis according to the M1 method described by Beal [139], with the assumption that they would not significantly influence model development. The

sample above the upper limit of quantification was temporarily excluded during the different model development stages, to avoid potential bias, but was later returned at the final stage of each key model for a realistic assessment of the model parameters. After exclusion of the below LLOQ samples, the number of observations was reduced to 939, representing 257 individuals.

CRP was sampled during treatment cycles 1, 2, 3, and at the EOT visit. Sampling frequency varied across the different cycles; with a total of 255 patients having baseline CRP concentrations and only 170 patients having an EOT CRP concentration. Few patients ($\approx 30\%$) with day 1 CRP samples, had additional measurements on day 2 during the first two cycles (Fig. 3.7).

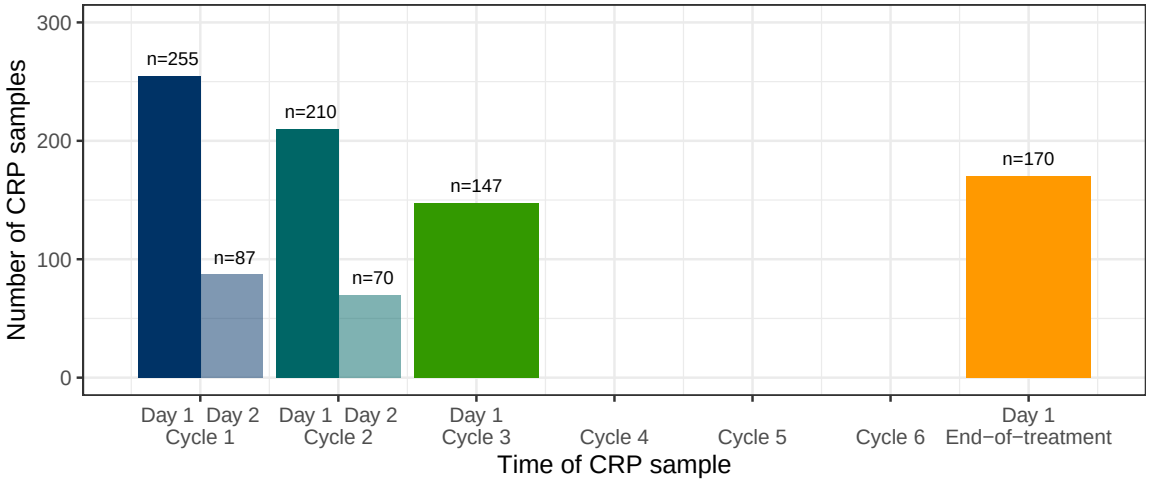

Figure 3.7: CRP sampling frequency across study time. CRP: C-reactive protein.

The distribution of the CRP concentrations was strongly right-skewed (Fig. 3.8, distributions of CRP concentrations stratified by the sampling time are provided in Appendix Fig. B.4.1) with a mean of 32.6 mg/L and a median of 13.4 mg/L, spanning a wide concentration range between 0.320 mg/L and 529 mg/L (CV%: 151%). This was suggestive of the need to use logarithmically transformed data for modelling purposes to account for the wide range of concentrations.

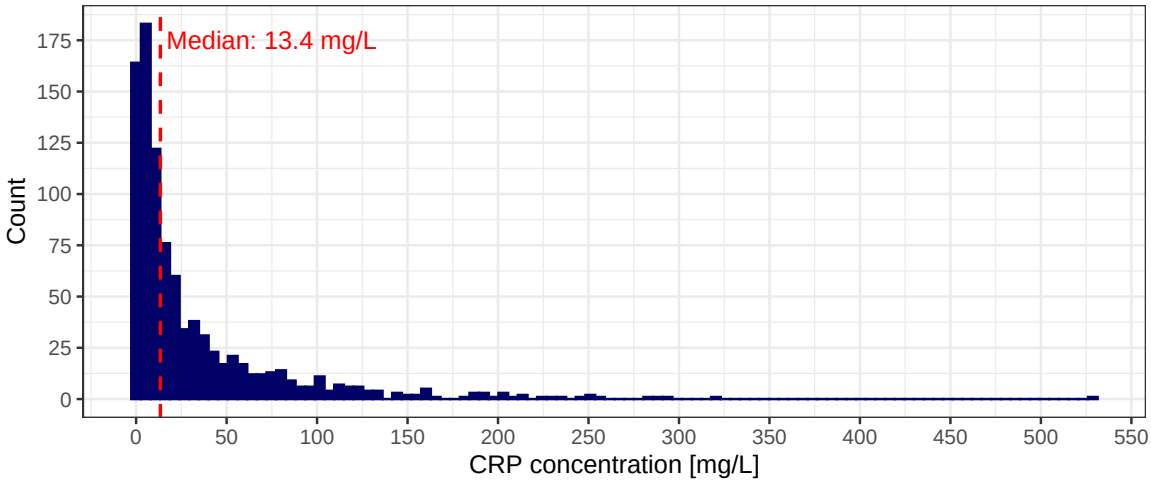

Figure 3.8: Histogram of CRP concentrations. Dashed red line: median CRP concentration. CRP: C-reactive protein.

The graphical visualisation of CRP concentrations over time revealed highly variable trends across patients, where baseline CRP concentrations ranged from 0.32 mg/L up to 281 mg/L. In general, CRP concentration showed an initial decrease over time and then stabilised staring from treatment cycle 3. Moreover, the time of the EOT sample was very heterogeneous and reflected the time whenever the patient exited the study (Fig. 3.9).

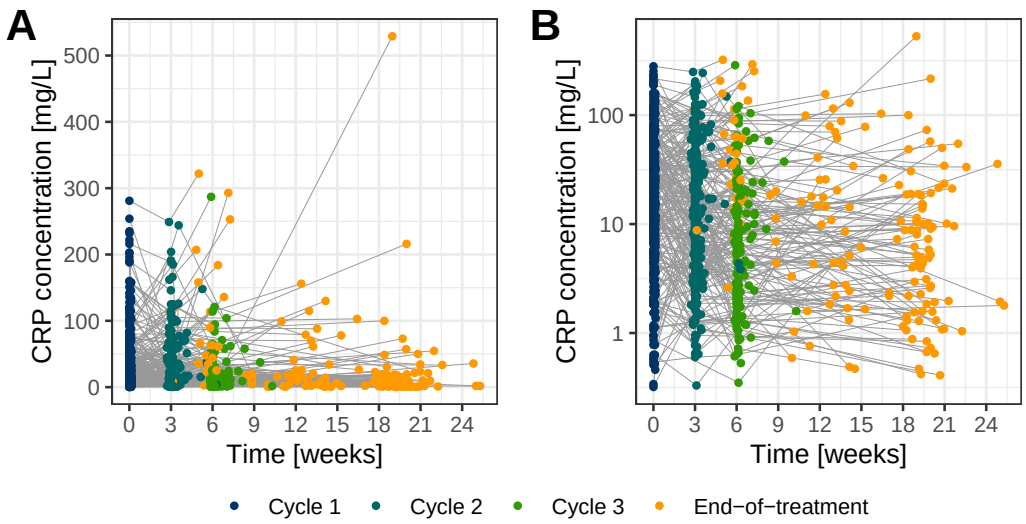

Figure 3.9: CRP concentrations versus time, colour-coded by sample time on **[A]** linear scale and **[B]** logarithmic scale. Coloured dots: CRP concentrations; grey lines: connected data points per individual. CRP: C-reactive protein.

A large magnitude of both positive and negative change from baseline was observed, between treatment cycles 1 and 2. This magnitude decreased as we went further in time (i.e. from cycle 2 to 3) and was almost steady from cycle 3 to EOT (Fig. 3.10).

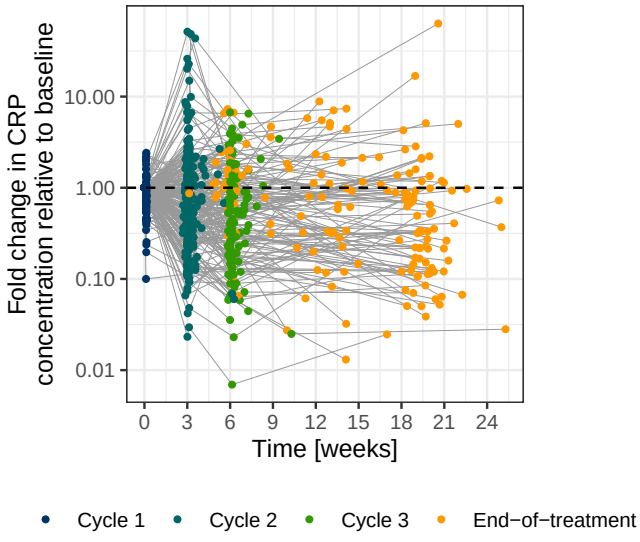

Figure 3.10: Fold change in CRP concentrations relative to baseline versus time on logarithmic scale, colour-coded by sample time. Coloured dots: CRP concentrations; grey lines: connected data points per individual. CRP: C-reactive protein.

Relation between CRP concentration and tumour size

Because of the lack of simultaneous times of CRP sampling and tumour size assessments, a direct investigation of a potential correlation between CRP concentrations and tumour size was not possible. Instead, the relation between baseline CRP concentration and baseline tumour size was investigated since both variables had baseline measurements. Baseline tumour size showed a positive correlation with baseline CRP concentrations (Fig. 3.11). Similarly, on a longitudinal level, the median tumour size measurements followed a similar trajectory compared with the median CRP concentrations over time, with an initial decline followed by a plateau (Fig. 3.12). Hence, our data supported the hypothesis that CRP concentration could be a reflection of tumour load and disease aggressiveness.

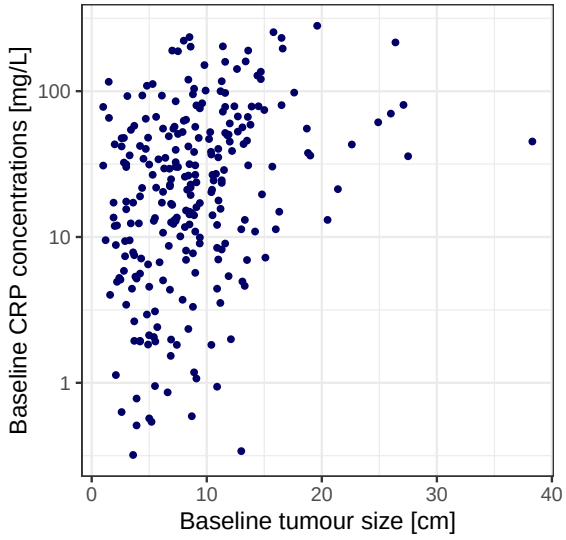

Figure 3.11: Baseline CRP concentrations (on a log-scale) versus baseline tumour size. CRP: C-reactive protein

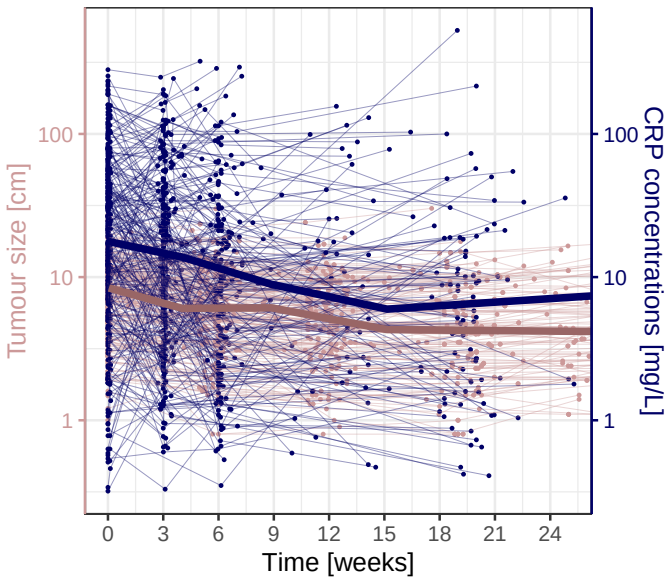

Figure 3.12: CRP concentrations (blue) and tumour size measurements (nude) versus time, colour-coded by variable type. Solid coloured line: median of observed CRP concentrations (blue) or tumour size (nude). CRP: C-reactive protein

Relation between baseline CRP concentration and time-to-event efficacy endpoints: Time-to-progression, progression-free survival and overall survival

Initial nonparametric (Fig. 3.13) and semi-parametric (Table 3.2) analyses were performed to assess the potential prognostic value of baseline CRP concentration on progression and survival endpoints. A strong significant difference was observed between patients with baseline CRP concentrations above and below median baseline CRP concentration (= 23.5 mg/L) with regards to PFS and OS (Fig. 3.13, B and C, respectively). A weaker, but still significant, influence was also observed with regards to TTP (Fig. 3.13, A).

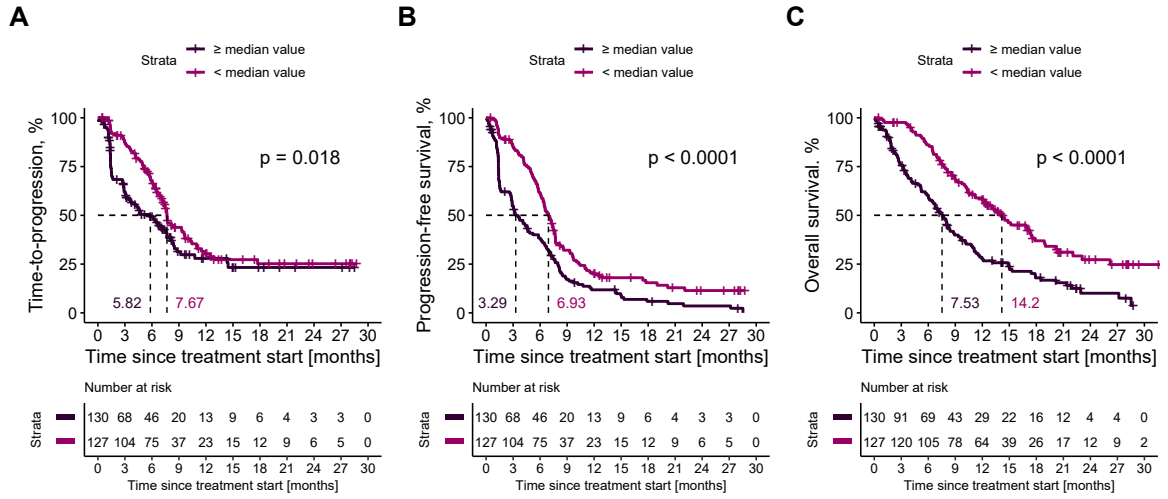

Figure 3.13: Kaplan-Meier plots of **[A]** time-to-progression, **[B]** progression-free survival and **[C]** overall survival, stratified by median baseline CRP concentration (23.5 mg/L). Cross: censored events. n = 257 based on the final analysis dataset (section 3.2.1.3).

Log-transformed baseline CRP concentration significantly influenced progression and survival endpoints as depicted by the results of the CoxPH model i.e. hazard ratio above 1 and 95% CI not including 1 (Table 3.2). Generally, higher log baseline CRP concentrations were associated with a worse prognosis; a one unit increase in log baseline CRP concentration was associated with an increased hazard of 21%, 31%, and 43% in TTP, PFS and OS, respectively.

Table 3.2: Results of Cox proportional hazards model of log-transformed baseline CRP concentration in relation to time-to-progression, progression-free survival and overall survival

Event	HR	95% CI	p-value*
Time-to-progression	1.21	1.06–1.37	$2.83 \cdot 10^{-3}$
Progression-free survival	1.31	1.18–1.46	$1.89 \cdot 10^{-7}$
Overall survival	1.43	1.27–1.63	$1.99 \cdot 10^{-9}$

CI: confidence interval; HR: hazard ratio
*p-value estimated based on likelihood ratio test

Results of both the nonparametric and semi-parametric analyses indicated that baseline CRP concentration was a strong predictor of efficacy endpoints. Most importantly, these results laid the basis and motivation for the further investigation of the prognostic value of longitudinal CRP concentration in a parametric TTE model and within our proposed modelling framework.

3.2.2 Modelling framework to identify predictors of efficacy in NSCLC patients

3.2.2.1 Development of a coupled tumour dynamics - CRP model to characterise circulating CRP concentration

To characterise the quantitative relationship between drug exposure, tumour dynamics and CRP concentration, a coupled tumour dynamics - CRP model was developed (Fig. 2.12, A). This was achieved by coupling the paclitaxel exposure-driven tumour size over time (Fig. 2.12, A, subsection 1) to the developed CRP model (Fig. 2.12, A, subsection 2), allowing the influence of tumour dynamics on CRP production to be quantified and the circulating CRP concentration over time to be characterised (Fig. 2.12, A, subsections 2–3).

Model development to characterise tumour dynamics

Tumour growth was estimated to occur linearly at a rate of 0.742 mm/month, while paclitaxel-induced tumour decay occurred at a first-order rate of $2.30{\cdot}10^{-5}$ $(\mu\text{mol}{\cdot}\text{h/L})^{-1}{\cdot}\text{h}^{-1}$. According to the developed tumour dynamics model (Eq. 2.62), a paclitaxel AUC_{cycle} of 4.49 μmol·h/L was required at the start of treatment to induce tumour shrinkage. Baseline paclitaxel effect declined exponentially with time at a rate of 0.021 day^{-1} to reach only 41.4% and 17.1% by the end of treatment cycles 2 and 4, respectively [67]. The detailed description of the performance and evaluation of the paclitaxel-induced tumour dynamics model has been previously published and reported in [67, 211].

Table 3.3 shows the final population tumour dynamics parameter estimates from both treatment arms of the whole patient population (n = 356) as well as the descriptive statistics of the estimated individual tumour dynamics parameter estimates, summarised from the tumour dynamics model developed within the multiple imputation framework—as described in section 2.6.3.1. As can be seen in Table 3.3, the median values of our derived individual tumour dynamics parameter estimates were closely in line with the previously reported final population tumour dynamics parameter estimates [67, 211].

The derived individual tumour sizes over time based on the individual tumour dynamics parameter estimates and their associated uncertainty (SE) are shown in Fig. 3.14 against the respective observed individual tumour size measurements for our patient population (i.e. subset of patients with CRP information; n = 257). No systematic bias or trends were observed as indicated by the regression trendline and the distribution of data points around the line of identity. Hence, the estimated individual tumour sizes could be reliably leveraged to further inform the circulating CRP concentration (Fig. 2.12, A, subsection 2).

Model development to characterise CRP concentration

The developed CRP turnover base model estimated both the population and individual CRP concentrations under the *steady-state* assumption. A population CRP production rate constant ($K_{in,CRP}$) of 0.508 $\text{mg}{\cdot}\text{L}^{-1}{\cdot}\text{h}^{-1}$ (95% CI: 0.433–0.594) was estimated with a large IIV of 123% (Table 3.4) and, based on a fixed CRP half-life of 19 h [97], corresponded to a population steady-state baseline CRP concentration of 13.9 mg/L.

Table 3.3: Population and individual tumour dynamics parameters (n = 365)

Parameter [unit]	Population parameters* Estimate	(RSE, %)	Individual parameters Estimate median (range)	SE median,% (range)
Fixed-effects parameters[a]				
K_{growth} [cm/h]	$1.03{\cdot}10^{-4}$	(43.2)	$1.04{\cdot}10^{-4}$ $(2.79 \cdot 10^{-5}$–$3.88 \cdot 10^{-3})$	1.17 (0–1.25)
β $[(\mu\text{mol}{\cdot}\text{h/L})^{-1}{\cdot}\text{h}^{-1}]$	$2.30{\cdot}10^{-5}$	(12.8)	$2.30{\cdot}10^{-5}$ $(1.11 \cdot 10^{-5}$–$2.11 \cdot 10^{-4})$	0.612 (0–0.853)
λ [1/h]	$8.75{\cdot}10^{-4}$	(15.4)	$8.75{\cdot}10^{-4}$ $(2.42 \cdot 10^{-4}$–$1.22 \cdot 10^{-3})$	0.494 (0–0.538)
Interindividual variability parameters				
$\omega_{K_{growth}}$ [CV,%]	117	(41.3)	—	—
ω_{β} [CV,%]	86.3	(14.7)	—	—
ω_{λ} [CV,%]	53.9	(38.2)	—	—
Residual variability parameter				
${\sigma_{exp}}^{b}$ [CV,%]	17.9	(7.17)	—	—

* as reported in [67]
[a] as described in Eq. 2.62
[b] estimated as additive residual variability on log-scale
β: paclitaxel area under the concentration-time curve from the start to end of a cycle-driven tumour decay rate constant at the start of treatment (t = 0); CV: coefficient of variation; K_{growth}: linear tumour growth rate constant; λ: rate constant for exponential decline in drug effect over time; RSE: relative standard error; SE: standard error.

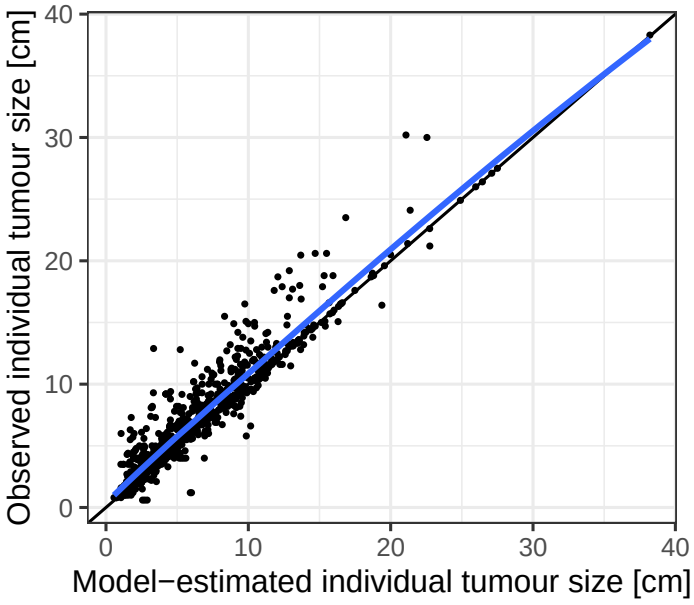

Figure 3.14: Observed individual tumour size versus model-estimated individual tumour size derived from the tumour dynamics model within the multiple imputation approach ($n_{observations} = 987$ from 257 patients). Solid black line: line of identity of slope 1, blue solid line: loess regression.

The visual inspection of the graphical exploration of $K_{in,CRP}$-associated variability against different covariates showed a potential trend with baseline tumour size, baseline IL-6, smoking status, NSCLC histology, disease stage, liver lesions, and baseline ECOG status (Appendix Fig. B.5.2 and B.5.1). Therefore, based on clinical relevance and graphical exploration, the following covariates were pre-selected to be tested for a potential impact on $K_{in,CRP}$:

- Covariates that reflect disease aggressiveness:
  - *Continuous covariates:* baseline tumour size i.e. baseline sum of diameters
  - *Categorical covariates*: disease stage, baseline ECOG status, presence/absence of brain lesions, number of target lesions, number of non-target lesions, and sum of target and non-target lesions
- Covariates that reflect an inflammatory status:
  - *Continuous covariates:* BMI as marker of obesity
  - *Categorical covariates:* smoking status
- Covariates of physiological relevance to CRP synthesis by hepatocytes:
  - *Continuous covariates:* ALT and AST as markers of liver injury, baseline IL-6 as cytokine precursor to CRP production
  - *Categorical covariates:* presence/absence of liver lesions as marker of liver injury
- Covariates which showed correlation with the individual variability of $K_{in,CRP}$ in the graphical exploration (Appendix Fig. B.5.2 and B.5.1):
  - *Categorical covariates:* NSCLC histology

On the other hand, no visible correlations were observed in the exploratory plots between age, weight, height, BSA, sex and $K_{in,CRP}$-associated variability (Appendix Fig. B.5.2 and B.5.1). Therefore, the potential impact of these covariates was not further explored.

Before exploring the impact of the pre-selected covariates and to avoid a biased estimation of the parameter-covariate relationship, the developed base model was modified where the outlying CRP concentration (section 3.2.1.3) and the individual with the spurious baseline IL-6 concentration of 34.9-fold higher concentration than the median baseline IL-6 concentration of 2.57 pg/mL (Appendix Fig. B.5.1) were *temporarily* excluded during the covariate model development steps but were later re-introduced to the final covariate model. The pre-selected covariates were then investigated for potential correlations (correlation matrix shown in Appendix Fig. B.6.1, number of identified correlations = 3), such that if one of the strongly correlated covariates (i.e. $r > \pm 0.6$) was included, the corresponding correlated pair was excluded from further investigation.

Following the forward inclusion steps of the SCM procedure, only baseline IL-6, baseline tumour size, disease stage, smoking status, and ALT were found to significantly impact $K_{in,CRP}$ and were included in the CRP turnover model. These covariates contributed to a significant reduction in OFV of 95.8, 14.9, 9.23, 10.7 and 5.66 points, respectively (threshold > 3.84 for $\alpha = 0.05$, 1 *df*). However, following the backward deletion step, ALT was removed from the model as the OFV increased by 5.66 points (threshold > 6.64 for $\alpha = 0.001$, 1 *df*)—further details of the SCM steps are provided in Appendix A.6, Tables A.6.1 and A.6.2. On the other hand, baseline ECOG status, NSCLC histology, BMI, AST, presence of liver lesions, presence of brain lesions, number of non-target lesions, number of target lesions, and sum of target and non-target lesions did not explain the variability between patients (i.e. IIV) associated with the CRP production rate constant, $K_{in,CRP}$.

Thus, the final covariate CRP turnover model only included baseline IL-6 described by a linear function, baseline tumour size described by an exponential function, and disease stage and smoking status described by a fractional change model on $K_{in,CRP}$ (Eq. 3.1 and Fig. 3.15). Inclusion of these covariates explained 27.2 % of the variability associated with $K_{in,CRP}$ (IIV_{kin}) and resulted in a 110 point drop in OFV (p-value $= 4.097 \cdot 10^{-22}$, 5 df) from the base model (Table 3.4). Hence, the equation for $K_{in,CRP}$ was given by (Eq. 3.1),

$$K_{in,CRP} = \theta_{Kin,CRP} \cdot \big(1 + \theta_{IL6} \cdot (IL6 - 2.57)\big) \cdot exp\big(\theta_{BLTS} \cdot (BLTS - 8.25)\big) \cdot \begin{cases} 1, & \text{Non-smokers} \\ 1 + \theta_{SMK2}, & \text{Former smokers} \\ 1 + \theta_{SMK3}, & \text{Current smokers} \end{cases} \cdot \begin{cases} 1, & \text{Stage IV} \\ 1 + \theta_{stage}, & \text{Stage IIIB} \end{cases} \cdot exp(\eta_{kin,CRP}) \tag{3.1}$$

where $\theta_{Kin,CRP}$ is the $K_{in,CRP}$ of a non-smoker with disease stage IV, baseline IL-6 of 2.57 pg/mL, and baseline tumour size of 8.25 cm (median values); θ_{IL6} is the fractional change in $K_{in,CRP}$ per unit change in baseline IL-6 from the median value of 2.57 pg/mL; θ_{BLTS} is the exponent reflecting the change in the natural log of $K_{in,CRP}$ per unit change in baseline tumour size from the median value of 8.25 cm; θ_{SMK2} and θ_{SMK3} are the fractional changes in $K_{in,CRP}$ in former and current smokers, respectively, compared with non-smokers; and θ_{stage} is the fractional change in $K_{in,CRP}$ in patients with disease stage IIIB compared with patients with disease stage IV.

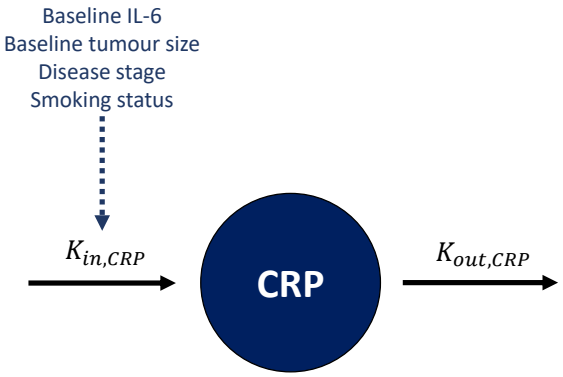

Figure 3.15: Schematic diagram of the final covariate CRP turnover model.
IL-6: interleukin 6; $K_{in,CRP}$: CRP zero-order production rate constant; $K_{out,CRP}$: CRP first-order degradation rate constant. CRP: C-reactive protein.

The final covariate CRP turnover model estimated $K_{in,CRP}$ for a non-smoker with stage IV NSCLC when baseline IL-6 is 2.57 pg/mL and tumour size is 8.25 cm to be 0.297 $\text{mg}\cdot\text{L}^{-1}\cdot\text{h}^{-1}$ (95% CI: 0.204–0.429, Table 3.4) corresponding to a population steady-state baseline CRP concentration of 8.14 mg/L.

Table 3.4 shows the parameter estimates of both the base and final covariate CRP turnover models. In general, parameters were precisely estimated (RSE $< 30\%$), with the exception of the effect of the former smoking status on $K_{in,CRP}$ whose RSE was 57.8% (95% CI: 0.020–1.415). The models were also robust as shown from the high proportion of the successful bootstrap runs. Plots showing individual and population steady-state predictions per individual, as predicted by the final covariate CRP turnover model, are provided in Appendix Fig. B.7.1.

Table 3.4: Parameter estimates of the base and covariate CRP turnover models (n = 257)

Parameter [unit]	Base model Estimate (RSE, %)	95% CI[a]	Covariate model Estimate (RSE, %)	95% CI[b]
Fixed-effects parameters				
$K_{in,CRP}$ [mg·L^{-1}·h^{-1}]	0.508 (8.30)	0.433–0.594	0.297 (17.9)	0.204–0.429
$K_{out,CRP}$ [h^{-1}]	0.0365*		0.0365*	
Covariate-effects parameters on $K_{in,CRP}$[c]				
Baseline IL-6	—	—	0.263 (14.0)	0.175–0.324
Baseline tumour size	—	—	0.0432 (28.0)	0.017–0.070
Disease stage IIIB	—	—	-0.401 (28.7)	-0.596–-0.102
Former smokers	—	—	0.536 (57.8)	0.020–1.415
Current smokers	—	—	1.11 (40.0)	0.378–2.272
Interindividual variability parameters				
$\omega_{Kin,CRP}$ [CV,%]	123 (4.30)	112–132	95.3 (7.60)	80.2–109
Residual variability parameters				
σ_{exp}[d] [SD, mg/L]	0.892 (3.70)	0.822–0.955	0.889 (3.70)	0.818–0.953

OFV for base model = 1235.857; OFV for covariate model = 1125.852
[a] 95% CI obtained from 1000 bootstrap runs (successful minimisation = 100%)
[b] 95% CI obtained from 1000 bootstrap runs (successful minimisation = 99.6%)
[c] as described in Eq. 3.1
[d] estimated as additive residual variability on log-scale
* fixed to literature value [97]
CI: confidence interval; CRP: C-reactive protein; CV: coefficient of variation; IL-6: interleukin 6; $K_{in,CRP}$: CRP zero-order production rate constant; $K_{out,CRP}$: CRP first-order degradation rate constant; OFV: objective function value; RSE: relative standard error; SD: standard deviation.

The forest plot in Fig. 3.16 depicts the univariate influence of each of the included covariates on $K_{in,CRP}$. It shows that the upper (95$^{\text{th}}$ percentile) and lower (5$^{\text{th}}$ percentile) percentiles of baseline IL-6 were associated with a 325% increase and a 56.2% decrease in $K_{in,CRP}$, respectively, with an estimated 9.7-fold change between those two percentiles. Similarly, the upper (95$^{\text{th}}$ percentile) and lower (5$^{\text{th}}$ percentile) percentiles of baseline tumour size were associated with a 50% increase and a 22.9% decrease in $K_{in,CRP}$, respectively, with an estimated 1.95-fold change between those two percentiles. Compared to non-smokers, a smoking status, either former or current, was associated with a positive impact on CRP synthesis with both former and current smokers having a 52% and a 109% increase in $K_{in,CRP}$, respectively (1.38-fold change). Finally, a less aggressive disease state, depicted by the less advanced disease stage IIIB, was associated with a 39.6% decrease in $K_{in,CRP}$, compared to disease stage IV.

A sub-population of current smokers, with disease stage IV, high baseline IL-6 (14.9 pg/mL i.e. 95$^{\text{th}}$ percentile) and high tumour load (baseline tumour size: 17.8 cm, i.e. 95$^{\text{th}}$ percentile) would be estimated to have a 68.3-fold higher $K_{in,CRP}$ and consequently a less favourable higher inflammatory level i.e. baseline CRP at steady state, compared to a sub-population of non-smokers with a less aggressive disease stage (stage IIIB), lower tumour load (baseline tumour size: 2.20 cm i.e. 5$^{\text{th}}$ percentile) and low baseline IL-6 (0.438 pg/mL i.e. 5$^{\text{th}}$ percentile).

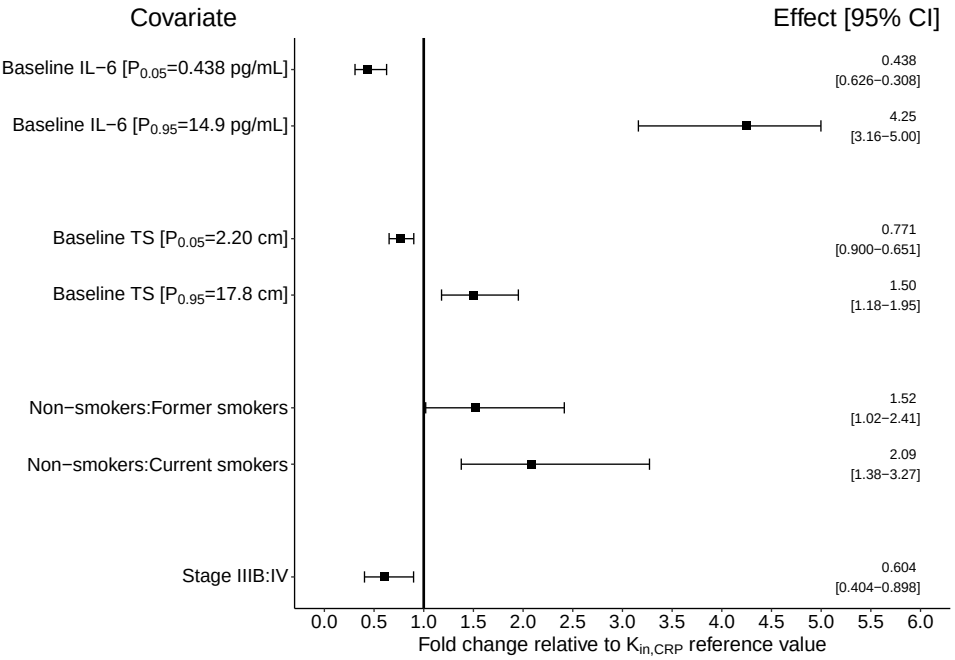

Figure 3.16: Forest plot of the impact of significant covariates on C-reactive protein production rate constant ($K_{in,CRP}$) relative to the reference value 1 (bold vertical line) representing the reference patient (non-smoker with disease stage IV, baseline tumour size = 8.25 cm, and baseline IL-6 = 2.57 pg/mL).
Effects of the continuous covariates (i.e. baseline IL-6, baseline tumour size) are shown at the 5^{th} and 95^{th} percentiles of the respective covariate distribution and the effects of the categorical covariates (i.e. smoking status, disease stage) are shown relative to the reference category of the respective categorical covariate. Black boxes: covariate effects; horizontal lines: 95% confidence intervals (CI); IL-6: interleukin 6; TS: tumour size.

Linking tumour dynamics to CRP

Before informing the CRP turnover model with the longitudinal model-derived individual tumour size data, the model was subjected to few modifications. First, since baseline tumour size was inherently considered within the longitudinal tumour size data, the former was removed from the model (as a covariate) to avoid its redundant inclusion and allow an unbiased, and now more informative estimation of the relationship between tumour size and CRP production. Second, to ensure a basal physiological level of CRP production regardless of the tumour size impact, a basal unperturbed $K_{in,CRP}$, $K_{in,CRP,basal}$, was included. $K_{in,CRP,basal}$ was fixed to 0.0109 mg/L and reflected a basal CRP concentration of 0.3 mg/L, corresponding to the LLOQ of CRP concentration. Thus, Fig. 3.15 could now be depicted as represented in Fig. 3.17.

Amongst the different models and the different tumour size-related metrics that were explored (section 2.6.3.1), only three different structural models fulfilled the selection criteria whereas others failed to successfully minimise or showed misspecifications in their predictive performance. These three structural models included:

- *Model A* with a power function relating the tumour size relative to baseline tumour size with $K_{in,CRP}$; associated with a 114 point drop in AIC.
- *Model B* with an exponential function relating the change (difference) between tumour size and baseline tumour size with $K_{in,CRP}$; associated with a 100 point drop in AIC.
- *Model C* with a linear model relating tumour size relative to baseline tumour size with $K_{in,CRP}$; associated with a 110 point drop in AIC.

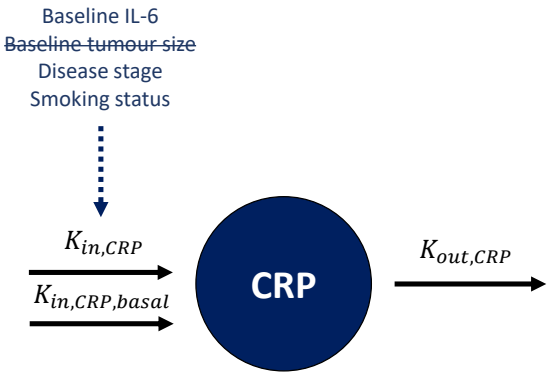

Figure 3.17: Updated schematic diagram of the final covariate CRP turnover model to allow for the investigation of the influence of longitudinal tumour size data on $K_{in,CRP}$. Compared to the original schematic diagram depicted in Fig. 3.15, baseline tumour size was removed from the set of influential covariates on $K_{in,CRP}$ and a basal production rate constant, $K_{in,CRP,basal}$, was added and fixed to 0.0109 mg/L (corresponding to the lower limit of quantification of CRP concentration of 0.3 mg/L). IL-6: interleukin 6; $K_{in,CRP}$: CRP zero-order production rate constant; $K_{out,CRP}$: CRP first-order degradation rate constant; CRP: C-reactive protein.

Inclusion of an IIV component to the function describing the relation between tumour size-related metric and $K_{in,CRP}$ for each of the above-mentioned models resulted in no improvement to (i) model A (IIV was unidentifiable) or to (ii) model B (Δ AIC = 2). However, on the contrary, the linear model relating tumour size relative to baseline tumour size with $K_{in,CRP}$ (i.e. model C) was associated with a 12.8 point drop in AIC. Therefore, the latter was not only numerically superior compared with the other models but also more mechanistically sound in characterising the diverse patient profiles and was consequently selected as the best model to describe the relation between tumour dynamics, represented by tumour size, and CRP concentration (AIC = 993; base model AIC = 1115, Δ AIC = -122).

Therefore, the relation between tumour dynamics and $K_{in,CRP}$ was characterised by a linear model relating the ratio of tumour size to baseline tumour size at any given time (i.e. x-fold change in tumour size from baseline tumour size: $\frac{Tumour\,size\,(t)}{Baseline\,tumour\,size}$), to CRP production rate constant, $K_{in,CRP}$ (Fig. 3.18). In this model, tumour size-dependent CRP production rate constant ($K_{in,CRP,TS}$) positively and linearly changed per unit change in the ratio of tumour size relative to baseline tumour size, by a factor of 0.819 as a function of the magnitude of $K_{in,CRP}$ (i.e. $K_{in,CRP,TS} = K_{in,CRP} \cdot 0.819 \cdot \frac{Tumour\,size\,(t)}{Baseline\,tumour\,size}$) (Fig. 3.19, A). In other words: at a specific magnitude of $K_{in,CRP}$, $K_{in,CRP,TS}$ positively and linearly changed by a factor of 0.819 per unit change in the product of $K_{in,CRP}$ and the ratio of tumour size to baseline tumour size (Fig. 3.19, B). Such that at time point, t, for a patient with a typical $K_{in,CRP}$ of 0.39 $mg{\cdot}L^{-1}{\cdot}h^{-1}$ (Table 3.5) and a ratio of tumour size to baseline tumour size of 1, a change in that ratio from 1 to 0.5 (i.e. 50% tumour shrinkage) would linearly decrease $K_{in,CRP,TS}$ from 0.319 $mg{\cdot}L^{-1}{\cdot}h^{-1}$ to 0.159 $mg{\cdot}L^{-1}{\cdot}h^{-1}$ i.e. 50% reduction in CRP production rate constant and consequently (after accounting for $K_{in,CRP,basal}$) result in a 51.5% reduction in CRP concentration. A large IIV of 60.4%CV was associated with this linear relationship and explained the variability associated with the patients CRP profiles and their diverse individual tumour sizes.

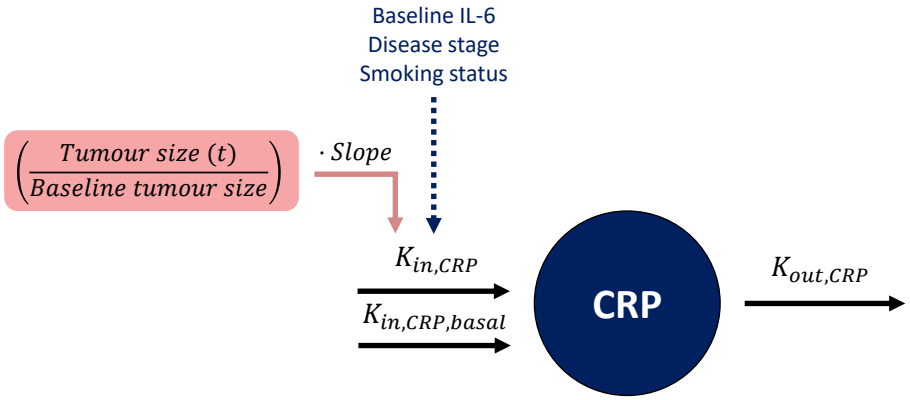

Figure 3.18: Schematic diagram of the coupled tumour size - CRP model. In the CRP turnover model (dark blue), CRP concentration was influenced by two production rate constants ($K_{in,CRP}$) and ($K_{in,CRP,basal}$), the latter being unperturbed by covariates or tumour size to ensure a basal level of CRP concentration. $K_{in,CRP}$ was influenced by baseline interleukin 6 (IL-6), disease stage and smoking status. The tumour dynamics model-derived longitudinal ratio of tumour size to baseline tumour size (pink) informed CRP production through a linear relationship (i.e. slope parameter).
CRP: C-reactive protein; $K_{out,CRP}$: CRP first-order degradation rate constant.

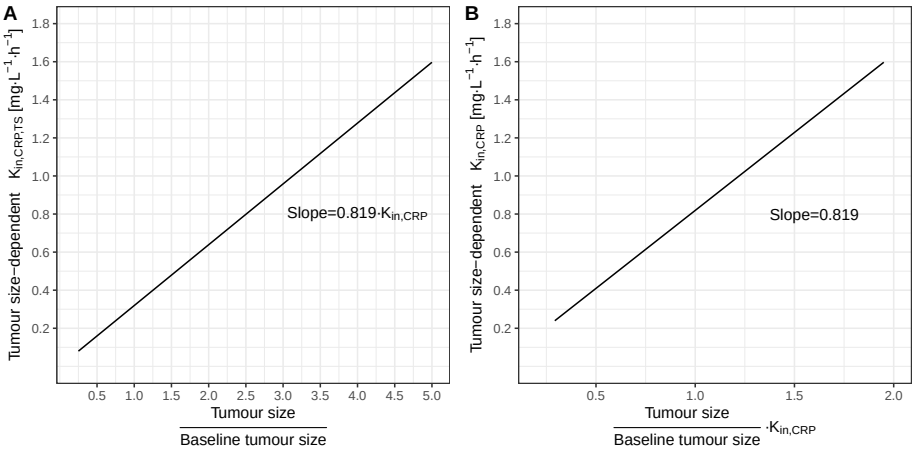

Figure 3.19: Visualisation of the relationship between tumour size-related metric and CRP production. This relationship can be described as either **[A]** tumour size-dependent $K_{in,CRP}$ i.e. $K_{in,CRP,TS}$ positively and linearly changes per unit change in the ratio of tumour size relative to baseline tumour size by a factor of 0.819 as a function of the magnitude of $K_{in,CRP}$ or **[B]** that at a specific magnitude of $K_{in,CRP}$, CRP production would positively change by a linear factor of 0.819 per unit change in the product of $K_{in,CRP}$ and the ratio of tumour size to baseline tumour size.
Note that this visualisation does not consider the basal unperturbed CRP zero-order production rate constant ($K_{in,CRP,basal} = 0.0109$ mg·L^{-1}·h^{-1}) which should be added to the magnitude of $K_{in,CRP,TS}$ for a total estimate of CRP production rate constant. CRP: C-reactive protein.

Re-inclusion of the outliers (i.e. the outlying CRP measurement and the individual with high baseline IL-6), which were temporarily excluded, increased the residual variability by 0.005 mg/L (standard deviation scale). Nevertheless, this did not compromise on the precision with which parameters were estimated (RSE < 26%). Table 3.5 shows the final parameter estimates of the coupled tumour size - CRP turnover covariate model. The model estimated both the fixed- and random-effects parameter estimates with high precision; RSE < 26% and < 15.2%, respectively (Table 3.5). Moreover, the nonparametric bootstrap confirmed the model's robustness (successful minimisation > 90%) and calculated narrow 95% CIs reflective of the level of parameter uncertainty

(Table 3.5). Graphical evaluation of the model showed no misspecification (Fig. 3.20) and an adequate predictive performance (Fig. 3.21): Plots of model predictions versus observations of CRP concentration showed a uniform and random distribution of the data points around the line of identity, indicative of good model prediction (Fig. 3.20 A, B); and the plots of residuals showed a uniform and random distribution around zero (Fig. 3.20 C, D). Finally, the VPC also showed an adequate predictive performance across the different CRP percentiles over time in which the median trajectory of the observed CRP concentration (i.e. 50^{th} percentile) as well as the 5^{th} and 95^{th} percentiles overlay the respective percentiles of the simulated data and fell within the 95% CIs of the simulated percentiles (Fig. 3.21). Plots showing individual and population predictions per individual are provided in Appendix Fig. B.8.1 and could be compared to the steady-state predictions previously predicted by the CRP turnover model (Appendix Fig. B.7.1).

Table 3.5: Parameter estimates of the coupled tumour dynamics - CRP turnover model (n = 257)

Parameter [unit]	Estimate	(RSE, %)	95% CI[a]
Fixed-effects parameters			
$K_{in,CRP}$ [mg·L^{-1}·h^{-1}]	0.39	(0.60)	0.252–0.602
$K_{in,CRP,basal}$ [mg·L^{-1}·h^{-1}]	0.0109!		—
$K_{out,CRP}$ [h^{-1}]	0.0365*		—
Slope (linear model linking tumour size to CRP)	0.819	(6.70)	0.711–0.952
Covariate-effects parameters on $K_{in,CRP}$[b]			
Baseline IL-6	0.315	(8.20)	0.244–0.363
Disease stage IIIB	-0.392	(26.0)	-0.598–-0.097
Former smokers	0.645	(12.1)	0.0353–1.64
Current smokers	1.26	(19.0)	0.398–2.56
Interindividual variability parameters [CV,%]			
$\omega_{Kin,CRP}$	92.1	(7.40)	74.2–107
ω_{slope}	60.4	(15.2)	40.3–77.9
$\omega_{K_{growth}}$	100#		—
ω_{β}	100#		—
ω_{λ}	100#		—
$\omega_{\text{baseline tumour size}}$	100#		—
Residual variability parameters			
σ_{exp}[c] [SD, mg/L]	0.763	(1.70)	0.686–0.831

OFV = 1011.738
[a] 95% CI obtained from 1000 bootstrap runs (successful minimisation = 97.9%)
[b] as described in Eq. 3.1
[c] estimated as additive residual variability on log-scale
! fixed to corresponding lower limit of quantification of CRP concentration (0.3 mg/L)
* fixed to literature value [97]
# fixed, variability derived from reported uncertainty as described in [218]
β: paclitaxel area under the concentration-time curve from the start to the end of a cycle-driven tumour decay rate constant at start of treatment (t = 0); CI: confidence interval; CRP: C-reactive protein; CV: coefficient of variation; IL-6: interleukin 6; $K_{in,CRP}$: CRP zero-order production rate constant of a non-smoker with disease stage IV, and baseline IL-6 of 2.57 pg/mL (median value); $K_{in,CRP,basal}$: CRP basal unperturbed zero-order production rate constant; $K_{out,CRP}$: CRP first-order degradation rate constant; K_{growth}: linear tumour growth rate constant; λ: rate constant for exponential decline in drug effect over time; OFV: objective function value; RSE: relative standard error; SD: standard deviation.

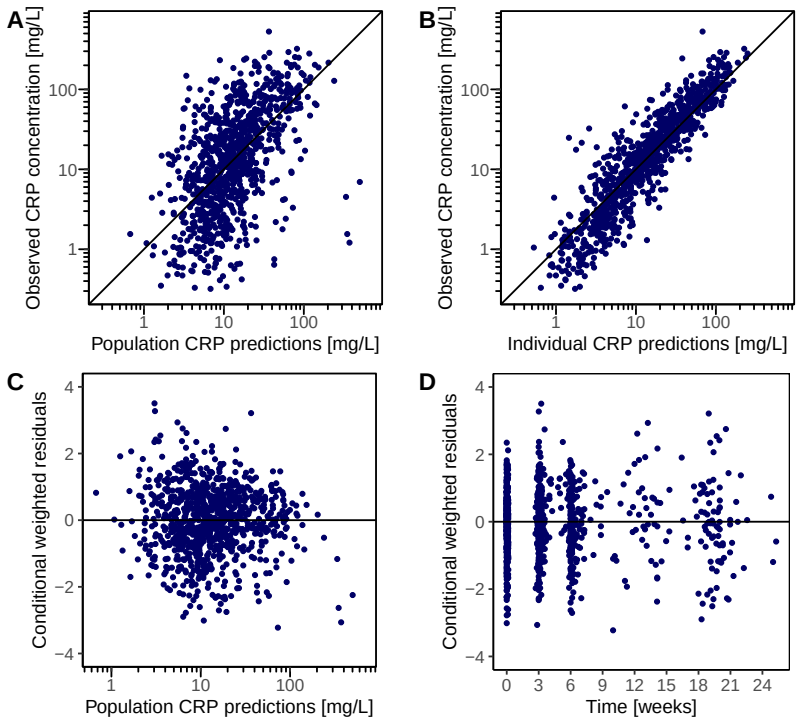

Figure 3.20: Goodness-of-fit plots of the coupled tumour size - CRP model for model-predicted CRP concentration and for residuals.
Blue dots: data points; solid black line: line of identity of slope 1 or zero; CRP: C-reactive protein.

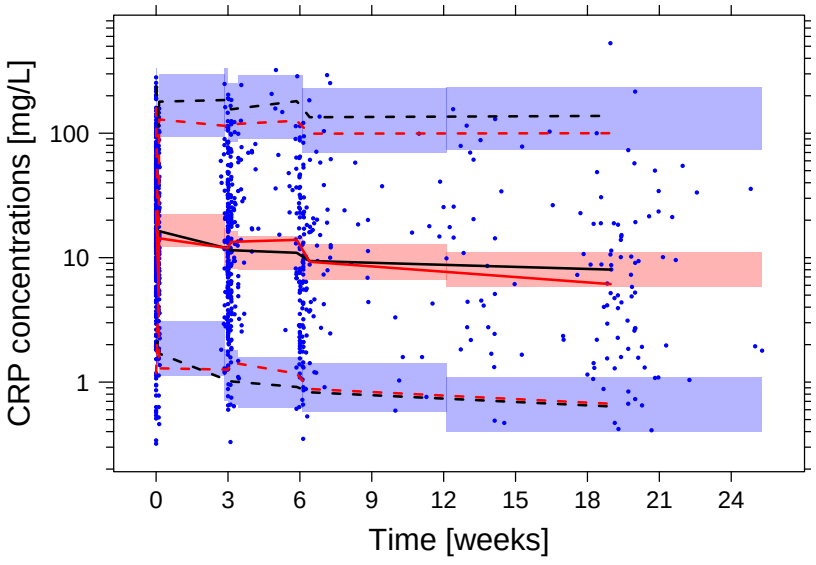

Figure 3.21: Visual predictive check (n = 250 simulations) of the coupled tumour size-CRP model.
Dots: observed CRP concentrations; solid red and black lines: median of the observed and simulated CRP concentrations, respectively; upper and lower dotted red and black lines: 5^{th} and 95^{th} percentiles of the observed and simulated CRP concentrations, respectively. Shaded areas: 95% confidence interval of the simulated percentiles. The number of bins was set to 8 with equal number of CRP observations. CRP: C-reactive protein.

3.2.2.2 Characterisation of efficacy endpoints and impact of predictors

Model-derived CRP- and tumour size-related metrics, along with other metrics of inflammation, and baseline patient and disease characteristics (Fig. 2.12, A, subsections 4–6, Table 2.6) were explored as potential predictors of the efficacy endpoints within a parametric TTE modelling framework, to identify the significant predictors of the characterised PFS and OS (Fig. 2.12, A, subsections 7 and 8), and explore their impact on the median TTE.

Characterisation of progression-free survival

Due to the chosen landmark time, only patients who survived without progression up to at least treatment cycle 3 (i.e. day 42 from the start of treatment) were included in the PFS TTE analysis. Thus, 203 of the 257 patients with CRP measurements i.e. 78.9% of patients (median PFS 7.07 months) were included in the analysis to characterise PFS.

The TTE base model with a parametric log-normal hazard function (Eq. 3.2) describing an initial increase in hazard followed by a decrease (i.e. log-normal distribution of the PFS event time) best described the observed PFS over time (Fig. 3.22, E). The other parametric hazard functions showed a strong underprediction of PFS events after 6 months and either failed to capture the distribution of PFS events between 6 months and 16 months (Fig. 3.22, B: Weibull, D: log-logistic) or between 6 months and 21 months (Fig. 3.22, A: exponential, C: Gompertz). On the other hand, the log-normal hazard function better predicted the events between 7 months and 14 months which were closely aligned with the 90% CI of the simulated profile (Fig. 3.22, E). In line with that, the log-normal hazard function was associated with superior model performance compared with the exponential (Δ AIC = +60.6), Weibull (Δ AIC = +37.8), Gompertz (Δ AIC = +60.5) or log-logisitc (Δ AIC = +36.6) hazard functions. Therefore, the log-normal hazard function with a standard deviation of 0.77 and a mean of 8.53 of the underlying normal distribution (Table 3.6, left part) was chosen as the base model to characterise PFS events and for the subsequent exploration of predictors.

Table 3.6: Parameter estimates of the progression-free survival parametric time-to-event base and covariate models

Parameter	Base model		Covariate model	
[unit]	Estimate (RSE, %)	95% CI[a]	Estimate (RSE, %)	95% CI[b]
Fixed-effects parameters[c]				
σ [unitless]	0.77 (5.90)	0.675–0.857	0.906 (8.80)	0.755–1.14
μ [unitless]	8.53 (0.70)	8.42–8.64	9.11 (2.50)	8.76–9.86
Covariate-effects parameters on $h_0(t)$[d]				
CRP_{cycle3}	—	—	0.109 (55.4)	0.0348–0.445
$CRP_{cycle3-2}$	—	—	-0.26 (37.2)	-0.461–-0.0637

OFV for base model = 3346.744; OFV for covariate model = 3319.203
[a] 95% CI obtained from 1000 bootstrap runs (successful minimisation = 100%)
[b] 95% CI obtained from 1000 bootstrap runs (successful minimisation = 99.9%)
[c] time unit of the log-normal hazard function was [hour]
[d] as described in Eq. 3.3
CI: confidence interval; CRP_{cycle3}: CRP concentrations at treatment cycle 3; $CRP_{cycle3-2}$: change (difference) in CRP concentrations between treatment cycle 3 and 2; μ: mean of the log-normal hazard function; OFV: objective function value; RSE: relative standard error; σ: standard deviation of the log-normal hazard function.

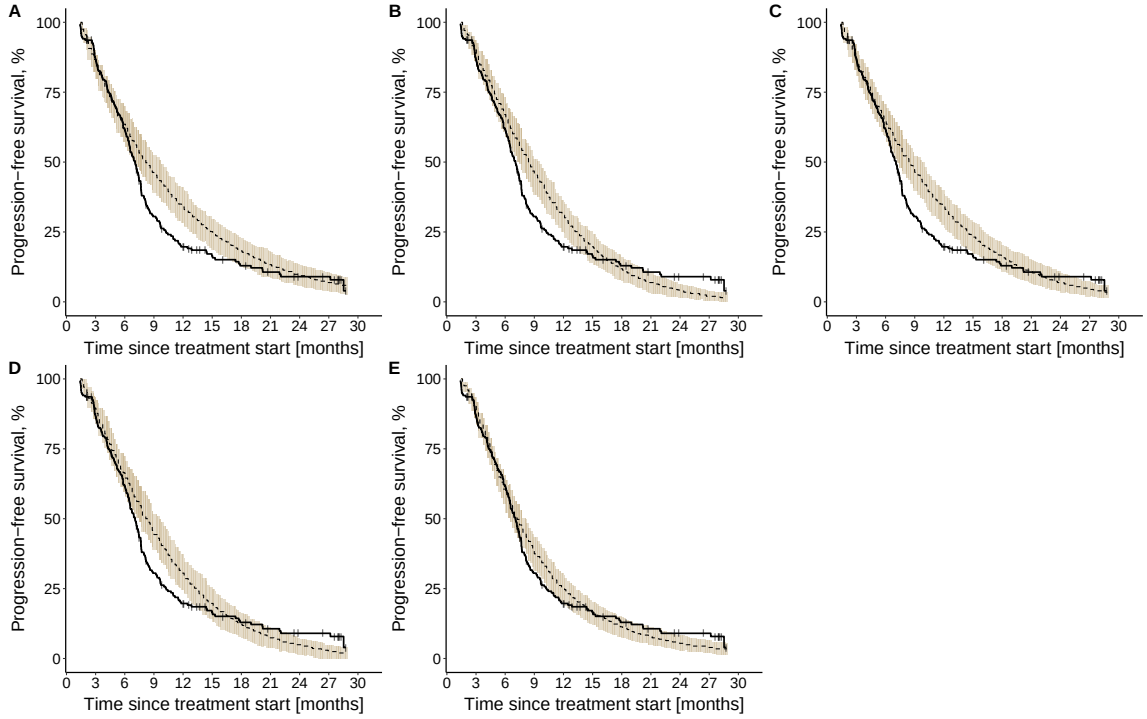

Figure 3.22: Kaplan-Meier visual predictive checks (n = 250) of progression-free survival time-to-event base models with **[A]** exponential, **[B]** Weibull, **[C]** Gompertz, **[D]** Log-logistic, and **[E]** Log-normal hazard functions.
Solid line: observed progression-free survival data (thin vertical lines represent censoring times corresponding to the time of the patient's last participation in the study), dashed line: median model predicted profile, with 90% confidence interval (beige shade).

Potential predictors (Table 2.6) were explored on the log-normal hazard function describing PFS. Appendix Fig. B.9.1 shows the identified correlations (n = 15) between these pre-selected predictors, which were taken into consideration during the SCM process to avoid inclusion of two strongly correlated predictors (i.e. $r > \pm 0.6$). Following the forward inclusion steps of the SCM procedure, CRP concentrations at treatment cycle 3 (CRP_{cycle3}), difference in CRP concentrations between treatment cycle 3 and 2 ($CRP_{cycle3-2}$), and disease stage significantly affected the hazard and contributed to a reduction of 20.2, 7.38, and 4.94 points in OFV, respectively (threshold > 3.84 for $\alpha = 0.05$, 1 df). Only disease stage was removed during the backward deletion step as it contributed to a 4.94 point increase in OFV when excluded (threshold > 6.64 for $\alpha = 0.001$, 1 df). Further details of the SCM steps are provided in Appendix A.7, Tables A.7.1 and A.7.2.

Thus, the final covariate PFS TTE model included the inflammatory level at treatment cycle 3 i.e. CRP_{cycle3}, described by a linear function, and the extent of reduction in inflammatory level between treatment cycle 3 and cycle 2 i.e. $CRP_{cycle3-2}$ described by a power function on the log-normal hazard function, as significant predictors of PFS (Eq. 3.3). Inclusion of these predictors resulted in 27.5 points drop in OFV (p-value $= 1.068 \cdot 10^{-6}$, 2 df), compared with the base model (Table 3.6). Thus, the final PFS model was described as in Eqs. 3.2 and 3.3,

$$h_0(t) = \frac{1}{t \cdot \sqrt{2\pi\sigma^2}} \cdot \frac{e^{-\frac{1}{2}\cdot(\frac{log(t)-\mu}{\sigma})^2}}{1-\phi(\frac{log(t)-\mu}{\sigma})} \tag{3.2}$$

$$h(t) = h_0(t) \cdot (1 + \theta_{CRP_{cycle3}} \cdot CRP_{cycle3}) \cdot (CRP_{cycle3-2}{}^{\theta_{CRP_{cycle3-2}}}) \tag{3.3}$$

where $h_0(t)$ is the baseline log-normal hazard function parameterised by the mean (μ) and standard deviation (σ) of the underlying normal distribution, as well as the standard normal cumulative distribution function (ϕ); $\theta_{CRP_{cycle3}}$ and $\theta_{CRP_{cycle3-2}}$ are parameters relating the effect of CRP_{cycle3} and $CRP_{cycle3-2}$ to the hazard, respectively; and $h(t)$ is the modified log-normal hazard after inclusion of the predictors.

Based on simulations from the base and final covariate PFS TTE models, the KM VPC depicted in Fig. 3.23 show the predictive performance of the model after inclusion of the two CRP-related predictors (Fig. 3.23, B) compared to the base model (Fig. 3.23, A). Despite the numerical improvement associated with the covariate model, improvement in the model prediction as shown in the KM VPC was minimal. A slight improvement could mainly be observed in the later time (i.e. >15 months) in which predictions fell closer to the 90% CI compared with the base model. Nevertheless, in general, the final covariate PFS model adequately predicted the distribution of PFS events over time.

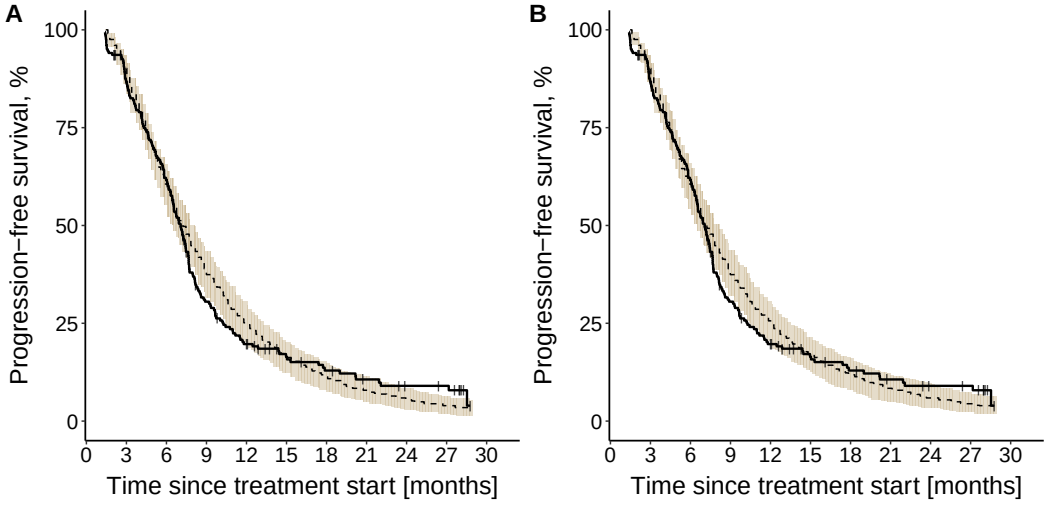

Figure 3.23: Kaplan-Meier visual predictive checks (n = 250) of progression-free survival time-to-event models with log-normal hazard function: **[A]** base model, **[B]** final covariate model.
Solid line: observed progression-free survival data (thin vertical lines represent censoring times corresponding to time of patient's last participation in the study), dashed line: median model predicted profile, with 90% confidence interval (beige shade).

Assessment of the impact of the identified predictors of progression-free survival

The univariate impact of each of the two identified CRP-related predictors on the median PFS is depicted in Fig. 3.24: The inflammatory level at treatment cycle 3 showed a bigger impact (i.e. 7.47 months difference between the 5^{th} and 95^{th} percentile of CRP_{cycle3}) on median PFS compared with the reduction in inflammatory level between treatment cycle 3 and cycle 2 (i.e. 4.66 months difference between the 5^{th} and 95^{th} percentile of $CRP_{cycle3-2}$). Whereas investigation of the combined influence of both predictors (Fig. 3.25) revealed that for a patient cohort with low inflammatory level at treatment cycle 3 and high reduction in inflammatory level between treatment cycle 3 and cycle 2 (i.e. favourable condition), median PFS was 16.5 months whereas for a patient cohort with a high inflammatory level at treatment cycle 3 and low reduction in inflammatory level between cycle 3 and cycle 2 (i.e. less favourable condition), median PFS was 13.4 months shorter, and reached only 3.1 months.

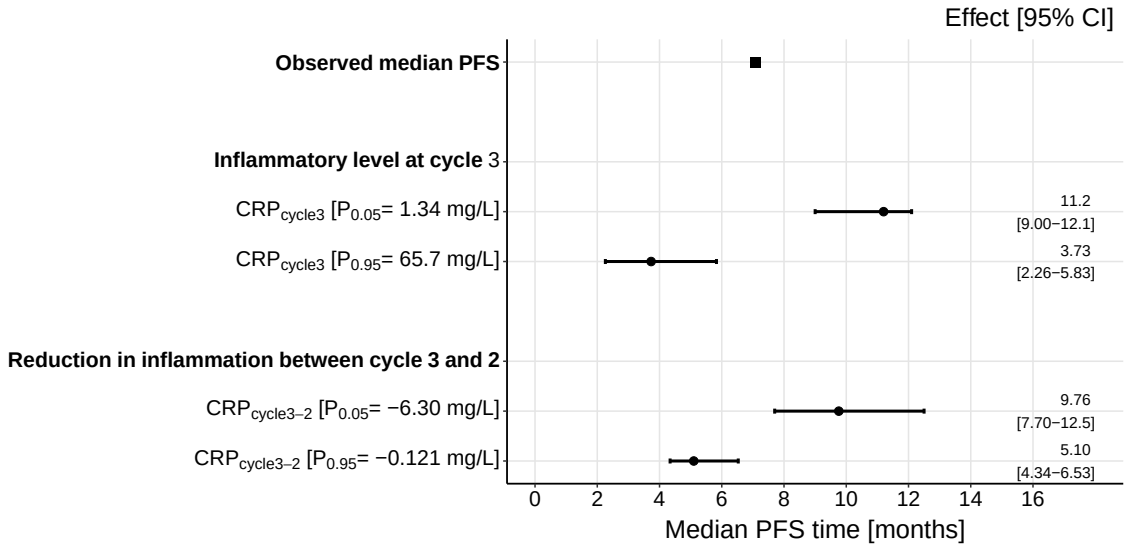

Figure 3.24: Forest plot of the impact of the identified significant predictors on median progression-free survival (PFS).
Effects of the continuous predictors (i.e. CRP_{cycle3}, $CRP_{cycle3-2}$) are shown at the 5^{th} and 95^{th} percentiles of the respective predictor distribution. Black dots: predictor effects; horizontal lines: 95% confidence intervals (CI). CRP: C-reactive protein.

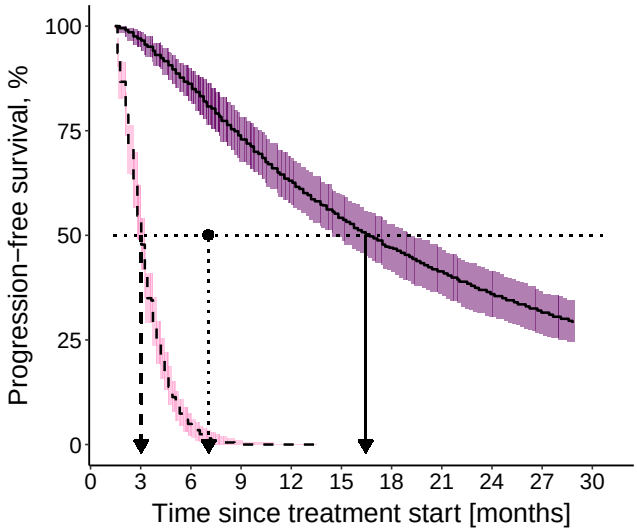

Figure 3.25: Kaplan-Meier plots of simulated (n = 250) progression-free survival profiles under the combined effect of the 5^{th} percentile of the two continuous predictors: CRP_{cycle3}: 1.34 mg/L and $CRP_{cycle3-2}$: -6.30 mg/L (solid black line with 90% confidence interval as purple shade) and the 95^{th} percentile of the continuous predictors: CRP_{cycle3}: 65.7 mg/L and $CRP_{cycle3-2}$: -0.121 mg/L (dashed black line with 90% confidence interval as pink shade).
Dotted horizontal line: 50% progression-free survival; dashed vertical line: median progression-free survival time at 95^{th} percentile values of predictors; dotted vertical line: observed median progression-free survival time; solid vertical line: median progression-free survival time at 5^{th} percentile values of predictors.

Characterisation of overall survival

Due to the chosen landmark time and similar to the TTE analysis of PFS, only patients who survived up to at least treatment cycle 3 (i.e. day 42 from the start of treatment) were included in the TTE analysis. Thus, 235 of the 257 patients with CRP measurements i.e. 91.4% of patients (median OS 11.1 months) were included in the analysis to characterise OS.

The TTE base model with a parametric Weibull hazard function, parameterised with scale (λ) and shape (α) parameters, best described the observed OS over time (Fig. 3.26, B), compared with the exponential (Fig. 3.26, A) and Gompertz (Fig. 3.26, C) hazard functions. α was estimated to be >1 (Table 3.7), meaning that the risk of death (i.e. hazard) was increasing with time. Despite an overprediction of survival between 6 months and 17 months, and a slight underprediction after 24 months, the Weibull TTE model still showed an adequate description in the early months (<6 months) and between 17 and 24 months (Fig. 3.26, B). On the other hand, although the exponential and Gompertz TTE models adequately predicted the risk of death during the first 6 months (Fig. 3.26, A, C), the exponential TTE model showed a consistent overprediction of survival across the entire time-course (Fig. 3.26, A), whereas the Gompertz TTE model alternated between an overprediction of survival (between 6 and 24 months) followed by an underprediction (>24 months) (Fig. 3.26, C). In line with that, both the exponential and Gompertz TTE models did not provide a significant improvement in model performance compared with the Weibull TTE model (Δ AIC $= +20.1$ and $+16.1$, respectively). Therefore, the Weibull TTE model was adopted as the base model for the subsequent exploration of predictors.

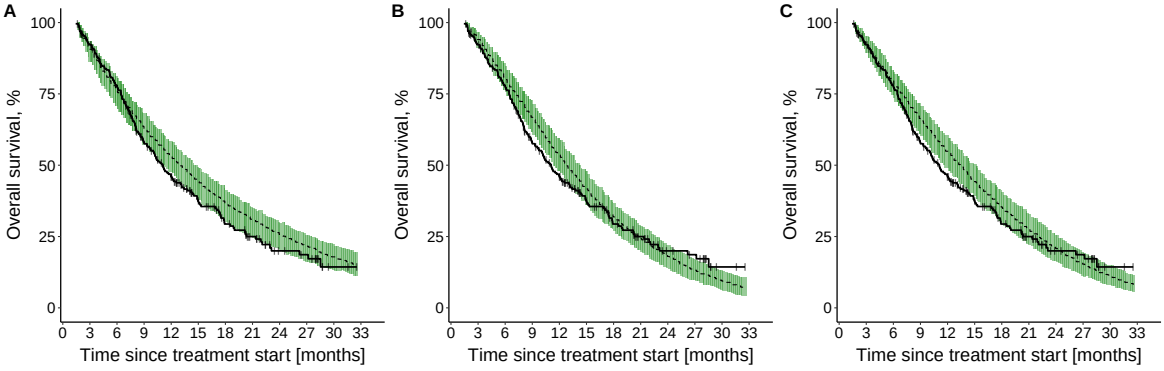

Figure 3.26: Kaplan-Meier visual predictive checks (n = 250) of overall survival time-to-event base models with **[A]** exponential, **[B]** Weibull, and **[C]** Gompertz hazard functions.
Solid line: observed survival data (thin vertical lines represent censoring times corresponding to time of patient's last participation in the study), dashed line: median model predicted profile, with 90% confidence interval (green shade).

The same potential predictors (Table 2.6) which were explored on the log-normal hazard function describing PFS, were also explored on the Weibull hazard function describing OS. Following the forward inclusion steps of the SCM procedure, and as with PFS, CRP_{cycle3} and $CRP_{cycle3-2}$ significantly affected the hazard, in addition to the presence of liver lesions and baseline tumour size. These predictors contributed to a reduction of 53.0, 10.2, 12.8, and 12.4 points in OFV, respectively (threshold >3.84 for $\alpha = 0.05$, 1 *df*). None of the identified predictors was removed during the backward deletion step since all of them were associated with a significant worsening of model performance when excluded (threshold >6.64 for $\alpha = 0.001$, 1 *df*). Further details of the SCM steps are provided in Appendix A.8, Table A.8.1.

Table 3.7: Parameter estimates of the overall survival parametric time-to-event base and covariate models

Parameter [unit]	Base model Estimate (RSE, %)	95% CI[a]	Covariate model Estimate (RSE, %)	95% CI[b]
Fixed-effects parameters				
λ [1/h]	$8.6{\cdot}10^{-5}$ (6.20)	$7.69{\cdot}10^{-5}$–$9.78{\cdot}10^{-5}$	$1.6{\cdot}10^{-5}$ (25.7)	$9.32{\cdot}10^{-6}$–$2.65{\cdot}10^{-5}$
α [unitless]	1.38 (5.20)	1.25–1.54	1.68 (5.30)	1.54–1.92
Covariate-effects parameters on $h_0(t)$[c]				
CRP_{cycle3}	—	—	0.781 (12.8)	0.595–0.999
$CRP_{cycle3-2}$	—	—	-0.392 (24.9)	-0.606–-0.185
Baseline TS	—	—	0.491 (33.2)	0.201–0.881
Liver lesions	—	—	1.02 (36.3)	0.374–2.03

OFV for base model = 3302.668; OFV for covariate model = 3214.314
[a] 95% CI obtained from 1000 bootstrap runs (successful minimisation = 100%)
[b] 95% CI obtained from 1000 bootstrap runs (successful minimisation = 98.7%)
[c] as described in Eq. 3.5
CI: confidence interval; TS: tumour size; OFV: objective function value; RSE: relative standard error; λ: scale parameter of the Weibull hazard function; α: shape parameter of the Weibull hazard function; CRP_{cycle3}: CRP concentrations at treatment cycle 3; $CRP_{cycle3-2}$: change (difference) in CRP concentrations between treatment cycle 3 and 2.

Thus, the final covariate OS TTE model included the inflammatory level at treatment cycle 3 i.e. CRP_{cycle3} described by a power function, the extent of reduction in inflammatory level between treatment cycle 3 and cycle 2 i.e. $CRP_{cycle3-cycle2}$ described by a power function, tumour load i.e. baseline tumour size described by a power function and presence of liver lesions described by a fractional change model on the Weibull hazard function, as significant predictors of OS. Inclusion of these four predictors resulted in 88.4 points drop in OFV (p-value = $2.95 \cdot 10^{-18}$, 4 df), compared with the base model (Table 3.7). Thus, the final OS model was described as in Eqs. 3.4 and 3.5,

$$h_0(t) = \lambda\alpha(\lambda t)^{\alpha-1} \tag{3.4}$$

$$h(t) = h_0(t) \cdot {CRP_{cycle3}}^{\theta_{CRP_{cycle3}}} \cdot {CRP_{cycle3-2}}^{\theta_{CRP_{cycle3-2}}} \cdot BLTS^{\theta_{BLTS}} \cdot \begin{cases} 1, & \text{No liver lesions} \\ 1+\theta_{liver}, & \text{With liver lesions} \end{cases} \tag{3.5}$$

where $h_0(t)$ is the baseline Weibull hazard function parameterised by the the scale parameter (λ) and the shape parameter (α); $\theta_{CRP_{cycle3}}$, $\theta_{CRP_{cycle3-2}}$, θ_{BLTS}, and θ_{liver} are parameters relating the effect of CRP_{cycle3}, $CRP_{cycle3-2}$, baseline tumour size ($BLTS$), and presence of liver lesions to the Weibull hazard, respectively; and $h(t)$ is the modified Weibull hazard after inclusion of the predictors.

Based on simulations from the base and final covariate OS TTE models, the KM VPC (Fig. 3.27) shows the improvement in model prediction after inclusion of the identified predictors (Fig. 3.27, B) compared to the base model (Fig. 3.27, A). An improved model prediction is observed especially between 7 months and 17 months, and at later times (> 24 months) where predictions fell closer to the 90% CI compared with the base model.

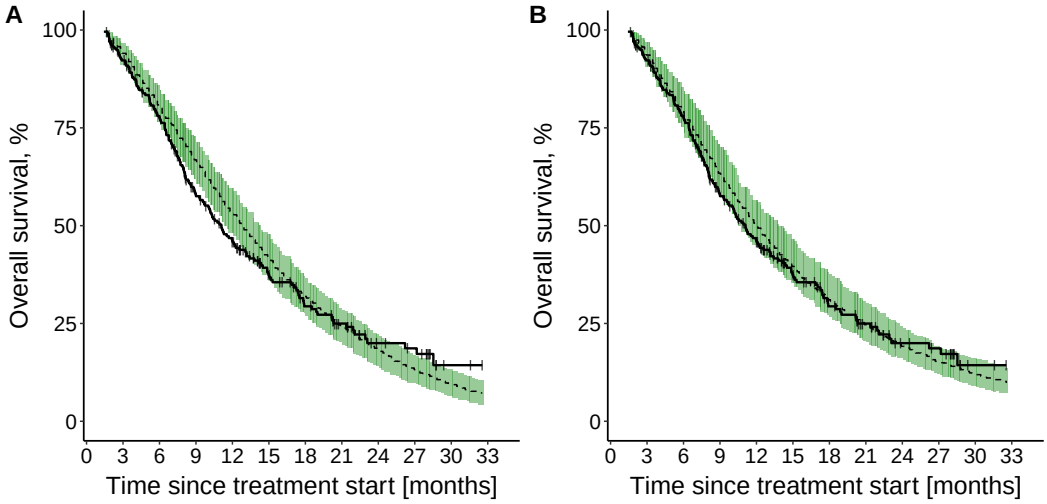

Figure 3.27: Kaplan-Meier visual predictive checks (n = 250) of overall survival time-to-event models with Weibull hazard function: **[A]** base model, **[B]** final covariate model.
Solid line: observed survival data (thin vertical lines represent censoring times corresponding to time of patient's last participation in the study), dashed line: median model predicted profile, with 90% confidence interval (green shade).

Assessment of the impact of the identified predictors of overall survival

The univariate impact of each of the identified predictors on the median OS is depicted in Fig. 3.28. The inflammatory level at treatment cycle 3 showed by far the largest impact on median OS (i.e. 25.6 months difference between the 5^{th} and 95^{th} percentile of CRP_{cycle3}) compared with the reduction in inflammatory level between treatment cycle 3 and cycle 2 (i.e. 11.2 months difference between the 5^{th} and 95^{th} percentile of $CRP_{cycle3-2}$) or tumour load (i.e. 7.24 months difference between the 5^{th} and 95^{th} percentile of baseline tumour size). The investigation of the combined influence of all predictors (Fig. 3.29) revealed that for a patient cohort with favourable characteristics i.e. low inflammatory level at treatment cycle 3, high reduction in inflammatory level between treatment cycle 3 and cycle 2, low tumour load and absence of liver lesions; median OS was not reached until the end of the observed time span of 33 months whereas for a patient cohort with a high inflammatory level at treatment cycle 3, low reduction in inflammatory level between treatment cycle 3 and cycle 2, high tumour load and presence of liver lesions (i.e. less favourable condition), median OS was only 2.24 months.

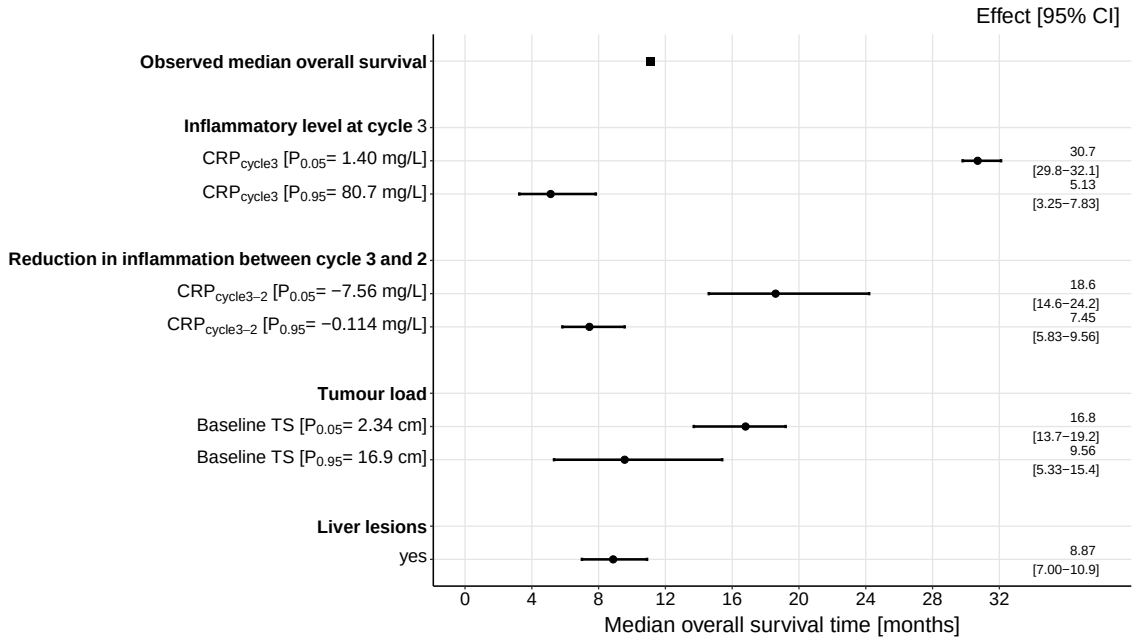

Figure 3.28: Forest plot of the impact of identified significant predictors on median overall survival.
Effects of the continuous predictors (i.e. CRP_{cycle3}, $CRP_{cycle3-2}$, baseline tumour size) are shown at the 5^{th} and 95^{th} percentiles of the respective predictor distribution and effects of the categorical predictors (i.e. liver lesions) are shown relative to the reference category (i.e. no liver lesions). Black dots: predictor effects; horizontal lines: 95% confidence intervals (CI). CRP: C-reactive protein; TS: tumour size.

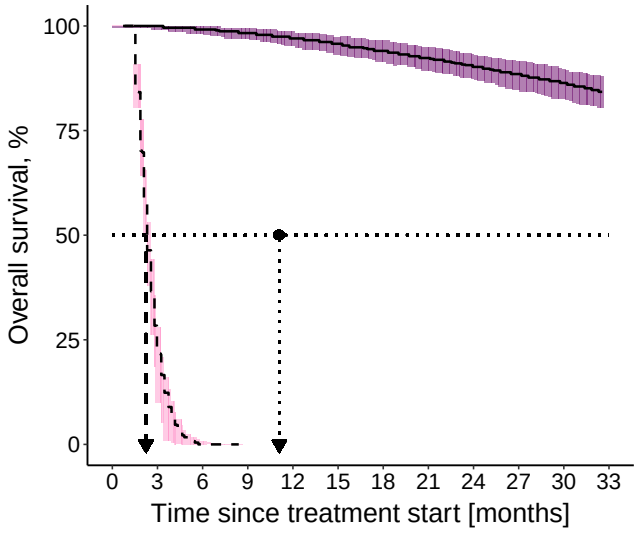

Figure 3.29: Kaplan-Meier plots of simulated (n = 250) overall survival profiles under the combined effect of the 5^{th} percentile of the continuous predictors: CRP_{cycle3}: 1.40 mg/L, $\text{CRP}_{cycle3-2}$: -7.56 mg/L and baseline tumour size: 2.34 cm in absence of liver lesions (solid black line with 90% confidence interval as purple shade) and the 95^{th} percentile of the continuous predictors: CRP_{cycle3}: 80.7 mg/L, $\text{CRP}_{cycle3-2}$: -0.114 mg/L and baseline tumour size: 16.9 cm in presence of liver lesions (dashed black line with 90% confidence interval as pink shade).
Dotted horizontal line: 50% overall survival; dashed vertical line: median overall survival time at 95^{th} percentile values of predictors in presence of liver lesions; dotted vertical line: observed median overall survival time; solid vertical line: median progression-free survival time at 5^{th} percentile of the predictors in absence of liver lesions.

3.2.2.3 Impact of different levels of inflammation on efficacy endpoints

Since the inflammatory level at treatment cycle 3 was the most impactful predictor of both PFS and OS, the influence of the different inflammatory levels i.e. as depicted by CRP_{cycle3} percentiles, on these efficacy endpoints was systematically explored by simulations. An even impact was observed for both PFS and OS against the different inflammatory levels (Fig. 3.30 A and B, respectively). As expected, increased inflammation was associated with shorter median PFS and OS.

Comparison of these simulated profiles against observed PFS and OS events stratified by similar percentile intervals of model-estimated CRP_{cycle3} is shown in Fig. 3.30 C and D, respectively. In the PFS KM plot (Fig. 3.30, C), the upper and lower percentile intervals showed a similar pattern to the simulated profiles (Fig. 3.30, A) and the upper three percentile intervals showed a comparable median PFS—although there was an observed overlap of the intermediate CRP_{cycle3} percentile intervals. For OS, Fig. 3.30, D depicted a very similar pattern to the simulation results (Fig. 3.30, B) with a corresponding median OS for all the percentile intervals—except for the lowest two percentile intervals which were simulated to have a longer median OS compared with what was observed (18 months vs 15 months and 30 months vs 21 months median OS).

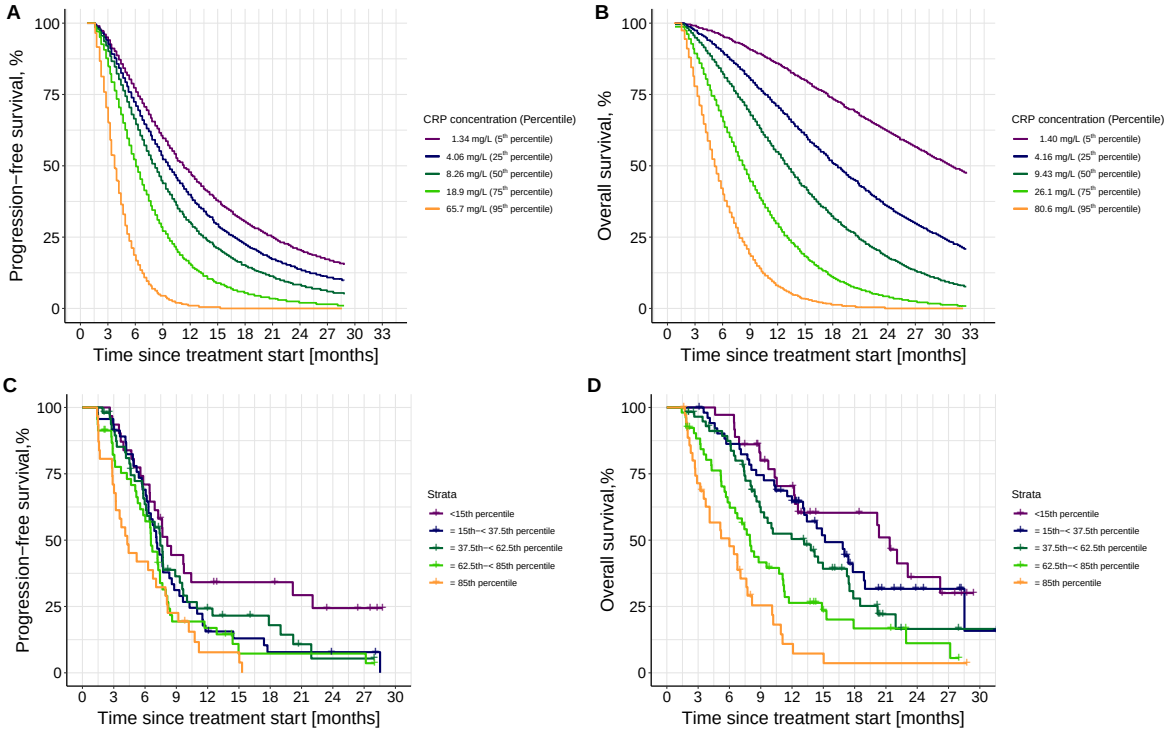

Figure 3.30: *Upper panel* Kaplan-Meier visual predictive checks (n = 250) of simulated **[A]** progression-free survival and **[B]** overall survival events at different concentrations (i.e. percentiles) of C-reactive protein (CRP) concentration at start of treatment cycle 3 (CRP_{cycle3}). Different colours in the upper panel indicate different percentiles of CRP_{cycle3}. *Lower panel* Kaplan-Meier plots of observed distribution of **[C]** progression-free survival and **[D]** overall survival events stratified by model-estimated CRP_{cycle3}, colour-coded by the different percentile intervals. Percentile intervals were chosen so that their median corresponds to the percentiles of the simulated profiles in the upper panel i.e. 5^{th}, 25^{th}, 50^{th}, 75^{th} , and 95^{th}. Solid coloured line in the lower panel: observed progression-free survival or overall survival; crosses in the lower panel: censoring times corresponding to time of patient's last participation in the study; different colours in the lower panel indicate different percentile intervals of CRP_{cycle3}.

4 | Discussion

Throughout the two projects presented in this thesis, we have successfully demonstrated how pharmacometrics can be leveraged to optimise early clinical decision-making in both clinical drug development and at bedside, within the oncology setting.

With the advances in model-based approaches to analysis dose-finding studies, it was essential to understand and evaluate the performance of the newly proposed cLRT approach, which leveraged the principles of NLME, versus the MCP approach, which has gained both EMA and FDA qualifications as "fit-for-purpose", for more informative implementation when it comes to analysing dose-finding studies of compounds within the early clinical drug development in the oncology setting. While at bedside, the high need for early decision-making during cancer treatment and specifically with lung cancer as a high burden disease of poor prognosis, motivated the use of population approaches to amalgamate the different oncology endpoints: drug exposure, tumour dynamics, biomarkers, and survival. Through the characterisation of the relationship among the former three and the identification of early predictors of efficacy endpoints (PFS, OS), patients at risk of poor prognosis could be identified, early on, for a timely clinical decision-making.

The following sections discuss each project in further details.

4.1 Project I: Simulation-based evaluation of cLRT and MCP approaches in oncology under different clinical trial scenarios

In *project I*, we compared the performance of cLRT, a recently proposed approach based on the LRT and utilising the principles of NLME modelling, to MCP, one of the typical approaches used to analyse dose-finding phase II studies and which has been recommended by the EMA and FDA. Further, we investigated the performance of both approaches within the oncology setting while considering different study design variables and drug strengths typically used in oncology trials. Our simulation results showed that compared to MCP, cLRT had higher power identifying presence of a dose-response. However, this was at the cost of an inflated type I error. Moreover, the ability of cLRT to discriminate between the impact of the different study design variables, compared to MCP, was low.

4.1.1 Simulation study

Our proposed simulation study design and data-generating dose-response true models aimed to mimic an oncology phase II clinical study setting.

First, in contrast to a fixed number of patients per arm i.e. sample size per arm —regardless of the total number of arms, we chose a fixed *total* number of patients, i.e. sample size, of 24 and 36 that would be equally balanced amongst the different dose levels i.e. number of arms. This was driven by the fact that clinical trials are commonly dependent on specific resources that translate to a target/feasible sample size (i.e. a maximum number of recruited patients) and finding an optimal allocation as testing more doses with fewer individuals per arm or testing fewer doses with more individuals per arm, and determining a sufficiently powered study design given these prerequisites are usually the aim.

Second, the different number of study arms aimed not only to explore the conventional design, with few dose levels, but also the extreme scenario with e.g. six dose levels, to prove the limited benefit of multiple dose levels which would consequently be associated with a lower dose-response detection power *if* not met by a large sample size.

Third, we considered dose levels ranging from 0 mg to 60 mg. Although a dose level of zero is a hypothetical scenario in oncology trials, we considered the placebo arm to have received no drug for the sake of simplicity (further discussion on how to mitigate this under section 4.1.4). The selection of the intermediate doses was calculated to ensure an even distribution across the whole dose-response shape (Appendix Fig. B.1.1).

Fourth, we acknowledge that within the study design, the sampling times (every 7 days) were more frequent than what one would expect in clinical practice. However, given our preliminary investigation, we opted for frequent sampling to test the performance under a setting that would provide more stability to the simulations/re-estimations. For the chosen data-generating dose-response true models, we aimed to simulate the different endpoints that would be of relevance in defining PFS and assessing efficacy in an oncology trial—in contrast to only focusing on the tumour dynamics as the sole determinant of progression. However, for the sake of simplicity within the cLRT framework the occurrence of any of the other different endpoints (probability of appearance of new lesions, probability of progression of non-target lesions, and death events prior to progression) was considered a drop out event in the candidate models. This took into consideration the probability of dropping out of clinical studies with time—a key characteristic feature of oncology trials. For MCP, although the simulated clinical trial data were based on the same data-generating dose-response true model as with cLRT, a time invariant endpoint i.e. response, had to be chosen since MCP does not consider longitudinal data. A direct comparison between time invariant endpoints and longitudinal data was not feasible. Nevertheless, this endpoint was chosen to be the best change in tumour size from baseline to closely mimic cLRT in considering the change in tumour size over time.

Lastly, to assess the performance of both cLRT and MCP under different dose-response relationships, we explored relationships that ranged in complexity from linear to sigmoidal.

4.1.2 Evaluation of cLRT simulations/re-estimations

The investigation of the intermediate steps of cLRT showed that the cLRT procedure was not always robust (stable) due to the inclusion of multiple stochastic simulations (section 3.1.1). This was similarly experienced at a different extent in the work by Chasseloup and Karlsson [220]. The N simulations under the null hypothesis (here N was set to 100) did not always result in 100% successful M simulations-estimations i.e. 100% successful executions after model-fitting and selection, under the null hypothesis. This was a sign that one of the simulations may not have resulted in a plausible combination of variables in the simulated dataset due to the stochastic simulation. Consequently, this simulated dataset was not supported by the model, resulting in a failed M simulation, model-fitting and selection steps (Fig. 2.9, sections 3.1–3.4). For this reason, individual simulated datasets which did not result in at least a 90% success rate of their respective M simulations-estimations were excluded to avoid inclusion of unreliable estimates.

Further, unrealistic values were observed in t_{cLRT} and $t_{cLRT_{Ho}}$, and extreme outliers were observed in the $t_{cLRT_{Ho}}$ distribution (section 3.1.1). Since, the test statistic was defined as the difference between the OFV of the best-fitting candidate model and that of the flat dose-response model, the test statistic was expected to be negative with a reasonable difference in the Δ OFV, reflecting the improvement in the model when a dose-response relationship was considered (since it was assumed that the candidate model should always outperform the flat). Consequently, t_{cLRT} and $t_{cLRT_{Ho}}$ falling in the positive domain were excluded as they indicated that the candidate model was worse, probably due to the lack of a successful minimisation of the candidate model to the best set of parameter estimates (i.e. local minima). On a similar note, $t_{cLRT_{Ho}}$ falling in the extreme negative domain ($<$-60) were also excluded from the $t_{cLRT_{Ho}}$ distribution under the null hypothesis, as they did not translate to a realistic improvement (i.e. Δ OFV) and were attributed to the flat model lacking the best-set of parameter estimates (i.e. local minima) such that the difference was too wide between the candidate and flat model due to the bad performance of the flat. The latter two cases were mainly driven by the variability encountered from the stochastic simulations in which both the candidate and the flat models did not well-fit the simulated data (section 4.1.5).

Data exclusion was thus done to avoid introducing bias that would have been driven by opposing assumptions, extreme outliers, or non-robust simulations to the results. A cut-off value of 90% successful M simulations-estimations was set as a criterion to assess the robustness of the null hypothesis. Whereas a cut-off value of $t_{cLRT_{Ho}} <$-60 was chosen based on the visual inspection of the $t_{cLRT_{Ho}}$ distribution of the different scenarios (similar results were obtained for a cut-off value of $t_{cLRT_{Ho}} <$-100).

4.1.3 Evaluation of cLRT performance versus MCP performance

Although cLRT had a high power identifying presence of a dose-response (average power 98.5% for strong drug effect and 81.0% for weak drug effect, Fig. 3.2), this was achieved at the cost of an inflated type I error (average type I error 13.0%, Fig. 3.1). On the contrary, in MCP, whose type I error was low (average type I error 2.67%, Fig. 3.1), its power was much reduced (average power 36.4% for strong drug effect and 17.7% for weak drug effect, Fig. 3.2). The difference in the observed power could be attributed to the principles underlying each approach: cLRT leveraged the full longitudinal information, in our case change in tumour size over time, whereas MCP relied on a chosen time-invariant endpoint, here: the best change in tumour size from baseline. Moreover, the use of the LRT and the NLME-based approaches have been previously shown to be associated with high power [126, 220]. The inflated type I error with cLRT could have been due to our limited number of simulations ($N = 100$) and the small sample size i.e. number of patients to which MCP was apparently more robust. Nevertheless, a recently published investigation [220] of cLRT performance using the same number of simulations as presented here ($N = 100$) but with a larger sample size has also concluded, in line with our results and in contrast to previous simulations [126], that cLRT lacked control of type I error and primarily attributed this to AIC selection bias and potential misspecification in the placebo model which was relied upon for simulations under the null hypothesis to derive the critical value.

In cLRT, the lack of a clearly increasing power with fewer dose levels and a larger sample size per arm i.e. more patients per arm—in contrast to a clearer trend with MCP—requires further systematic exploration of the robustness of this method and its sensitivity when it comes to the impact of the different study designs on its performance (section 4.1.4). However, it could be argued that because of the high power of cLRT associated with the strong drug effect ($\approx 100\%$), which was in line with the previous work [126], the impact of these study design variables was no longer prominent. This argument is supported by the faint trend that appears with the weaker drug effect (Fig. 3.2). It is also noticed that with MCP, the strength of the observed trend was dependent on the drug effect, indicating that in presence of a weak drug effect the study design will not have a big impact on the power of detecting a dose-response relationship (Fig. 3.2). The clear trend observed with MCP with fewer dose levels is expected and could be explained by the presence of more individuals at the informative doses compared with fewer number of individuals at each dose level with increased number of dose levels [125]—within the range of the observed dose levels. These results were also supported by the simulations performed by Chen *et al.* [221], although they were not as strong when they explored the case of a fixed total number of patients.

Our simulation results showed that a larger sample size per arm was not associated with higher probability of detecting the true underlying dose-response model for both cLRT and MCP—in contrast to what was reported by Chen *et al.* for MCP [221]. This discrepancy might have been due to the type of endpoint on which their MCP analysis was based on and/or the different study design. Chen *et al.* [221] relied on a binary endpoint and explored a sample size per arm of 10, 20, 40 and 80 patients with a non-fixed total sample size which was much larger than the sample size per arm of 4, 6, 8, 9 and 12 patients that we explored at a fixed total sample size.

In line with what was reported by Chen *et al.* [221] for MCP and a sample size of 10 patients per arm, the most commonly chosen candidate model was to a great extent independent of the true shape of the data-generating model and was inclined towards a simple model i.e. linear/log-linear, with the more

complex models being less likely to be selected even if they were the true data-generating models. However, in the previous work investigating the performance of cLRT [126], a tendency towards choosing the simple models was only evident with their investigated weak drug effect, whereas in case of a strong drug effect the true shape was identified with the highest proportion. A reason for our slightly deviating results with cLRT, in which selection of the simplest model was observed for scenarios under both the strong and weak drug effect, could be attributed to the chosen drug effect with respect to the therapeutic area. A "stronger" drug effect could have been needed to select the true shape with the highest proportion and maximum power. Nevertheless, it is indeed difficult to define what is the extent of a strong drug effect since in some therapeutic areas/indications a modest drug effect would be considered a satisfactory objective for a drug effect to be efficacious.

4.1.4 Limitations of work

Due to the computational capacity, a main limitation of this work was the low number of simulations which we believe may have inflated our type I error and the impact of noise from the stochastic simulations. We suggest the implementation of this framework under larger number of simulations e.g. $N = 500$ or 1000. However, it is worth noting that the cLRT framework is a computationally intensive and demanding procedure since derivation of the critical value requires simulations and (re-)estimations for its approximation. Therefore, careful thought with regards to the computing time and the available computational power may be needed. Further modification of the execution script could also be a possibility where certain execution tasks could be paralleled or accelerated for a reduced computational burden.

A second limitation that could challenge the implementation of cLRT in the oncology setting is the need for placebo arm data to derive $q_{\alpha,cLRT}$ under the null hypothesis of a no drug effect (section 2.5.2). Clinical oncology studies usually lack conventional placebo arm data due to ethical concerns and standard of care treatment arms would not provide information on a no-drug effect to represent the null hypothesis. To overcome this, an external source of data has to be used as no exposure level. We therefore suggest (i) the use of historical controls from literature or real-world data, (ii) simulation of placebo arm patients from developed disease progression models, or (iii) use of data from previous clinical trials whose experimental drug has failed (i.e. did not show efficacy) or which used very low dose levels assuming no relevant exposure.

A third limitation is our sample size. We acknowledge that our sample size was smaller than what was investigated in the works of Buatois *et al.* [126] for cLRT and Chen *et al.* [221] for MCP. It may have even impacted our type I error and led to the observed unstable results as previously predicted by Buatois *et al.* [126]. Nevertheless, we believe that opting for a small sample size was more realistic and better mimicked a conventional sample size of phase II trials. However, it would still be useful to investigate a larger sample size to help decipher the reason(s) for the inflated type I error and the instability that we observed.

Further limitations are the exclusion of some of the simulations/re-estimations within the cLRT framework due to the observed lack of robustness (section 4.1.2) as well as the frequent observation times in the study design that do not closely align with clinical practice (section 4.1.1). As mentioned earlier, the reason for the latter was that we sought a setting that would provide more stability during the automation process.

4.1.5 Outlook and next steps

As mentioned under section 4.1.4, we recommend the implementation of this framework under a larger simulation scenario to better assess the type I error and impact of outliers. We also recommend a thorough and more systematic investigation of the reasons behind the non-robustness observed within the cLRT framework, with respect to the potential impact of sample size, model complexity, placebo model, and extent of variability within the stochastic simulations, to identify reasons for the instability of certain simulations/re-estimations.

Once reasons for the observed cLRT instability are identified, the stability of the cLRT framework is controlled, and the automated framework is made more robust; there would be a huge potential for further systematic exploration of the impact of each of our selected variables and/or model components on the performance of cLRT. For example, it would be useful to explore the impact of (i) the different dose allocations i.e. varying the distribution of doses within each dose level rather than sticking to an even distribution, (ii) the extent of variability in the model components, (iii) the complexity of the model, (iv) a more limited and realistic sampling scheme that better mimics clinical practice, and (v) a small vs. large sample size.

Moreover, we believe that the impact of different drug effects should be further investigated. Drug effects should be selected to span the expected range of an acceptable therapeutic benefit within the respective therapeutic area, and in our case different cancer types. It would also be interesting to investigate if identification of the "true" dose-response relationship with high proportion was dependent on the drug effect as in [126] or a combination of multiple other factors e.g. sample size per arm as in [221].

As a later step and based on the identified dose-response relationship, expansion of cLRT and MCP to the "Mod" step should follow, to select the optimal dose for further testing in the confirmatory phase III trial.

In conclusion, through a simulation-based framework, this work has investigated the performance of cLRT—as an alternative methodology to MCP approach—to identify presence of a significant drug effect in dose-ranging clinical studies by leveraging NLME principles. Specifically, we focused on its performance in oncology and investigated different study designs to identify variables critical to its optimal performance. Although cLRT showed high power, it was non-robust and was associated with an inflated type I error. More systematic investigations are needed to identify reasons of its non-robustness and justify its computational needs.

4.2 Project II: Identification of predictors of efficacy in NSCLC patients: A tumour dynamics - CRP modelling framework

In *project II*, we successfully identified early predictors of efficacy in advanced NSCLC patients. We also demonstrated that not only a lower inflammatory level (i.e. lower concentrations of CRP_{cycle3}) but also a larger reduction in the inflammatory level across treatment cycles, specifically larger drop between CRP_{cycle3} and CRP_{cycle2}, were associated with longer PFS. In addition to the latter two CRP-related metrics, a smaller tumour load (i.e. baseline tumour size) and absence of liver lesions were associated with longer survival.

4.2.1 Clinical study data: The CEPAC-TDM study

The CEPAC-TDM data exploration was the initial step to support our modelling framework and assumptions. The observed CRP's kinetic change and correlation to tumour size measurements (Fig. 3.11 and 3.12) motivated the characterisation of the quantitative relationship between paclitaxel exposure, tumour size, and CRP concentrations. Furthermore, the significant impact of baseline CRP on the TTP, PFS and OS (Fig. 3.13) encouraged the in-depth exploration of the potential of longitudinal (non-baseline) CRP concentrations as prognostic predictors in NSCLC patients—expanding on the commonly reported approaches investigating *only* the prognostic value of baseline measurements [67–74, 74–80, 80–88, 104, 105].

4.2.2 Modelling framework to identify predictors of efficacy in NSCLC patients

4.2.2.1 Development of coupled tumour dynamics - CRP model to characterise circulating CRP concentration

Our coupled tumour dynamics - CRP model successfully predicted circulating CRP concentrations (section 3.2.2.1). Moreover, it characterised the quantitative relationship between paclitaxel exposure; tumour size, a metric of tumour dynamics; and CRP concentration, an inflammatory serum biomarker and marker of disease aggressiveness. A change in paclitaxel exposure-driven tumour size relative to baseline tumour size, was found to linearly and positively affect CRP production (Fig. 3.19). This coupled model later paved the way to explore the predictive value of model-derived variables e.g. early tumour response and non-baseline CRP concentration, along with patient and disease characteristics on PFS and OS in patients with advanced NSCLC.

Characterisation of tumour dynamics

Since in the clinical study tumour assessments were scarce and did not coincide with the CRP sampling times, we leveraged our previously developed tumour dynamics model [67] as a source of dense (i.e. more frequent) individual tumour sizes over time. The tumour dynamics model was developed based on the same study population and helped explore different tumour size metrics that could best reflect tumour dynamics and its influence on CRP production.

The adopted tumour dynamics model which we developed [67] successfully characterised the tumour size over time for advanced NSCLC patients with adequate uncertainty and predictive performance. Inclusion of a drug exposure metric and accounting for the development of resistance offered a realistic description of the change in tumour size over time especially in presence of individualised paclitaxel treatment (i.e. dose adaptations) that would consequently change drug exposure across the treatment cycles.

In the leveraged tumour dynamics model, a linear tumour growth rate constant, assuming a zero-order tumour growth rate, best characterised the change in tumour size over time, a first-order paclitaxel-induced tumour decay best characterised the drug-induced tumour decay, and an exponential decline in paclitaxel effect accounted for the development of resistance.

Our estimated linear tumour growth rate constant of 0.742 mm/month was similar to what has been previously reported in atezolizumab-treated NSCLC patients by Netterberg *et al.* [222], with a value of 0.724 mm/month (RSE 17%). However, it was different in other works which implemented different tumour growth functions than what was implemented in our case. In specific, it was lower (almost half) to what Chigusta *et al.* [223] have reported (1.47 mm/month, 95% CI 1.05–2.04) for necitumumab/gemcitabine/cisplatin-treated squamous NSCLC patients and almost 8-fold lower than the 6 mm/month (SE 0.04%) reported for paclitaxel/carboplatin-treated NSCLC patients by Wang *et al.* [224]. The reason for these similarities/discrepancies could be attributed to the components of the respective tumour dynamics model; in the former [222], a similar structural model was developed, accounting for the development of resistance as well as tumour shrinkage as a function of drug exposure (AUC_{cycle}) and interleukin 18, whereas in the latter two [223, 224], only a shrinkage component was considered (first-order in [223] and exponential in [224]) which could have impacted how the linear tumour growth rate constant parameter was estimated and have made the comparison to our result unjustified. Tumour doubling time—based on our median baseline tumour size of 8.3 cm and our estimated tumour growth rate constant—was calculated to be ≈ 2.5 years [211], which was within the range of 0.148–6.18 years that has been previously reported for adenocarcinoma and squamous cell carcinoma histological subtypes of NSCLC [225].

Although tumour decay can be described independent of the drug [223, 224]; here, drug exposure (specifically paclitaxel AUC per treatment cycle) was the driving force for the tumour decay. As a result, the impact of the individualised dosing i.e. changing drug exposure, on tumour growth could be assessed and simulation-based exploration of different doses and/or dosing regimens on tumour growth profile could be investigated—as reported in [211]. Although different metrics of paclitaxel exposure: paclitaxel concentration-time profile or paclitaxel AUC_{cycle} were explored; the latter was associated with a significant numerical improvement in the model fit, as reported in [211].

The estimated paclitaxel AUC_{cycle}-driven tumour decay rate constant at the start of treatment of $2.30{\cdot}10^{-5}\,(\mu\text{mol}{\cdot}\text{h/L})^{-1}{\cdot}\text{h}^{-1}$ translated to that a paclitaxel AUC_{cycle} of $4.49\,\mu\text{mol}{\cdot}\text{h/L}$ was needed for tumour shrinkage to occur, given a median baseline tumour size of 8.3 cm and a K_{growth} of 0.742 mm/month [211]. Nevertheless, drug effect exponentially declined with time at a rate of $8.75{\cdot}10^{-4}\,\text{h}^{-1}$ reflecting the emergence of resistance to paclitaxel treatment with time. Tumour resistance to paclitaxel treatment could be attributed to the (i) overexpression of the multidrug resistance gene (i.e. efflux pump); (ii) mutations in the tubulin subunit(s) or alterations in the binding regions; (iii) reduced function of apoptopic proteins; or (iv) alterations in cytokine expression (e.g. IL-6) [45–47] (section 1.3.2.5)—any of which would eventually lead to a reduced

drug effect with time. As mentioned, accounting for the development of resistance as well as tumour shrinkage as a function of drug exposure (AUC_{cycle}) was previously applied in atezolizumab-treated NSCLC patients [222]. Hence, our individual tumour size profiles were driven by the individual-specific paclitaxel AUC_{cycle} and the respective IIV. It is worth to mention that evaluation of the impact of both paclitaxel and the platinum drug was not feasible. Since all patients received a combination therapy, it was therefore difficult to discriminate between the independent effects. Moreover, since the dose of the platinum drugs was unchanged and only the paclitaxel dose was individualised, we can assume that the (changing) drug effect on tumour growth was adequately reflected by the (changing) paclitaxel exposure [67].

A limitation of the clinical data was the lack of paclitaxel exposure in almost 50% of the study population. To overcome this, the multiple imputation approach was adopted in this thesis. The multiple imputation leveraged the previously developed paclitaxel PK model [210]—based on the CEPAC-TDM patients with PK samples—to stochastically simulate *complete* datasets with imputed PK information to which the developed tumour dynamics model was then fitted. Rubin's rule [217], applied on both the population *and* individual level data, allowed to summarise the results while accounting for the variability and uncertainty associated with the imputed data and encountered from the multiple imputation framework within which the tumour dynamics model was developed. Population tumour dynamics parameter estimates derived from the multiple imputation were estimated with acceptable precision (RSE $< 43.2\%$, Table 3.3). Moreover, the descriptive statistics of the estimated individual tumour dynamics parameter estimates were in line with the respective population estimates and showed a small standard error ranging only from 0.0494% to 1.17% (Table 3.3).

To inform circulating CRP concentrations, dense (i.e. more frequent) individual tumour sizes over time had to be estimated. To estimate the individual tumour size, leveraging the individual tumour dynamics parameter estimates was preferred over the use of the population tumour dynamics parameter estimates. The former offered a more precise and less biased approach compared to the latter (similar to the sequential PK/PD modelling approaches where use of individual PK information is often preferred rather than to rely on the population PK parameters) [218]. Additionally, for a more improved precision that would only be closely achieved with simultaneous modelling of tumour size and CRP concentration, we accounted for the uncertainty of the individual tumour dynamics parameter estimates, encountered from the multiple imputation, to further inform the estimation of individual tumour sizes [218]. Simultaneous modelling of the tumour dynamics with CRP was practically not feasible, due to the development of the tumour dynamics model within the multiple imputation framework. A simultaneous approach would have necessitated expansion of this framework to include the coupled tumour dynamics - CRP model. This would have been both computationally exhaustive and impractical when it comes to identifying the most appropriate coupled tumour dynamics - CRP model. If the multiple imputation had not been needed i.e. no missed PK information, it would have been indeed useful to methodologically compare the performance of the framework under the simultaneous approach versus the use of the individual parameters with their uncertainty since earlier work had reported similar performance between both approaches [218, 226].

Characterisation of CRP concentration

The CRP turnover model well characterised the steady-state CRP concentration. Moreover, it identified baseline IL-6, baseline tumour size, disease stage, and smoking status as significant covariates impacting the CRP production rate constant.

The structural turnover model has been commonly applied to characterise different biomarkers with respect to different types of cancer [222, 227–244]. However, limited work has sought to characterise CRP in general [229, 245–247], of which only one was in cancer patients with breast cancer [229], and none was in lung cancer or NSCLC. Therefore, to our knowledge, this work is the first not only to characterise longitudinal CRP in advanced NSCLC patients but also to integrate it within a tumour-biomarker modelling framework.

The previous work characterising CRP spanned a broad spectrum of patient population (e.g. patients with rheumatoid arthritis, Crohn's disease, and systemic lupus erythromatosus [246], Gram-positive bacterial infections [247], trauma [245] or breast cancer [229]) with the majority reporting an inhibitory drug effect on CRP production or impact by its precursor IL-6. As a result, a direct comparison of our CRP model parameters to the reported results was not meaningful since they were very heterogeneous and work was not comparable. The covariate analysis was performed to identify influential factors on CRP concentration irrespective of the tumour dynamics. Since $K_{in,CRP}$ was the only estimable parameter and associated with IIV, covariates were investigated only on the CRP production rate constant, $K_{in,CRP}$. Our identified covariates on CRP production were all plausible and reflected not only the influence of CRP's physiological cytokine precursor but also CRP's association with disease aggressiveness and inflammatory status. The four identified covariates were baseline IL-6, baseline tumour size, disease stage, and smoking status. IL-6 is an inflammatory cytokine and the main precursor (inducer) of CRP gene expression [248]. Previous work has interestingly linked IL-6 surge to occur two days prior to CRP surge [229]. Unfortunately, linking IL-6 kinetics to CRP production, in a similar manner, was not feasible in our case due to the lack of longitudinal IL-6 measurements. Nevertheless, we managed to leverage the available baseline IL-6 concentration to inform CRP production. Although, it would have indeed been useful to link IL-6 kinetics, as precursor of CRP production, to the CRP kinetic profile; we do not expect that an earlier prediction of the expected CRP concentration (at the time scale of days) would be of a large impact with respect to prognostic predictions in real life (at the time scale of months). The subsequent two identified covariates: baseline disease stage and baseline tumour size, are metrics of disease aggressiveness and reflect the previously observed CRP's positive correlation with advanced disease stage and metastasis [92]. Lastly, in line with our findings, previous research has linked smoking with a positive increase in CRP concentration [249, 250]. This positive association has been hypothesised to be mediated through the chronic systemic inflammatory response and oxidative stress caused by smoking [251, 252]. Although former smokers had lower CRP concentration compared with current smokers, they still showed higher levels compared with non-smokers, indicating a long-term tissue damage and persistence of a low-grade inflammatory response as a result of smoking [253, 254].

Linking tumour dynamics to CRP

Our coupled tumour dynamics - CRP model characterised the change in CRP concentration over time, in which CRP was influenced by the fold change in tumour size relative to baseline tumour size and the previously identified covariates: baseline IL-6, disease stage, and smoking status.

Different structural models were investigated to identify the most reliable characterisation of the link between tumour dynamics and CRP production. Compared to a direct effect, a delay in the effect between tumour dynamics and CRP production, potentially mediated by IL-6-induced CRP activation, did not prove superior in terms of model performance. This could be attributed to the fact that the chronic inflammatory state induced by the advanced disease stage masked the lag time reflecting CRP synthesis, since CRP is in principle constantly elevated. The linear model linking the tumour size relative to baseline tumour size with CRP production translated to the direct, positive, and proportional impact the inflammatory status, associated with the malignant tumour, has on triggering the release of the inflammatory cytokines. Which, in turn, activates CRP production and release [92].

A similar framework, linking tumour dynamics to circulating serum biomarkers was previously applied in SCLC patients [240]. In that framework, a hypothetical tumour dynamic compartment informed by the potential impact of the different treatment strategies (i.e. radiotherapy and chemotherapy) and the development of resistance to chemotherapy, influenced the circulating biomarker concentrations. However, no impact of drug concentration or tumour size measurement was considered. In contrast to that framework, in our work, our developed coupled tumour dynamics - CRP model rather leveraged our previously developed tumour dynamics model to link drug exposure and dense tumour sizes to biomarker concentration with the knowledge that the predicted tumour sizes accounted for the impact of treatment (i.e. paclitaxel exposure) and development of resistance—given the structural model of the tumour dynamics model (Eq. 2.62). Model evaluation showed no misspecification in the GOF plots and an adequate predictive performance was observed in the VPC especially in the terminal part excluding the presence of bias for not considering drop outs. Hence, post-treatment dense tumour sizes and CRP concentrations were successfully estimated for all patients, across the whole duration of treatment, allowing a rich assessment of non-baseline tumour sizes and CRP concentrations, at later time points, as potential predictors of efficacy endpoints.

4.2.2.2 Characterisation of efficacy endpoints

The efficacy endpoints (PFS and OS) were selected as they represent the commonly used clinical endpoints for oncological assessments [57]. While OS needs longer follow-up, PFS is assessed earlier and takes into consideration not only the tumour response but also survival/death. Thus, it can be considered a surrogate for OS—if earlier assessments are required—and is preferred over TTP as it better correlates with survival/death events.

The choice to model the efficacy endpoints by means of TTE analysis was mainly driven by the fact that time was a critical factor to be considered. A binary logistic regression model [255, 256] which would dichotomise patients to progressed/not progressed or alternatively dead/alive would clearly exclude time but most importantly, as a consequence, (i) not account for patients who progress early and (ii) unfairly pool all patients who have progressed—regardless *when*—into one cohort. Since the majority of patients progressed/died, given their advanced disease state, the choice of this modelling approach would have been an oversimplification and inaccurate reflection of the nature of the data.

An alternative modelling approach could have been a multi-state model to account for the sequence of the different states a patient could fall into e.g. alive, not-progressed, progressed, dead etc. Although this is becoming a popular approach [257–261], the small sample size of our data hindered a reliable and powered assessment of the different states (previous investigation on the same clinical data communicated personally with Joachim Grevel, 12 October 2022). Therefore, a TTE analysis and specifically parametric TTE models, best fitted the nature of the data and our objectives to identify predictors of the efficacy endpoints and evaluate their impact on the time of occurrence of the event. Since it was previously reported that no significant difference in PFS or OS existed between both treatment arms (p-value=0.228 and p-value=0.682, respectively [33]), no stratification was adopted and both arms were pooled for a joint characterisation of PFS and OS.

Landmark survival analysis within a parametric TTE model has been commonly applied before in NSCLC [224, 262, 263] as a reliable method to investigate post/non-baseline predictors e.g. tumour dynamics. The landmark survival analysis ensures that events are predicted *after* the monitoring time of the predictor, hence, avoiding the *immortal time bias* as described by Khandelwal *et al.* [264] (i.e. no prediction of PFS or OS event before assessment of predictor e.g. CRP at the start of treatment cycle 3). It also allowed all covariates measured prior to our selected landmark time to be eligible potential predictors for further investigation. While landmark analysis has the advantage of a straightforward application and easy interpretation of results, it is worth to mention that results should be interpreted in the context of the landmark population and are only generalisable to patients who have survived up until the landmark time (in our case the start of treatment cycle 3). Moreover, common limitations of landmark analysis which include exclusion of the events occurring before the landmark time, and dependency of the results on the choice of the landmark time, can be easily mitigated [177]. To overcome these limitations: (i) a choice of an early landmark time can avoid exclusion of a high proportion of events and a consequent loss of power as a result of the decreased sample size, and (ii) an *a priori* choice of the landmark time together with a sensitivity analysis at e.g. repeated different landmark points, to demonstrate consistent results, can be performed. Although the application of a landmark analysis comes at the expense of excluding potential predictors measured after the landmark time [264], this was not applicable in our case since, *a priori*, we were focused on the identification of *early* predictors.

Characterisation of PFS and OS by means of a parametric TTE model and the chosen landmark time at the start of treatment cycle 3 demonstrated that for PFS, there was an initial increase in hazard followed by a later decrease over time (i.e. log-normal distribution of hazard). This distribution, as inferred from Fig. 3.6, was probably initially dominated by progression and death events and later dominated by only the death events where the progression events reach a plateau. For OS, a Weibull distribution characterised an increasing risk of death with time. The TTE models for both PFS and OS were able to successfully describe the observed median TTE of the landmark population. The simulated median PFS was 7.13 months (90% CI: 7.09–7.17 months) in line with the observed median PFS of 7.07 months, whereas the simulated median OS was 11.8 months (90% CI: 11.7–11.9 months) in line with the observed median OS of 11.1 months. It is also worth noting that compared to our patient population, longer median PFS and OS were expected in the landmark patient population because it included only patients surviving beyond treatment cycle 3.

Although modelling the risk, probability and/or reasons of dropping out from a study are usually encouraged to avoid biased parameter estimates—especially for further realistic simulations [265], our framework in *project II* did not consider patient dropouts. In our study context, a patient dropout did not necessarily mean being lost to follow-up but it rather referred to neither being enrolled in the study nor receiving further study treatment any longer. Therefore the main reason for not accounting for the probability of dropping out from the study in our TTE models was that not completing all six treatment cycles did not impact the data collection for the time-to-event endpoints. In other words, the majority of patients who did not complete all six treatment cycles were still followed up for information on progression and survival, and a similar ratio of right censored vs uncensored patients was observed between the patient cohort who completed the study duration vs those who did not (section 3.2.1.2). Therefore, an early exit from the study before completing all six treatment cycles was not considered to be a competing event (i.e. did not hinder a progression/death event to be monitored) nor was it considered a censored event on its own (i.e. patients continued to be followed-up). Additionally, no misspecifications were observed in any of the diagnostic plots (especially in the later time points) suggesting that the model described the data well without bias (Figs 3.27 and 3.23). Therefore, in our case, dropouts with respect to *not* completing all six treatment cycles could be described as "ignorable" [265] since the interpretation of the efficacy endpoints would still be valid in their absence.

Identified significant predictors and their impact

Our focus on *early* prognostic metrics was the main motivation for the pre-selection of potential predictors; to allow for an early identification of patients' prognosis, timely decision-making, and for future translation into clinical practice where early metrics are necessary. For this reason, we identified the first three treatment cycles to be of relevance and consequently chose a landmark time defined at the start of treatment cycle 3 for our investigation, and did not explore later predictors or the impact of later landmark times.

Our coupled tumour dynamics - CRP model offered a dense prediction of individual tumour sizes and CRP concentrations across the whole duration of treatment of 18 weeks—even for patients with observed short survival or lack of dense measurements. Hence, post-treatment tumour size and CRP concentrations were estimated for all patients and allowed for a rich assessment of non-baseline CRP concentrations and tumour shrinkage at different time points. With a focus on early predictors and since one of the main objectives was leveraging longitudinal data for their potential prognostic value, different metrics were derived (e.g. fold change, relative change, absolute difference) based on the longitudinal data of the first 3 treatment cycles. Derivation of non-time varying metrics based on longitudinal data for TTE analysis in NSCLC has been commonly applied before with respect to tumour size-related metrics e.g. tumour shrinkage at week 8 [67, 224]. However, none to our knowledge has similarly leveraged longitudinal biomarker concentration. Investigation of these longitudinal predictors as time-varying covariates was not considered since in clinical practice the whole time profile of the respective predictor e.g. CRP concentration or tumour size would not be available early on during treatment.

Due to the challenges described earlier with the multiple imputation (section 4.2.2.1), and the numerical complexity and computational intensity required with simultaneous modelling i.e. jointly characterising tumour dynamics - CRP - survival, a two-stage approach was adopted in which

metrics derived from the coupled tumour dynamics - CRP model were utilised in a sequential manner as potential predictors in the TTE model. Besides the less computational need, the two-stage approach was found to be similar in terms of parameter estimates, covariate effects, and prediction of OS in the context of NSCLC when compared with the joint modelling approach in a recent investigation [266]. Therefore, it reliably suited our objective and analysis plan.

Based on our analysis, the inflammatory level at the start of treatment cycle 3 (CRP_{cycle3}) and the difference in the inflammatory level between the start of treatment cycle 3 and cycle 2 ($CRP_{cycle3-2}$) were significant predictors of both PFS and OS. Additionally, tumour load (baseline tumour size) and presence of liver lesions were predictors of OS.

Previous work has already extensively reported a positive correlation between baseline CRP concentration and poor prognosis in advanced NSCLC patients receiving different treatment modalities [67–88, 90, 104]; however, none to our knowledge has investigated longitudinal CRP data—apart from Xiao *et al.* [68] who reported two CRP measurements before and after chemotherapy but unfortunately without specifying when exactly the second measurement was and without leveraging both measurements in their reported analysis. Exploration of the wide panel of CRP-related metrics revealed superior predictive potential for non-baseline CRP concentrations compared with baseline CRP concentrations. Specifically, CRP_{cycle3} was the strongest predictor of both PFS and OS compared to the other time points e.g. baseline CRP or CRP_{cycle2}. This indicated that as we move further in time (i.e. treatment cycles), CRP becomes more representative of the patient disease state and co-morbid condition especially after receiving treatment, compared with the baseline concentration, as an example, which only reflects the patient's condition prior to treatment. Moreover, $CRP_{cycle3-2}$ was also a strong predictor of PFS and OS. This adds to the fact that not only the inflammatory level i.e. absolute CRP concentration, at a later time point was predictive of the patient's prognosis but also the dynamic change of CRP and its degree of reduction across treatment cycles. Our results reported that compared to the observed median PFS in the respective landmark population, median PFS was 58.4% and 38.0% longer with the 5^{th} percentile of CRP_{cycle3} and $CRP_{cycle3-2}$, respectively and 47.2% and 27.9% shorter with the 95^{th} percentile of CRP_{cycle3} and $CRP_{cycle3-2}$, respectively. For OS, and compared to the observed median OS in the respective landmark population, median OS was 177% and 67.6% substantially longer with the 5^{th} percentile of CRP_{cycle3} and $CRP_{cycle3-2}$, respectively and 53.8% and 32.9% shorter with the 95^{th} percentile of CRP_{cycle3} and $CRP_{cycle3-2}$, respectively. Regrettably, no similar predictors were previously identified in the literature to offer comparison to our results.

In addition to the CRP-related metrics, presence of liver lesions and a high tumour load were associated with shorter OS i.e. worse prognosis. Our results are in line with previous work [267–275] which reported worse prognosis for NSCLC patients with metastatic liver lesions regardless of the type of treatment, with an even worse prognosis for patients receiving cytotoxic therapy compared to those receiving targeted therapy or immunotherapy [268]. Moreover, patients with liver lesions had the worst prognosis compared to patients with lesions in other metastatic sites as the brain or bone [270–272, 275]. For NSCLC patients with liver lesions, a median OS ranging from 3 months to 9.3 months was reported [267, 268, 270–273, 275]—in line with our reported median OS of 8.87 months (Fig. 3.28). Similarly, tumour size is a well-known prognostic factor of survival in patients with NSCLC [276]. Specifically, baseline tumour size, as a measure of baseline tumour burden, has been reported to be a prognostic marker in NSCLC patients, in which a larger baseline

tumour size was consistent with poor prognosis [211, 224, 277–280]. Although the biological explanation for this relationship is unclear, it is hypothesised that a higher tumour burden could be linked to the development of more chemotherapy-resistant tumour cells, and/or decreased sensitivity to therapy due to the presence of hypoxia and poor perfusion within the tumour microenvironment [277]. In line with the current knowledge, our results showed that compared to the observed median OS in the respective landmark population, median OS was 51.3% longer and 13.8% shorter at the 5^{th} and 95^{th} percentiles of baseline tumour size, respectively. Despite the fact that tumour shrinkage at week 8 has been an additional tumour-related metric commonly linked to poor prognosis [67, 224, 263]—especially in comparison with tumour shrinkage at weeks 4 and 6 [224], our results only identified baseline tumour size as a tumour-related metric when tested along with tumour shrinkage at week 7 (exploring the impact of tumour shrinkage at week 8 was not applicable due to the chosen landmark time at the start of treatment cycle 3). The reason for this could be attributed to the fact that in previous work, only baseline characteristics were explored besides the longitudinal tumour-size changes i.e. tumour shrinkage, whereas in our presented work longitudinal CRP-related metrics as well as longitudinal neutrophil-to-lymphocyte ratio-related metrics were additionally explored—as markers of inflammation; which could have masked the impact of tumour shrinkage i.e. the significance of tumour shrinkage in a univariate assessment ceases to be the case when jointly evaluated with longitudinal CRP-related metrics. This was an indication that longitudinal CRP metrics were more significant, as prognostic predictors, than longitudinal tumour size when both were jointly tested.

Challenging the limitations of the landmark analysis

As mentioned earlier, despite the advantages and ease of application of the landmark survival analysis, it is nevertheless associated with limitations.

First, to minimise exclusion of a high proportion of events and a consequent loss of power as a result of the decreased sample size, and to identify early metrics of efficacy, the start of treatment cycle 3 was chosen *a priori* as the landmark time of interest (mid-treatment regimen time) and no later time points were considered. A landmark time at the start of treatment cycle 3 resulted in a total sample size of 203 (78.9%) and 235 (91.4%) patients for the characterisation of PFS and OS, respectively.

Second, to challenge the dependency of our identified predictors on our chosen landmark time at the start of treatment cycle 3 and demonstrate consistent results, a covariate analysis was undertaken while defining a landmark time at treatment cycle 2 instead. Considering the exclusion of metrics involving treatment cycle 3 (e.g. CRP_{cycle3} and $CRP_{cycle3-2}$), the analysis identified very similar predictors but with a weaker impact (Figs. 4.1 and 4.2). CRP at later treatment cycles was always a dominating significant predictor; the most recent CRP concentration i.e. CRP_{cycle2} was the most significant predictor on PFS, and along with liver lesions on OS (for OS, baseline tumour size was significant at p-value 0.05 but not 0.01). This indicated the robustness of our results with a landmark time at treatment cycle 3. Whereas on one hand, a landmark time at the start of treatment cycle 3 offered the opportunity to identify more impactful predictors and account for the extent of reduction in CRP concentrations by measuring $CRP_{cycle3-2}$, an even earlier monitoring time—if needed for deciding on a specific prognostic outcome for a patient—still identified CRP to be dominating and CRP_{cycle2} to be the strongest predictor on efficacy endpoints, despite a weaker magnitude of impact compared with CRP_{cycle3}.

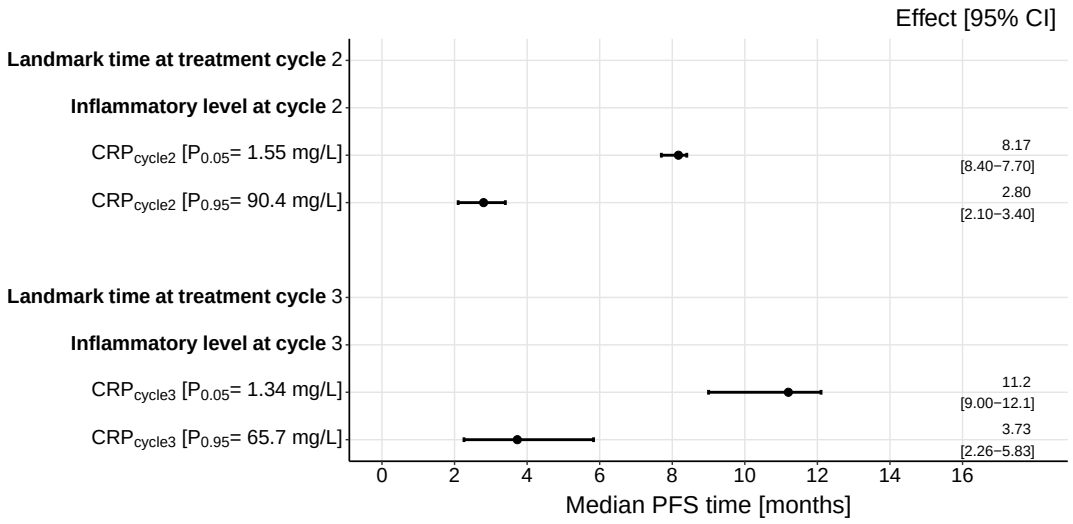

Figure 4.1: Forest plot of the impact of CRP_{cycle2} identified at a landmark time at the start of treatment cycle 2 and CRP_{cycle3} identified at a landmark time at the start of treatment cycle 3 on median progression-free survival (PFS). Effects of the continuous predictors (i.e. CRP_{cycle2}, CRP_{cycle3}) are shown at the 5^{th} and 95^{th} percentiles of the respective predictor distribution. Black dots: predictor effect; horizontal lines: 95% confidence intervals (CI) derived based on the parameter uncertainty hence the difference in magnitude. CRP: C-reactive protein.

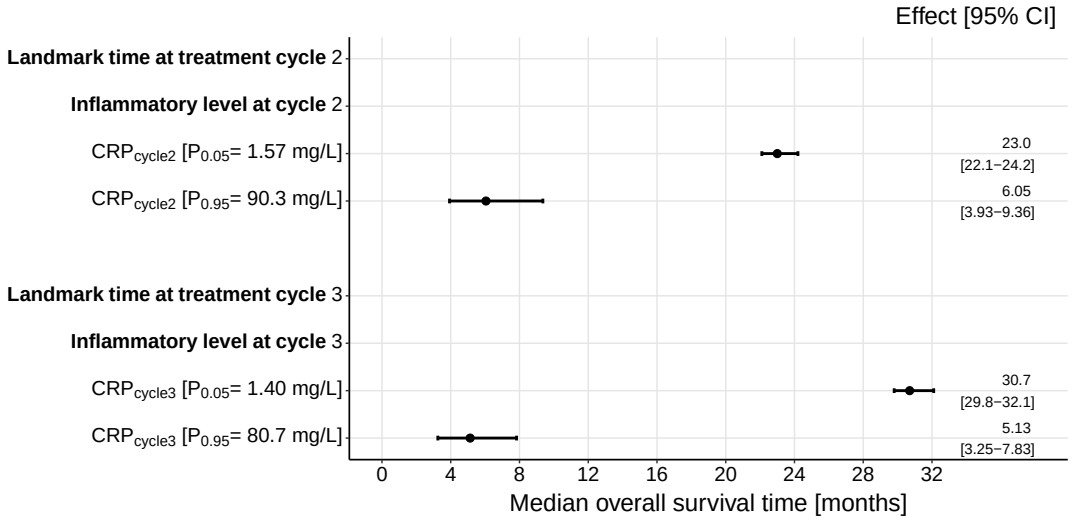

Figure 4.2: Forest plot of the impact of CRP_{cycle2} identified at a landmark time at the start of treatment cycle 2 and CRP_{cycle3} identified at a landmark time at the start of treatment cycle 3 on median overall survival. Effects of the continuous predictors (i.e. CRP_{cycle2}, CRP_{cycle3}) are shown at the 5^{th} and 95^{th} percentiles of the respective predictor distribution. Black dots: predictor effect; horizontal lines: 95% confidence intervals (CI) derived based on the parameter uncertainty hence the difference in magnitude. CRP: C-reactive protein.

4.2.2.3 Evaluation of the modelling framework

To determine our model's reproducibility and robustness, evaluation of our modelling framework was essential. The evaluation of our modelling framework with an external dataset (section 2.1.2.4) was unfortunately not feasible. Given our specific patient cohort (advanced NSCLC patients receiving first-line chemotherapy with paclitaxel in combination with a platinum drug) and the unique characteristics of our dataset (tumour size assessments, CRP concentrations, progression/death events), identifying an external dataset of similar characteristics was not possible. Alternatively, splitting of our data into training and testing datasets was not an optimal approach because of the small number of patients and the risk of loss of power. Nevertheless, to overcome these hurdles, we alternatively compared our simulation-based results to our observed data across the different percentiles of CRP_{cycle3}—as presented in Fig. 3.30. The KM plots of our observed data (Fig. 3.30, C and D) showed a similar pattern to our simulation-based results (Fig. 3.30, A and B): the upper three percentile intervals of model-estimated CRP_{cycle3} across the observed PFS events (Fig. 3.30, C) had a comparable median PFS to our simulation results (Fig. 3.30, A) whereas for OS all the different percentile intervals (Fig. 3.30, D)—apart from the lowest two—had a corresponding median OS compared to our simulation results (Fig. 3.30, B). Although it was clear in Fig. 3.30 that higher inflammatory levels were positively associated with poor prognosis, our work did not focus on the impact of modulating inflammation on efficacy endpoints i.e. whether blocking CRP release or administration of anti-inflammatory medications would reflect a favourable prognosis. Alternatively, we sought to investigate the potential of monitoring inflammation as a potential prognostic marker through the measurements of specific inflammatory biomarkers. Thus, based on our results, we can claim that monitoring CRP at treatment cycles 2 and 3 in advanced NSCLC patients can offer a prognostic value for predicting PFS and OS.

4.2.2.4 Outlook and application of our modelling framework

The strongest potential of our modelling framework successfully lies in the identification of advanced NSCLC patients at risk of poor prognosis by defining early prognostic predictors of PFS and OS in order to determine patients who have high probability of treatment benefit and those who are at risk of treatment failure. Hence, optimising patient response by informing treatment decisions and avoiding the needless exposure of patients to toxic chemotherapy.

Currently, for a more cost-effective and easy-to-apply framework in clinical practice, we would suggest to evaluate our framework for optimal CRP sampling strategies. This can be achieved by evaluating CRP concentration across different time points to identify the few optimal and most informative time points that still provide the most precise predictions. The aim would be to (i) minimise the challenges associated with longitudinal sampling e.g. more laborious work, extra cost compared to single sampling, and patient compliance when it comes to repeated sampling, while (ii) providing precise predictions of our identified predictors and consequently for the endpoint of interest in clinical practice. As a subsequent step, the framework could then be integrated in an interactive platform e.g. an R Shiny application in which patient's characteristics and key tumour size and CRP measurements can be imported for prognostic predictions and a seamless application in the clinical setting and the real-world of NSCLC.

Furthermore, this work can be expanded to investigate the impact of different dosing regimens or linked to dosing recommendations for a model-informed precision dosing and therapy optimisation in clinical practice. It could also identify an optimal treatment strategy if linked to the different toxicities and adverse events e.g. neuropathy and/or neutropaenia for a realistic amalgamation of the pillars of a successful oncology treatment, in which not only survival and lack of progression are the ultimate aims but also a good quality-of-life for the patient that equally balances treatment efficacy with its associated toxicity in what is known as a "clinical utility index" [281, 282].

Although this modeling framework was applied to a specific patient cohort, treatment protocol, and study design, and consequently the impact of the identified predictors is in the first place only applicable to this specific setting; our clinical data reflected the real-world target population along with the encountered challenges of missing data. Therefore, based on our modelling framework and after its successful evaluation, we are confident that our model(s) could be applied in the real-world of NSCLC and that a similar modelling framework can generally be applied to further explore the potential of different biomarkers (such as molecular biomarkers which are becoming more promising especially with the advances in molecular testing) and across different treatment modalities within the clinical setting (such as immunotherapy, targeted therapy)—within or outside the scope of oncology.

5 | Conclusion and perspectives

In this thesis, we have focused on applying pharmacometric methodologies for optimised early clinical decision-making in oncology at both early clinical drug development and at bedside.

Project I was driven by the need for dose-optimisation in oncology to better identify safer and more tolerable doses for the patients and based on the recent Food and Drug Administration (FDA)'s initiative (Project Optimus) to better characterise the dose-response relationship. The objective of *project I* was to better understand the performance of model-based approaches which are recommended to inform the design of phase II dose-finding studies and to optimally leverage their data, for the identification of a drug effect and the expected dose-response relationship. A simulation-based exploratory assessment was performed within an oncology setting to evaluate the newly proposed combined Likelihood Ratio Test (cLRT) approach versus the Multiple Comparison Procedure (MCP) approach which received the FDA's "fit-for-purpose" qualification. Based on the simulation results, the power of identifying a drug effect was higher for cLRT compared with MCP, due to leveraging of the full information of the longitudinal data. Nevertheless, cLRT was indeed more computationally demanding, less robust, and associated with an inflated type I error (i.e. false positive). Therefore, further investigations under models spanning different complexities and study designs are needed, to further understand reasons for cLRT's lack of robustness and justify the computational capacity it needs to identify a dose-response relationship.

Only if reasons behind cLRT's inflated type I error are identified and cLRT's robustness is established, then cLRT can offer a more reliable and more powerful alternative to MCP, which has shown better robustness and better control of type I error but with weaker power. Through a robust and powerful method, with a well-controlled type I error, that allows leveraging the complexity of dose-finding clinical data, successful characterisation of the dose-response relationship can be achieved and dose/dosing needed for the intended response in confirmatory clinical trials can be recommended. In line with the goals of "Project Optimus", a better informed dose selection would maximise the drug's efficacy, safety and tolerability. This would in return (i) increase chances of successful phase III trials, hence saving time, effort, and money; (ii) prevent/minimise exposure of patients to more toxic or less efficacious doses; and (iii) would eventually reflect on faster approvals of promising therapies to the intended patient population [3, 4].

Of equal importance to finding safe and efficacious doses/dosing for therapeutic compounds, is the need for early monitoring of patient response at bedside. Monitoring of patient response allows to predict treatment outcomes and consequently allows the efficient use of therapeutic compounds through the best treatment decisions. Non-small cell lung cancer (NSCLC) represents a case example of a cancer disease with high burden and poor prognosis. Thus, timely and early prediction of prognosis is recommended to spare NSCLC patients unnecessary toxicities from ineffective treatments. For this reason, *Project II* aimed to identify early predictors of efficacy in advanced NSCLC patients treated with first-line combination chemotherapy of paclitaxel and either carboplatin or cisplatin, with a special focus on leveraging longitudinal C-reactive protein (CRP)

concentration, as an inflammatory and routine serum biomarker associated with tumour progression and aggressiveness, to identify early prognostic metrics. The successfully developed coupled tumour dynamics - CRP modelling framework characterised the relationship between drug exposure, tumour dynamics and circulating CRP concentrations. Moreover, it offered a rich source of longitudinal tumour sizes and CRP concentrations for assessment as predictors of the common efficacy endpoints, PFS and OS. Thus, by leveraging different tumour size- and CRP-related metrics and considering a wide panel of patient- and disease-related characteristics as potential predictors, we identified the strongest prognostic factors of PFS to be the inflammatory level at treatment cycle 3 (CRP_{cycle3}) and the extent of reduction in the inflammatory level between treatment cycle 3 and cycle 2. Besides the latter two CRP-related predictors, baseline tumour size and presence of liver lesions, were of additional prognostic value for OS. However, amongst all of the identified predictors, CRP_{cycle3} was by far of the highest impact. Therefore, strengthening the fact that non-baseline measurements are more representative of the patient status throughout the course of treatment and should be better leveraged compared to baseline measurements.

Identification of CRP as a prognostic predictor represents a minimally invasive, and an easily and reliably assessed inflammatory biomarker. Although CRP is a non-specific marker of inflammation and tumour aggressiveness, it nevertheless showed a strong prognostic value for PFS and OS and a strong reflection of the patient (co-)morbid state. As a result, monitoring of the inflammatory level through measurements of circulating CRP concentration can reflect the aggressiveness of the malignant state and presents a promising prognostic marker to better and early-on guide treatment decisions. Even though the use of laboratory-based serum biomarkers for prediction of prognosis and consequently informing treatment decisions is still not fully established in clinical practice—due to the lack of reproducibility, standardised collection, and consistent cut-off values, we believe that with the prognostic value shown with CRP and the presence of reproducible, accurate, precise, and reliable methods for its assay [92,94], CRP presents an attractive biomarker to motivate, promote, and initiate the use of laboratory-based serum biomarkers and their integration in clinical practice.

In conclusion, the work presented in this thesis has applied pharmacometric modelling and simulation principles to help inform early clinical decision-making at different stages within the oncology setting. We began with focusing on the assessment of model-based approaches to early on identify a dose-response relationship during the development phase of therapeutic compounds. We recommended that for the cLRT approach leveraging longitudinal data, despite its high power to identify a drug effect, more systematic assessments are needed to investigate its robustness and justify its computational demands. Later, we focused on the early prediction of patient response and prognostic outcome, and identified CRP at treatment cycle 3 as the most prognostic predictor of PFS and OS in patients with advanced NSCLC. Thus, longitudinal measurements of CRP can help clinicians not only to decide on the most useful treatment decision but also to early on alter treatment to spare patients unnecessary toxicities. Application of our developed modelling framework goes beyond NSCLC. It can be applied in different settings within oncology e.g. molecular biomarkers, treatment modalities and/or cancer types, to identify reliable and early prognostic markers when their relationship to the tumour is identified, for better therapy optimisation and patient monitoring. Thus, allowing early clinical decisions to be based on all the available knowledge and an optimal leverage of oncology data and endpoints through the use of pharmacometrics to better optimise the need for early clinical decision-making in oncology.

References

[1] H. Sung, J. Ferlay, R. L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, F. Bray. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: A Cancer Journal for Clinicians 71: 209–249 (2021).

[2] J. Fourie Zirkelbach, M. Shah, J. Vallejo, J. Cheng, A. Ayyoub, J. Liu, R. Hudson, R. Sridhara, G. Ison, L. Amiri-Kordestani, S. Tang, T. Gwise, A. Rahman, R. Pazdur, M. R. Theoret. Improving dose-optimization processes used in oncology drug development to minimize toxicity and maximize benefit to patients. Journal of Clinical Oncology 40: 3489–3500 (2022).

[3] L. V. Sacks, H. H. Shamsuddin, Y. I. Yasinskaya, K. Bouri, M. L. Lanthier, R. E. Sherman. Scientific and regulatory reasons for delay and denial of FDA approval of initial applications for new drugs, 2000-2012. JAMA 311: 378–384 (2014).

[4] J. Cross, H. Lee, A. Westelinck, J. Nelson, C. Grudzinskas, C. Peck. Postmarketing drug dosage changes of 499 FDA-approved new molecular entities, 1980–1999. Pharmacoepidemiology and Drug Safety 11: 439–446 (2002).

[5] C. Le Tourneau, J. J. Lee, L. L. Siu. Dose escalation methods in phase I cancer clinical trials. JNCI: Journal of the National Cancer Institute 101: 708–720 (2009).

[6] Food and Drug Administration. Optimizing the Dosage of Human Prescription Drugs and Biological Products for the Treatment of Oncologic Diseases: Guidance for Industry (2023). `https://www.fda.gov/media/164555/download`. (Accessed on: 08.12.2023).

[7] J. M. Bullock, A. Rahman, Q. Liu. Lessons learned: Dose selection of small molecule–targeted oncology drugs. Clinical Cancer Research 22: 2630–2638 (2016).

[8] Food and Drug Administration. Project Optimus: Reforming the dose optimization and dose selection paradigm in oncology. `https://www.fda.gov/about-fda/oncology-center-excellence/project-optimus`. (Accessed on: 08.12.2023).

[9] Y. K. Lyauk, D. M. Jonker, T. M. Lund. Dose finding in the clinical development of 60 US Food and Drug Administration–approved drugs compared with learning vs. confirming recommendations. Clinical and Translational Science 12: 481–489 (2019).

[10] S. Fudio, A. Sellers, L. Pérez Ramos, B. Gil-Alberdi, A. Zeaiter, M. Urroz, A. Carcas, R. Lubomirov. Anti-cancer drug combinations approved by US FDA from 2011 to 2021: Main design features of clinical trials and role of pharmacokinetics. Cancer Chemotherapy and Pharmacology 90: 285–299 (2022).

[11] European Medicines Agency. ICH Topic E 4: Dose Response Information to Support Drug Registration (1994). `https://www.ema.europa.eu/en/documents/scientific-guideline/ich-e-4-dose-response-information-support-drug-registration-step-5_en.pdf`. (Accessed on: 08.12.2023).

[12] Food and Drug Administration. Statistical Review and Evaluation: Qualification of Statistical Approach. `https://www.fda.gov/media/77169/download`. (Accessed on: 08.12.2023).

[13] European Medicines Agency. Qualification Opinion of MCP-Mod as an efficient statistical methodology for model-based design and analysis of Phase II dose finding studies under model uncertainty (2014). `https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/qualification-opinion-mcp-mod-efficient-statistical-methodology-model-based-design-and-analysis-phase-ii-dose-finding-studies-under-model-uncertainty_en.pdf`. (Accessed on: 2023-12-12).

[14] F. T. Musuamba, E. Manolis, N. Holford, S. A. Cheung, L. E. Friberg, K. Ogungbenro, M. Posch, J. Yates, S. Berry, N. Thomas, S. Corriol-Rohou, B. Bornkamp, F. Bretz, A. C. Hooker, P. H. van der Graaf, J. F. Standing, J. Hay, S. Cole, V. Gigante, K. Karlsson, T. Dumortier, N. Benda, F. Serone, S. Das, A. Brochot, F. Ehmann, R. Hemmings, I. S. Rusten. Advanced methods for dose and regimen finding during drug development: Summary of the EMA/EFPIA workshop on dose finding (London 4–5 December 2014). CPT: Pharmacometrics & Systems Pharmacology 6: 418–429 (2017).

[15] S. Mills, P. Donnan, D. Buchanan, B. H. Smith. Age and cancer type: Associations with increased odds of receiving a late diagnosis in people with advanced cancer. BMC Cancer 23: 1174 (2023).

[16] J. P. van Meerbeeck, C. Franck. Lung cancer screening in Europe: Where are we in 2021? Translational Lung Cancer Research 10: 2407 (2021).

[17] Robert Koch Institute. Cancer in Germany 2019/2020 (2024). https://www.krebsdaten.de/Krebs/EN/Content/Publications/Cancer_in_Germany/cancer_chapters_2019_2020/cancer_germany_2019_2020.pdf?__blob=publicationFile. (Accessed on: 27.05.2024).

[18] F. Galli, S. Rohrmann, M. Lorez. Lung cancer survival in Switzerland by histology, TNM stage and age at diagnosis. Schweizer Krebs-Bulletin 39: 69–73 (2019).

[19] Cancer Australia. National cancer stage at diagnosis data (2018). https://ncci.canceraustralia.gov.au/features/national-cancer-stage-diagnosis-data. (Accessed on: 28.12.2023).

[20] P. M. Ellis, R. Vandermeer. Delays in the diagnosis of lung cancer. Journal of Thoracic Disease 3: 183 (2011).

[21] Cancer Research UK. Five reasons the UK must diagnose cancer earlier–and four ways to do it (2015). https://news.cancerresearchuk.org/2015/08/10/five-reasons-the-uk-must-diagnose-cancer-earlier-and-four-ways-to-do-it/. (Accessed on: 28.12.2023).

[22] National Cancer Institute. Stage at Diagnosis. https://progressreport.cancer.gov/diagnosis/stage. (Accessed on: 28.12.2023).

[23] World Health Organization. Global cancer burden growing, amidst mounting need for services (2024). https://www.who.int/news/item/01-02-2024-global-cancer-burden-growing--amidst-mounting-need-for-services. (Accessed on: 2024-05-27).

[24] C. Li, S. Lei, L. Ding, Y. Xu, X. Wu, H. Wang, Z. Zhang, T. Gao, Y. Zhang, L. Li. Global burden and trends of lung cancer incidence and mortality. Chinese Medical Journal 136: 1583–1590 (2023).

[25] American Cancer Society. Cancer Facts & Figures 2023 (2023). https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/annual-cancer-facts-and-figures/2023/2023-cff-special-section-lung-cancer.pdf. (Accessed on: 2023-07-06).

[26] American Lung Association. State of Lung cancer: 2022 Report (2022). https://www.lung.org/getmedia/647c433b-4cbc-4be6-9312-2fa9a449d489/solc-2022-print-report. (Accessed on: 2023-07-06).

[27] D. Karadoğan, M. Ünsal. Prognosis and Prognostic Factors of Lung Cancer. In: Airway diseases. Springer. 1–16 (2023).

[28] American Cancer Society. Tests for Lung Cancer (2024). https://www.cancer.org/cancer/types/lung-cancer/detection-diagnosis-staging/how-diagnosed.html. (Accessed on: 2024-06-01).

[29] S. Mirsadraee, D. Oswal, Y. Alizadeh, A. Caulo, E. J. van Beek. The 7^{th} lung cancer TNM classification and staging system: Review of the changes and implications. World Journal of Radiology 4: 128–134 (2012).

[30] M. Zheng. Classification and pathology of lung cancer. Surgical Oncology Clinics 25: 447–468 (2016).

[31] D. Planchard, S. Popat, K. Kerr, S. Novello, E. F. Smit, C. Faivre-Finn, T. S. Mok, M. Reck, P. E. van Schil, M. D. Hellmann, S. Peters, on behalf of the ESMO Guidelines Committe. Metastatic non-small cell

lung cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up (2019). https://www.esmo.org/content/download/227453/3874538/file/ESMO-CPG-mNSCLC-18SEPT2019.pdf. (Accessed on: 2023-07-07).

[32] G. Goeckenjan, H. Sitter, M. Thomas, D. Branscheid, M. Flentje, F. Griesinger, N. Niederle, M. Stuschke, T. Blum, K. Deppermann, J. Ficker, L. Freitag, A. S. Lübbe, T. Reinhold, E. Späth-Schwalbe, D. Ukena, M. Wickert, M. Wolf, S. Andreas, T. Auberger, R. P. Baum, B. Baysal, J. Beuth, H. Bickeböller, A. Böcking, R. M. Bohle, I. Brüske, O. Burghuber, N. Dickgreber, S. Diederich, H. Dienemann, W. Eberhardt, S. Eggeling, T. Fink, B. Fischer, M. Franke, G. Friedel, T. Gauler, S. Gütz, H. Hautmann, A. Hellmann, D. Hellwig, F. Herth, C. Heussel, W. Hilbe, F. Hoffmeyer, M. Horneber, R. M. Huber, J. Hübner, H. U. Kauczor, K. Kirchbacher, D. Kirsten, T. Kraus, S. Lang, U. Martens, A. Mohn-Staudner, K. Müller, J. Müller-Nordhorn, D. Nowak, U. Ochmann, B. Passlick, I. Petersen, R. Pirker, B. Pokrajac, M. Reck, S. Riha, C. Rübe, A. Schmittel, N. Schönfeld, W. Schütte, M. Serke, G. Stamatis, M. Steingräber, M. Steins, E. Stoelben, L. Swoboda, H. Teschler, H. W. Tessen, M. Weber, A. Werner, H. E. Wichmann, W. E. Irlinger, C. Witt, H. Worth, German Respiratory Society and German Cancer Society. Prevention, Diagnosis, Therapy, and Follow-up of Lung Cancer. Pneumologie 65: 39–59 (2011).

[33] M. Joerger, J. Von Pawel, S. Kraff, J. R. Fischer, W. Eberhardt, T. C. Gauler, L. Mueller, N. Reinmuth, M. Reck, M. Kimmich, F. Mayer, H. G. Kopp, D. M. Behringer, Y. D. Ko, R. A. Hilger, M. Roessler, C. Kloft, A. Henrich, B. Moritz, M. C. Miller, S. J. Salamone, U. Jaehde. Open-label, randomized study of individualized, pharmacokinetically (PK)-guided dosing of paclitaxel combined with carboplatin or cisplatin in patients with advanced non-small-cell lung cancer (NSCLC). Annals of Oncology 27: 1895–1902 (2016).

[34] M. A. Jordan, L. Wilson. Microtubules as a target for anticancer drugs. Nature Reviews Cancer 4: 253–265 (2004).

[35] O. van Tellingen, M. Huizing, V. Panday, J. Schellens, W. Nooijen, J. Beijnen. Cremophor EL causes (pseudo-) non-linear pharmacokinetics of paclitaxel in patients. British Journal of Cancer 81: 330–335 (1999).

[36] L. Gianni, C. M. Kearns, A. Giani, G. Capri, L. Viganó, A. Lacatelli, G. Bonadonna, M. J. Egorin. Nonlinear pharmacokinetics and metabolism of paclitaxel and its pharmacokinetic/pharmacodynamic relationships in humans. Journal of Clinical Oncology 13: 180–190 (1995).

[37] M. T. Huizing, G. Giaccone, L. J. van Warmerdam, H. Rosing, P. J. Bakker, J. B. Vermorken, P. E. Postmus, N. van Zandwijk, M. G. Koolen, W. W. ten Bokkel Huinink, W. J. van der Vijgh, F. J. Bierhorst, A. Lai, O. Dalesio, H. M. Pinedo, C. H. Veenhof, J. H. Beijnen. Pharmacokinetics of paclitaxel and carboplatin in a dose-escalating and dose-sequencing study in patients with non-small-cell lung cancer. The European Cancer Centre. Journal of Clinical Oncology 15: 317–329 (1997).

[38] M. Joerger, A. D. Huitema, D. J. Richel, C. Dittrich, N. Pavlidis, E. Briasoulis, J. B. Vermorken, E. Strocchi, A. Martoni, R. Sorio, H. P. Sleeboom, M. A. Izquierdo, D. I. Jodrell, H. Calvert, A. V. Boddy, H. Hollema, R. Fety, W. J. F. van der Vijgh, G. Hempel, E. Chatelut, M. Karlsson, J. Wilkins, B. Tranchand, A. H. G. J. Schrijvers, C. Twelves, J. H. Beijnen, J. H. M. Schellens. Population pharmacokinetics and pharmacodynamics of paclitaxel and carboplatin in ovarian cancer patients: A study by the European organization for research and treatment of cancer-pharmacology and molecular mechanisms group and new drug development group. Clinical Cancer Research 13: 6410–6418 (2007).

[39] A. A. Miller, G. L. Rosner, M. J. Egorin, D. Hollis, S. M. Lichtman, M. J. Ratain, Cancer and Leukemia Group B. Prospective evaluation of body surface area as a determinant of paclitaxel pharmacokinetics and pharmacodynamics in women with solid tumors: Cancer and Leukemia Group B Study 9763. Clinical Cancer Research 10: 8325–8331 (2004).

[40] T. Ohtsu, Y. Sasaki, T. Tamura, Y. Miyata, H. Nakanomyo, Y. Nishiwaki, N. Saijo. Clinical pharmacokinetics and pharmacodynamics of paclitaxel: A 3-hour infusion versus a 24-hour infusion. Clinical Cancer Research: An official journal of the American Association for Cancer Research 1: 599–606 (1995).

[41] M. Nakajima, Y. Fujiki, S. Kyo, T. Kanaya, M. Nakamura, Y. Maida, M. Tanaka, M. Inoue, T. Yokoi. Pharmacokinetics of paclitaxel in ovarian cancer patients and genetic polymorphisms of CYP2C8, CYP3A4, and MDR1. The Journal of Clinical Pharmacology 45: 674–682 (2005).

[42] D. R. Mould, G. F. Fleming, K. M. Darcy, D. Spriggs. Population analysis of a 24-h paclitaxel infusion in advanced endometrial cancer: A gynaecological oncology group study. British Journal of Clinical Pharmacology 62: 56–70 (2006).

[43] M. Kobayashi, K. Oba, J. Sakamoto, K. Kondo, N. Nagata, T. Okabayashi, T. Namikawa, K. Hanazaki. Pharmacokinetic study of weekly administration dose of paclitaxel in patients with advanced or recurrent gastric cancer in Japan. Gastric Cancer 10 (2007).

[44] M. Jiko, I. Yano, E. Sato, K. Takahashi, H. Motohashi, S. Masuda, M. Okuda, N. Ito, E. Nakamura, T. Segawa, T. Kamoto, O. Ogawa, K. Inui. Pharmacokinetics and pharmacodynamics of paclitaxel with carboplatin or gemcitabine, and effects of CYP3A5 and MDR1 polymorphisms in patients with urogenital cancers. International Journal of Clinical Oncology 12: 284–290 (2007).

[45] R. Yusuf, Z. Duan, D. Lamendola, R. Penson, M. Seiden. Paclitaxel resistance: Molecular mechanisms and pharmacologic manipulation. Current Cancer Drug Targets 3: 1–19 (2003).

[46] E. K.-H. Han, L. Gehrke, S. Tahir, R. Credo, S. Cherian, H. Sham, S. Rosenberg, S.-C. Ng. Modulation of drug resistance by α-tubulin in paclitaxel-resistant human lung cancer cell lines. European Journal of Cancer 36: 1565–1571 (2000).

[47] A. M. Barbuti, Z.-S. Chen. Paclitaxel through the ages of anticancer therapy: Exploring its role in chemoresistance and radiation therapy. Cancers 7: 2360–2371 (2015).

[48] S. Dasari, P. B. Tchounwou. Cisplatin in cancer therapy: Molecular mechanisms of action. European Journal of Pharmacology 740: 364–378 (2014).

[49] E. K. Rowinsky, M. Gilbert, W. McGuire, D. Noe, L. Grochow, A. Forastiere, D. Ettinger, B. Lubejko, B. Clark, S. Sartorius. Sequences of taxol and cisplatin: A phase I and pharmacologic study. Journal of Clinical Oncology 9: 1692–1703 (1991).

[50] Carboplatin: Summary of product characteristics (2011). `https://docetp.mpa.se/LMF/Carboplatin%20Pfizer%20concentrate%20for%20solution%20for%20infusion%20ENG%20SmPC_09001bee807a14c4.pdf`. (Accessed on: 2023-06-23).

[51] T. B. de Castria, E. M. da Silva, A. F. Gois, R. Riera. Cisplatin versus carboplatin in combination with third-generation drugs for advanced non-small cell lung cancer. Cochrane Database of Systematic Reviews (2013).

[52] V. F. Vasconcellos, G. N. Marta, E. M. da Silva, A. F. Gois, T. B. de Castria, R. Riera. Cisplatin versus carboplatin in combination with third-generation drugs for advanced non-small cell lung cancer. Cochrane Database of Systematic Reviews (2020).

[53] R. E. Sanborn. Cisplatin versus carboplatin in NSCLC: Is there one "best" answer? Current Treatment Options in Oncology 9: 326–342 (2008).

[54] R. Santana-Davila, A. Szabo, C. Arce-Lara, C. D. Williams, M. J. Kelley, J. Whittle. Cisplatin versus carboplatin-based regimens for the treatment of patients with metastatic lung cancer. An analysis of Veterans Health Administration data. Journal of Thoracic Oncology 9: 702–709 (2014).

[55] M. M. Oken, R. H. Creech, D. C. Tormey, J. Horton, T. E. Davis, E. T. McFadden, P. P. Carbone. Toxicity and response criteria of the Eastern Cooperative Oncology Group. American Journal of Clinical Oncology 5: 649–656 (1982).

[56] E. A. Eisenhauer, P. Therasse, J. Bogaerts, L. H. Schwartz, D. Sargent, R. Ford, J. Dancey, S. Arbuck, S. Gwyther, M. Mooney, L. Rubinstein, L. Shankar, L. Dodd, R. Kaplan, D. Lacombe, J. Verweij. New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). European Journal of Cancer 45: 228–247 (2009).

[57] Food and Drug Administration. Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics: Guidance for Industry (2018). `https://www.fda.gov/media/71195/download`. (Accessed on: 11.11.2022).

[58] Biomarkers Definitions Working Group. Biomarkers and surrogate endpoints: Preferred definitions and conceptual framework. Clinical Pharmacology & Therapeutics 69: 89–95 (2001).

[59] M. Šutić, A. Vukić, J. Baranašić, A. Försti, F. Džubur, M. Samaržija, M. Jakopović, L. Brčić, J. Knežević. Diagnostic, predictive, and prognostic biomarkers in non-small cell lung cancer (NSCLC) management. Journal of Personalized Medicine 11: 1102 (2021).

[60] M. Paesmans. Prognostic and predictive factors for lung cancer. Breathe 9: 112–121 (2012).

[61] M. D. Brundage, D. Davies, W. J. Mackillop. Prognostic factors in non-small cell lung cancer: A decade of progress. Chest 122: 1037–1057 (2002).

[62] I. Trulson, S. Holdenrieder. Prognostic value of blood-based protein biomarkers in non-small cell lung cancer: A critical review and 2008–2022 update. Tumor Biology 1–51 (2023).

[63] T. Deng, J. Zhang, Y. Meng, Y. Zhou, W. Li. Higher pretreatment lactate dehydrogenase concentration predicts worse overall survival in patients with lung cancer. Medicine 97 (2018).

[64] M. Sung, W. S. Jang, H. R. Kim, J. A. Park, S. M. Lim, H. R. Kim, B. C. Cho, Y. R. Park, M. H. Hong. Prognostic value of baseline and early treatment response of neutrophil-lymphocyte ratio, C-reactive protein, and lactate dehydrogenase in non-small cell lung cancer patients undergoing immunotherapy. Translational Lung Cancer Research 12: 1506 (2023).

[65] A. Tjokrowidjaja, S. J. Lord, T. John, C. R. Lewis, P.-S. Kok, I. C. Marschner, C. K. Lee. Pre-and on-treatment lactate dehydrogenase as a prognostic and predictive biomarker in advanced non–small cell lung cancer. Cancer 128: 1574–1583 (2022).

[66] C. Liao, Z. Yu, W. Guo, Q. Liu, Y. Wu, Y. Li, L. Bai. Prognostic value of circulating inflammatory factors in non-small cell lung cancer: A systematic review and meta-analysis. Cancer Biomarkers 14: 469–481 (2014).

[67] F. W. Ojara, A. Henrich, N. Frances, Y. M. Nassar, W. Huisinga, N. Hartung, K. Geiger, S. Holdenrieder, M. Joerger, C. Kloft. A prognostic baseline blood biomarker and tumor growth kinetics integrated model in paclitaxel/platinum treated advanced non-small cell lung cancer patients. CPT: Pharmacometrics & Systems Pharmacology 12: 1714–1725 (2023).

[68] X. Xiao, S. Wang, G. Long. C-reactive protein is a significant predictor of improved survival in patients with advanced non-small cell lung cancer. Medicine 98 (2019).

[69] P. Chan, M. Marchand, K. Yoshida, S. Vadhavkar, N. Wang, A. Lin, B. Wu, M. Ballinger, N. Sternheim, J. Y. Jin, R. Bruno. Prediction of overall survival in patients across solid tumors following atezolizumab treatments: A tumor growth inhibition–overall survival modeling framework. CPT: Pharmacometrics & Systems Pharmacology 10: 1171–1182 (2021).

[70] H. Scott, D. McMillan, L. Forrest, D. Brown, C. McArdle, R. Milroy. The systemic inflammatory response, weight loss, performance status and survival in patients with inoperable non-small cell lung cancer. British Journal of Cancer 87: 264–267 (2002).

[71] A. Koch, H. Fohlin, S. Sörenson. Prognostic significance of C-reactive protein and smoking in patients with advanced non-small cell lung cancer treated with first-line palliative chemotherapy. Journal of Thoracic Oncology 4: 326–332 (2009).

[72] B. Gagnon, M. Abrahamowicz, Y. Xiao, M. Beauchamp, N. MacDonald, G. Kasymjanova, H. Kreisman, D. Small. Flexible modeling improves assessment of prognostic value of C-reactive protein in advanced non-small cell lung cancer. British Journal of Cancer 102: 1113–1122 (2010).

[73] G. Soussi, N. B. Alaya, N. Chaouch, H. Racil. Development and validation of a prognostic index for survival in non-small cell lung cancer: Results from a Tunisian cohort study. Cancer Epidemiology 53: 111–118 (2018).

[74] D. Machado, C. Marques, M. Dias, S. Campainha, A. Barroso. Inflammatory prognostic biomarkers in advanced non-small cell lung cancer. Pulmonology 25: 181–183 (2019).

[75] B. Agoram, A. C. Heatherington, M. R. Gastonguay. Development and evaluation of a population pharmacokinetic-pharmacodynamic model of darbepoetin alfa in patients with nonmyeloid malignancies undergoing multicycle chemotherapy. The AAPS Journal 8: E552–E563 (2006).

[76] Y. Seo, W. Eo, S. Kim, B. Shim, S. Lee. Can nutritional status predict overall survival in patients with advanced non-small cell lung cancer? Nutrition and Cancer 71: 1108–1117 (2019).

[77] O. Fiala, M. Pesek, J. Finek, O. Topolcan, J. Racek, M. Minarik, L. Benesova, Z. Bortlicek, A. Poprach, T. Buchler. High serum level of C-reactive protein is associated with worse outcome of patients with advanced-stage NSCLC treated with erlotinib. Tumor Biology 36: 9215–9222 (2015).

[78] L. Zhu, X. Li, Y. Shen, Y. Cao, X. Fang, J. Chen, Y. Yuan. A new prognostic score based on the systemic inflammatory response in patients with inoperable non-small-cell lung cancer. OncoTargets and Therapy 4879–4886 (2016).

[79] O. Fiala, P. Hosek, M. Pesek, J. Finek, J. Racek, T. Buchler, A. Poprach, K. Hejduk, R. Chloupková, O. Sorejs, M. Ecksteinova, M. Vitovec, K. Cizkova, R. Kucera, O. Topolcan. Prognostic role of serum C-reactive protein in patients with advanced-stage NSCLC treated with pemetrexed. Neoplasma 64: 605–610 (2017).

[80] J. Tanizaki, K. Haratani, H. Hayashi, Y. Chiba, Y. Nakamura, K. Yonesaka, K. Kudo, H. Kaneda, Y. Hasegawa, K. Tanaka, M. Takeda, A. Ito, K. Nakagawa. Peripheral blood biomarkers associated with clinical outcome in non-small cell lung cancer patients treated with nivolumab. Journal of Thoracic Oncology 13: 97–105 (2018).

[81] R. Cehreli, T. Yavuzsen, H. Ates, T. Akman, H. Ellidokuz, I. Oztop. Can inflammatory and nutritional serum markers predict chemotherapy outcomes and survival in advanced stage nonsmall cell lung cancer patients? BioMed Research International 2019 (2019).

[82] A. M. Hopkins, G. Kichenadasse, E. Garrett-Mayer, C. S. Karapetis, A. Rowland, M. J. Sorich. Development and validation of a prognostic model for patients with advanced lung cancer treated with the immune checkpoint inhibitor atezolizumab. Clinical Cancer Research 26: 3280–3286 (2020).

[83] M. J. Fidler, C. L. Fhied, J. Roder, S. Basu, S. Sayidine, I. Fughhi, M. Pool, M. Batus, P. Bonomi, J. A. Borgia. The serum-based VeriStrat® test is associated with proinflammatory reactants and clinical outcome in non-small cell lung cancer patients. BMC Cancer 18: 1–9 (2018).

[84] J. U. Lim, C. D. Yeo, H. S. Kang, C. K. Park, J. S. Kim, J. W. Kim, S. J. Kim, S. H. Lee. Elevated pretreatment platelet-to-lymphocyte ratio is associated with poor survival in stage IV non-small cell lung cancer with malignant pleural effusion. Scientific Reports 9: 4721 (2019).

[85] S. Iivanainen, J. Ahvonen, A. Knuuttila, S. Tiainen, J. P. Koivunen. Elevated CRP levels indicate poor progression-free and overall survival on cancer patients treated with PD-1 inhibitors. ESMO Open 4: e000531 (2019).

[86] M. Kogo, T. Sunaga, S. Nakamura, T. Akita, T. Kurihara, Y. Shikama, H. Nakajima, T. Tobe, K. Yoneyama, Y. Kiuchi. Prognostic index for survival in patients with advanced non-small-cell lung cancer treated with third-generation agents. Chemotherapy 62: 239–245 (2017).

[87] Y. Adachi, A. Tamiya, Y. Taniguchi, T. Enomoto, K. Azuma, S. Kouno, Y. Inagaki, N. Saijo, K. Okishio, S. Atagi. Predictive factors for progression-free survival in non-small cell lung cancer patients receiving nivolumab based on performance status. Cancer Medicine 9: 1383–1391 (2020).

[88] Y. Jin, Y. Sun, X. Shi, J. Zhao, L. Shi, X. Yu. Prognostic value of circulating C-reactive protein levels in patients with non-small cell lung cancer: A systematic review with meta-analysis. Journal of Cancer Research and Therapeutics 10: C160–C166 (2014).

[89] S. Shrotriya, D. Walsh, N. Bennani-Baiti, S. Thomas, C. Lorton. C-reactive protein is an important biomarker for prognosis tumor recurrence and treatment response in adult solid tumors: A systematic review. PloS One 10: e0143080 (2015).

[90] S. Badovinac, M. Korsic, D. Mursic, M. Samarzija, B. Cucevic, M. Roglic, M. Jakopovic. Cancer-related inflammation as predicting tool for treatment outcome in locally advanced and metastatic non-small cell lung cancer. Journal of Thoracic Disease 8: 1497 (2016).

[91] C.-L. Han, G.-X. Meng, Z.-N. Ding, Z.-R. Dong, Z.-Q. Chen, J.-G. Hong, L.-J. Yan, H. Liu, B.-W. Tian, L.-S. Yang, J. S. Xue, T. Li. The predictive potential of the baseline C-reactive protein levels for the efficiency of immune checkpoint inhibitors in cancer patients: A systematic review and meta-analysis. Frontiers in Immunology 13: 827788 (2022).

[92] P. C. Hart, I. M. Rajab, M. Alebraheem, L. A. Potempa. C-reactive protein and cancer—diagnostic and therapeutic insights. Frontiers in Immunology 11: 595835 (2020).

[93] H. Aref, S. Refaat. CRP evaluation in non-small cell lung cancer. Egyptian Journal of Chest Diseases and Tuberculosis 63: 717–722 (2014).

[94] M. B. Pepys, G. M. Hirschfield. C-reactive protein: A critical update. The Journal of Clinical Investigation 111: 1805–1812 (2003).

[95] W. Ansar, S. Ghosh. C-reactive protein and the biology of disease. Immunologic Research 56: 131–142 (2013).

[96] Food and Drug Administration. Review Criteria for Assessment of C-Reactive Protein (CRP), High Sensitivity C-Reactive Protein (hsCRP) and Cardiac C-Reactive Protein (cCRP) Assays (2005). `https://www.fda.gov/media/71337/download`. (Accessed on: 26.05.2023).

[97] D. M. Vigushin, M. B. Pepys, P. N. Hawkins. Metabolic and scintigraphic studies of radioiodinated human C-reactive protein in health and disease. The Journal of Clinical Investigation 91: 1351–1357 (1993).

[98] H. K. Meier-Ewert, P. M. Ridker, N. Rifai, N. Price, D. F. Dinges, J. M. Mullington. Absence of diurnal variation of C-reactive protein concentrations in healthy human subjects. Clinical Chemistry 47: 426–430 (2001).

[99] G. Nilsonne, M. Lekander, T. Åkerstedt, J. Axelsson, M. Ingre. Diurnal variation of circulating interleukin-6 in humans: A meta-analysis. PloS one 11: e0165799 (2016).

[100] L. A. Potempa, I. M. Rajab, M. E. Olson, P. C. Hart. C-reactive protein and cancer: Interpreting the differential bioactivities of its pentameric and monomeric, modified isoforms. Frontiers in Immunology 12: 744129 (2021).

[101] K. Heikkilä, S. Ebrahim, D. A. Lawlor. A systematic review of the association between circulating concentrations of C reactive protein and cancer. Journal of Epidemiology & Community Health 61: 824–833 (2007).

[102] N. Katsumata, K. Eguchi, M. Fukuda, N. Yamamoto, Y. Ohe, F. Oshita, T. Tamura, T. Shinkai, N. Saijo. Serum levels of cytokines in patients with untreated primary lung cancer. Clinical Cancer Research: An official journal of the American Association for Cancer Research 2: 553–559 (1996).

[103] T. Matsuguchi, S. Okamura, C. Kawasaki, K. Shimoda, F. Omori, S. Hayashi, N. Kimura, Y. Niho. Constitutive production of granulocyte colony-stimulating factor and interleukin-6 by a human lung cancer cell line, KSNY: gene amplification and increased mRNA stability. European Journal of Haematology 47: 128–133 (1991).

[104] X. Jing, C. Huang, H. Zhou, C. Li, L. Fan, J. Chen, G. Zhang, Y. Liu, Z. Cui, D. Qi, J. Ma. Association between serum C-reactive protein value and prognosis of patients with non-small cell lung cancer: A meta-analysis. International Journal of Clinical and Experimental Medicine 8: 10633 (2015).

[105] S. Holdenrieder, D. Nagel, P. Stieber. Estimation of prognosis by circulating biomarkers in patients with non-small cell lung cancer. Cancer Biomarkers 6: 179–190 (2010).

[106] P. J. Williams, E. I. Ette. Pharmacometrics: Impacting drug development and pharmacotherapy. In: Pharmacometrics: The science of quantitative pharmacology. Wiley Online Library. 1–21 (2007).

[107] D. R. Mould, R. Upton. Basic concepts in population modeling, simulation, and model-based drug development. CPT: Pharmacometrics & Systems Pharmacology 1: 1–14 (2012).

[108] R. Madabushi, P. Seo, L. Zhao, M. Tegenge, H. Zhu. Role of model-informed drug development approaches in the lifecycle of drug development and regulatory decision-making. Pharmaceutical Research 39: 1669–1680 (2022).

[109] J. Y. Lee, C. E. Garnett, J. V. Gobburu, V. A. Bhattaram, S. Brar, J. C. Earp, P. R. Jadhav, K. Krudys, L. J. Lesko, F. Li, J. Liu, R. Madabushi, A. Marathe, N. Mehrotra, C. Tornoe, Y. Wang, H. Zhu. Impact of pharmacometric analyses on new drug approval and labelling decisions: A review of 198 submissions between 2000 and 2008. Clinical Pharmacokinetics 50: 627–635 (2011).

[110] L.-F. Hsu. A Survey of Population Pharmacokinetic Reports Submitted to the USFDA: An Analysis of Common Issues in NDA and BLA from 2012 to 2021. Clinical Pharmacokinetics 61: 1697–1703 (2022).

[111] Food and Drug Administration. Population Pharmacokinetics: Guidance for Industry (2022). `https://www.fda.gov/media/128793/download`. (Accessed on: 2023-10-09).

[112] European Medicines Agency. Guideline on reporting the results of population pharmacokinetic analyses (2007). `https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-reporting-results-population-pharmacokinetic-analyses_en.pdf`. (Accessed on: 2023-10-09).

[113] L. B. Sheiner. The population approach to pharmacokinetic data analysis: Rationale and standard data analysis methods. Drug Metabolism Reviews 15: 153–171 (1984).

[114] L. B. Sheiner, S. L. Beal. Evaluation of methods for estimating population pharmacokinetic parameters. I. Michaelis-Menten model: Routine clinical pharmacokinetic data. Journal of Pharmacokinetics and Biopharmaceutics 8: 553–571 (1980).

[115] E. I. Ette, P. J. Williams. Population pharmacokinetics II: Estimation methods. Annals of Pharmacotherapy 38: 1907–1915 (2004).

[116] L. B. Sheiner, S. L. Beal. Evaluation of methods for estimating population pharmacokinetic parameters. III. Monoexponential model: Routine clinical pharmacokinetic data. Journal of Pharmacokinetics and Biopharmaceutics 11: 303–319 (1983).

[117] S. S. Shord, H. Zhu, J. Liu, A. Rahman, B. Booth, I. Zineh. US Food and Drug Administration embraces using innovation to identify optimized dosages for patients with cancer. CPT: Pharmacometrics & Systems Pharmacology 12: 1573 (2023).

[118] K. Venkatakrishnan, L. E. Friberg, D. Ouellet, J. Mettetal, A. Stein, I. Trocóniz, R. Bruno, N. Mehrotra, J. Gobburu, D. Mould. Optimizing oncology therapeutics through quantitative translational and clinical pharmacology: Challenges and opportunities. Clinical Pharmacology & Therapeutics 97: 37–54 (2015).

[119] A. Ruiz-Garcia, P. Baverel, D. Bottino, M. Dolton, Y. Feng, I. González-García, J. Kim, S. Robey, I. Singh, D. Turner, S. P. Wu, D. Yin, D. Zhou, H. Zhu, P. Bonate. A comprehensive regulatory and

industry review of modeling and simulation practices in oncology clinical drug development. Journal of Pharmacokinetics and Pharmacodynamics 50: 147–172 (2023).

[120] X. Su, Y. Li, P. Müller, C.-W. Hsu, H. Pan, K.-A. Do. A semi-mechanistic dose-finding design in oncology using pharmacokinetic/pharmacodynamic modeling. Pharmaceutical Statistics 21: 1149–1166 (2022).

[121] A. S. Zandvliet, J. H. Schellens, J. H. Beijnen, A. D. Huitema. Population pharmacokinetics and pharmacodynamics for treatment optimization in clinical oncology. Clinical Pharmacokinetics 47: 487–513 (2008).

[122] Food and Drug Administration. Drug Development Tools: Fit-for-Purpose Initiative (2022). `https://www.fda.gov/drugs/development-approval-process-drugs/drug-development-tools-fit-purpose-initiative`. (Accessed on: 12.12.2023).

[123] F. Bretz, J. C. Pinheiro, M. Branson. Combining multiple comparisons and modeling techniques in dose-response studies. Biometrics 61: 738–748 (2005).

[124] J. Pinheiro, B. Bornkamp, E. Glimm, F. Bretz. Model-based dose finding under model uncertainty using general parametric models. Statistics in Medicine 33: 1646–1661 (2014).

[125] Food and Drug Administration. FDA Qualification of MCP-Mod Method (2015). `https://www.fda.gov/media/99313/download`. (Accessed on: 07.12.2023).

[126] S. Buatois, S. Ueckert, N. Frey, S. Retout, F. Mentré. cLRT-Mod: An efficient methodology for pharmacometric model-based analysis of longitudinal Phase II dose finding studies under model uncertainty. Statistics in Medicine 40: 2435–2451 (2021).

[127] H. Dette, S. Titoff, S. Volgushev, F. Bretz. Dose response signal detection under model uncertainty. Biometrics 71: 996–1008 (2015).

[128] A. Yin, D. J. A. Moes, J. G. van Hasselt, J. J. Swen, H.-J. Guchelaar. A review of mathematical models for tumor dynamics and treatment resistance evolution of solid tumors. CPT: Pharmacometrics & Systems Pharmacology 8: 720–737 (2019).

[129] L. Claret, P. Girard, P. M. Hoff, E. van Cutsem, K. P. Zuideveld, K. Jorga, J. Fagerberg, R. Bruno. Model-based prediction of phase III overall survival in colorectal cancer on the basis of phase II tumor dynamics. Journal of Clinical Oncology 27: 4103–4108 (2009).

[130] A. Bernard, H. Kimko, D. Mital, I. Poggesi. Mathematical modeling of tumor growth and tumor growth inhibition in oncology drug development. Expert Opinion on Drug Metabolism & Toxicology 8: 1057–1069 (2012).

[131] B. Ribba, N. H. Holford, P. Magni, I. Trocóniz, I. Gueorguieva, P. Girard, C. Sarr, M. Elishmereni, C. Kloft, L. E. Friberg. A review of mixed-effects models of tumor growth and effects of anticancer drug treatment used in population analysis. CPT: Pharmacometrics & Systems Pharmacology 3: 1–10 (2014).

[132] A. A. Suleiman, L. Nogova, U. Fuhr. Modeling NSCLC progression: Recent advances and opportunities available. The AAPS Journal 15: 542–550 (2013).

[133] D. Mould, A.-C. Walz, T. Lave, J. Gibbs, B. Frame. Developing exposure/response models for anticancer drug treatment: Special considerations. CPT: Pharmacometrics & Systems Pharmacology 4: 12–27 (2015).

[134] N. L. Dayneka, V. Garg, W. J. Jusko. Comparison of four basic models of indirect pharmacodynamic responses. Journal of Pharmacokinetics and Biopharmaceutics 21: 457–478 (1993).

[135] R. Upton, D. Mould. Basic concepts in population modeling, simulation, and model-based drug development: Part 3—introduction to pharmacodynamic modeling methods. CPT: Pharmacometrics & Systems Pharmacology 3: 1–16 (2014).

[136] A. Sharma, W. J. Jusko. Characteristics of indirect pharmacodynamic models and applications to clinical drug responses. British Journal of Clinical Pharmacology 45: 229–239 (1998).

[137] B. C. Bender, E. Schindler, L. E. Friberg. Population pharmacokinetic–pharmacodynamic modelling in oncology: A tool for predicting clinical response. British Journal of Clinical Pharmacology 79: 56–71 (2015).

[138] M. Bergstrand, M. O. Karlsson. Handling data below the limit of quantification in mixed effect models. The AAPS Journal 11: 371–380 (2009).

[139] S. L. Beal. Ways to fit a PK model with some data below the quantification limit. Journal of Pharmacokinetics and Pharmacodynamics 28: 481–504 (2001).

[140] D. R. Mould, R. N. Upton. Basic concepts in population modeling, simulation, and model-based drug development—Part 2: Introduction to pharmacokinetic modeling methods. CPT: Pharmacometrics & Systems Pharmacology 2: 1–14 (2013).

[141] P. L. Bonate. Nonlinear Mixed Effects Models: Theory. In: Pharmacokinetic-pharmacodynamic modeling and simulation. Springer, New York. 233–301 (2011).

[142] J. S. Owen, J. Fiedler-Kelly. Model Building. In: Introduction to Population Pharmacokinetic/Pharmacodynamic Analysis with Nonlinear Mixed Effects Models. Wiley Online Books. 90–177 (2014).

[143] Pharmacometrics and PKPD Research Groups. Time-varying covariates. Department of Pharmacy, Uppsala University. Uppsala, Sweden (2022).

[144] U. Wählby, A. H. Thomson, P. A. Milligan, M. O. Karlsson. Models for time-varying covariates in population pharmacokinetic-pharmacodynamic analysis. British Journal of Clinical Pharmacology 58: 367–377 (2004).

[145] E. N. Jonsson, M. O. Karlsson. Automated covariate model building within NONMEM. Pharmaceutical Research 15: 1463–1468 (1998).

[146] PsN. SCM user guide (2019). `https://uupharmacometrics.github.io/PsN/docs.html`. (Accessed on: 2023-03-20).

[147] R. J. Svensson, E. N. Jonsson. Efficient and relevant stepwise covariate model building for pharmacometrics. CPT: Pharmacometrics & Systems Pharmacology 11: 1210–1222 (2022).

[148] M. R. Gastonguay. A full model estimation approach for covariate effects: Inference based on clinical importance and estimation precision. AAPS J 6: W4354 (2004).

[149] M. O. Karlsson. A full model approach based on the covariance matrix of parameters and covariates. In: 21^{st} Populationn Approach Group Europe (PAGE) Meeting, Abstract 2455. Venice, Italy (2012). `www.page-meeting.org/?abstract=2455`.

[150] G. Yngman, H. Bjugård Nyberg, J. Nyberg, E. N. Jonsson, M. O. Karlsson. An introduction to the full random effects model. CPT: Pharmacometrics & Systems Pharmacology 11: 149–160 (2022).

[151] X. S. Xu, M. Yuan, H. Zhu, Y. Yang, H. Wang, H. Zhou, J. Xu, L. Zhang, J. Pinheiro. Full covariate modelling approach in population pharmacokinetics: Understanding the underlying hypothesis tests and implications of multiplicity. British Journal of Clinical Pharmacology 84: 1525–1534 (2018).

[152] J. Ribbing, J. Nyberg, O. Caster, E. N. Jonsson. The lasso—a novel method for predictive covariate model building in nonlinear mixed effects models. Journal of Pharmacokinetics and Pharmacodynamics 34: 485–517 (2007).

[153] E. Sibieude, A. Khandelwal, J. S. Hesthaven, P. Girard, N. Terranova. Fast screening of covariates in population models empowered by machine learning. Journal of Pharmacokinetics and Pharmacodynamics 48: 597–609 (2021).

[154] E. Sibieude, A. Khandelwal, P. Girard, J. S. Hesthaven, N. Terranova. Population pharmacokinetic model selection assisted by machine learning. Journal of Pharmacokinetics and Pharmacodynamics 49: 257–270 (2022).

[155] D. Fisher, S. Shafer. Fisher/Shafer NONMEM Workshop Pharmacokinetic and Pharmacodynamic Analysis with NONMEM. 7–11 March, Ghent, Belgium (2007).

[156] P. L. Bonate. Nonlinear Mixed Effects Models: Theory. In: Pharmacokinetic-pharmacodynamic modeling and simulation. Springer, New York. 205–266 (2006).

[157] R. M. Savic, M. C. Kjellsson, M. O. Karlsson. Evaluation of the nonparametric estimation method in NONMEM VI. European Journal of Pharmaceutical Sciences 37: 27–35 (2009).

[158] C. Dansirikul, H. E. Silber, M. O. Karlsson. Approaches to handling pharmacodynamic baseline responses. Journal of Pharmacokinetics and Pharmacodynamics 35: 269–283 (2008).

[159] P. L. Bonate. The art of modeling. In: Pharmacokinetic-pharmacodynamic modeling and simulation. Springer, New York. 1–60 (2011).

[160] J. S. Owen, J. Fiedler-Kelly. Introduction to model evaluation. In: Introduction to Population Pharmacokinetic/Pharmacodynamic Analysis with Nonlinear Mixed Effects Models. Wiley Online Books. 212–231 (2014).

[161] A.-G. Dosne, M. Bergstrand, K. Harling, M. O. Karlsson. Improving the estimation of parameter uncertainty distributions in nonlinear mixed effects models using sampling importance resampling. Journal of Pharmacokinetics and Pharmacodynamics 43: 583–596 (2016).

[162] R. M. Savic, M. O. Karlsson. Importance of shrinkage in empirical bayes estimates for diagnostics: Problems and solutions. The AAPS Journal 11: 558–569 (2009).

[163] A. C. Hooker, C. E. Staatz, M. O. Karlsson. Conditional Weighted Residuals (CWRES): A Model Diagnostic for the FOCE Method. Pharmaceutical Research 24: 2187–2197 (2007).

[164] E. N. Jonsson, J. Nyberg. A quantitative approach to the choice of number of samples for percentile estimation in bootstrap and visual predictive check analyses. CPT: Pharmacometrics & Systems Pharmacology 11: 673–686 (2022).

[165] M. Karlsson, N. Holford. A Tutorial on Visual Predictive Checks. In: 17^{th} Population Approach Group Europe (PAGE) Meeting, Abstract 1434. Marseille, France (2008). `www.page-meeting.org/?abstract=1434`.

[166] J. S. Owen, J. Fiedler-Kelly. Simulation basics. In: Introduction to Population Pharmacokinetic/Pharmacodynamic Analysis with Nonlinear Mixed Effects Models. Wiley Online Books. 265–284 (2014).

[167] T. Dey, S. R. Lipsitz, Z. Cooper, Q.-D. Trinh, M. Krzywinski, N. Altman. Survival analysis-time-to-event data and censoring. Nature Methods 19: 906–908 (2022).

[168] S. Prinja, N. Gupta, R. Verma. Censoring in clinical trials: Review of survival analysis techniques. Indian Journal of Community Medicine: Official publication of Indian Association of Preventive & Social Medicine 35: 217 (2010).

[169] D. Collett. Some non-parametric procedures. In: Modelling survival data in medical research. CRC press. 15–53 (2003).

[170] M. K. Goel, P. Khanna, J. Kishore. Understanding survival analysis: Kaplan-Meier estimate. International Journal of Ayurveda Research 1: 274 (2010).

[171] Cox Proportional-Hazards Model. `http://www.sthda.com/english/wiki/cox-proportional-hazards-model`. (Accessed on: 2023-03-07).

[172] V. Bewick, L. Cheek, J. Ball. Statistics Review 12: Survival analysis. Critical Care 8: 1–6 (2004).

[173] Cox Model Assumptions. http://www.sthda.com/english/wiki/cox-model-assumptions. (Accessed on: 2023-03-07).

[174] R. Tibshirani. A plain man's guide to the proportional hazards model. Clinical and Investigative Medicine. Medecine Clinique et Experimentale 5: 63–68 (1982).

[175] R. C. Van Wijk, U. S. Simonsson. Finding the right hazard function for time-to-event modeling: A tutorial and Shiny application. CPT: Pharmacometrics & Systems Pharmacology 11: 991–1001 (2022).

[176] C. J. Morgan. Landmark analysis: A primer. Journal of Nuclear Cardiology 26: 391–393 (2019).

[177] U. Dafni. Landmark analysis at the 25-year landmark point. Circulation: Cardiovascular Quality and Outcomes 4: 363–371 (2011).

[178] J. M. Bland, D. G. Altman. The logrank test. BMJ 328: 1073 (2004).

[179] H. Wickham, R. François, L. Henry, K. Müller. Dplyr: A Grammar of Data Manipulation (2021). https://CRAN.R-project.org/package=dplyr.

[180] H. Wickham. The Split-Apply-Combine Strategy for Data Analysis. Journal of Statistical Software 40: 1–29 (2011).

[181] A. Zeileis, G. Grothendieck. Zoo: S3 Infrastructure for Regular and Irregular Time Series. Journal of Statistical Software 14: 1–27 (2005).

[182] H. Wickham. Reshaping Data with the reshape Package. Journal of Statistical Software 21: 1–20 (2007).

[183] K. Müller, H. Wickham. Tibble: Simple Data Frames (2021). https://CRAN.R-project.org/package=tibble.

[184] H. Wickham, L. Henry. Tidyr: Tidy Messy Data (2020). https://CRAN.R-project.org/package=tidyr.

[185] H. Wickham. Stringr: Simple, Consistent Wrappers for Common String Operations (2022). https://CRAN.R-project.org/package=stringr.

[186] S. M. Bache, H. Wickham. Magrittr: A Forward-Pipe Operator for R (2020). https://CRAN.R-project.org/package=magrittr.

[187] H. Wickham. Ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York (2016).

[188] E. N. Jonsson, M. O. Karlsson. Xpose–an S-PLUS based population pharmacokinetic/pharmacodynamic model building aid for NONMEM. Computer Methods and Programs in Biomedicine 58: 51–64 (1999).

[189] A. Kassambara. Ggpubr: 'ggplot2' Based Publication Ready Plots (2020). https://CRAN.R-project.org/package=ggpubr.

[190] T. L. Pedersen. Ggforce: Accelerating 'ggplot2' (2019). https://CRAN.R-project.org/package=ggforce.

[191] H. Wickham, L. Henry, D. Vaughan. Vctrs: Vector Helpers (2023). https://CRAN.R-project.org/package=vctrs.

[192] T. Wei, V. Simko. R package "corrplot": Visualization of a Correlation Matrix (2017). https://github.com/taiyun/corrplot.

[193] R. Keizer. Vpc: Create Visual Predictive Checks (2021). https://CRAN.R-project.org/package=vpc.

[194] R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria (2019). https://www.R-project.org/.

[195] C. O. Wilke. Ggtext: Improved Text Rendering Support for 'ggplot2' (2020). https://CRAN.R-project.org/package=ggtext.

[196] B. Bornkamp. DoseFinding: Planning and Analyzing Dose Finding Experiments (2019). https://CRAN.R-project.org/package=DoseFinding.

[197] A. Kassambara, M. Kosinski, P. Biecek. Survminer: Drawing Survival Curves using 'ggplot2' (2020). https://CRAN.R-project.org/package=survminer.

[198] T. M. Therneau. A Package for Survival Analysis in R (2021). https://CRAN.R-project.org/package=survival.

[199] T. M. Therneau, P. M. Grambsch. Modeling Survival Data: Extending the Cox Model. Springer, New York (2000).

[200] PsN. Bootstrap user guide (2016). https://uupharmacometrics.github.io/PsN/docs.html. (Accessed on: 2023-03-20).

[201] PsN. SSE user guide (2018). https://uupharmacometrics.github.io/PsN/docs.html. (Accessed on: 2023-05-15).

[202] PsN. VPC and NPC user guide (2018). https://uupharmacometrics.github.io/PsN/docs.html. (Accessed on: 2023-03-20).

[203] S. M. Krishnan, L. E. Friberg, J. Nyberg, C. Sarr, R. Niebecker, S. Wang, A. Pérez Pitarch. A framework for prediction of progression free survival based on modelling of sub-endpoints. In: 30^{th} Population Approach Group Europe (PAGE) Meeting, Abstract 10071. Ljubljana, Slovenia (2022). www.page-meeting.org/?abstract=10071.

[204] Y. M. Nassar, F. W. Ojara, A. Pérez-Pitarch, K. Geiger, W. Huisinga, N. Hartung, R. Michelet, S. Holdenrieder, M. Joerger, C. Kloft. C-Reactive Protein as an Early Predictor of Efficacy in Advanced Non-Small-Cell Lung Cancer Patients: A Tumor Dynamics-Biomarker Modeling Framework. Cancers 15: 5429 (2023).

[205] US Department of Health and Human Services. Common terminology criteria for adverse events (CTCAE) version 4.0. National Institutes of Health, National Cancer Institute (2009).

[206] A. Henrich. Pharmacometric modelling and simulation to optimise paclitaxel combination therapy based on pharmacokinetics, cumulative neutropenia and efficacy. Ph.D. thesis. Freie Universität Berlin (2017).

[207] M. Joerger, S. Kraff, A. D. Huitema, G. Feiss, B. Moritz, J. H. Schellens, J. H. Beijnen, U. Jaehde. Evaluation of a pharmacology-driven dosing algorithm of 3-weekly paclitaxel using therapeutic drug monitoring: A pharmacokinetic-pharmacodynamic simulation study. Clinical Pharmacokinetics 51: 607–617 (2012).

[208] L. Z. López, A. A. Pastor, J. M. A. Beitia, J. A. Velilla, J. G. Deiró. Determination of docetaxel and paclitaxel in human plasma by high-performance liquid chromatography: Validation and application to clinical pharmacokinetic studies. Therapeutic Drug Monitoring 28: 199–205 (2006).

[209] Bundesärtztekammer. Richtlinie der Bundesärztekammer zur Qualitätssicherung laboratoriumsmedizinischer Untersuchungen (2023). https://www.bundesaerztekammer.de/fileadmin/user_upload/BAEK/Themen/Qualitaetssicherung/_Bek_BAEK_RiLi_BAEK_ONLINE_FINAL_VERS_26_05_2023.pdf. (Accessed on: 2024-05-06).

[210] A. Henrich, M. Joerger, S. Kraff, U. Jaehde, W. Huisinga, C. Kloft, Z. P. Parra-Guillen. Semimechanistic bone marrow exhaustion pharmacokinetic/pharmacodynamic model for chemotherapy-induced cumulative neutropenia. Journal of Pharmacology and Experimental Therapeutics 362: 347–358 (2017).

[211] F. W. Ojara. Pharmacometric-based characterisation of peripheral neuropathy, tumour growth kinetics and overall survival to identify predictors of toxicity and efficacy to optimise paclitaxel/platinum chemotherapy. Ph.D. thesis. Freie Universität Berlin (2022).

[212] F. W. Ojara, A. Henrich, N. Hartung, M. Joerger, M. Roessler, J. Von Pawel, W. Huisinga, C. Kloft. Examining the relationship between paclitaxel exposure and peripheral neuropathy in non-small cell lung cancer. In: 26^{th} Population Approach Group Europe (PAGE) Meeting, Abstract 7278. Budapest, Hungary (2017). www.page-meeting.org/?abstract=7278.

[213] D. J. Irby, M. E. Ibrahim, A. M. Dauki, M. A. Badawi, S. M. Illamola, M. Chen, Y. Wang, X. Liu, M. A. Phelps, D. R. Mould. Approaches to handling missing or "problematic" pharmacology data: Pharmacokinetics. CPT: Pharmacometrics & Systems Pharmacology 10: 291–308 (2021).

[214] Å. M. Johansson, M. O. Karlsson. Multiple imputation of missing covariates in NONMEM and evaluation of the method's sensitivity to η-shrinkage. The AAPS Journal 15: 1035–1042 (2013).

[215] F. W. Ojara, A. Henrich, N. Frances, W. Huisinga, N. Hartung, M. Joerger, C. Kloft. Time-to-event analysis of paclitaxel-associated peripheral neuropathy in advanced non–small-cell lung cancer highlighting key influential treatment/patient factors. Journal of Pharmacology and Experimental Therapeutics 375: 430–438 (2020).

[216] E. M. Svensson, R. J. Svensson, L. H. Te Brake, M. J. Boeree, N. Heinrich, S. Konsten, G. Churchyard, R. Dawson, A. H. Diacon, G. S. Kibiki, L. T. Minja, N. E. Ntingiya, I. Sanne, S. H. Gillespie, M. Hoelscher, P. P. J. Phillips, U. S. H. Simonsson, R. Aarnoutse. The potential for treatment shortening with higher rifampicin doses: Relating drug exposure to treatment response in patients with pulmonary tuberculosis. Clinical Infectious Diseases 67: 34–41 (2018).

[217] D. B. Rubin. Multiple imputation for nonresponse in surveys. John Wiley & Sons (1987).

[218] B. D. Lacroix, L. E. Friberg, M. O. Karlsson. Evaluation of IPPSE, an alternative method for sequential population PKPD analysis. Journal of Pharmacokinetics and Pharmacodynamics 39: 177–193 (2012).

[219] L. B. Sheiner, D. R. Stanski, S. Vozeh, R. D. Miller, J. Ham. Simultaneous modeling of pharmacokinetics and pharmacodynamics: Application to d-tubocurarine. Clinical Pharmacology & Therapeutics 25: 358–371 (1979).

[220] E. Chasseloup, M. O. Karlsson. Comparison of Seven Non-Linear Mixed Effect Model-Based Approaches to Test for Treatment Effect. Pharmaceutics 15: 460 (2023).

[221] J. Chen, T. Liu. Statistical considerations on implementing the MCP-Mod method for binary endpoints in clinical trials. Contemporary Clinical Trials Communications 19: 100641 (2020).

[222] I. Netterberg, C.-C. Li, L. Molinero, N. Budha, S. Sukumaran, M. Stroh, E. N. Jonsson, L. E. Friberg. A PK/PD analysis of circulating biomarkers and their relationship to tumor response in atezolizumab-treated non-small cell lung cancer patients. Clinical Pharmacology & Therapeutics 105: 486–495 (2019).

[223] E. Chigutsa, A. Long, J. Wallin. Exposure–Response Analysis of Necitumumab Efficacy in Squamous Non-Small Cell Lung Cancer Patients. CPT: Pharmacometrics & Systems Pharmacology 6: 560–568 (2017).

[224] Y. Wang, C. Sung, C. Dartois, R. Ramchandani, B. Booth, E. Rock, J. Gobburu. Elucidation of relationship between tumor size and survival in non-small-cell lung cancer patients can aid early decision making in clinical drug development. Clinical Pharmacology & Therapeutics 86: 167–174 (2009).

[225] J. A. Mackintosh, H. M. Marshall, I. A. Yang, R. V. Bowman, K. M. Fong. A retrospective study of volume doubling time in surgically resected non-small cell lung cancer. Respirology 19: 755–762 (2014).

[226] S. M. Krishnan, L. E. Friberg. Bayesian forecasting of tumor size metrics and overall survival. CPT: Pharmacometrics & Systems Pharmacology 11: 1604–1613 (2022).

[227] S.-f. Ye, J. Li, S.-m. Ji, H.-h. Zeng, W. Lu. Dose-biomarker-response modeling of the anticancer effect of ethaselen in a human non-small cell lung cancer xenograft mouse model. Acta Pharmacologica Sinica 38: 223–232 (2017).

[228] A. Puszkiel, B. You, L. Payen, J. Lopez, J. Guitton, P. Rousset, J. Fontaine, J. Péron, D. Maillet, S. Tartas, N. Bonnin, V. Trillet-Lenoir, O. Colomban, D. Augu-Denechere, G. Freyer, M. Tod. A PK-PD model linking biomarker dynamics to progression-free survival in patients treated with everolimus and sorafenib combination therapy, EVESOR phase I trial. Cancer Chemotherapy and Pharmacology 91: 413–425 (2023).

[229] I. Netterberg, M. O. Karlsson, E. I. Nielsen, A. L. Quartino, H. Lindman, L. E. Friberg. The risk of febrile neutropenia in breast cancer patients following adjuvant chemotherapy is predicted by the time course of interleukin-6 and C-reactive protein by modelling. British Journal of Clinical Pharmacology 84: 490–500 (2018).

[230] L.-H. Lin, M. Ghasemi, S. M. Burke, C. K. Mavis, J. R. Nichols, P. Torka, D. E. Mager, F. J. Hernandez-Ilizaliturri, A. K. Goey. Population Pharmacokinetics and Pharmacodynamics of Carfilzomib in Combination with Rituximab, Ifosfamide, Carboplatin, and Etoposide in Adult Patients with Relapsed/Refractory Diffuse Large B Cell Lymphoma. Targeted Oncology 1–11 (2023).

[231] B. P. Solans, A. López-Díaz de Cerio, A. Elizalde, L. J. Pina, S. Inogés, J. Espinós, E. Salgado, L. D. Mejías, I. F. Trocóniz, M. Santisteban. Assessing the impact of the addition of dendritic cell vaccination to neoadjuvant chemotherapy in breast cancer patients: A model-based characterization approach. British Journal of Clinical Pharmacology 85: 1670–1683 (2019).

[232] J. H. Ahmed, E. Makonnen, R. K. Bisaso, J. K. Mukonzo, A. Fotoohi, A. Aseffa, R. Howe, M. Hassan, E. Aklillu. Population pharmacokinetic, pharmacogenetic, and pharmacodynamic analysis of cyclophosphamide in Ethiopian breast cancer patients. Frontiers in Pharmacology 11: 406 (2020).

[233] S. C. Tate, T. F. Burke, D. Hartman, P. Kulanthaivel, R. P. Beckmann, D. M. Cronier. Optimising the combination dosing strategy of abemaciclib and vemurafenib in BRAF-mutated melanoma xenograft tumours. British Journal of Cancer 114: 669–679 (2016).

[234] E. Schindler, M. Amantea, M. O. Karlsson, L. Friberg. PK-PD modeling of individual lesion FDG-PET response to predict overall survival in patients with sunitinib-treated gastrointestinal stromal tumor. CPT: Pharmacometrics & Systems Pharmacology 5: 173–181 (2016).

[235] E. Schindler, M. A. Amantea, M. O. Karlsson, L. E. Friberg. A pharmacometric framework for axitinib exposure, efficacy, and safety in metastatic renal cell carcinoma patients. CPT: Pharmacometrics & Systems Pharmacology 6: 373–382 (2017).

[236] Y. Yao, Z. Wang, L. Yong, Q. Yao, X. Tian, T. Wang, Q. Yang, C. Hao, T. Zhou. Longitudinal and time-to-event modeling for prognostic implications of radical surgery in retroperitoneal sarcoma. CPT: Pharmacometrics & Systems Pharmacology 11: 1170–1182 (2022).

[237] R. J. Keizer, Y. Funahashi, T. Semba, J. Wanders, J. Beijnen, J. Schellens, A. Huitema. Evaluation of α 2-integrin expression as a biomarker for tumor growth inhibition for the investigational integrin inhibitor E7820 in preclinical and clinical studies. The AAPS Journal 13: 230–239 (2011).

[238] E. Hansson, G. Ma, M. Amantea, J. French, P. Milligan, L. E. Friberg, M. O. Karlsson. PKPD modeling of predictors for adverse effects and overall survival in sunitinib-treated patients with GIST. CPT: Pharmacometrics & Systems Pharmacology 2: 1–9 (2013).

[239] E. Hansson, M. Amantea, P. Westwood, P. Milligan, B. Houk, J. French, M. O. Karlsson, L. E. Friberg. PKPD modeling of VEGF, sVEGFR-2, sVEGFR-3, and sKIT as predictors of tumor dynamics and overall survival following sunitinib treatment in GIST. CPT: Pharmacometrics & Systems Pharmacology 2: 1–9 (2013).

[240] N. Buil-Bruna, J.-M. López-Picazo, M. Moreno-Jiménez, S. Martín-Algarra, B. Ribba, I. F. Trocóniz. A population pharmacodynamic model for lactate dehydrogenase and neuron specific enolase to predict tumor progression in small cell lung cancer patients. The AAPS Journal 16: 609–619 (2014).

[241] P. B. Pierrillas, E. Henin, J. Ogier, L. Kraus-Berthier, M. Chenel, F. Bouzom, M. Tod. Tumor growth inhibition modelling based on receptor occupancy and biomarker activity of a new Bcl-2 inhibitor in mice. Journal of Pharmacology and Experimental Therapeutics 367: 414–424 (2018).

[242] S. Yamazaki. Translational pharmacokinetic-pharmacodynamic modeling from nonclinical to clinical development: A case study of anticancer drug, crizotinib. The AAPS Journal 15: 354–366 (2013).

[243] S. Yamazaki, J. L. Lam, H. Y. Zou, H. Wang, T. Smeal, P. Vicini. Mechanistic understanding of translational pharmacokinetic-pharmacodynamic relationships in nonclinical tumor models: A case study of orally available novel inhibitors of anaplastic lymphoma kinase. Drug Metabolism and Disposition 43: 54–62 (2015).

[244] L. Bueno, D. P. de Alwis, C. Pitou, J. Yingling, M. Lahn, S. Glatt, I. F. Trocóniz. Semi-mechanistic modelling of the tumour growth inhibitory effects of LY2157299, a new type I receptor TGF-β kinase antagonist, in mice. European Journal of Cancer 44: 142–150 (2008).

[245] S. Yang, T. P. Dumitrescu. Population pharmacokinetics and pharmacodynamics modelling of dilmapimod in severe trauma subjects at risk for acute respiratory distress syndrome. Drugs in R&D 17: 145–158 (2017).

[246] C. Li, S. Shoji, J. Beebe. Pharmacokinetics and C-reactive protein modelling of anti-interleukin-6 antibody (PF-04236921) in healthy volunteers and patients with autoimmune disease. British Journal of Clinical Pharmacology 84: 2059–2074 (2018).

[247] C. Ogami, Y. Tsuji, Y. Muraki, A. Mizoguchi, M. Okuda, H. To. Population pharmacokinetics and pharmacodynamics of teicoplanin and C-reactive protein in hospitalized patients with gram-positive infections. Clinical Pharmacology in Drug Development 9: 175–188 (2020).

[248] N. R. Sproston, J. J. Ashworth. Role of C-reactive protein at sites of inflammation and infection. Frontiers in Immunology 9: 754 (2018).

[249] M. Ohsawa, A. Okayama, M. Nakamura, T. Onoda, K. Kato, K. Itai, Y. Yoshida, A. Ogawa, K. Kawamura, K. Hiramori. CRP levels are elevated in smokers but unrelated to the number of cigarettes and are decreased by long-term smoking cessation in male smokers. Preventive Medicine 41: 651–656 (2005).

[250] J.-H. Cho, D.-C. Lee, H.-J. Lee. Association of Duration of Smoking Cessation or Cumulative Smoking Amount with Serum hs-CRP Level in Korean Adults: A Nationwide Population-Based Cross-Sectional Study. Toxics 10: 533 (2022).

[251] R. P. Tracy, B. M. Psaty, E. Macy, E. G. Bovill, M. Cushman, E. S. Cornell, L. H. Kuller. Lifetime smoking exposure affects the association of C-reactive protein with cardiovascular disease risk factors and subclinical disease in healthy elderly subjects. Arteriosclerosis, Thrombosis, and Vascular biology 17: 2167–2176 (1997).

[252] S. Tonstad, J. L. Cowan. C-reactive protein as a predictor of disease in smokers and former smokers: A review. International Journal of Clinical Practice 63: 1634–1641 (2009).

[253] S. Gallus, A. Lugo, P. Suatoni, F. Taverna, E. Bertocchi, R. Boffi, A. Marchiano, D. Morelli, U. Pastorino. Effect of tobacco smoking cessation on C-reactive protein levels in a cohort of low-dose computed tomography screening participants. Scientific Reports 8: 12908 (2018).

[254] D. G. Yanbaeva, M. A. Dentener, E. C. Creutzberg, G. Wesseling, E. F. Wouters. Systemic effects of smoking. Chest 131: 1557–1566 (2007).

[255] S. H. Walker, D. B. Duncan. Estimation of the probability of an event as a function of several independent variables. Biometrika 54: 167–179 (1967).

[256] J.-H. Lim, K.-E. Lee, K.-S. Hahn, K.-W. Park. Analyzing survival data as binary outcomes with logistic regression. Communications for Statistical Applications and Methods 17: 117–126 (2010).

[257] L. C. Cheung, P. S. Albert, S. Das, R. J. Cook. Multistate models for the natural history of cancer progression. British Journal of Cancer 127: 1279–1288 (2022).

[258] S. M. Krishnan, L. E. Friberg, R. Bruno, U. Beyer, J. Y. Jin, M. O. Karlsson. Multistate model for pharmacometric analyses of overall survival in HER2-negative breast cancer patients treated with docetaxel. CPT: Pharmacometrics & Systems Pharmacology 10: 1255–1266 (2021).

[259] J. G. Le-Rademacher, T. M. Therneau, F.-S. Ou. The utility of multistate models: A flexible framework for time-to-event data. Current Epidemiology Reports 9: 183–189 (2022).

[260] S. M. Krishnan, L. E. Friberg, F. Mercier, R. Zhang, B. Wu, J. Y. Jin, T. Hoang, M. Ballinger, R. Bruno, M. O. Karlsson. Multistate Pharmacometric Model to Define the Impact of Second-Line Immunotherapies on the Survival Outcome of the IMpower131 Study. Clinical Pharmacology & Therapeutics 113: 851–858 (2023).

[261] C.-W. Lin, M. Nagase, S. Doshi, S. Dutta. A multistate platform model for time-to-event endpoints in oncology clinical trials. CPT: Pharmacometrics & Systems Pharmacology (2023).

[262] A. A. Suleiman, S. Frechen, M. Scheffler, T. Zander, L. Nogova, M. Kocher, U. Jaehde, J. Wolf, U. Fuhr. A modeling and simulation framework for adverse events in erlotinib-treated non-small-cell lung cancer patients. The AAPS Journal 17: 1483–1491 (2015).

[263] L. Claret, J.-F. Lu, R. Bruno, C.-P. Hsu, Y.-J. Hei, Y.-N. Sun. Simulations using a drug–disease modeling framework and phase II data predict phase III survival outcome in first-line non–small-cell lung cancer. Clinical Pharmacology & Therapeutics 92: 631–634 (2012).

[264] A. Khandelwal, A.-M. Grisic, J. French, K. Venkatakrishnan. Pharmacometrics golems: Exposure-response models in oncology. Clinical Pharmacology & Therapeutics 112: 941–945 (2022).

[265] M. A. Björnsson, L. E. Friberg, U. S. Simonsson. Performance of nonlinear mixed effects models in the presence of informative dropout. The AAPS Journal 17: 245–255 (2015).

[266] A. Gonçalves, M. Marchand, P. Chan, J. Y. Jin, J. Guedj, R. Bruno. Comparison of two-stage and joint TGI-OS modeling using data from six atezolizumab clinical studies in patients with metastatic non-small cell lung cancer. CPT: Pharmacometrics & Systems Pharmacology (2024).

[267] J.-F. Wang, H.-D. Lu, Y. Wang, R. Zhang, X. Li, S. Wang. Clinical characteristics and prognosis of non-small cell lung cancer patients with liver metastasis: A population-based study. World Journal of Clinical Cases 10: 10882 (2022).

[268] M. G. Choi, C.-M. Choi, D. H. Lee, S.-W. Kim, S. Yoon, W. S. Kim, W. Ji, J. C. Lee. Different prognostic implications of hepatic metastasis according to front-line treatment in non-small cell lung cancer: A real-world retrospective study. Translational Lung Cancer Research 10: 2551 (2021).

[269] H. Xu, P. Ming, Z. Zhao, N. Zhao, D. Zhou, X. Tang, D. Cao. Assessing the Relationship Between Liver Metastases and the Survival of Patients With Non-Small Cell Lung Cancer After Immune Checkpoint Inhibitors Treatment: A Systematic Review and Meta-Analysis. Integrative Cancer Therapies 22: 15347354231164584 (2023).

[270] J. Li, H. Zhu, L. Sun, W. Xu, X. Wang. Prognostic value of site-specific metastases in lung cancer: A population based study. Journal of Cancer 10: 3079 (2019).

[271] B. Campos-Balea, J. de Castro Carpeño, B. Massutí, D. Vicente-Baz, D. Pérez Parente, P. Ruiz-Gracia, L. Crama, M. Cobo Dols. Prognostic factors for survival in patients with metastatic lung adenocarcinoma: An analysis of the SEER database. Thoracic Cancer 11: 3357–3364 (2020).

[272] Y. Ren, C. Dai, H. Zheng, F. Zhou, Y. She, G. Jiang, K. Fei, P. Yang, D. Xie, C. Chen. Prognostic effect of liver metastasis in lung cancer patients with distant metastasis. Oncotarget 7: 53245 (2016).

[273] S. Sridhar, L. Paz-Ares, H. Liu, K. Shen, C. Morehouse, N. Rizvi, N. H. Segal, X. Jin, Y. Zheng, R. Narwal, A. Gupta, P. A. Dennis, J. Ye, P. Mukhopadhyay, B. W. Higgs, K. Ranade. Prognostic significance of liver metastasis in durvalumab-treated lung cancer patients. Clinical Lung Cancer 20: e601–e608 (2019).

[274] T. Funazo, T. Nomizo, Y. H. Kim. Liver metastasis is associated with poor progression-free survival in patients with non–small cell lung cancer treated with nivolumab. Journal of Thoracic Oncology 12: e140–e141 (2017).

[275] Y. Liao, X. Fan, X. Wang. Effects of different metastasis patterns, surgery and other factors on the prognosis of patients with stage IV non-small cell lung cancer: A Surveillance, Epidemiology, and End Results (SEER) linked database analysis. Oncology Letters 18: 581–592 (2019).

[276] M. van Laar, W. A. van Amsterdam, A. S. van Lindert, P. A. de Jong, J. J. Verhoeff. Prognostic factors for overall survival of stage III non-small cell lung cancer patients on computed tomography: A systematic review and meta-analysis. Radiotherapy and Oncology 151: 152–175 (2020).

[277] D. Gerber, S. Dahlberg, A. Sandler, D. Ahn, J. Schiller, J. Brahmer, D. Johnson. Baseline tumour measurements predict survival in advanced non-small cell lung cancer. British Journal of Cancer 109: 1476–1481 (2013).

[278] C. Pusceddu, L. Melis, B. Sotgia, D. Guerzoni, A. Porcu, A. Fancellu. Usefulness of percutaneous microwave ablation for large non-small cell lung cancer: A preliminary report. Oncology Letters 18: 659–666 (2019).

[279] L. Claret, J. Y. Jin, C. Ferté, H. Winter, S. Girish, M. Stroh, P. He, M. Ballinger, A. Sandler, A. Joshi, A. Rittmeyer, D. Gandara, J. C. Soria, R. Bruno. A model of overall survival predicts treatment outcomes with atezolizumab versus chemotherapy in non–small cell lung cancer based on early tumor kinetics. Clinical Cancer Research 24: 3292–3298 (2018).

[280] M. Katsurada, T. Nagano, M. Tachihara, T. Kiriu, K. Furukawa, K. Koyama, T. Otoshi, R. Sekiya, D. Hazama, D. Tamura, K. Nakata, N. Katsurada, M. Yamamoto, K. Kobayashi, Y. Nishimura. Baseline tumor size as a predictive and prognostic factor of immune checkpoint inhibitor therapy for non-small cell lung cancer. Anticancer Research 39: 815–825 (2019).

[281] I. Winzenborg, A. M. Soliman, M. Shebley. A Personalized Medicine Approach Using Clinical Utility Index and Exposure-Response Modeling Informed by Patient Preferences Data. CPT: Pharmacometrics & Systems Pharmacology 10: 40–47 (2021).

[282] C. Fernández-Teruel, S. Fudio, R. Lubomirov. Integrated exposure–response analysis of efficacy and safety of lurbinectedin to support the dose regimen in small-cell lung cancer. Cancer Chemotherapy and Pharmacology 89: 585–594 (2022).

Publications

Original research articles (peer-reviewed)

- **Y.M. Nassar**, F.W. Ojara, A. Pérez-Pitarch, K. Geiger, W. Huisinga, N. Hartung, R. Michelet, S. Holdenrieder, M. Joerger*, C. Kloft*.
  C-reactive protein as an early predictor of efficacy in advanced non-small cell lung cancer patients: A tumor dynamics-biomarker modeling framework.
  Cancers 15: 5429 (2023).
  *shared senior authorship
  doi: 10.3390/cancers15225429

- F.W. Ojara, A. Henrich, N. Frances, **Y.M. Nassar**, W. Huisinga, N. Hartung, K. Geiger, S. Holdenrieder, M. Joerger, C. Kloft.
  A prognostic baseline serum biomarker and tumour growth kinetics integrated model in paclitaxel/platinum treated advanced non-small cell lung cancer patients.
  CPT Pharmacometrics Syst. Pharmacol. 12: 1714–1725 (2023).
  doi: 10.1002/psp4.12937

- **Y.M. Nassar**[#], N. Hohmann[#], R. Michelet, K. Gottwalt, A.D. Meid, J. Burhenne, W. Huisinga, W.E Haefeli, G. Mikus*, C. Kloft*.
  Quantification of the time course of CYP3A inhibition, activation, and induction using a population pharmacokinetic model of microdosed midazolam continuous infusion.
  Clin. Pharmacokinet. 61: 1595–1607 (2022).
  [#]shared first authorship; *shared senior authorship
  doi: 10.1007/s40262-022-01175-6

Review articles

- F.W. Ojara[#],**Y.M. Nassar**[#], K. Geiger, M. Joerger, M. van Dyk, S. Holdenrieder, C. Kloft.
  Prognostic impact of serum biomarkers for decision-making in the treatment of advanced non-small cell lung cancer patients: Review of existing approaches for modelling serum biomarker data and potential role for pharmacometrics
  [#]shared first authorship
  In preparation

Conference abstracts (oral/poster presentation)

- **Y.M. Nassar**, J. Schick, C. Kloft, S. Bossert, A. Pérez Pitarch.
  Simulation-based evaluation of combined Likelihood Ratio Test (cLRT) and Multiple Comparison Procedure (MCP) approaches to identify a dose-response relationship in the oncology setting under different study design variables.
  32nd Population Approach Group Europe (PAGE), Rome, Italy, 25–28 June 2024.
  PAGE 32: 10769 [www.page-meeting.org/?abstract=10769], (2024).

- **Y.M. Nassar**, F.W. Ojara, A. Pérez Pitarch, K. Geiger, W. Huisinga, R. Michelet, S. Holdenrieder, M. Joerger, C. Kloft.
  A joint tumour dynamics and C-reactive protein turnover model to identify early longitudinal prognostic predictors of overall survival in advanced non-small cell lung cancer patients.
  31st Population Approach Group Europe (PAGE), A Coruña, Spain, 27–30 June 2023.
  PAGE 31: 10304 [www.page-meeting.org/?abstract=10304], (2023).

- **Y.M. Nassar**, F.W. Ojara, A. Pérez Pitarch, S. Holdenrieder, K. Krueger, R. Michelet, W. Huisinga, M. Joerger, C. Kloft.
  Quantitative impact of different patient and disease factors on C-reactive protein (CRP) synthesis, a prognostic biomarker in advanced non-small cell lung cancer (NSCLC) patients.
  43rd European Organization of Research and Treatment of Cancer-Pharmacology and Molecular Mechanisms (EORTC-PAMM) meeting, Florence, Italy, 15–17 December 2022.
  PAMM oral communications & posters, 26
  [https://sites.altilab.com/files/CONGRES/2022/PAMM-ABSTRACTS-BOOKLET.pdf], (2022).

- **Y. Nassar**, Z.P. Parra-Guillén, K. Koster, W. Huisinga, M. Joerger, C. Kloft.
  Cachexia-associated anticancer drug toxicity is minimally mediated by alteration of drug's pharmacokinetics: Erlotinib as case study.
  30th Population Approach Group Europe (PAGE), Ljubljana, Slovenia, 28 June–1 July 2022.
  PAGE 30: 10073 [www.page-meeting.org/?abstract=10073], (2022).

- **Y. Nassar**, F. Ojara, S. Holdenrieder, K. Krueger, M. Joerger, R. Michelet, W. Huisinga, C. Kloft.
  Prognostic impact of five serum biomarkers in patients with advanced non-small cell lung cancer (NSCLC): An exploratory analysis.
  European Society of Medical Oncology (ESMO) Virtual Congress, 16–21 September 2021.
  Ann. Oncol. 32: S404 (2021).
  doi: 10.1016/j.annonc.2021.08.392

- **Y. Nassar**, G. Mikus, W. Huisinga, R. Michelet, C. Kloft.
  Quantifying the dynamics of the modulatory effect of CYP3A perpetrators on hepatic CYP3A activity using a nonlinear mixed-effects model of microdosed midazolam and its metabolite 1-hydroxymidazolam.
  29th Population Approach Group Europe (PAGE), Online, 2–3 and 6–7 September 2021. PAGE 29: 9619 [www.page-meeting.org/?abstract=9619], (2021).

Oral presentations (without abstracts)

- **Y.M. Nassar**.
  C-reactive protein as an early predictor of efficacy in advanced non-small cell lung cancer patients: A tumor dynamics-biomarker modeling framework.
  26^{th} Pharmacokinetics/Pharmacodynamics (PK/PD) Expert Meeting, Isny, Germany, 14–16 March 2024.

- **Y.M. Nassar**.
  A joint tumour size and C-reactive protein model to identify early longitudinal predictors of efficacy in lung cancer patients.
  12^{th} PharMetrX Symposium. Berlin/Potsdam, Germany. 20 April 2023.

Dissertation (Masters thesis)

- **Y.M. Nassar**. Pharmacometrics of treosulfan in children undergoing haematopoietic stem cell transplantation. University of Oxford, United Kingdom (2017).

Appendices

A | Tables

A.1 Dose adaptations in the CEPAC-TDM study

Table A.1.1: Dose adaptation for paclitaxel, carboplatin, and cisplatin in the CEPAC-TDM study

Drug	Condition	Dose adaptation	Treatment arm
Paclitaxel	Neutropaenia grade 4	20 % dose reduction	A
	Neutropaenia $<$ grade 3	Increase paclitaxel dose	B
	Chemotherapy-associated grade 2 peripheral neuropathy persisting at the start of the subsequent cycle	If cisplatin is given, switch to carboplatin first and if neuropathy persists, paclitaxel dose is reduced by 20%	A,B
	Non-haematological toxicity grade ≥ 3	20% dose reduction	A,B
Carboplatin	Renal function	Adapted every cycle according to the Cockroft-Gault formula	A,B
	Thrombocytopenia grade ≥ 3	75% dose reduction	A,B
Cisplatin	Chemotherapy-associated grade 2 peripheral neuropathy persisting at the start of the subsequent cycle	Change to carboplatin	A,B
	Chemotherapy-associated nephrotoxicity, if glomerular filtration rate remains $< 60\,\mathrm{mL/min}$	Change to carboplatin	A,B
	Cisplatin-associated ototoxicity	Change to carboplatin	A,B

In case of two simultaneous toxicities, dose adaptation occured according to the toxicity requiring higher dose reduction and not in an additive manner.

A.2 Datasets

Table A.2.1: Definition of dataset variables used in the development of the coupled tumour - CRP model

Variable	Description	Source	Unit	Possible values	Changes, imputations, assumptions, comments
ID	Patient identifier xx Study centre + 10 yyy ID within study centre Example: 16007	—	—	Numeric > 0	—
TIME	Relative time elapsed from start of 1^{st} paclitaxel administration/baseline CRP measurement	—	h	Numeric ≥ 0	All events before 1^{st} administration were set to 0.00
TALD	Time after last dose: relative time elapsed from last paclitaxel administration	—	h	Numeric ≥ 0	All events before 1^{st} administration were set to 0.00
DAY	Day of the study cycle (this is used as a marker and does *not* correspond to actual dates) 1 1^{st} day of cycle (drug administration/ baseline CRP measurement) or visit for OCC > 6 2 2^{nd} day of cycle (CRP measurement)	TIME	—	1, 2	If OCC/DAY and date were not consistent, date was assumed to be correct and OCC/DAY were adapted accordingly

Continued on next page

Table A.2.1: *Cont.* Definition of dataset variables used in the development of the coupled tumour - CRP model

Variable	Description	Source	Unit	Possible values	Changes, imputations, assumptions, comments
OCC	Occasion/treatment cycle 1 Cycle 1 including screening/baseline 2—6 Cycle 2—6 7 Starting from 3 weeks after last dose including end-of-treatment visit and follow-up	TIME (for occasion 7)	—	1, 2, 3, 4, 5, 6, 7	If OCC/DAY and date were not consistent, date was assumed to be correct and OCC/DAY were adapted accordingly
CMT	Compartment identifier 1 CRP 2 Tumour	—	—	1, 2	—
EVID	NONMEM event identification 0 Observation event (only DV) 1 Dosing event 2 Other event (e.g. dummy data point) 3 Reset event 4 Reset dosing event (e.g. cross over)	—	—	0, 1, 2, 3, 4	—
MDV	Missing dependent variable 0 Observation 1 Missing dependent variable	DV	—	0, 1	—
ARM	Study arm 1 BSA-guided treatment arm (Arm A) 2 PK-guided dosing arm (Arm B)	—	—	1, 2	—
FLAGARMA	Flag patients in arm A who have day 2 CRP samples i.e. no PK sampling on day 2 0 No 1 Yes	—	—	0, 1	Patients (n=5), CRP measurements (n=6)

Continued on next page

Table A.2.1: *Cont.* Definition of dataset variables used in the development of the coupled tumour - CRP model

Variable	Description	Source	Unit	Possible values	Changes, imputations, assumptions, comments
MAXOCC	Number of cycles a patient received in total during the study	OCC	—	1, 2, 3, 4, 5, 6	—
SEX	Sex 0 Female 1 Male	—	—	0, 1	—
SMOK	Smoking status at screening/baseline visit 1 Never 2 Former 3 Current	—	—	1, 2, 3	—
AGE	Age at screening/baseline visit	—	years	Numeric > 0	—
HT	Body height at screening/baseline visit	—	cm	Numeric > 0	—
WT	Body weight determined at the beginning of each cycle	—	kg	Numeric > 0	Imputation by last observation carried forward, for time points between measurements
BMI	Body mass index calculated as $\frac{weight\,[kg]}{height\,[m]^2}$	—	kg/m^2	Numeric > 0	—
BSA	Body surface area calculated by the DuBois-DuBois-formula: $0.007184 \cdot HT^{0.725} \cdot WT^{0.425}$	HT, WT	m^2	Numeric > 0	Missing values were imputed by applying the DuBois-DuBois-formula
ALT	Concentration of alanine aminotransferase at screening/baseline visit Reference interval female: <35 Reference interval male: <50	—	U/L	Numeric > 0	—

Continued on next page

Table A.2.1: *Cont.* Definition of dataset variables used in the development of the coupled tumour - CRP model

Variable	Description	Source	Unit	Possible values	Changes, imputations, assumptions, comments
AST	Concentration of aspartate aminotransferase at screening/baseline visit Reference interval female: <35 Reference interval male: <50	—	U/L	Numeric > 0	—
ECOG	ECOG performance status determined at the beginning of each cycle 0 Fully active 1 Restricted in physically strenuous activity but ambulatory 2 Ambulatory and capable of all self-care but unable to carry out any work activities 3 Capable of only limited self-care 4 Completely disabled 5 Dead	—	—	0, 1, 2, 3, 4, 5	Imputation by last observation carried forward, for time points between measurements
BLECOG	Baseline ECOG performance status 0 Fully active 1 Restricted in physically strenuous activity but ambulatory 2 Ambulatory and capable of all self-care but unable to carry out any work activities 3 Capable of only limited self-care 4 Completely disabled 5 Dead	ECOG, TIME, OCC	—	0, 1, 2, 3, 4, 5	—

Continued on next page

Table A.2.1: *Cont.* Definition of dataset variables used in the development of the coupled tumour - CRP model

Variable	Description	Source	Unit	Possible values	Changes, imputations, assumptions, comments
STAGE	Most recent tumour staging determined at screening/baseline visit 0 IIIB 1 IV	—	—	0, 1	—
LIVERMET	Presence of liver metastasis (lesions) 0 No 1 Yes	—	—	0, 1	Imputation by last observation carried forward, for time points between tumour assessment
BRAINMET	Presence of brain metastasis (lesions) 0 No 1 Yes	—	—	0, 1	Imputation by last observation carried forward, for time points between tumour assessment
NTARGLES	Number of target lesions max 5, according to RECIST	—	—	0, 1, 2, 3, 4, 5	—
NNTARGLES	Number of non-target lesions	—	—	Numeric ≥ 0	—
SUMNOLES	Sum of number of target and non-target lesions	NTARGLES, NNTARGLES	—	Numeric ≥ 0	—
SUMDIA	Sum of diameters of target lesions, according to RECIST	—	cm	Numeric ≥ 0	—
BLSD	Baseline (pre-dose) sum of diameters of target lesions	SUMDIA, TIME, OCC	cm	Numeric ≥ 0	For patients (n=1) with more than one measurement at baseline, record closer to the dosing time was picked

Continued on next page

Table A.2.1: *Cont.* Definition of dataset variables used in the development of the coupled tumour - CRP model

Variable	Description	Source	Unit	Possible values	Changes, imputations, assumptions, comments
HIST	NSCLC histology determined at screening/baseline visit 0 Carcinoma, not otherwise specified 1 Adenocarcinoma 2 Bronchioalveolar carcinoma 3 Squamous-cell carcinoma	—	—	0, 1, 2, 3	—
DV	Dependent variable i.e. CRP concentration sampled in cycles 1, 2, 3 and at end-of-treatment visit	—	mg/L	Numeric ≥ 0	—
LNDV	Natural logarithm of CRP concentration (DV)	DV	ln(mg/L)	Numeric ≥ 0	—
FLAGCRP	Flagged CRP samples 0 Observations with no queries 1 Observations with unusually high values but serial dilutions performed during analysis 2 BLQ observations	—	—	0, 1, 2	—
BLCRP	Baseline CRP concentration (cycle 1, day 1)	DV, TIME, OCC	mg/L	Numeric ≥ 0	—
FLAGBLCRP	Flag for baseline CRP concentration 0 Patients have baseline CRP 1 Patients do not have baseline CRP	—	—	0, 1	Missing baseline CRP concentrations (n=2)

Continued on next page

Table A.2.1: *Cont.* Definition of dataset variables used in the development of the coupled tumour - CRP model

Variable	Description	Source	Unit	Possible values	Changes, imputations, assumptions, comments
BLCRP2	Baseline CRP concentration (cycle 1 day 1) with missed concentrations imputed by median baseline CRP concentrations	DV, BLCRP	mg/L	Numeric ≥ 0	Missing baseline CRP concentration (n=2) imputed by median baseline CRP concentration (23.5 mg/L) calculated after excluding BLQ and missing CRP samples
BLIL6	Baseline IL-6 concentration (cycle 1, day 1)	—	pg/mL	Numeric ≥ 0	—
FLAGBLIL6	Flag for baseline IL-6 concentration 0 Patients have baseline IL-6 1 Patients do not have baseline IL-6 2 Patients with baseline IL-6 equal zero	—	—	0, 1, 2	No baseline IL-6 concentrations (n=23), baseline IL-6 concentrations equals zero (n=2)
NEWBLIL6	Baseline IL-6 concentration (cycle 1, day 1), with missed and zero concentrations imputed by median baseline IL-6 concentrations	—	pg/mL	Numeric ≥ 0	Missed (n=23) and zero concentrations (n=2) of baseline IL-6 concentrations replaced by median baseline IL-6 concentration (2.57 pg/mL) calculated after excluding patients with BLQ CRP and missing or zero baseline IL-6 concentration
SAMPLENO	CRP sample number within a cycle 1 Sample number 1 within cycle 2 Sample number 2 within cycle	—	—	1, 2	—
BM	Flag for record with CRP measurement 0 Not a CRP measurement 1 CRP measurement	—	—	0, 1	—

Continued on next page

Table A.2.1: *Cont.* Definition of dataset variables used in the development of the coupled tumour - CRP model

Variable	Description	Source	Unit	Possible values	Changes, imputations, assumptions, comments
IGR	Individual tumour growth rate constant for linear tumour growth derived from the multiple imputation framework	—	cm/h	Numeric ≥ 0	—
IBETA	Individual paclitaxel AUC_{cycle}-driven tumour decay rate constant at treatment start (time=0) derived from the multiple imputation framework	—	$1/(\mu mol/L{\cdot}h)/h$	Numeric ≥ 0	—
ILAMBDA	Individual rate constant for exponential decline in paclitaxel effect over time derived from the multiple imputation framework	—	1/h	Numeric ≥ 0	—
SEGR	Individual standard error (uncertainty) of linear tumour growth rate derived from the multiple imputation framework	—	cm/h	Numeric ≥ 0	—
SEBETA	Individual standard error (uncertainty) of drug-induced tumour decay rate constant derived from the multiple imputation framework	—	$1/(\mu mol/L{\cdot}h)/h$	Numeric ≥ 0	—
SELAMBDA	Individual standard error (uncertainty) of rate constant for drug effect derived from the multiple imputation framework	—	1/h	Numeric ≥ 0	—

Continued on next page

Table A.2.1: *Cont.* Definition of dataset variables used in the development of the coupled tumour - CRP model

Variable	Description	Source	Unit	Possible values	Changes, imputations, assumptions, comments
SEERR	Individual standard error (uncertainty) of baseline tumour size measurement derived from the multiple imputation framework	—	cm	Numeric ≥ 0	Error derived from modelling baseline tumour size according to the B2 method in the developed tumour growth inhibition model
MAUC	Paclitaxel area under the concentration-time curve from the start to the end of a cycle averaged per individual from the 50 replicates under the multiple imputation framework	—	μmol·h/L	Numeric ≥ 0	Paclitaxel area under the concentration-time curve in the last attended treatment cycle was carried forward to the follow-up records i.e. assuming an extended cycle duration

AUC_{cycle}: paclitaxel area under the concentration-time curve from the start to end of a cycle; BLQ: below limit of quantification; BSA: body surface area; CRP: C-reactive protein; ECOG: Eastern Cooperative Oncology Group; IL-6: interleukin 6; NSCLC: non-small cell lung cancer; PK: pharmacokinetic(s); RECIST: Response Evaluation Criteria in Solid Tumours

Table A.2.2: Definition of additional dataset variables used in the development of the time-to-event models

Variable	Description	Source	Unit	Possible values	Changes, imputations, assumptions, comments
DV	Dependent variable. For overall survival: 0 Alive/censored 1 Dead For progression-free survival: 0 Alive without progression/censored 1 Progression or death	PROGT, OST	—	0, 1	
SIM	Simulation flag 0 Estimation records 1 Dummy records for simulation	—	—	0, 1	Dummy records were added at a 7-day interval to allow for simulations across all possible time intervals
PROG	Tumour progression 0 No porgression detected 1 Time progression was detected 2 Time after progression	—	—	0, 1, 2	
PROGT	Time-to-progression calculated from administration of first dose until detection of tumour progression	PROG	Days	Numeric > 0	
DEATH	Death 0 Patient alive 1 Time of death	—	—	0, 1	
OST	Overall survival time calculated from administration of first dose until death	DEATH	Days	Numeric > 0	

Continued on next page

Table A.2.2: *Cont.* Definition of additional dataset variables used in the development of the time-to-event models

Variable	Description	Source	Unit	Possible values	Changes, imputations, assumptions, comments
CRP_{cycle1}	Model-predicted CRP concentration on day 1 of cycle 1	—	mg/mL	Numeric ≥ 0	Estimated from the coupled tumour-biomarker model
CRP_{cycle2}	Model-predicted CRP concentration on day 1 of cycle 2 or at time=504 h	—	mg/mL	Numeric ≥ 0	Estimated from the coupled tumour-biomarker model
CRP_{cycle3}	Model-predicted CRP concentration on day 1 of cycle 3 or at time=1008 h	—	mg/mL	Numeric ≥ 0	Estimated from the coupled tumour-biomarker model
$CRP_{\frac{cycle2-1}{cycle1}}$	Model-predicted relative change in CRP concentration between cycle 2 and cycle 1	CRP_{cycle2}, CRP_{cycle1}	—	Numeric ≥ 0	Derived from the coupled tumour-biomarker model as $\frac{CRP_{cycle2}-CRP_{cycle1}}{CRP_{cycle1}}$
$CRP_{\frac{cycle3-1}{cycle1}}$	Model-predicted relative change in CRP concentration between cycle 3 and cycle 1	CRP_{cycle3}, CRP_{cycle1}	—	Numeric ≥ 0	Derived from the coupled tumour-biomarker model as $\frac{CRP_{cycle3}-CRP_{cycle1}}{CRP_{cycle1}}$
$CRP_{\frac{cycle3-2}{cycle2}}$	Model-predicted relative change in CRP concentration between cycle 3 and cycle 2	CRP_{cycle3}, CRP_{cycle2}	—	Numeric ≥ 0	Derived from the coupled tumour-biomarker model as $\frac{CRP_{cycle3}-CRP_{cycle2}}{CRP_{cycle2}}$
$CRP_{cycle2-1}$	Model-predicted difference in CRP concentration between cycle 2 and cycle 1	CRP_{cycle2}, CRP_{cycle1}	mg/mL	Numeric ≥ 0	Derived from the coupled tumour-biomarker model as $CRP_{cycle2} - CRP_{cycle1}$
$CRP_{cycle3-1}$	Model-predicted difference in CRP concentration between cycle 3 and cycle 1	CRP_{cycle3}, CRP_{cycle1}	mg/mL	Numeric ≥ 0	Derived from the coupled tumour-biomarker model as $CRP_{cycle3} - CRP_{cycle1}$

Continued on next page

Table A.2.2: *Cont.* Definition of additional dataset variables used in the development of the time-to-event models

Variable	Description	Source	Unit	Possible values	Changes, imputations, assumptions, comments
$CRP_{cycle3-2}$	Model-predicted difference in CRP concentration between cycle 3 and cycle 2	CRP_{cycle3}, CRP_{cycle2}	mg/mL	Numeric ≥ 0	Derived from the coupled tumour-biomarker model as $CRP_{cycle3} - CRP_{cycle2}$
$CRP_{\frac{cycle2}{cycle1}}$	Model-predicted fold change in CRP concentration between cycle 2 and cycle 1	CRP_{cycle2}, CRP_{cycle1}	—	Numeric ≥ 0	Derived from the coupled tumour-biomarker model as $\frac{CRP_{cycle2}}{CRP_{cycle1}}$
$CRP_{\frac{cycle3}{cycle1}}$	Model-predicted fold change in CRP concentration between cycle 3 and cycle 1	CRP_{cycle3}, CRP_{cycle1}	—	Numeric ≥ 0	Derived from the coupled tumour-biomarker model as $\frac{CRP_{cycle3}}{CRP_{cycle1}}$
$CRP_{\frac{cycle3}{cycle2}}$	Model-predicted fold change in CRP concentration between cycle 3 and cycle 2	CRP_{cycle3}, CRP_{cycle2}	—	Numeric ≥ 0	Derived from the coupled tumour-biomarker model as $\frac{CRP_{cycle3}}{CRP_{cycle2}}$
$\mathrm{N/L}_{cycle1}$	Baseline (day 1, cycle 1) neutrophil-to-lymphocyte ratio	—	—	Numeric > 0	
$\mathrm{N/L}_{cycle2}$	Neutrophil-to-lymphocyte ratio at cycle 2	—	—	Numeric > 0	$\mathrm{N/L}_{cycle2}$ defined as measurements taken on day 20 (12.8% of patients) or day 21 (75.8% of patients) from start of treatment. Missing measurements (n=29, 11.3%) were imputed by the median value of $\mathrm{N/L}_{cycle2}$ of 3.895
$\mathrm{N/L}_{\frac{cycle2-1}{cycle1}}$	Relative change in N/L ratio between cycle 2 and cycle 1	$\mathrm{N/L}_{cycle2}$, $\mathrm{N/L}_{cycle1}$	—	Numeric > 0	Derived as $\frac{N/L_{cycle2}-N/L_{cycle1}}{N/L_{cycle1}}$

Continued on next page

Table A.2.2: *Cont.* Definition of additional dataset variables used in the development of the time-to-event models

Variable	Description	Source	Unit	Possible values	Changes, imputations, assumptions, comments
$N/L_{cycle2-1}$	Difference in N/L ratio between cycle 2 and cycle 1	N/L_{cycle2}, N/L_{cycle1}	—	Numeric 0	Derived as $N/L_{cycle2} - N/L_{cycle1}$
$N/L_{\frac{cycle2}{cycle1}}$	Fold change in N/L ratio between cycle 2 and cycle 1	N/L_{cycle2}, N/L_{cycle1}	—	Numeric 0	Derived as $\frac{N/L_{cycle2}}{N/L_{cycle1}}$
GR	Individual linear net tumour growth rate constant derived from the multiple imputation framework, after accounting for uncertainty	—	cm/h	Numeric ≥ 0	
RS7	Model-predicted change in tumour size at week 7 relative to baseline tumour size	BLSD	—	Numeric ≥ 0	Derived from the tumour size model as $\frac{\text{Tumour size}_{week7}}{\text{Baseline tumour size}} \cdot 100$
LM	Landmark time flag 0 Patients with events after landmark time (accepted) 1 Patients with events before landmark time (ignored)	—	—	0, 1	

AUC_{cycle}: paclitaxel area under the concentration-time curve from the start to end of a cycle; CRP: C-reactive protein.
Note: ID, TIME, EVID, ARM, FLAGARMA, MAXOCC, SMOK, HIST, BLECOG, STAGE, BRAINMET, LIVERMET, NTARGLES, NNTARGLES, SUMNOLES, BLCRP2, BLSD, and BLCRP were variables used in the time-to-event datasets as defined before in appendix Table A.2.1.

A.3 Evaluation of the robustness of cLRT and stability of the simulations

Table A.3.1: Proportion of non-robust N simulations under the null hypothesis i.e. not reaching 90% successful executions of M simulations-estimations after model-fitting and selection, per study design

Dose levels	Total number of patients	Number of arms (patients/arm)	Percentage of N simulations within which M simulations-estimations did not reach 90% successful executions
0, 12, 24, 36, 48, 60	24	6 (4)	1
0, 20, 40,60	24	4 (6)	1
0, 30, 60	24	3 (8)	1
0, 12, 24, 36, 48, 60	36	6 (6)	4
0, 20, 40,60	36	4 (9)	Null
0, 30, 60	36	3 (12)	1

$N = 100$
100% successful M executions after model-fitting and selection correspond to n = 100

Table A.3.2: Proportion of positive t_{cLRT} values per simulation scenario

Dose levels	Total number of patients	Number of arms (patients/arm)	Dose-response relationship	Percentage of N simulations with positive t_{cLRT} values	
				Weak drug effect	Strong drug effect
0, 12, 24, 36, 48, 60	24	6 (4)	Linear	3	—
			Log-linear	1	—
			E_{max}	5	—
			Sigmoidal	1	—
0, 20, 40, 60	24	4 (6)	Linear	6	—
			Log-linear	4	—
			E_{max}	1	—
			Sigmoidal	4	—
0, 30, 60	24	3 (8)	Linear	1	—
			Log-linear	2	—
			E_{max}	1	1
			Sigmoidal	2	—
0, 12, 24, 36, 48, 60	36	6 (6)	Linear	1	—
			Log-linear	2	—
			E_{max}	4	—
			Sigmoidal	3	—
0, 20, 40, 60	36	4 (9)	Linear	1	—
			Log-linear	—	—
			E_{max}	3	—
			Sigmoidal	2	—
0, 30, 60	36	3 (12)	Linear	1	—
			Log-linear	1	—
			E_{max}	3	—
			Sigmoidal	2	—

$N = 100$

Table A.3.3: Proportion of positive $t_{cLRT_{Ho}}$ values per simulation scenario

Dose levels	Total number of patients	Number of arms (patients/arm)	Range of percentage of positive $t_{cLRT_{Ho}}$ values within the $t_{cLRT_{Ho}}$ distribution of each N simulation
0, 12, 24, 36, 48, 60	24	6 (4)	14–44
0, 20, 40,60	24	4 (6)	11–42
0, 30, 60	24	3 (8)	14–44
0, 12, 24, 36, 48, 60	36	6 (6)	13–45
0, 20, 40,60	36	4 (9)	12–39
0, 30, 60	36	3 (12)	12–43

$N = 100$

Table A.3.4: Proportion of outlying $t_{cLRT_{Ho}}$ values (i.e. $<$-60) per simulation scenario

Dose levels	Total number of patients	Number of arms (patients/arm)	Percentage of N simulations with outlying $t_{cLRT_{Ho}}$ values	Range of percentage of outlying $t_{cLRT_{Ho}}$ values within the $t_{cLRT_{Ho}}$ distribution of each N simulation*
0, 12, 24, 36, 48, 60	24	6 (4)	9	1.3–38.9
0, 20, 40,60	24	4 (6)	16	1.2–31.8
0, 30, 60	24	3 (8)	14	1.3–23.4
0, 12, 24, 36, 48, 60	36	6 (6)	12	1.4–38.2
0, 20, 40,60	36	4 (9)	13	1.4–36.7
0, 30, 60	36	3 (12)	9	1.4–23.9

$N = 100$
*Further details are provided in Appendix Fig. B.2.1

A.4 Detailed information on the power of cLRT and MCP associated with each scenario

Table A.4.1: Power of cLRT and MCP approaches across the different study design variables stratified by strength of drug effect and dose-response relationship of the true data-generating dose-response model (cLRT (%)/MCP (%))

Dose-response relationship of true model	Total number of patients	Drug effect: Strong			Drug effect: Weak		
		Dose levels = 3	Dose levels = 4	Dose levels = 6	Dose levels = 3	Dose levels = 4	Dose levels = 6
Linear	36	98/52	99/48	100/43	83.7/23	77.8/22	73.7/24
	24	97/41	98/44	98/28	74.5/20	80.6/21	78.1/18
Log-linear	36	98/52	99/40	100/36	90.8/23	86/21	87.2/16
	24	98/44	98/33	100/20	82.5/19	81.1/16	84.7/13
E_{max}	36	97/42	99/36	100/30	79.2/20	78.4/20	84.9/12
	24	96.9/35	96/30	99/18	66.3/16	78.6/14	77.7/10
Sigmoidal	36	100/47	99/35	99/33	87.6/22	82.7/20	83.9/14
	24	99/38	97/31	100/17	75.3/16	83.2/14	84.7/11

A.5 Patient characteristics stratified by study arm

Table A.5.1: Baseline patient demographics and clinical characteristics stratified by study arm

Patient characteristic	Arm A (n = 126)	Arm B (n = 132)	All patients (n = 258)
Age [years]			
median	65.0	62.5	64.0
[range]	[41.0–76.0]	[43.0–78.0]	[41.0–78.0]
Sex, n (%)			
Female	41 (32.5)	50 (37.9)	91 (35.3)
Male	85 (67.5)	82 (62.1)	167 (64.7)
Body weight [kg]			
median	74.5	74.0	74.0
[range]	[44.0–135]	[42.0–117]	[42.0–135]
Body height [cm]			
median	172	171	171
[range]	[146–194]	[149–190]	[146–194]
Body mass index [kg/m^2]			
median	25.0	24.5	24.9
[range]	[16.8–41.7]	[17.2–37.4]	[16.8–41.7]
Body surface area [m^2]			
median	1.86	1.87	1.86
[range]	[1.39–2.49]	[1.34–2.32]	[1.34–2.49]
Neutrophil-to-lymphocyte ratio			
median	4.32	4.18	4.28
[range]	[0.798–94.0]	[1.47–19.1]	[0.798–94.0]
Alanine aminotransferase activity [U/L]			
median	24.5	22.0	23.0
[range]	[6.00–125]	[5.00–120]	[5.00–125]
Aspartate aminotransferase activity [U/L]			
median	21.0	21.0	21.0
[range]	[10.0–212]	[9.00–144]	[9.00–212]
Smoking status, n (%)			
Non-smokers	16 (12.7)	12 (9.09)	28 (10.9)
Former smokers	68 (54.0)	64 (48.5)	132 (51.2)
Current smokers	42 (33.3)	56 (42.4)	98 (38.0)
Disease stage, n (%)			
IIIB	14 (11.1)	27 (20.5)	41 (15.9)
IV	112 (88.9)	105 (79.5)	217 (84.1)
NSCLC histology, n (%)			
Adenocarcinoma	83 (65.9)	82 (62.1)	165 (64.0)
Bronchioalveolar carcinoma	—	1 (0.758)	1 (0.388)
Carcinoma, not otherwise specified	13 (10.3)	19 (14.4)	32 (12.4)
Squamous-cell carcinoma	30 (23.8)	30 (22.7)	60 (23.3)
Brain lesions, n (%)			
No	109 (86.5)	115 (87.1)	224 (86.8)
Yes	17 (13.5)	17 (12.9)	34 (13.2)
Liver lesions, n (%)			
No	108 (85.7)	99 (75.0)	207 (80.2)
Yes	18 (14.3)	33 (25.0)	51 (19.8)
Baseline ECOG performance status, n (%)			
0	70 (55.6)	65 (49.2)	135 (52.3)
1	48 (38.1)	57 (43.2)	105 (40.7)
2	8 (6.35)	10 (7.58)	18 (6.98)

Arm A: body surface area-guided paclitaxel dosing arm; Arm B: pharmacokinetic-guided paclitaxel dosing arm; ECOG: Eastern Cooperative Oncology Group; NSCLC: non-small cell lung cancer

A.6 Summary of SCM procedure for the CRP turnover model

Table A.6.1: Summary of forward inclusion steps of SCM procedure for the CRP turnover model

Forward step	Covariate included after each step	Covariate type	Functional relationship	ΔOFV threshold for retaining/ excluding	ΔOFV	*p*-value	*df*	Decrease in $\text{IIV}_{K_{in,CRP}}$, %
1	Baseline IL-6	Continuous	Linear	-3.84	-95.8	$1.27{\cdot}10^{-22}$	1	24
2	Baseline tumour size	Continuous	Exponential	-3.84	-14.9	0.000113	1	3.3
3	Disease stage	Categorical	Fractional change	-3.84	-9.23	0.00238	1	2.2
4	Smoking status	Categorical	Fractional change	-5.99	-10.7	0.00475	2	2.4
5	ALT	Continuous	Exponential	-3.84	-5.66	0.0174	1	0.7

ALT: alanine aminotransferase; *df*: degrees of freedom; $\text{IIV}_{Kin,CRP}$: interindividual variability associated with C-reactive protein production rate constant, $K_{in,CRP}$; IL-6: interleukin 6; OFV: objective function value.

Table A.6.2: Summary of backward deletion steps of SCM procedure for the CRP turnover model

Backward step	Covariate excluded after each step	Covariate type	Functional relationship	ΔOFV threshold for retaining/ excluding	ΔOFV	*p*-value	*df*	Increase in $\text{IIV}_{K_{in,CRP}}$, %
1	ALT	Continuous	Exponential	6.64	5.66	0.0174	1	0.7

ALT: alanine aminotransferase; *df*: degrees of freedom; $\text{IIV}_{Kin,CRP}$: interindividual variability associated with C-reactive protein production rate constant, $K_{in,CRP}$; OFV: objective function value.

A.7 Summary of SCM procedure for the time-to-event model of progression-free survival

Table A.7.1: Summary of forward inclusion steps of SCM procedure for the time-to-event model of progression-free survival

Forward step	Parameter	Covariate included after each step	Covariate type	Functional relationship	ΔOFV threshold for retaining/excluding	ΔOFV	*p*-value	*df*
1	$h_0(t)$	CRP_{cycle3}	Continuous	Linear	-3.84	-20.2	$7.1 \cdot 10^{-6}$	1
2	$h_0(t)$	$CRP_{cycle3-2}$	Continuous	Power	-3.84	-7.38	0.006614	1
3	$h_0(t)$	Disease stage	Categorical	Fractional change	-3.84	-4.94	0.0262	1

CRP_{cycle3}: CRP concentration at treatment cycle 3; $CRP_{cycle3-2}$: Absolute difference between CRP concentration at treatment cycle 3 and 2; *df*: degrees of freedom; $h_0(t)$: baseline log-normal hazard function; OFV: objective function value.

Table A.7.2: Summary of backward deletion step of SCM procedure for the time-to-event model of progression-free survival

Backward step	Parameter	Covariate excluded after each step	Covariate type	Functional relationship	ΔOFV threshold for retaining/excluding	ΔOFV	*p*-value	*df*
1	$h_0(t)$	Disease stage	Categorical	Fractional change	6.64	4.94	0.0262	1

df: degrees of freedom; $h_0(t)$: baseline log-normal hazard function; OFV: objective function value.

A.8 Summary of SCM procedure for the time-to-event model of overall survival

Table A.8.1: Summary of forward inclusion steps of SCM procedure for the time-to-event model of overall survival

Forward step	Parameter	Covariate included after each step	Covariate type	Functional relationship	ΔOFV threshold for retaining/excluding	ΔOFV	*p*-value	*df*
1	$h_0(t)$	CRP_{cycle3}	Continuous	Power	-3.84	-53.0	$3.33 \cdot 10^{-13}$	1
2	$h_0(t)$	Liver lesions	Categorical	Fractional change	-3.84	-12.8	$3.47 \cdot 10^{-4}$	1
3	$h_0(t)$	$CRP_{cycle3-2}$	Continuous	Power	-3.84	-10.2	$1.40 \cdot 10^{-3}$	1
3	$h_0(t)$	Baseline tumour size	Continuous	Power	-3.84	-12.4	$4.30 \cdot 10^{-4}$	1

CRP_{cycle3}: CRP concentration at treatment cycle 3; $CRP_{cycle3-2}$: Absolute difference between CRP concentration at treatment cycle 3 and 2; *df*: degrees of freedom; $h_0(t)$: baseline Weibull hazard function; OFV: objective function value.

Summary of backward deletion step: None of the identified predictors were removed during the backward deletion step.

B | Figures

B.1 Even distribution of dose levels

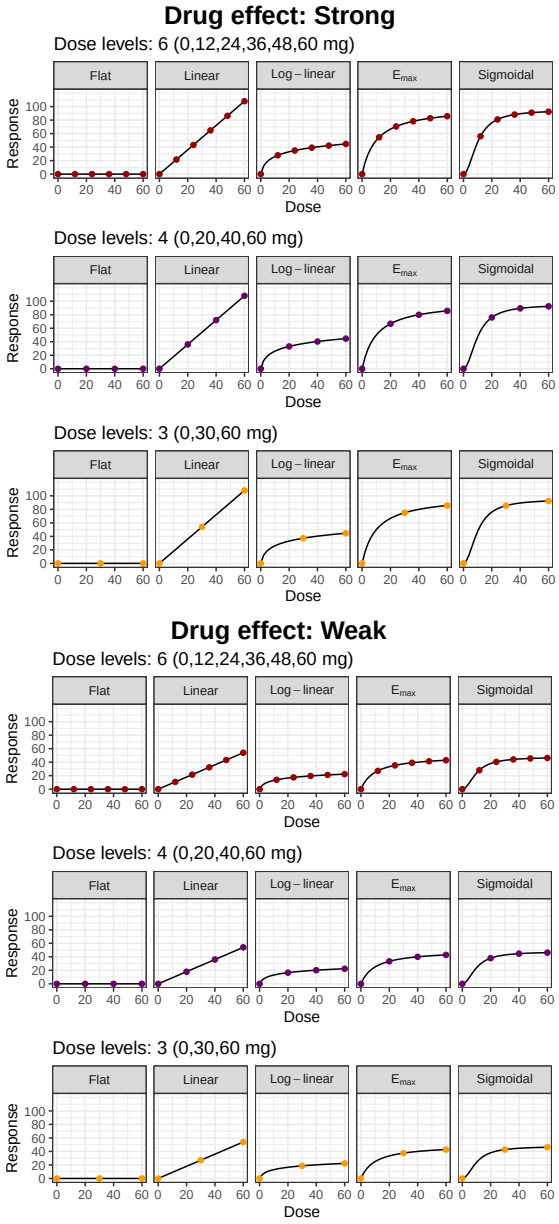

Figure B.1.1: Distribution of the dose levels (range: 0–60 mg) over the different dose-response relationships according to the number of study arms within the simulation scenarios of cLRT and MCP. Top panel: strong drug effect; bottom panel: weak drug effect.
Note: the magnitude of response is arbitrary (i.e. does not correspond to the change in tumour size) and only for the sake of illustrative comparison between the two drug strengths.

B.2 Outliers in the $t_{cLRT_{Ho}}$ distribution stratified per N simulations and study design

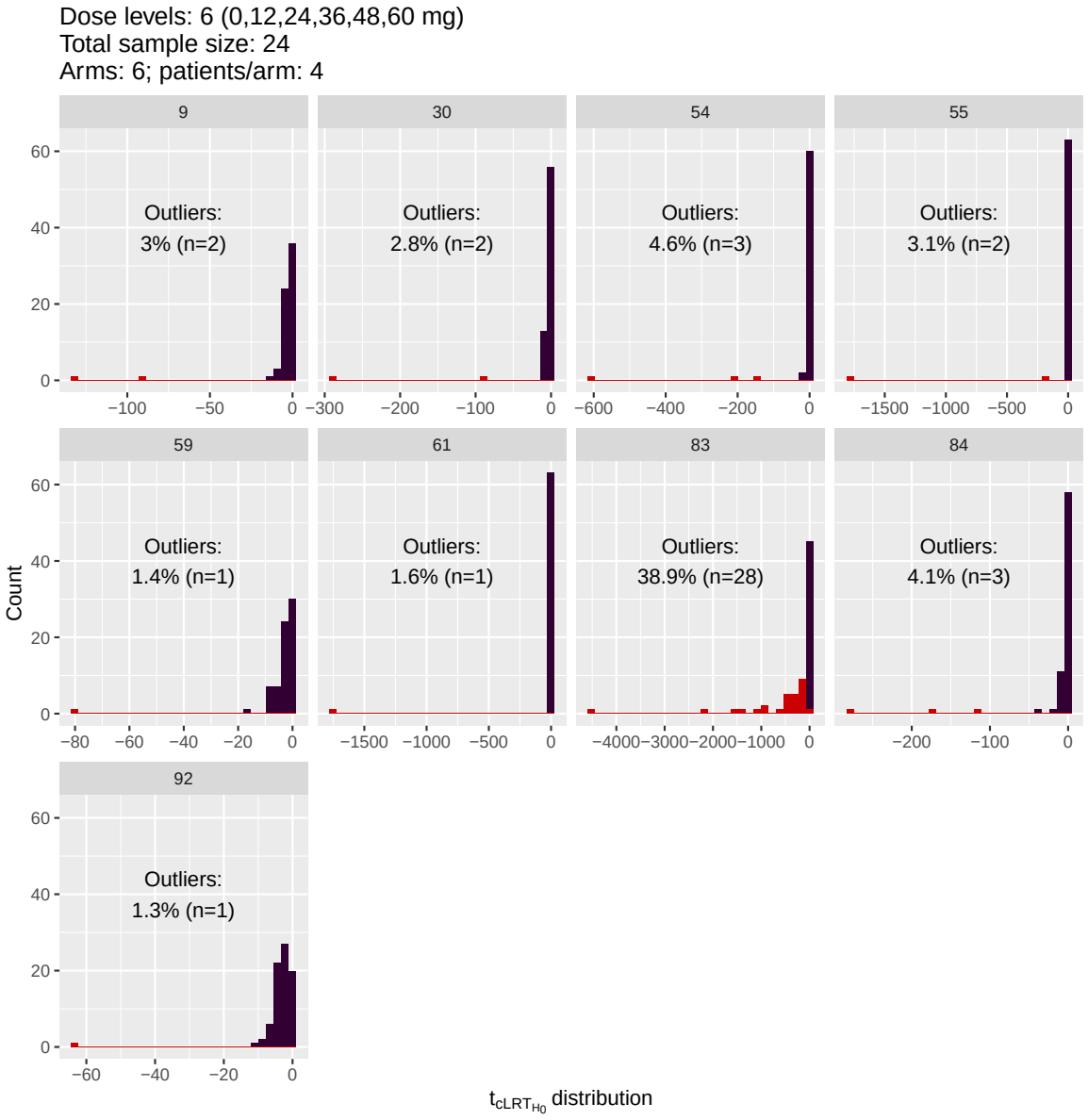

Figure B.2.1: Percentage (count) of outlying $t_{cLRT_{Ho}}$ values (i.e. <-60) within the $t_{cLRT_{Ho}}$ distribution of each N simulation.
Each panel represents the N simulation (number shown on top) with outlying $t_{cLRT_{Ho}}$ value(s) within its $t_{cLRT_{Ho}}$ distribution after excluding positive $t_{cLRT_{Ho}}$ values and N simulations within which M simulations-estimations did not reach 90% successful runs.
Red: excluded $t_{cLRT_{Ho}}$ value <-60.

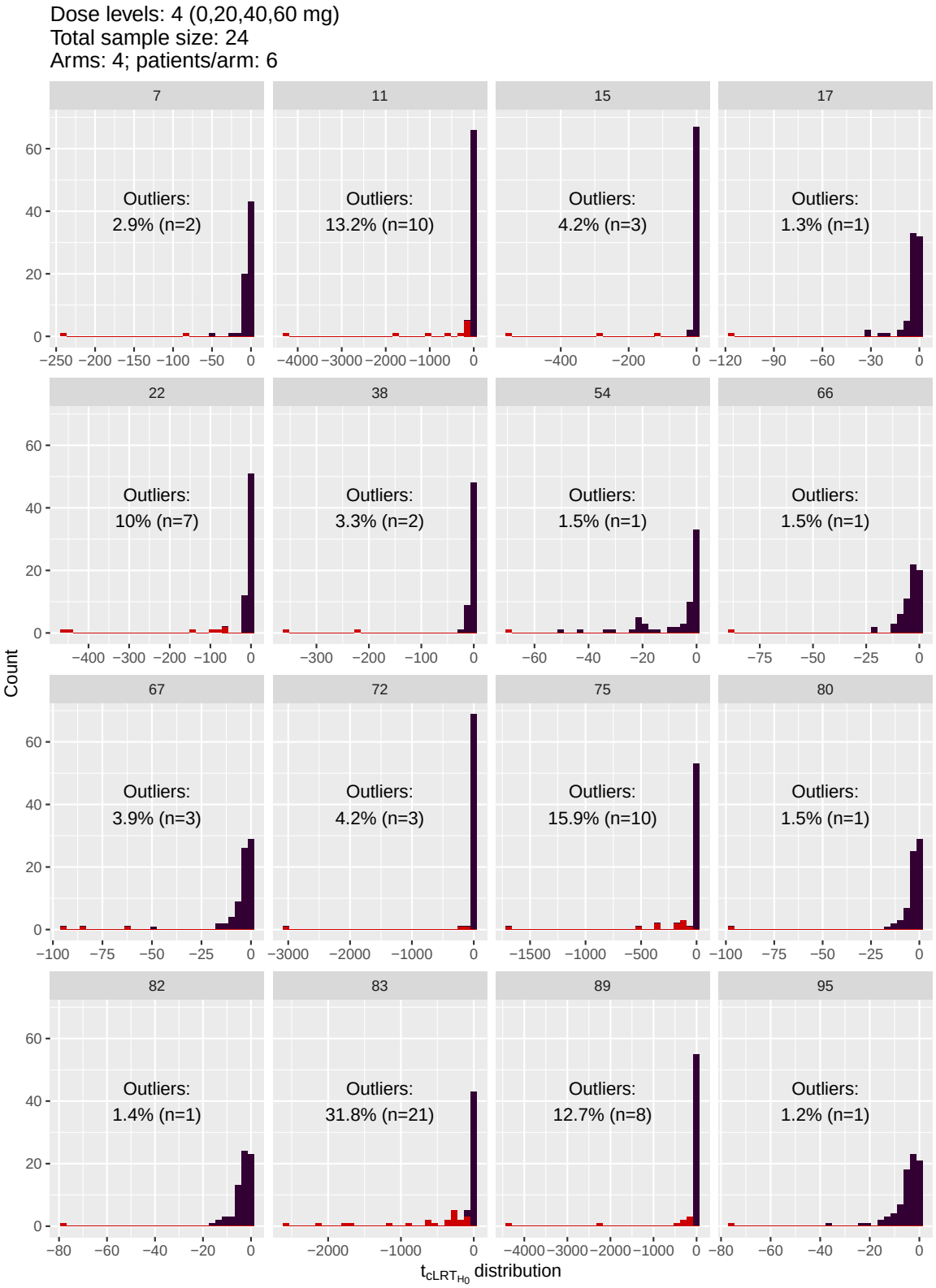

Figure B.2.1: *Cont.* Percentage (count) of outlying $t_{cLRT_{Ho}}$ values (i.e. <-60) within the $t_{cLRT_{Ho}}$ distribution of each N simulation.
Each panel represents the N simulation (number shown on top) with outlying $t_{cLRT_{Ho}}$ value(s) within its $t_{cLRT_{Ho}}$ distribution after excluding positive $t_{cLRT_{Ho}}$ values and N simulations within which M simulations-estimations did not reach 90% successful runs.
Red: excluded $t_{cLRT_{Ho}}$ value <-60.

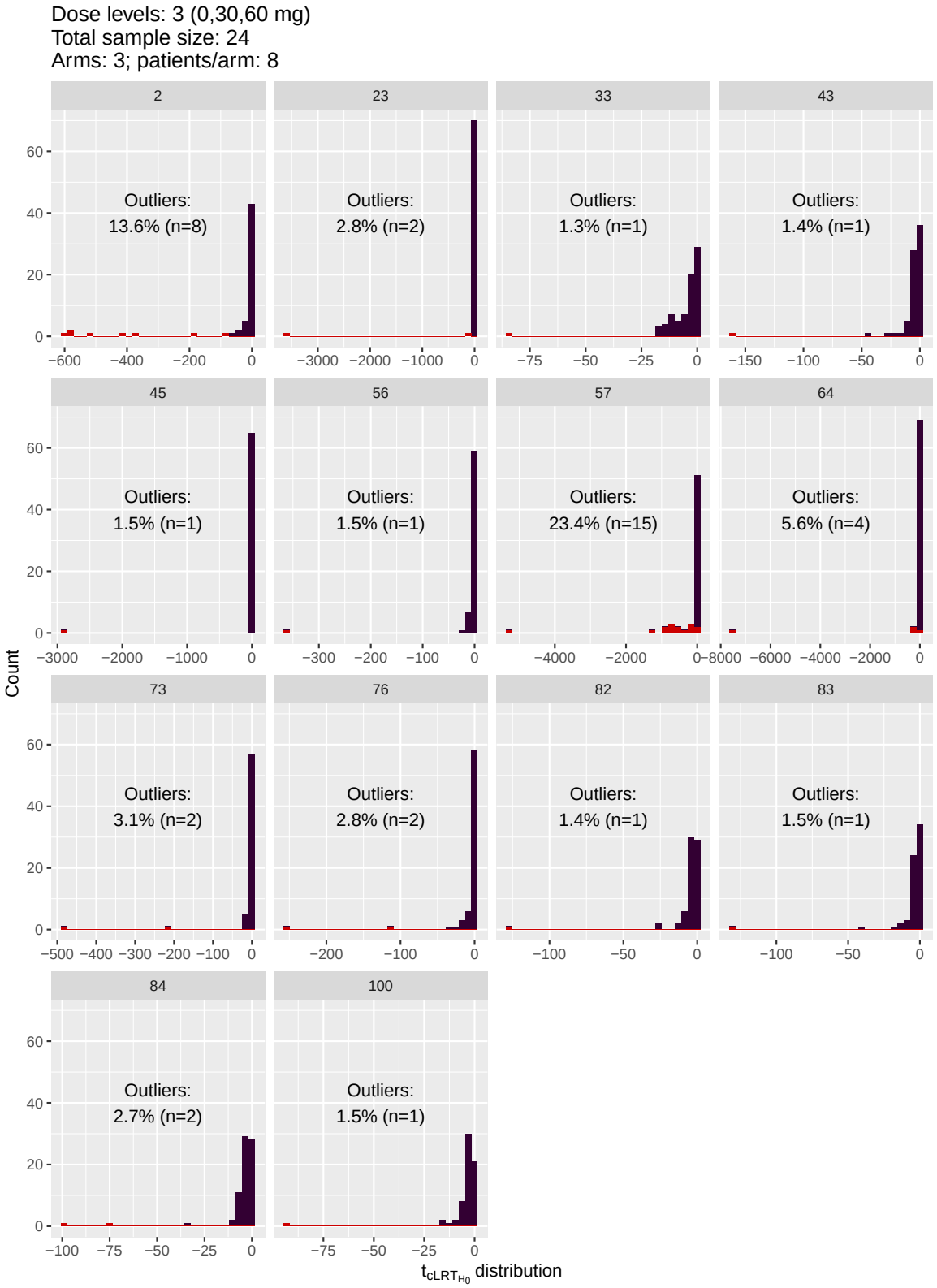

Figure B.2.1: *Cont.* Percentage (count) of outlying $t_{cLRT_{Ho}}$ values (i.e. <-60) within the $t_{cLRT_{Ho}}$ distribution of each N simulation.
Each panel represents the N simulation (number shown on top) with outlying $t_{cLRT_{Ho}}$ value(s) within its $t_{cLRT_{Ho}}$ distribution after excluding positive $t_{cLRT_{Ho}}$ values and N simulations within which M simulations-estimations did not reach 90% successful runs.
Red: excluded $t_{cLRT_{Ho}}$ value <-60.

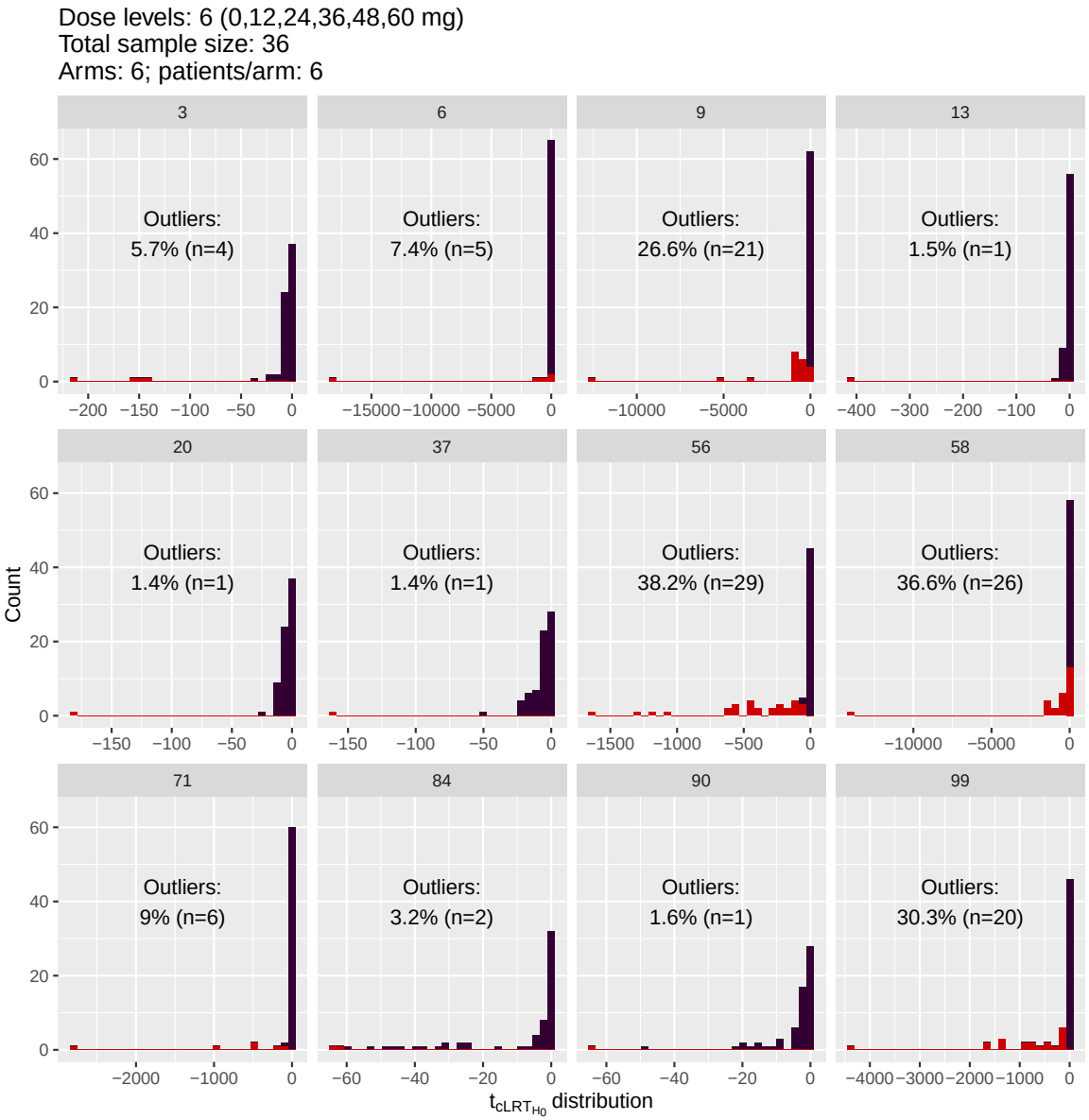

Figure B.2.1: *Cont.* Percentage (count) of outlying $t_{cLRT_{Ho}}$ values (i.e. <-60) within the $t_{cLRT_{Ho}}$ distribution of each N simulation.
Each panel represents the N simulation (number shown on top) with outlying $t_{cLRT_{Ho}}$ value(s) within its $t_{cLRT_{Ho}}$ distribution after excluding positive $t_{cLRT_{Ho}}$ values and N simulations within which M simulations-estimations did not reach 90% successful runs.
Red: excluded $t_{cLRT_{Ho}}$ value <-60.

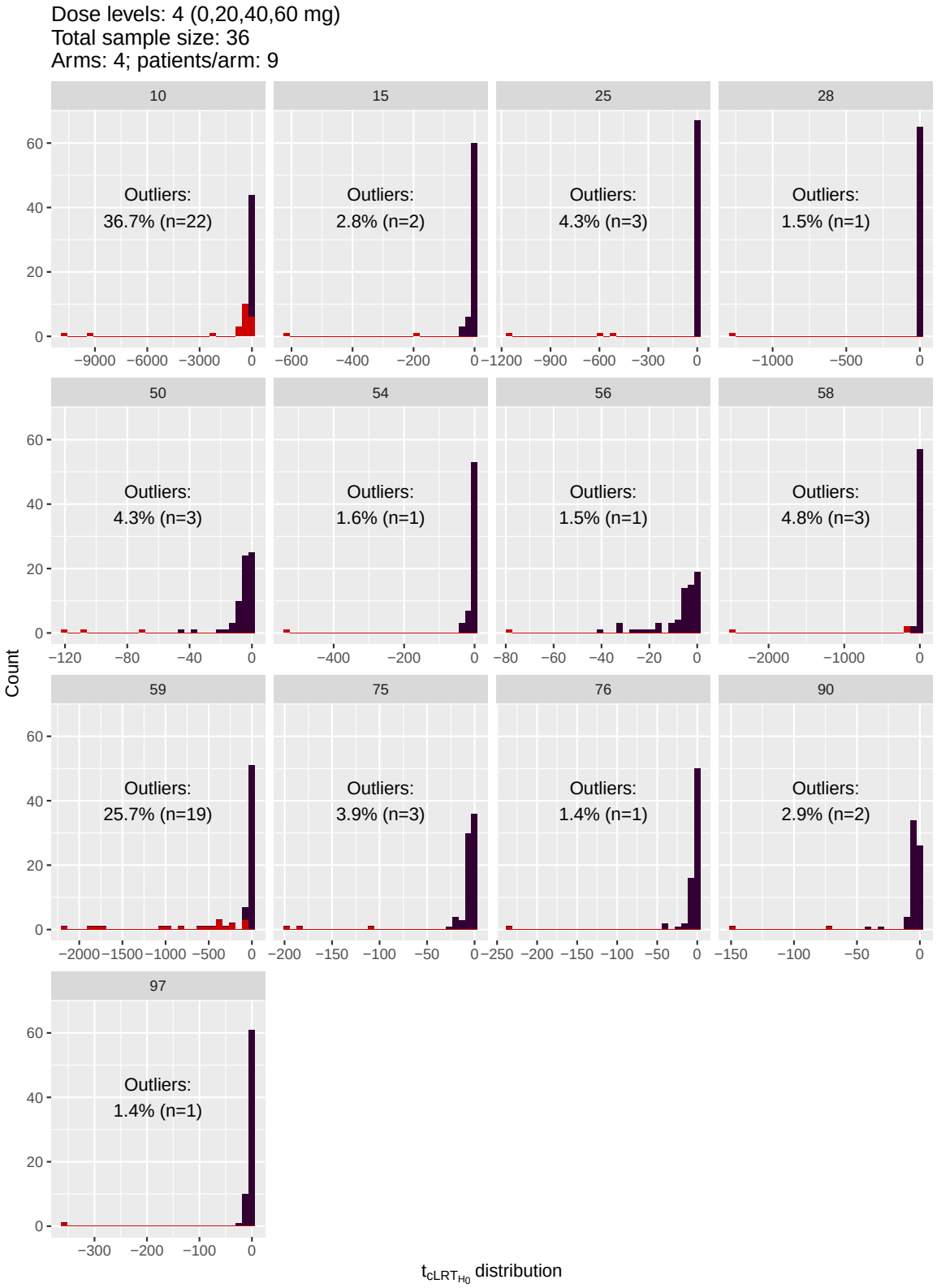

Figure B.2.1: *Cont.* Percentage (count) of outlying $t_{cLRT_{Ho}}$ values (i.e. $<$-60) within the $t_{cLRT_{Ho}}$ distribution of each N simulation.
Each panel represents the N simulation (number shown on top) with outlying $t_{cLRT_{Ho}}$ value(s) within its $t_{cLRT_{Ho}}$ distribution after excluding positive $t_{cLRT_{Ho}}$ values and N simulations within which M simulations-estimations did not reach 90% successful runs.
Red: excluded $t_{cLRT_{Ho}}$ value $<$-60.

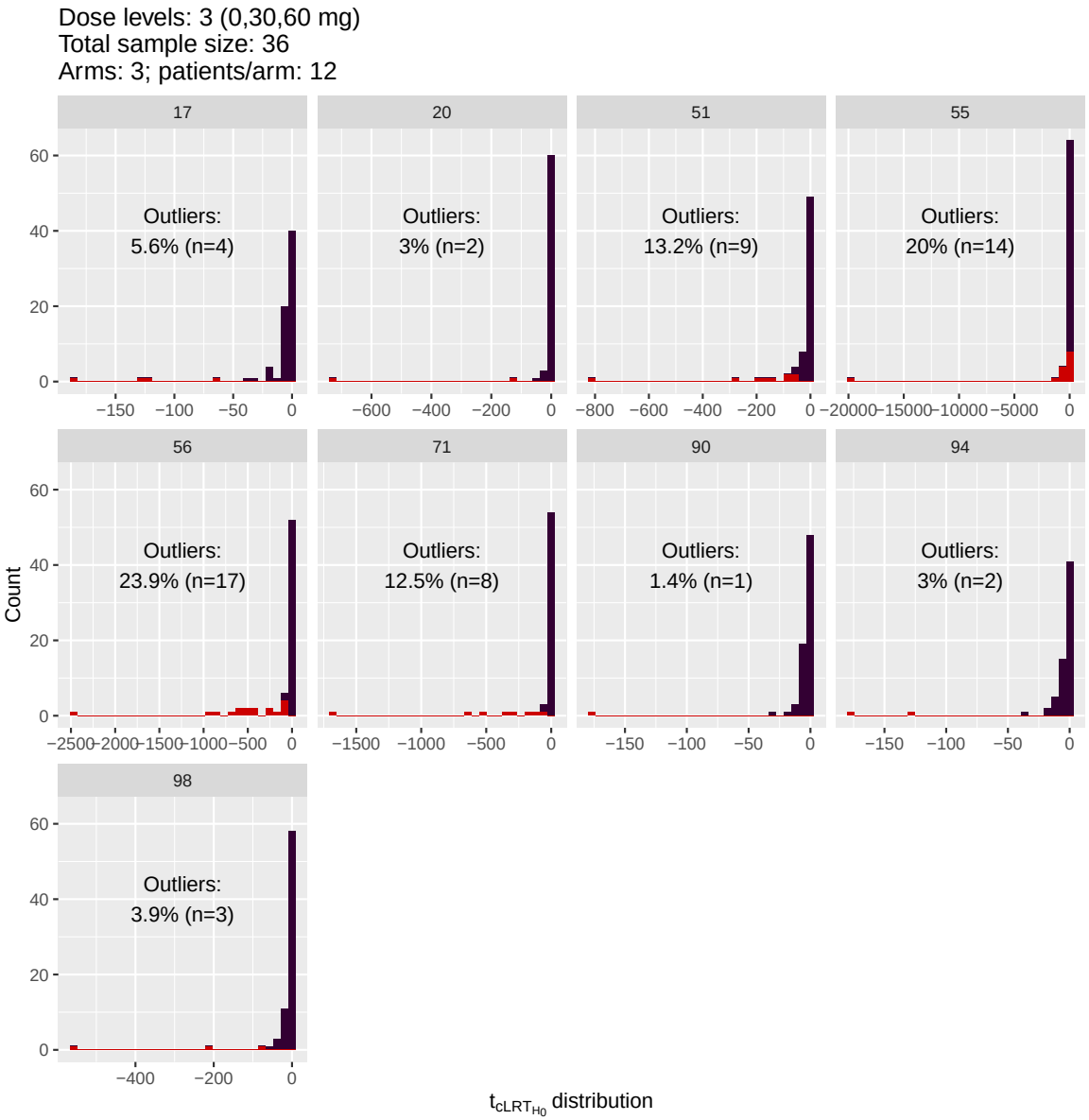

Figure B.2.1: *Cont.* Percentage (count) of outlying $t_{cLRT_{Ho}}$ values (i.e. $<$-60) within the $t_{cLRT_{Ho}}$ distribution of each N simulation.
Each panel represents the N simulation (number shown on top) with outlying $t_{cLRT_{Ho}}$ value(s) within its $t_{cLRT_{Ho}}$ distribution after excluding positive $t_{cLRT_{Ho}}$ values and N simulations within which M simulations-estimations did not reach 90% successful runs.
Red: excluded $t_{cLRT_{Ho}}$ value $<$-60.

B.3 Selection of best fitting candidate model and identification of true underlying dose-response model for MCP

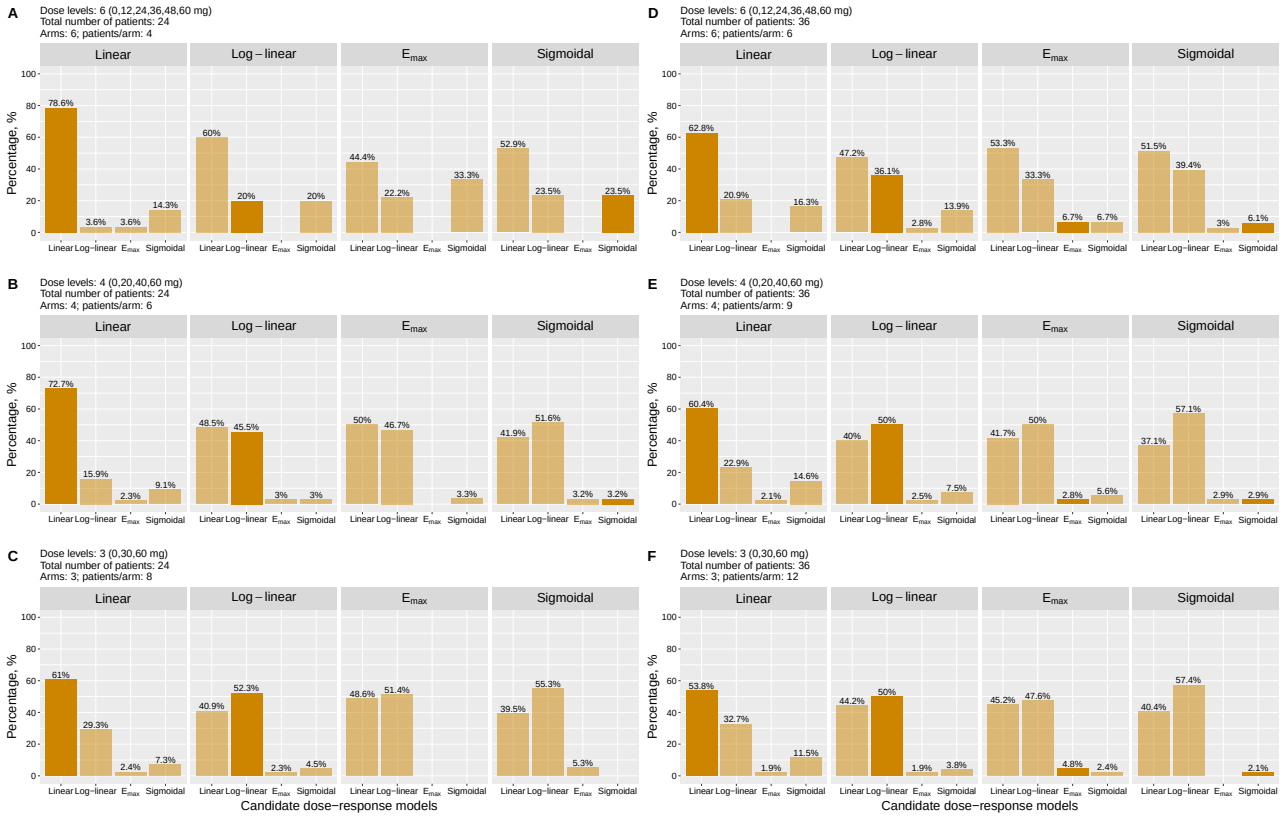

Figure B.3.1: Proportion of the selected best-fitting candidate dose-response models under the ***strong*** drug effect highlighting (darker bars) the proportion of the identified true underlying relationship of the dose-response model for multiple comparison procedure (MCP), per each group of scenarios (panels **A**–**F**) sharing the same set of study design variables (i.e. number of dose levels, total number of patients, number of arms) and split (**facets**) by the underlying dose-response relationship of the true data-generating model. Number of dose levels: top panel (**A, D**) 6, middle panel (**B, E**) 4, lower panel (**C, F**) 3; Total number of patients: left panel (**A**–**C**): 24, right panel (**D**–**F**): 36.

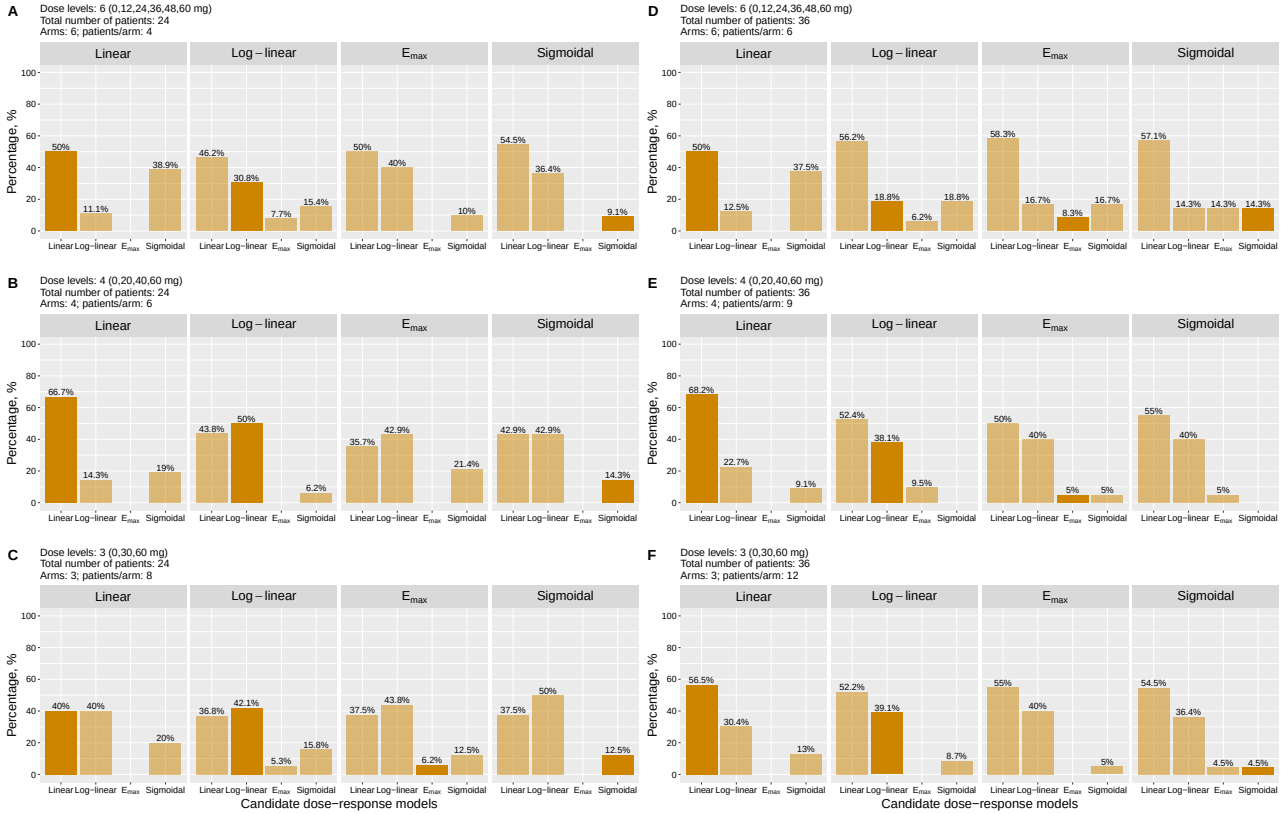

Figure B.3.2: Proportion of the selected best-fitting candidate dose-response models under the ***weak*** drug effect highlighting (darker bars) the proportion of the identified true underlying relationship of the dose-response model for multiple comparison procedure (MCP), per each group of scenarios (panels **A**–**F**) sharing the same set of study design variables (i.e. number of dose levels, total number of patients, number of arms) and split (**facets**) by the underlying dose-response relationship of the true data-generating model. Number of dose levels: top panel (**A, D**) 6, middle panel (**B, E**) 4, lower panel (**C, F**) 3; Total number of patients: left panel (**A**–**C**): 24, right panel (**D**–**F**): 36.

B.4 Histogram of CRP concentrations stratified by sampling time

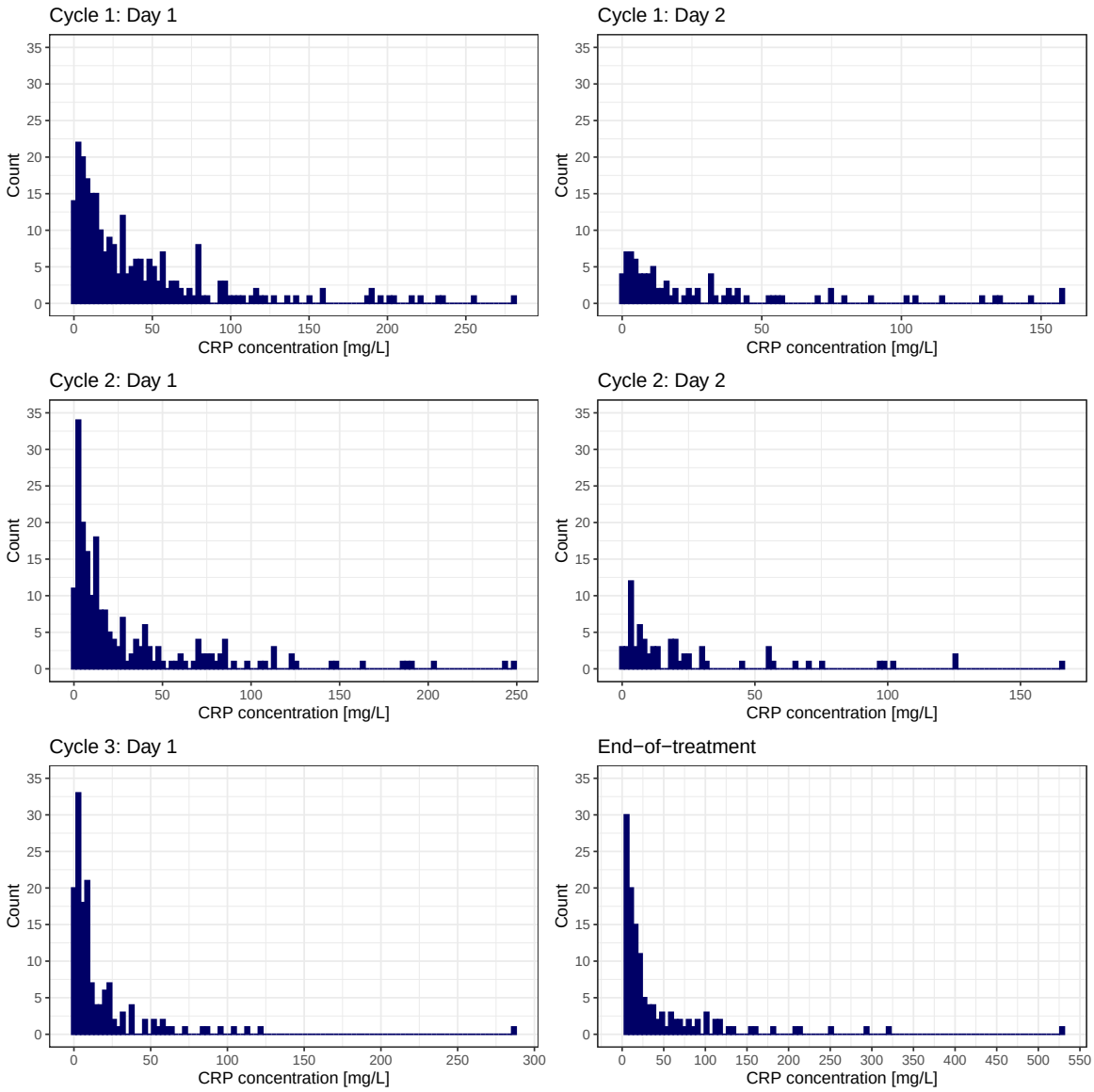

Figure B.4.1: Histogram of CRP concentrations stratified by sampling time. CRP: C-reactive protein.

B.5 Exploratory analyses of the influence of different continuous and categorical covariates on interindividual variability of $K_{in,CRP}$

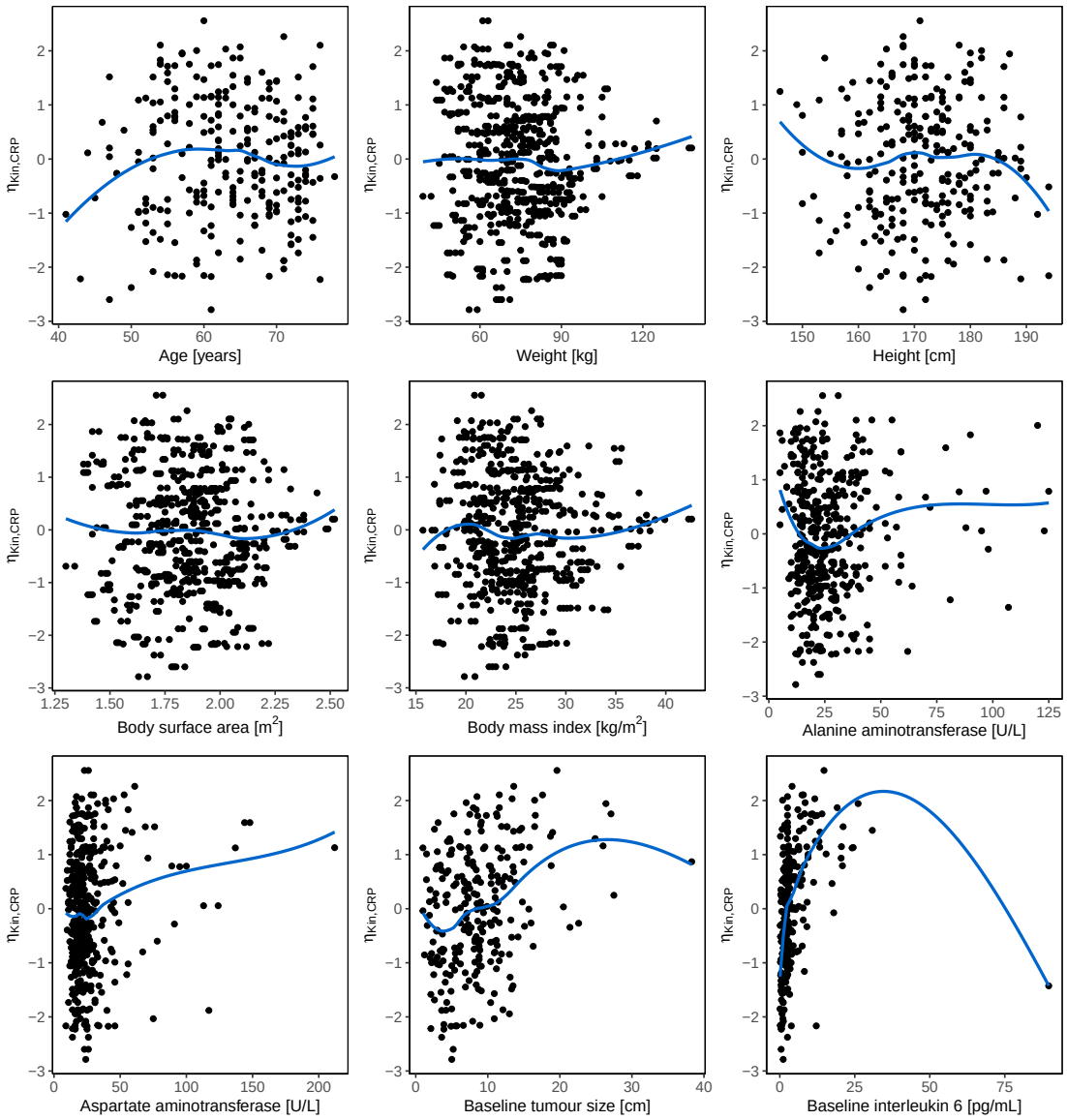

Figure B.5.1: Exploratory analyses of potential influence of different continuous covariates on interindividual variability of C-reactive protein production rate constant ($K_{in,CRP}$).
$\eta_{Kin,CRP}$: individual variability of $K_{in,CRP}$; blue line: loess regression line.
Note that the outlying baseline interleukin 6 data point was *temporarily* excluded during the covariate analysis to avoid distortion of the parameter-covariate relationship.

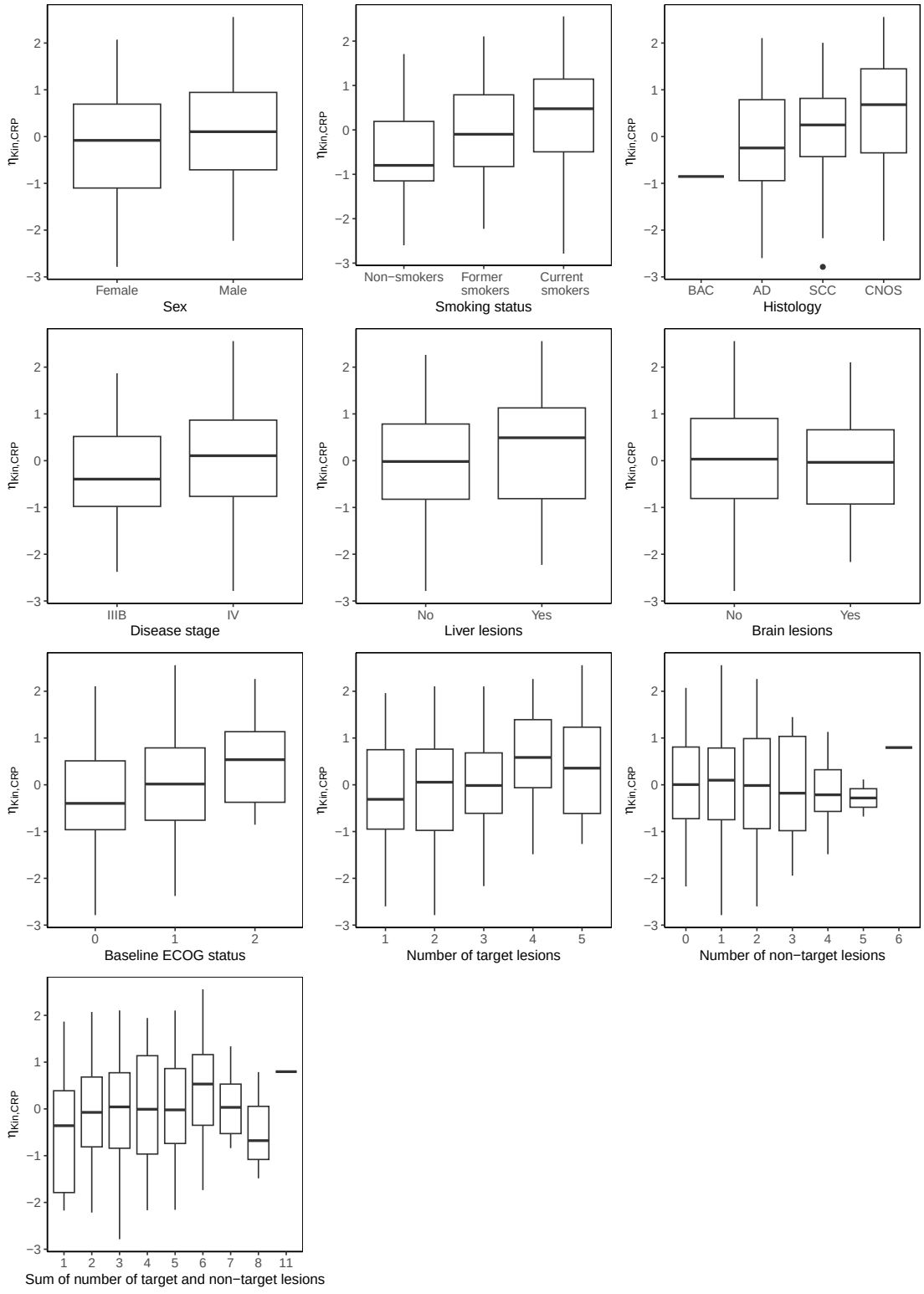

Figure B.5.2: Exploratory analyses of potential influence of different categorical covariates on interindividual variability of of C-reactive protein production rate constant ($K_{in,CRP}$).
$\eta_{Kin,CRP}$: individual variability of $K_{in,CRP}$; AD: adenocarcinoma; BAC: bronchioalveolar carcinoma; CNOS: carcinoma, not otherwise specified; ECOG: Eastern Cooperative Oncology Group; SCC: squamous-cell carcinoma. Boxes: interquartile range (IQR) including median; upper hinge: 25^{th} percentile, lower hinge: 75^{th} percentile; whiskers: range from box hinge to lowest/highest value within 1.5 IQR; data points: data beyond the whiskers.

B.6 Correlation matrix between pre-selected covariates tested in the CRP turnover model

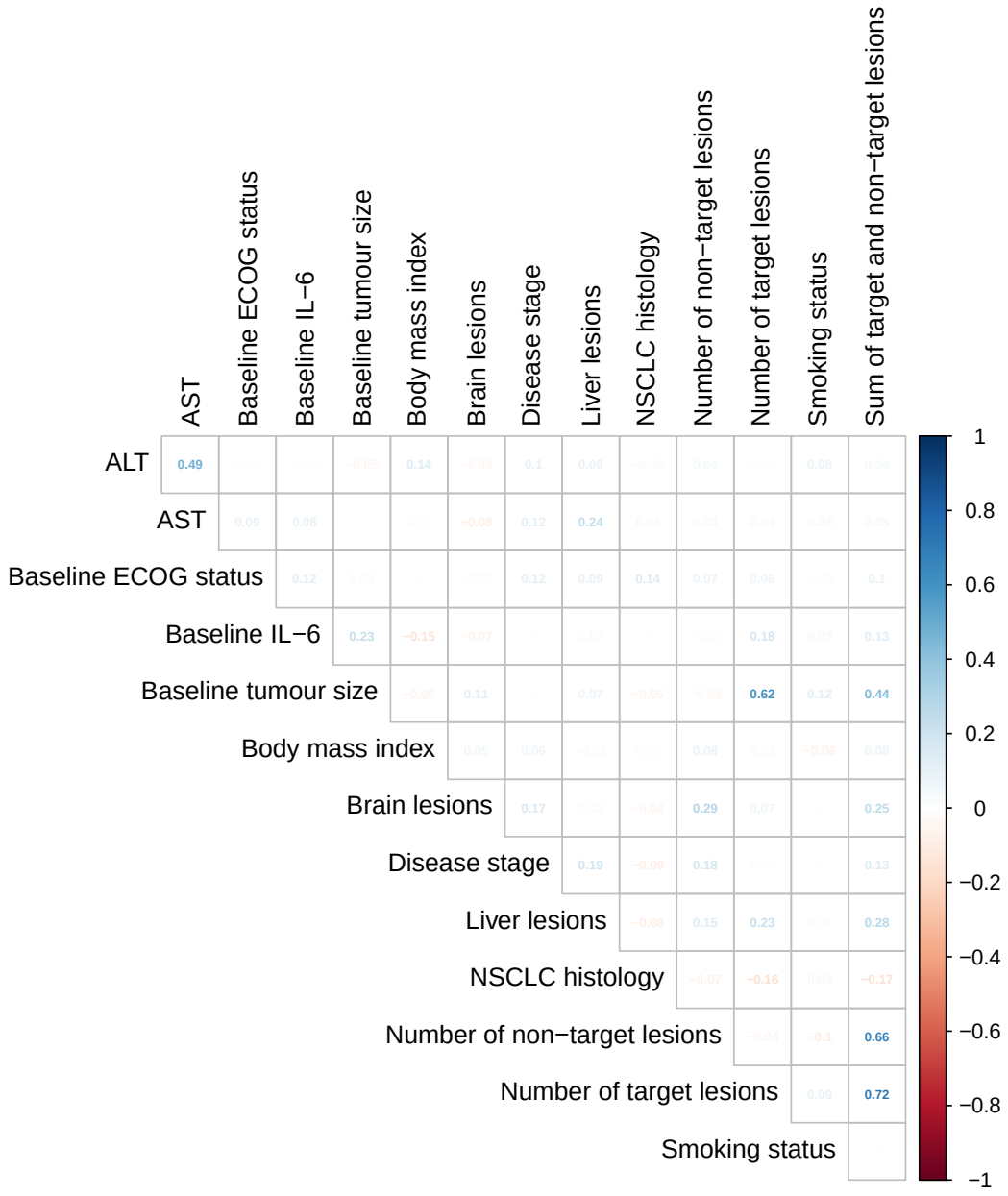

Figure B.6.1: Correlation matrix between pre-selected covariates tested in the CRP turnover model. Positive correlations (dark blue) were observed between number of target lesions and baseline tumour size ($r = 0.62$); number of target lesions and sum of target and non-target lesions ($r = 0.72$); and number of non-target lesions and sum of target and non-target lesions ($r = 0.66$).
ALT: alanine aminotransferase; AST: aspartate aminotransferase; ECOG: Eastern Cooperative Oncology Group; IL-6: interleukin 6; NSCLC: non-small cell lung cancer.
Note: Only strong positive (dark blue) or negative (dark orange) correlations (i.e. $r > \pm 0.6$) were of concern.

B.7 Individual plots from the CRP turnover model

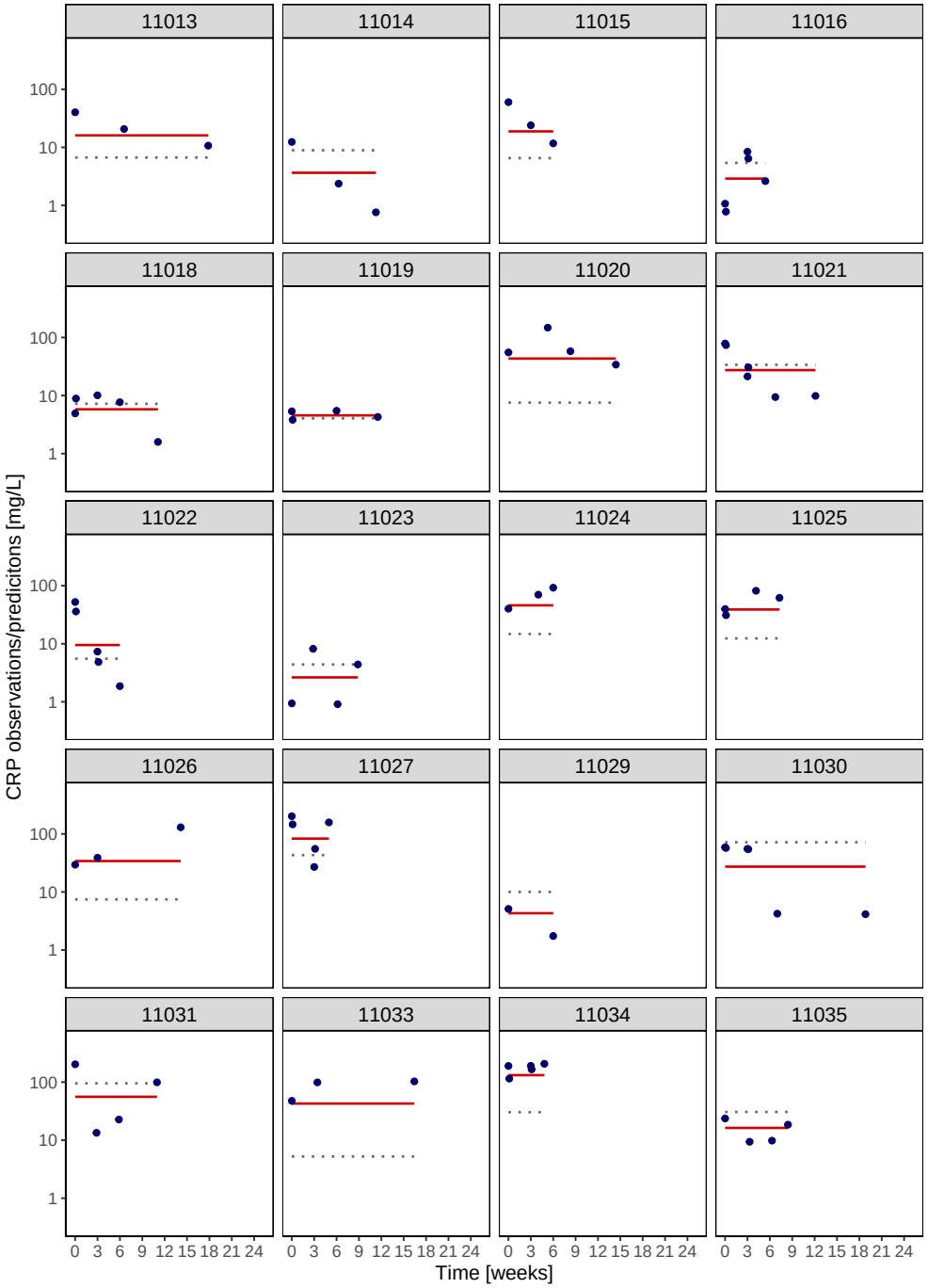

Figure B.7.1: Individual plots of CRP concentrations from the CRP covariate turnover model showing the observations (blue dots), individual predictions (solid red line) and population predictions (dotted grey line) in the logarithmic scale. Each panel represents one patient indicated by the shaded number on top. CRP: C-reactive protein.
Note: results are of the CRP covariate turnover model, hence the steady state predictions.

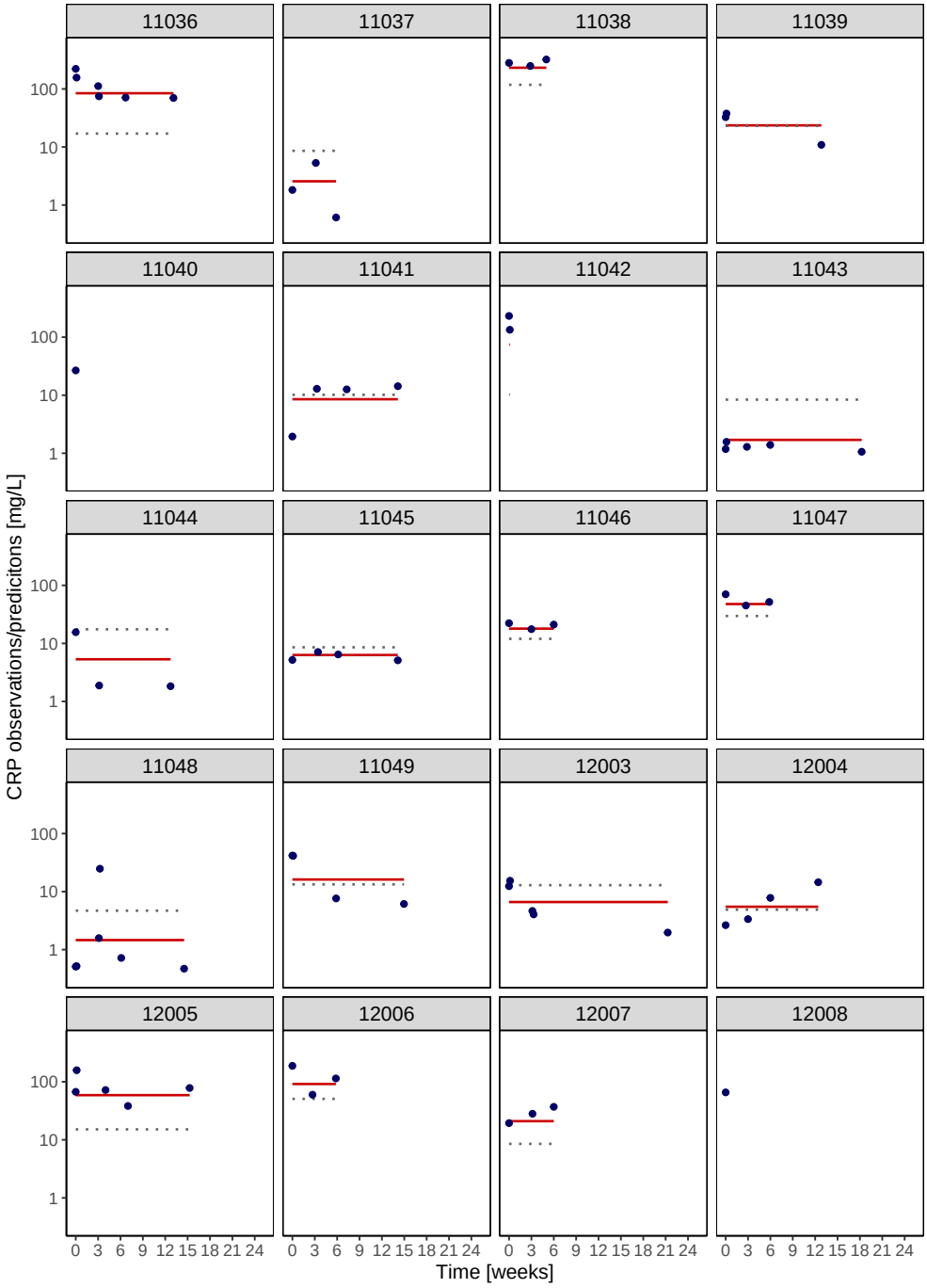

Figure B.7.1: *Cont.* Individual plots of CRP concentrations from the CRP covariate turnover model showing the observations (blue dots), individual predictions (solid red line) and population predictions (dotted grey line) in the logarithmic scale. Each panel represents one patient indicated by the shaded number on top. CRP: C-reactive protein.
Note: results are of the CRP covariate turnover model, hence the steady state predictions.

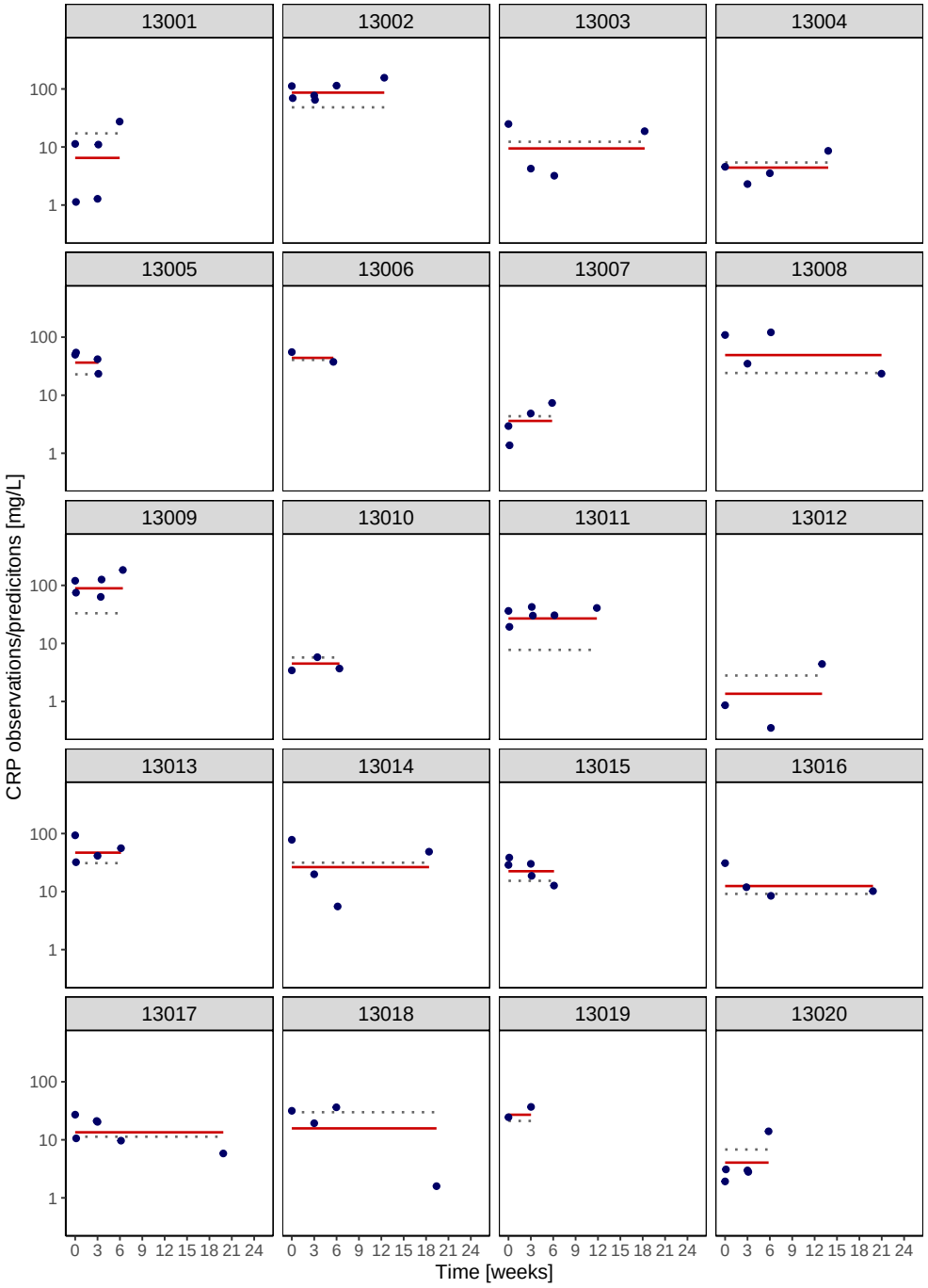

Figure B.7.1: *Cont.* Individual plots of CRP concentrations from the CRP covariate turnover model showing the observations (blue dots), individual predictions (solid red line) and population predictions (dotted grey line) in the logarithmic scale. Each panel represents one patient indicated by the shaded number on top. CRP: C-reactive protein.
Note: results are of the CRP covariate turnover model, hence the steady state predictions.

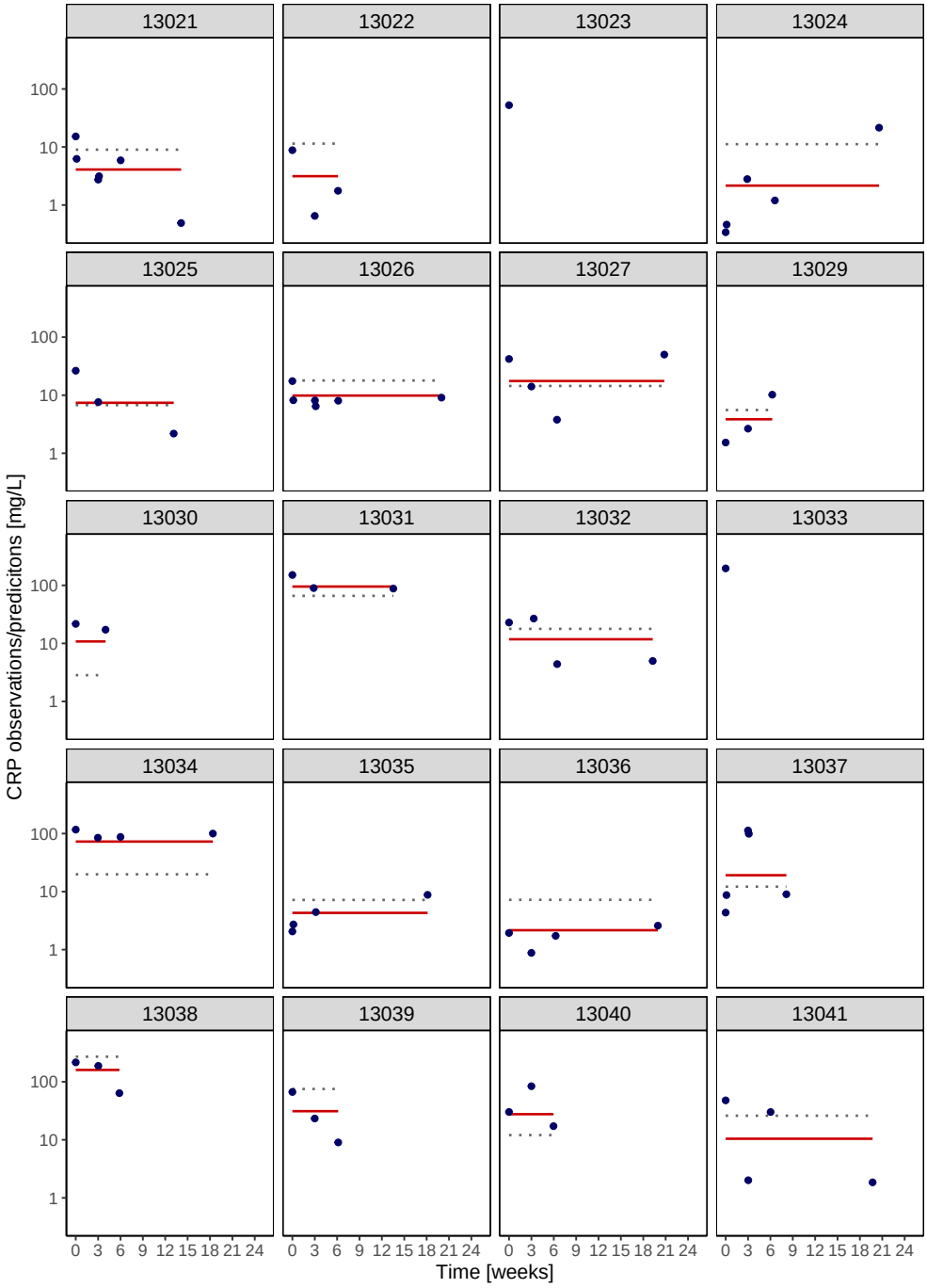

Figure B.7.1: *Cont.* Individual plots of CRP concentrations from the CRP covariate turnover model showing the observations (blue dots), individual predictions (solid red line) and population predictions (dotted grey line) in the logarithmic scale. Each panel represents one patient indicated by the shaded number on top. CRP: C-reactive protein.
Note: results are of the CRP covariate turnover model, hence the steady state predictions.

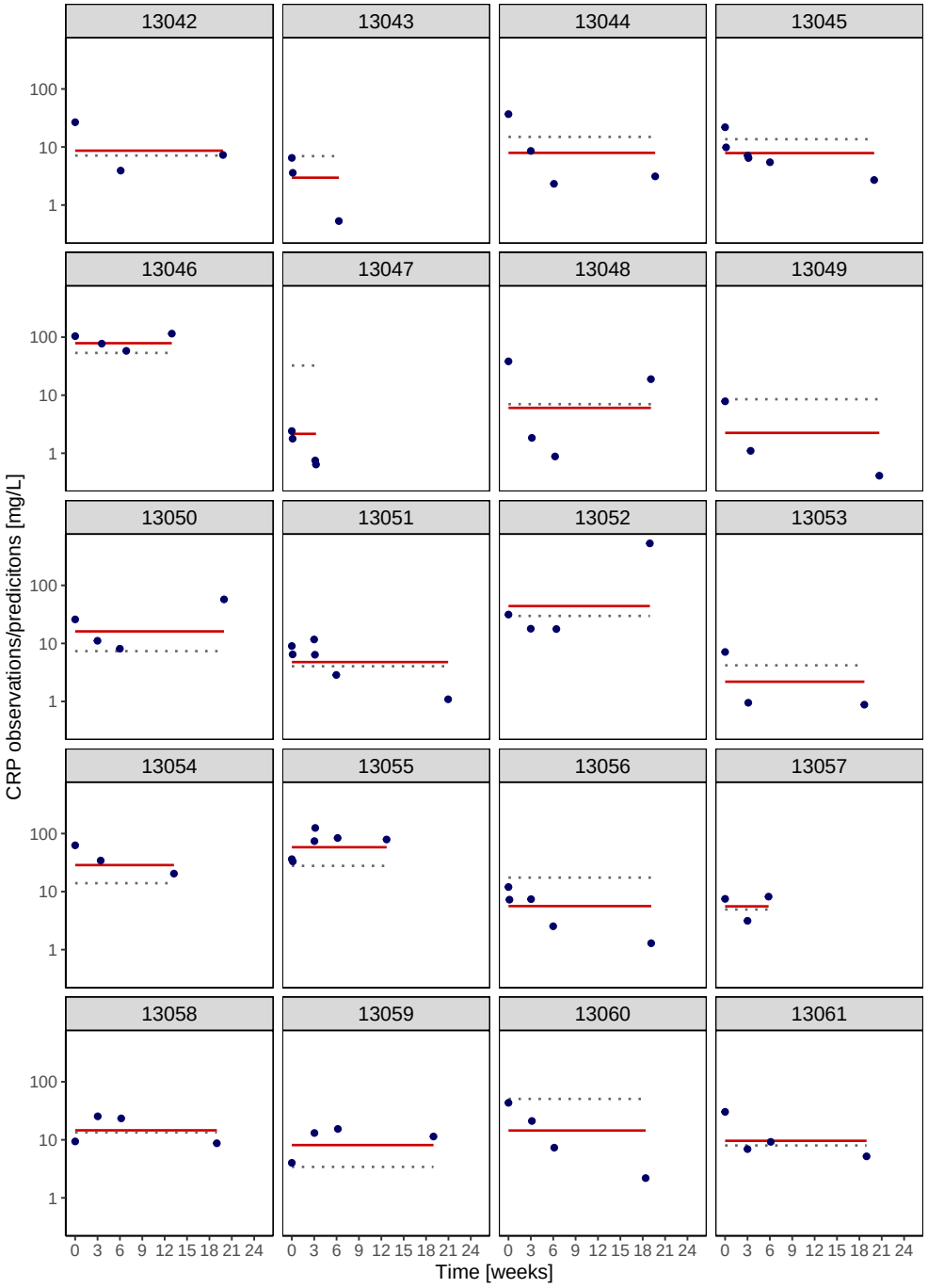

Figure B.7.1: *Cont.* Individual plots of CRP concentrations from the CRP covariate turnover model showing the observations (blue dots), individual predictions (solid red line) and population predictions (dotted grey line) in the logarithmic scale. Each panel represents one patient indicated by the shaded number on top. CRP: C-reactive protein.
Note: results are of the CRP covariate turnover model, hence the steady state predictions.

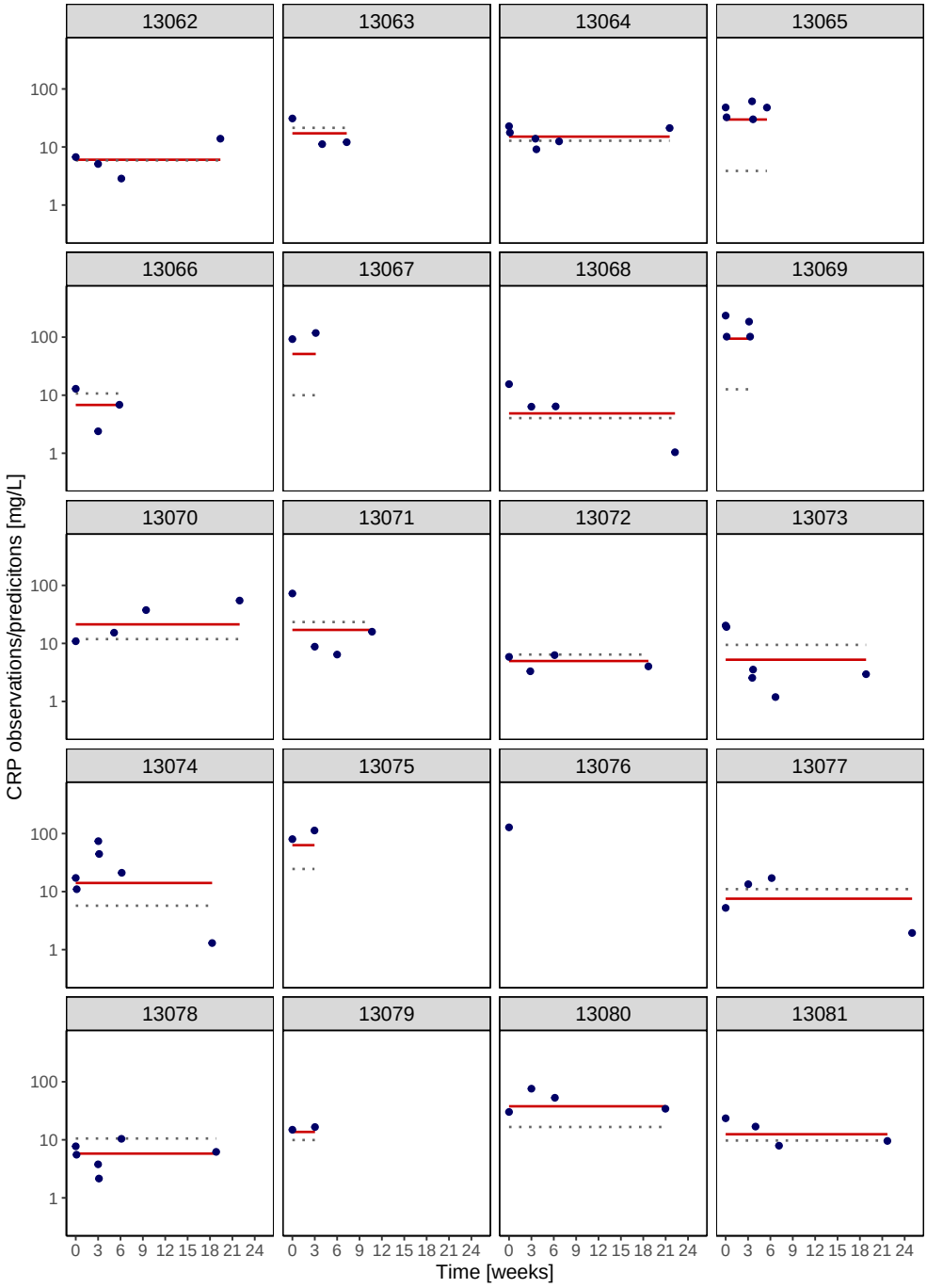

Figure B.7.1: *Cont.* Individual plots of CRP concentrations from the CRP covariate turnover model showing the observations (blue dots), individual predictions (solid red line) and population predictions (dotted grey line) in the logarithmic scale. Each panel represents one patient indicated by the shaded number on top. CRP: C-reactive protein.
Note: results are of the CRP covariate turnover model, hence the steady state predictions.

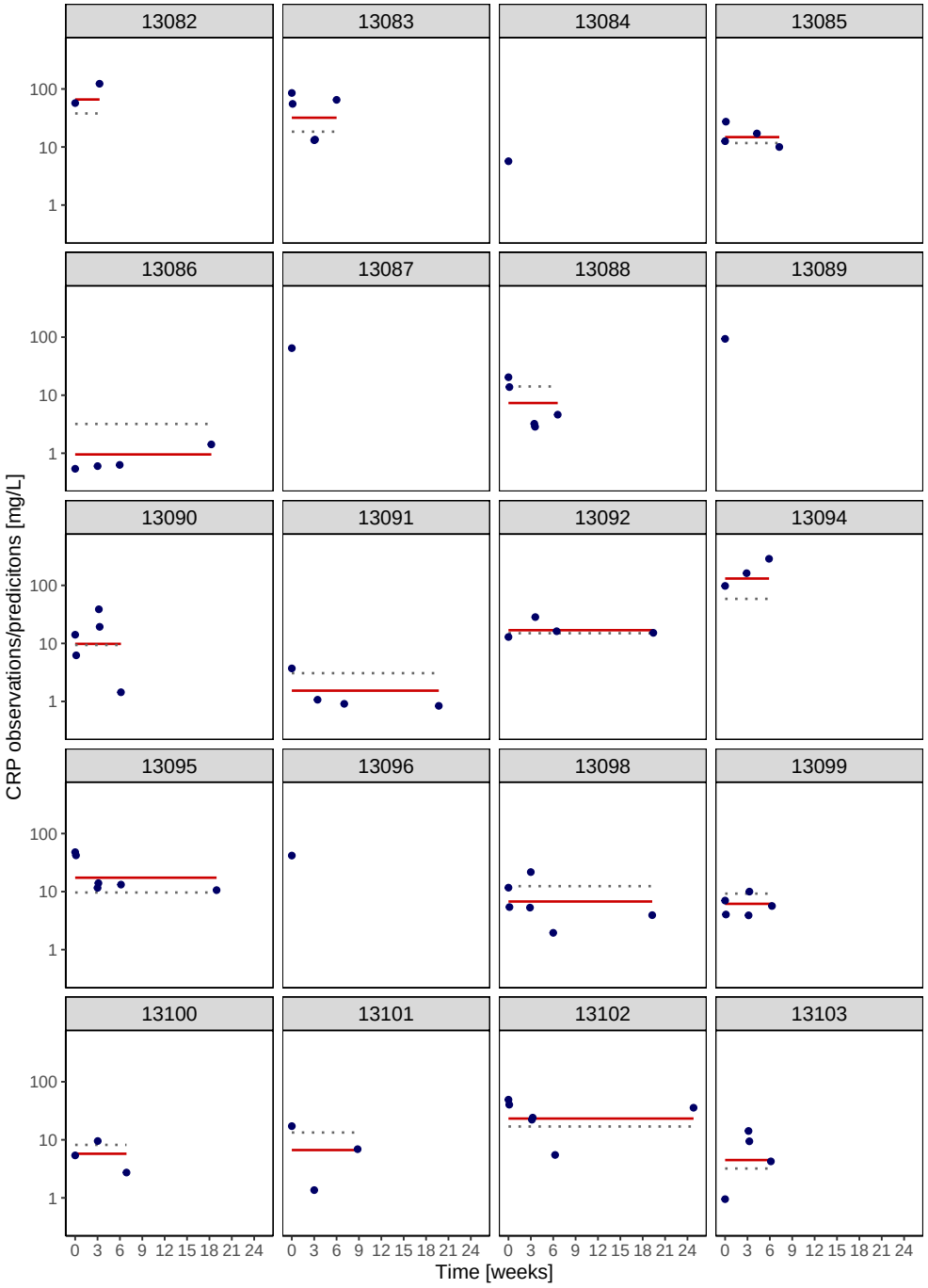

Figure B.7.1: *Cont.* Individual plots of CRP concentrations from the CRP covariate turnover model showing the observations (blue dots), individual predictions (solid red line) and population predictions (dotted grey line) in the logarithmic scale. Each panel represents one patient indicated by the shaded number on top. CRP: C-reactive protein.
Note: results are of the CRP covariate turnover model, hence the steady state predictions.

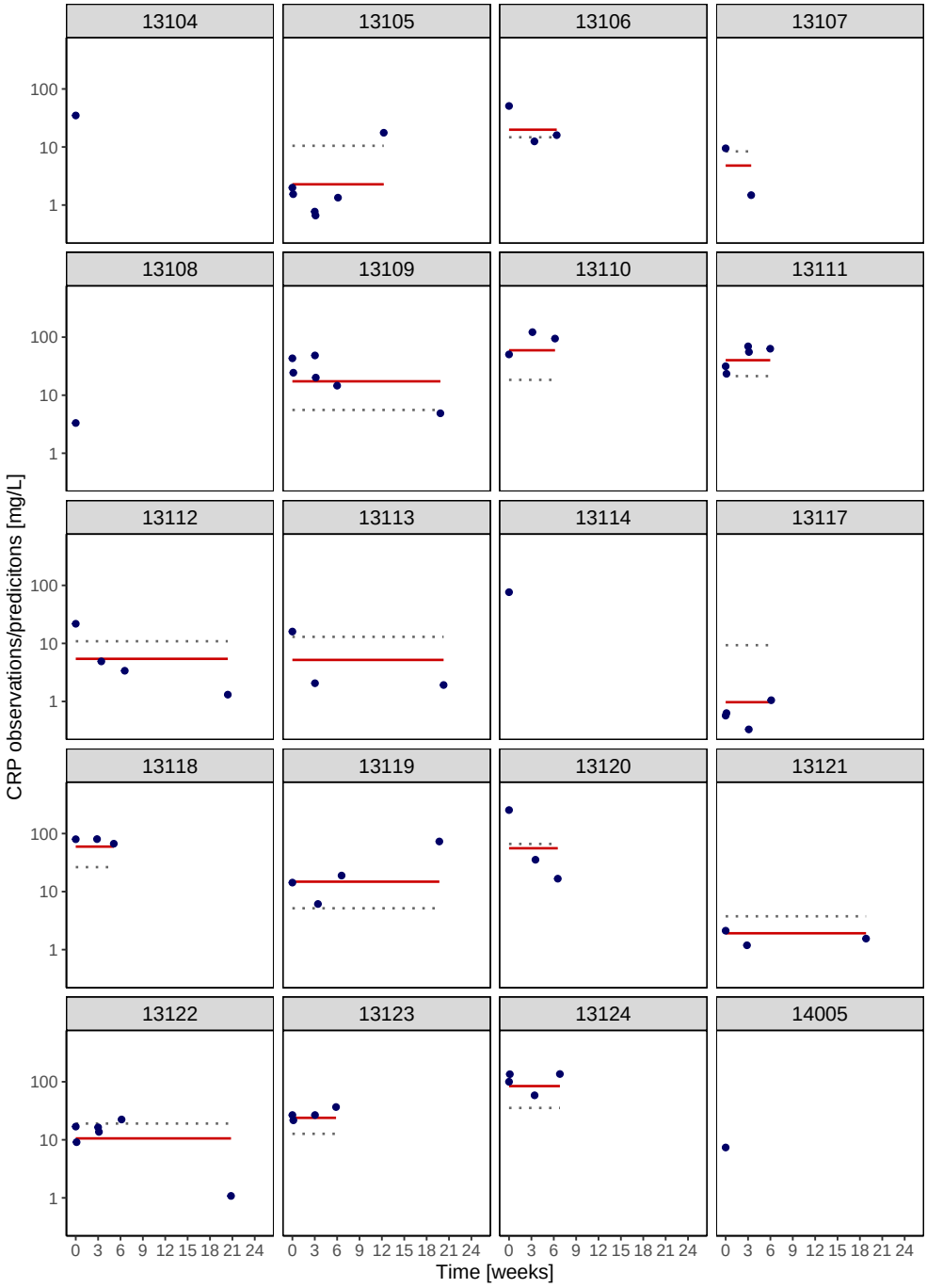

Figure B.7.1: *Cont.* Individual plots of CRP concentrations from the CRP covariate turnover model showing the observations (blue dots), individual predictions (solid red line) and population predictions (dotted grey line) in the logarithmic scale. Each panel represents one patient indicated by the shaded number on top. CRP: C-reactive protein.
Note: results are of the CRP covariate turnover model, hence the steady state predictions.

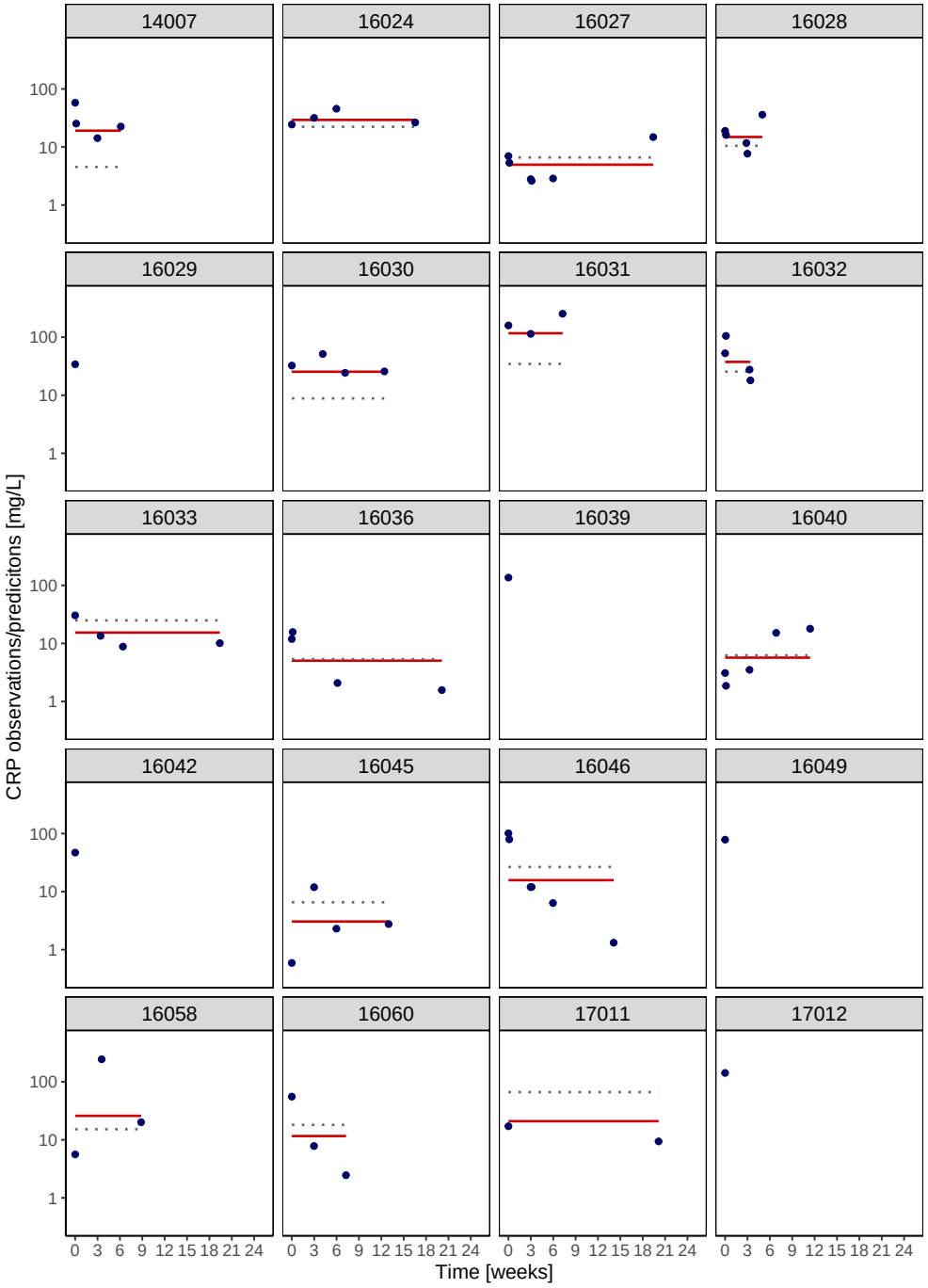

Figure B.7.1: *Cont.* Individual plots of CRP concentrations from the CRP covariate turnover model showing the observations (blue dots), individual predictions (solid red line) and population predictions (dotted grey line) in the logarithmic scale. Each panel represents one patient indicated by the shaded number on top. CRP: C-reactive protein.
Note: results are of the CRP covariate turnover model, hence the steady state predictions.

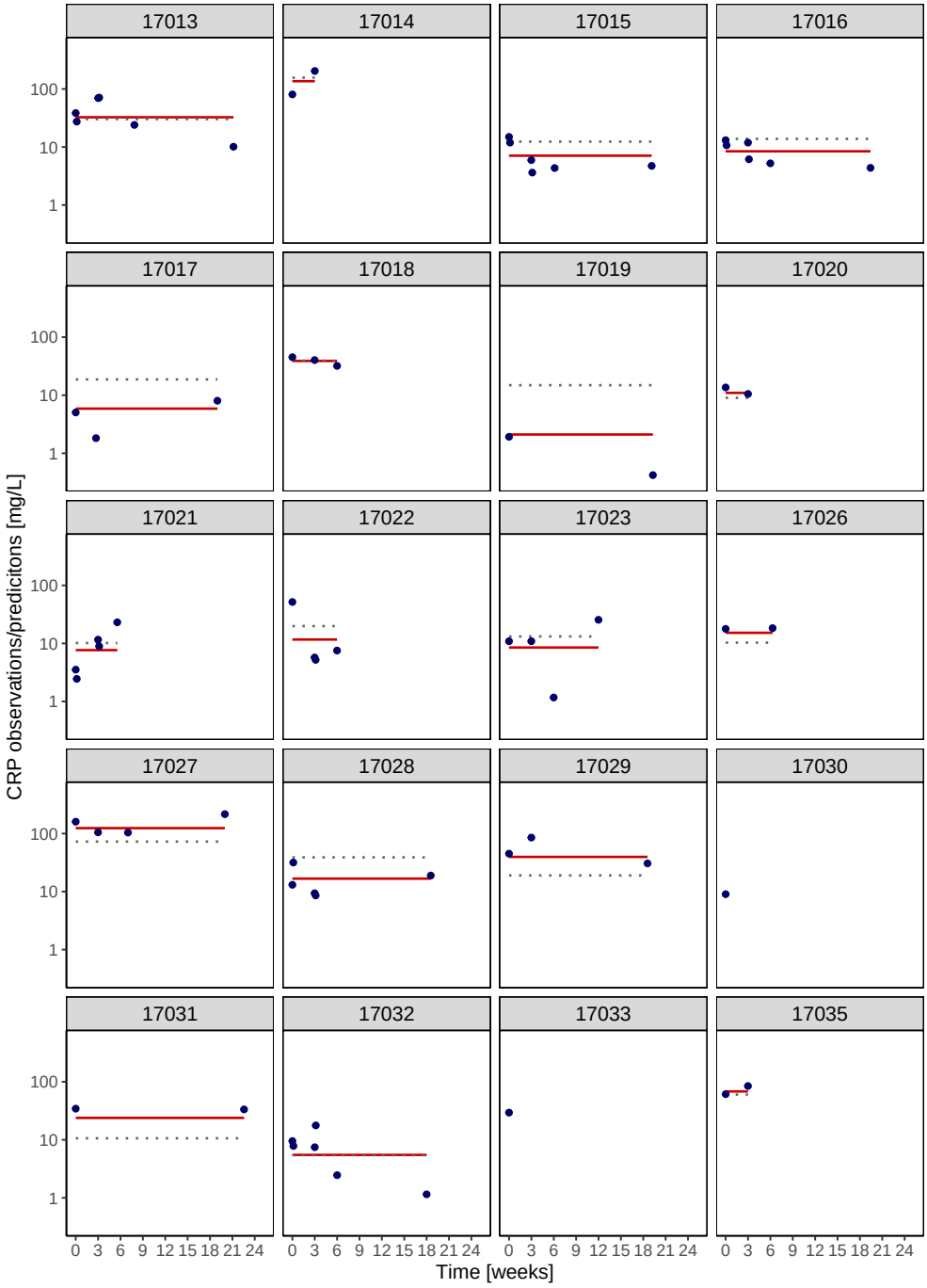

Figure B.7.1: *Cont.* Individual plots of CRP concentrations from the CRP covariate turnover model showing the observations (blue dots), individual predictions (solid red line) and population predictions (dotted grey line) in the logarithmic scale. Each panel represents one patient indicated by the shaded number on top. CRP: C-reactive protein.
Note: results are of the CRP covariate turnover model, hence the steady state predictions.

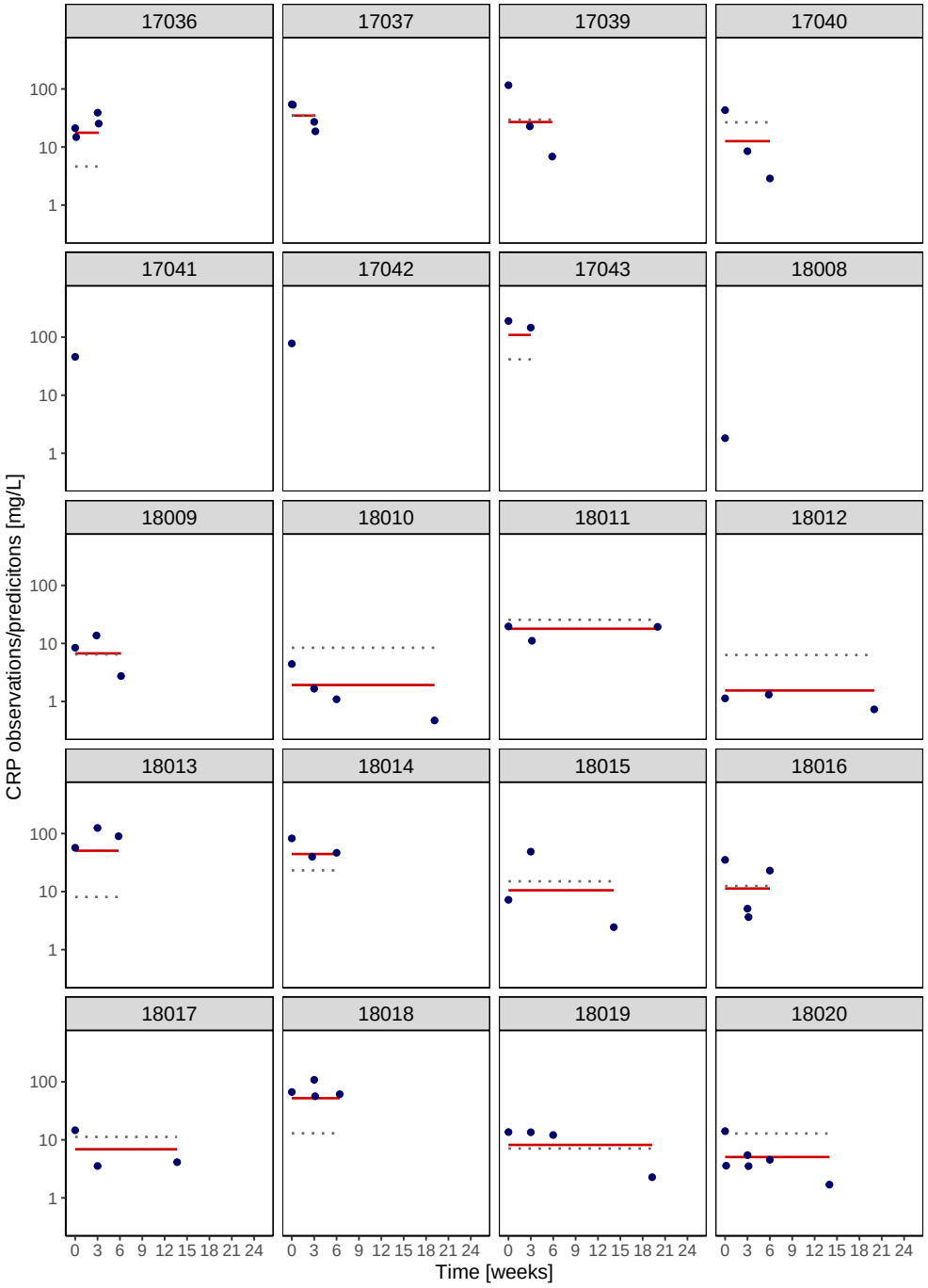

Figure B.7.1: *Cont.* Individual plots of CRP concentrations from the CRP covariate turnover model showing the observations (blue dots), individual predictions (solid red line) and population predictions (dotted grey line) in the logarithmic scale. Each panel represents one patient indicated by the shaded number on top. CRP: C-reactive protein.
Note: results are of the CRP covariate turnover model, hence the steady state predictions.

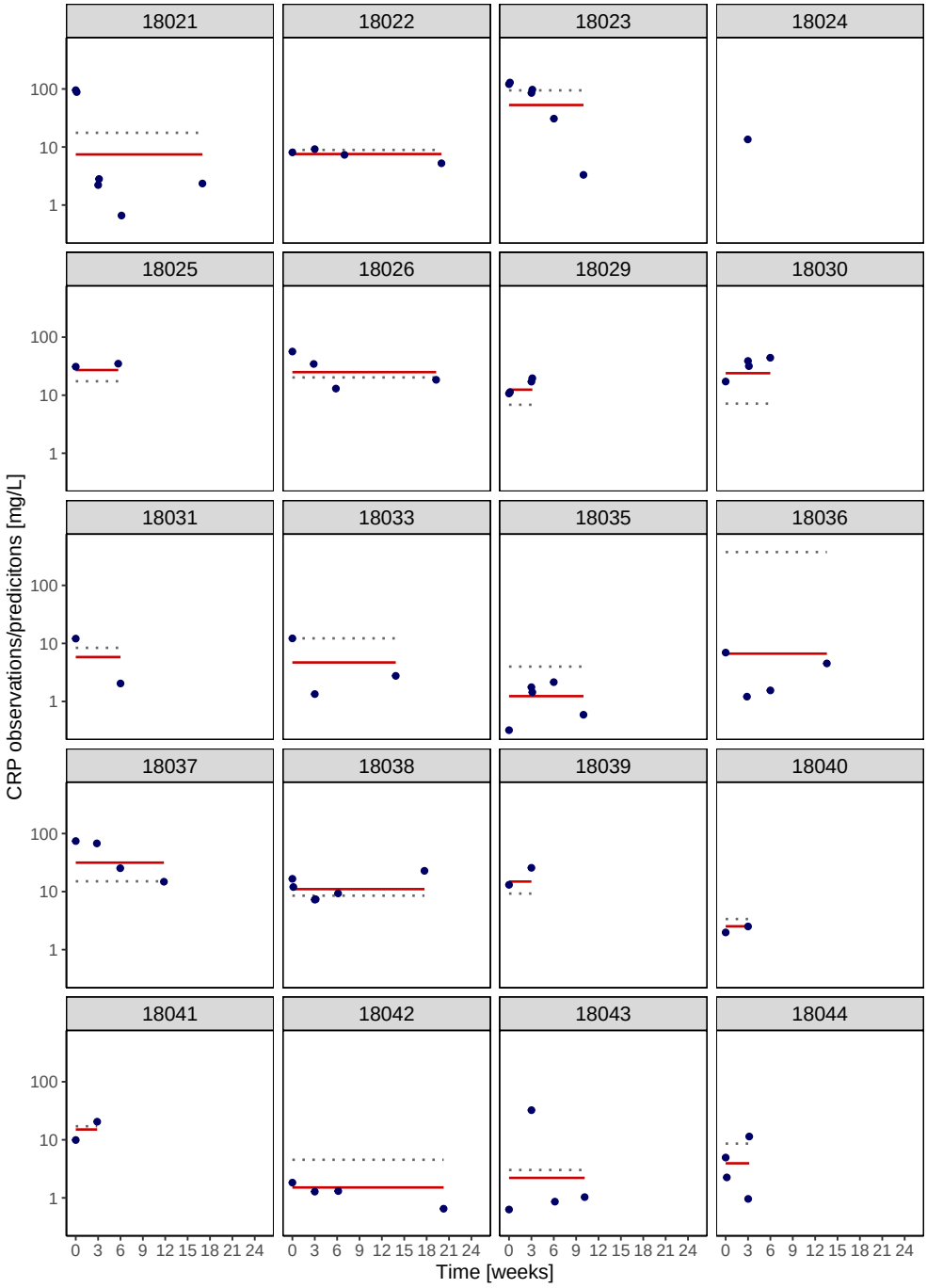

Figure B.7.1: *Cont.* Individual plots of CRP concentrations from the CRP covariate turnover model showing the observations (blue dots), individual predictions (solid red line) and population predictions (dotted grey line) in the logarithmic scale. Each panel represents one patient indicated by the shaded number on top. CRP: C-reactive protein.
Note: results are of the CRP covariate turnover model, hence the steady state predictions.

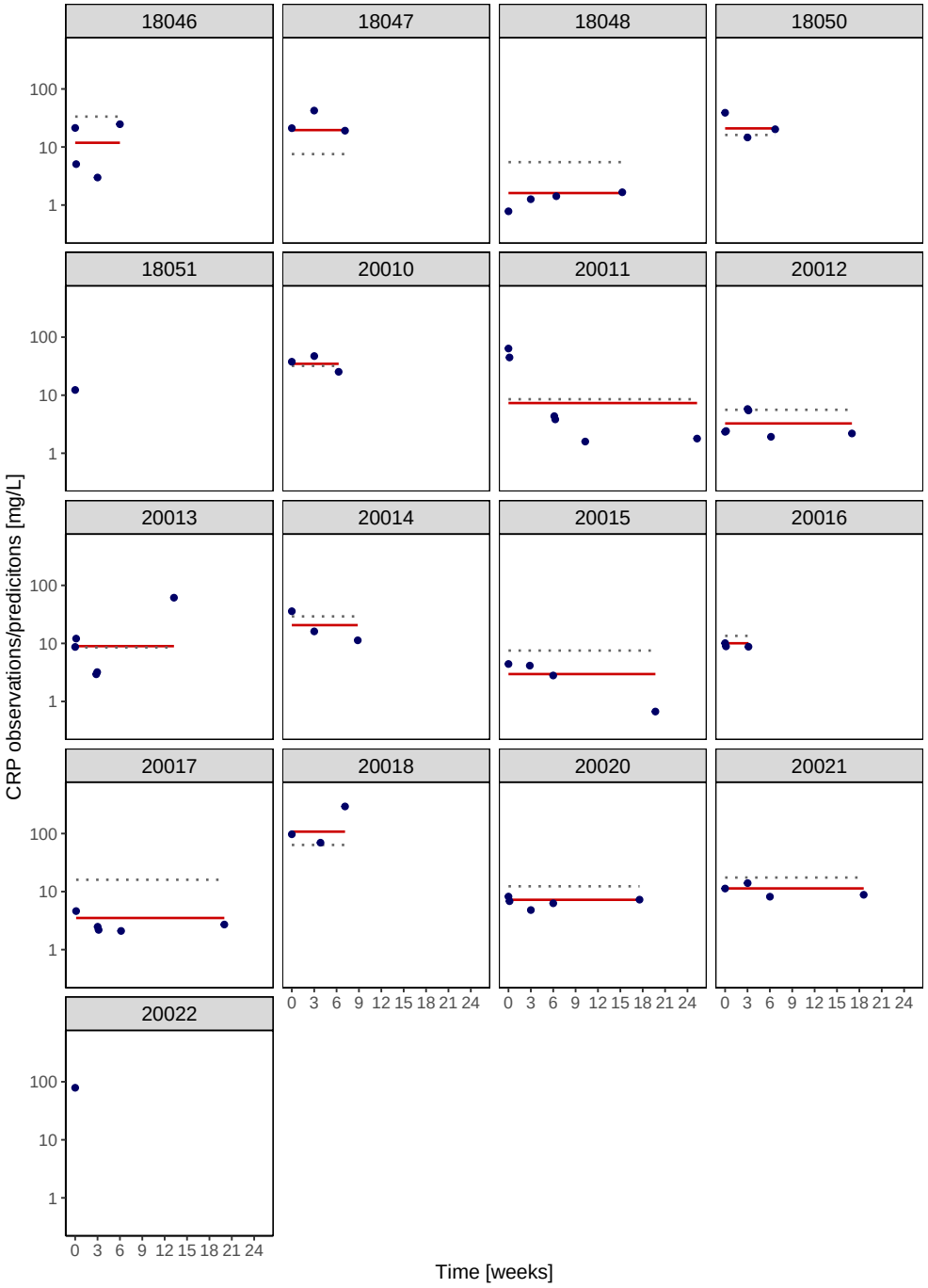

Figure B.7.1: *Cont.* Individual plots of CRP concentrations from the CRP covariate turnover model showing the observations (blue dots), individual predictions (solid red line) and population predictions (dotted grey line) in the logarithmic scale. Each panel represents one patient indicated by the shaded number on top. CRP: C-reactive protein.
Note: results are of the CRP covariate turnover model, hence the steady state predictions.

B.8 Individual plots from the coupled tumour dynamics - CRP model

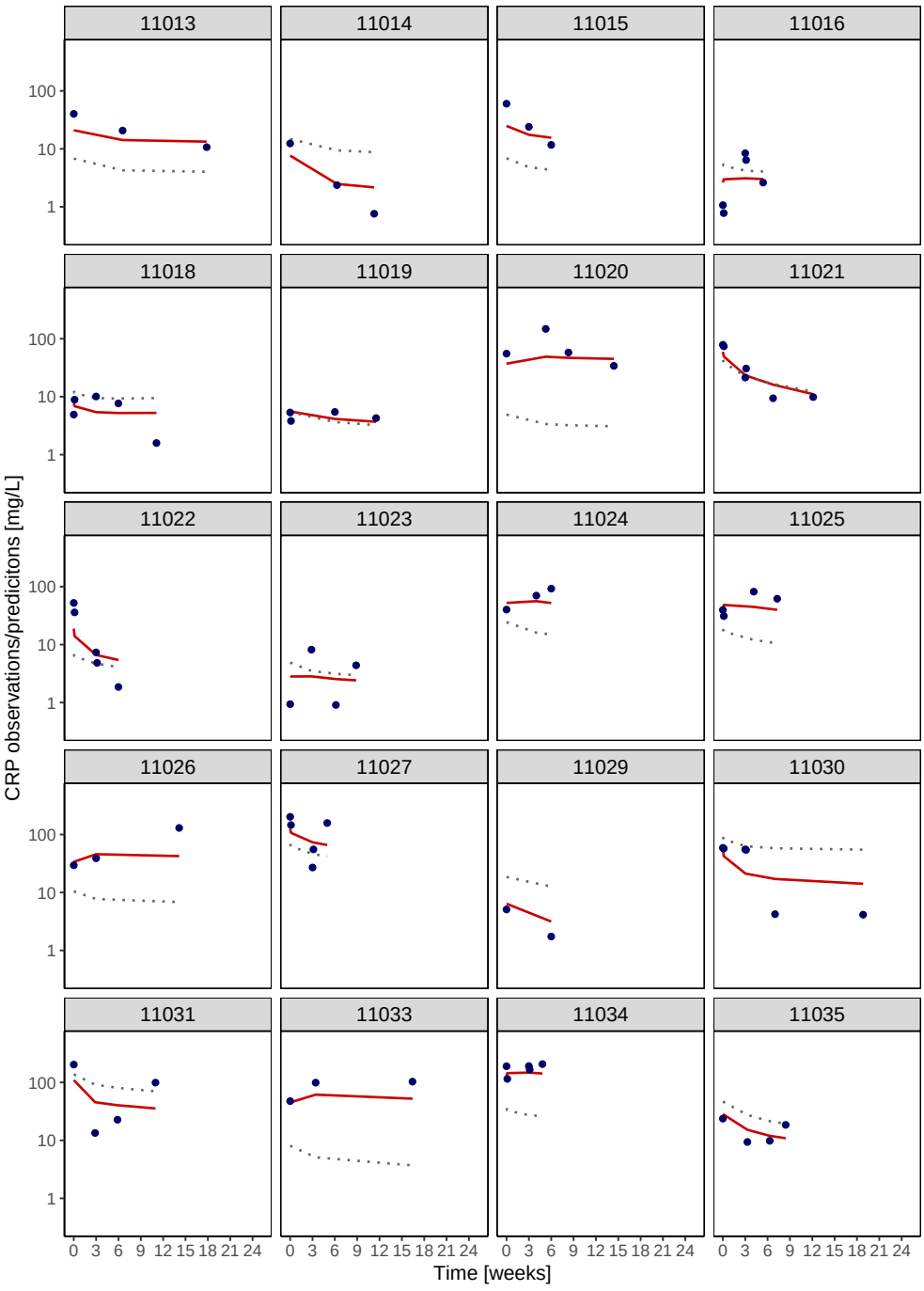

Figure B.8.1: Individual plots of CRP concentrations from the coupled tumour dynamics - CRP model showing the observations (blue dots), individual predictions (solid red line) and population predictions (dotted grey line) in the logarithmic scale. Each panel represents one patient indicated by the shaded number on top. CRP: C-reactive protein.
Note: results are of the coupled tumour dynamics - CRP model hence the deviation from the steady state predictions.

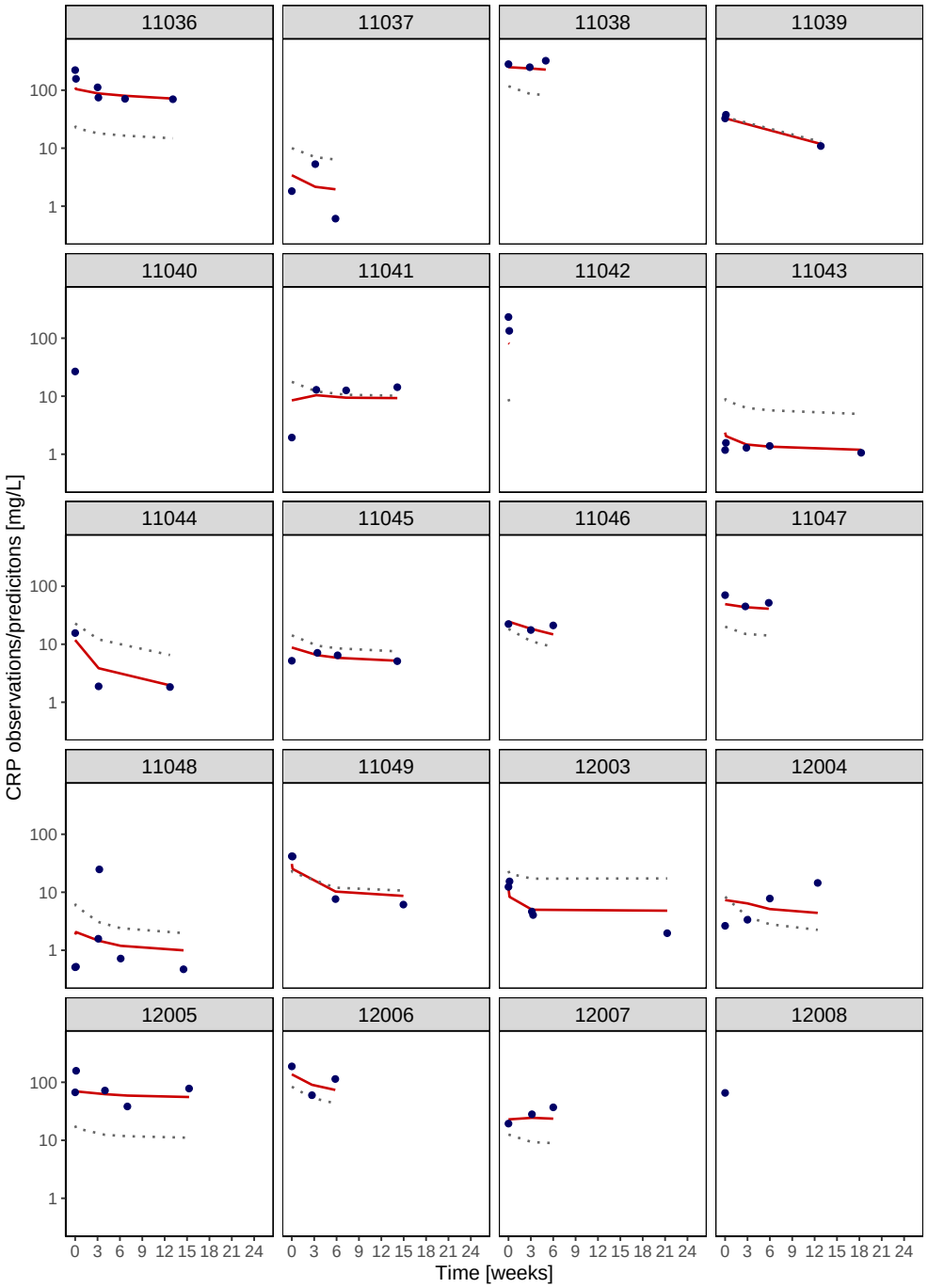

Figure B.8.1: *Cont.* Individual plots of CRP concentrations from the coupled tumour dynamics - CRP model showing the observations (blue dots), individual predictions (solid red line) and population predictions (dotted grey line) in the logarithmic scale. Each panel represents one patient indicated by the shaded number on top. CRP: C-reactive protein.
Note: results are of the coupled tumour dynamics - CRP model hence the deviation from the steady state predictions.

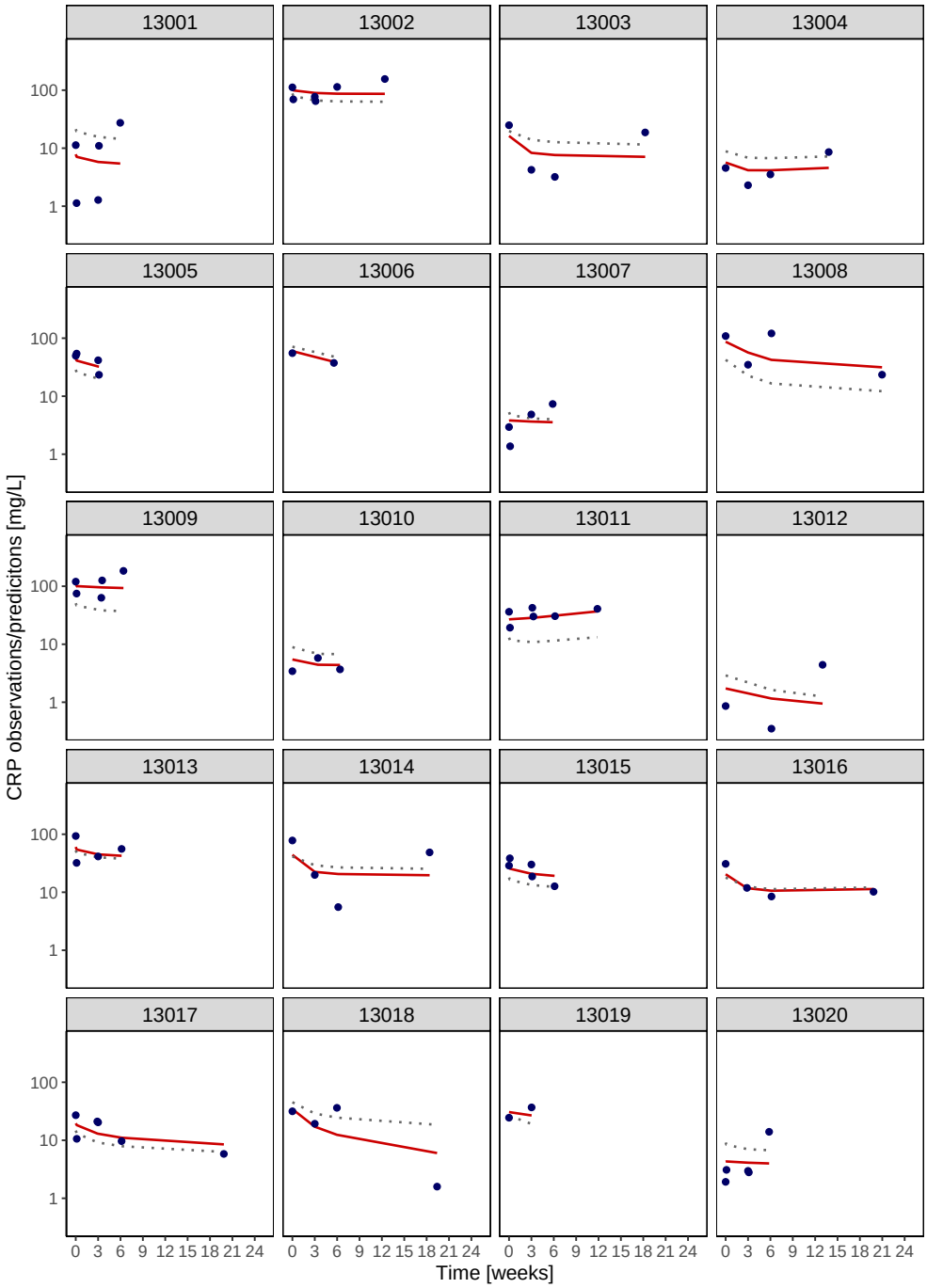

Figure B.8.1: *Cont.* Individual plots of CRP concentrations from the coupled tumour dynamics - CRP model showing the observations (blue dots), individual predictions (solid red line) and population predictions (dotted grey line) in the logarithmic scale. Each panel represents one patient indicated by the shaded number on top. CRP: C-reactive protein.
Note: results are of the coupled tumour dynamics - CRP model hence the deviation from the steady state predictions.

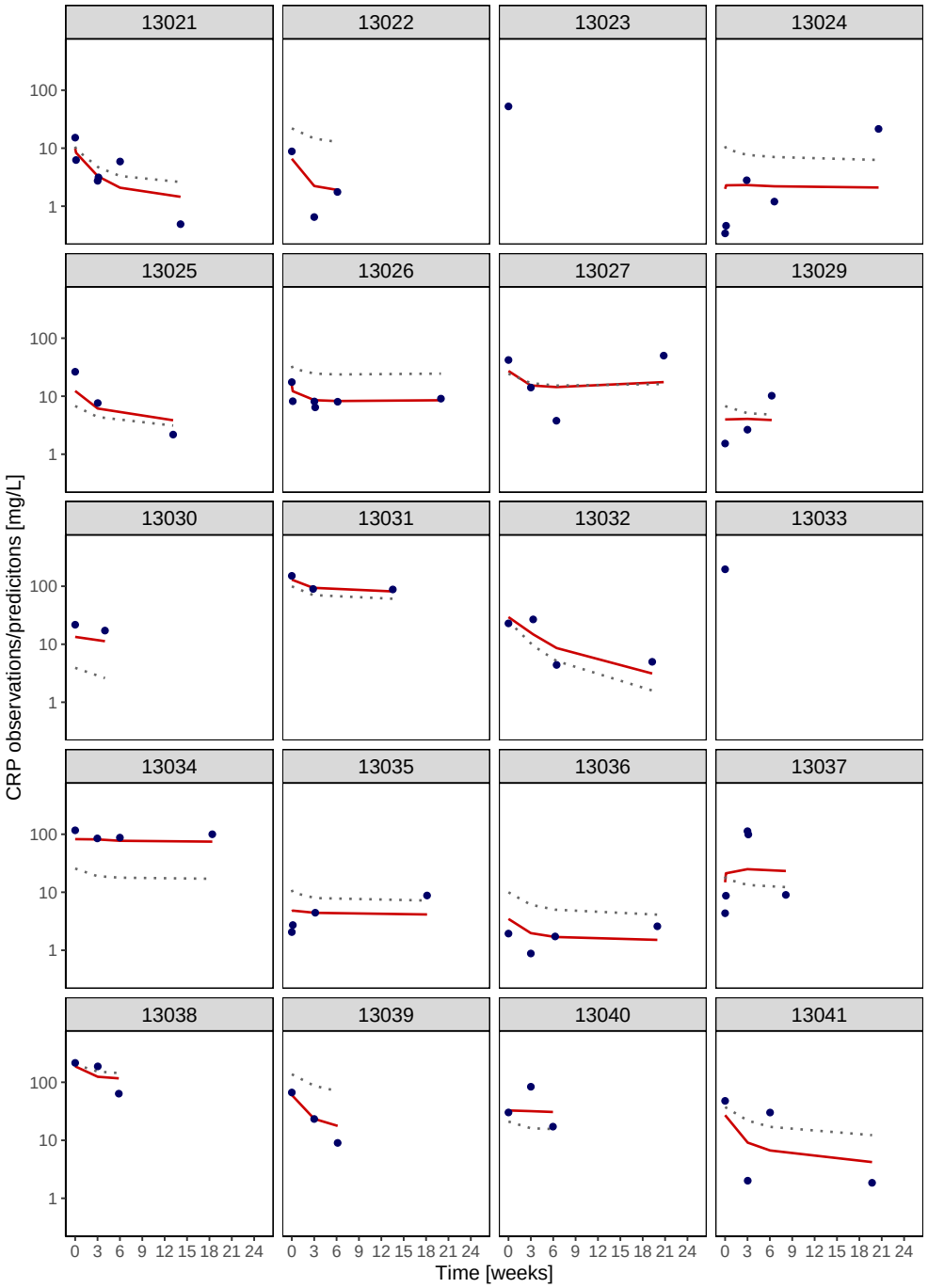

Figure B.8.1: *Cont.* Individual plots of CRP concentrations from the coupled tumour dynamics - CRP model showing the observations (blue dots), individual predictions (solid red line) and population predictions (dotted grey line) in the logarithmic scale. Each panel represents one patient indicated by the shaded number on top. CRP: C-reactive protein.
Note: results are of the coupled tumour dynamics - CRP model hence the deviation from the steady state predictions.

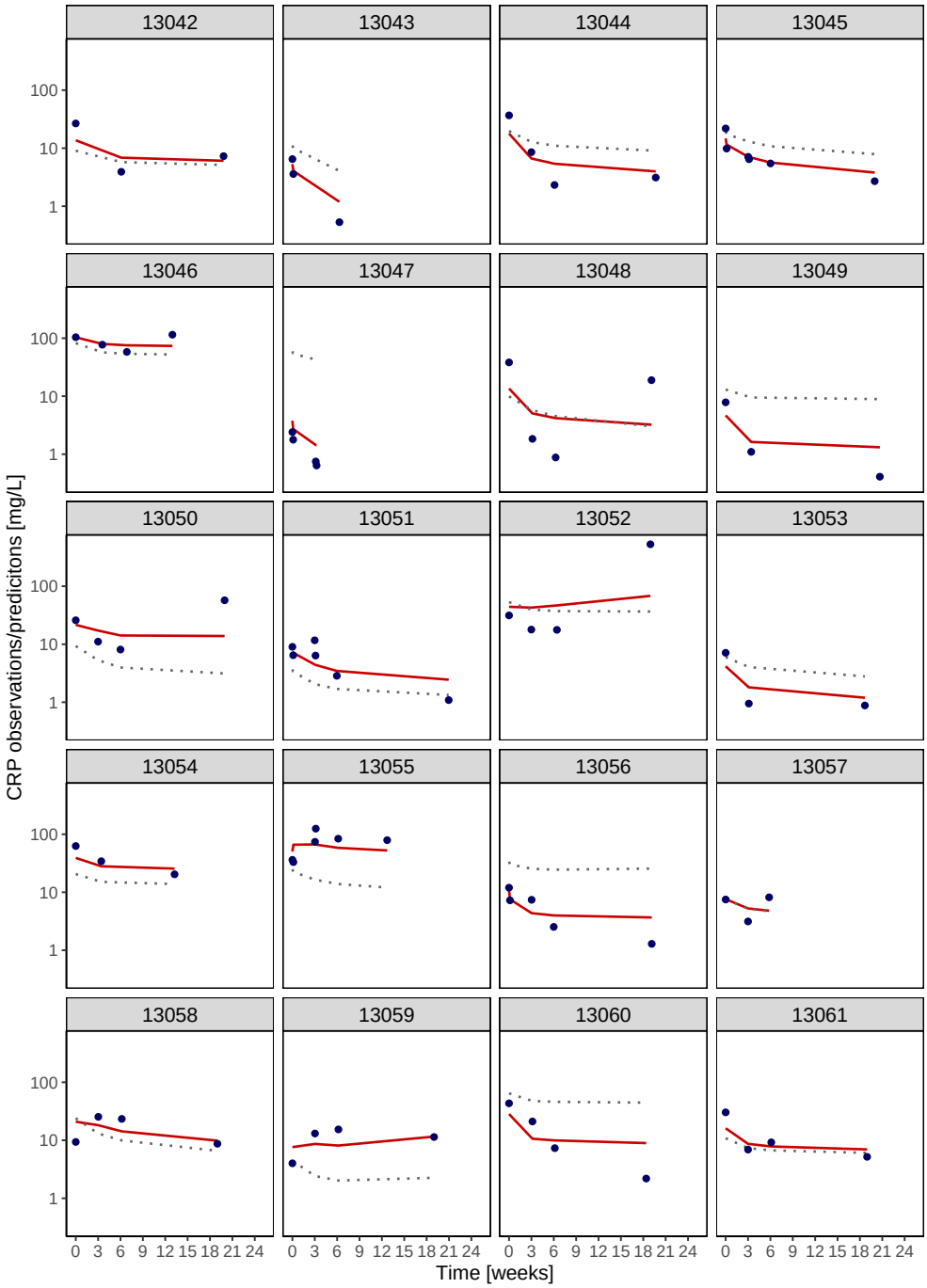

Figure B.8.1: *Cont.* Individual plots of CRP concentrations from the coupled tumour dynamics - CRP model showing the observations (blue dots), individual predictions (solid red line) and population predictions (dotted grey line) in the logarithmic scale. Each panel represents one patient indicated by the shaded number on top. CRP: C-reactive protein.
Note: results are of the coupled tumour dynamics - CRP model hence the deviation from the steady state predictions.

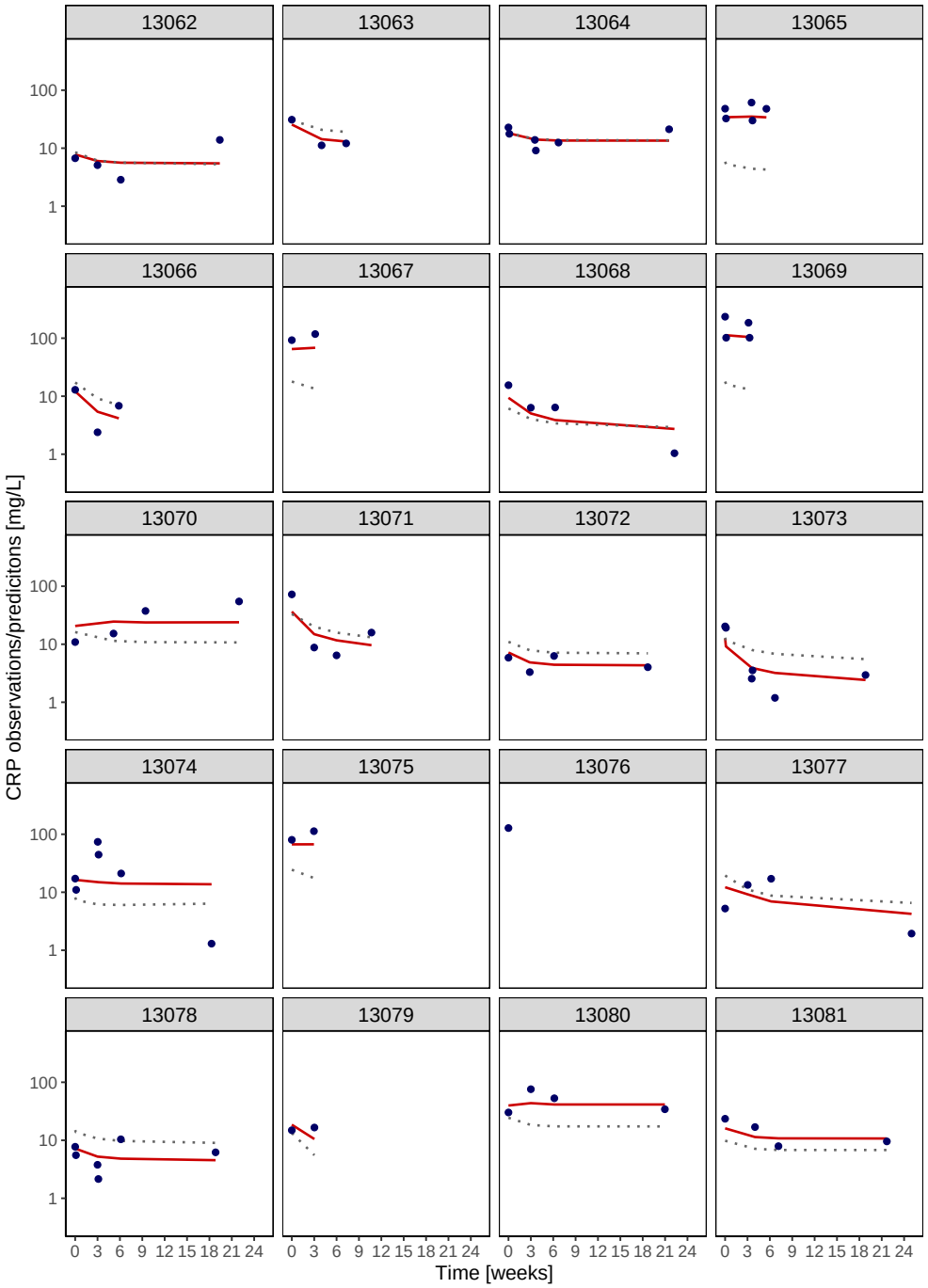

Figure B.8.1: *Cont.* Individual plots of CRP concentrations from the coupled tumour dynamics - CRP model showing the observations (blue dots), individual predictions (solid red line) and population predictions (dotted grey line) in the logarithmic scale. Each panel represents one patient indicated by the shaded number on top. CRP: C-reactive protein.
Note: results are of the coupled tumour dynamics - CRP model hence the deviation from the steady state predictions.

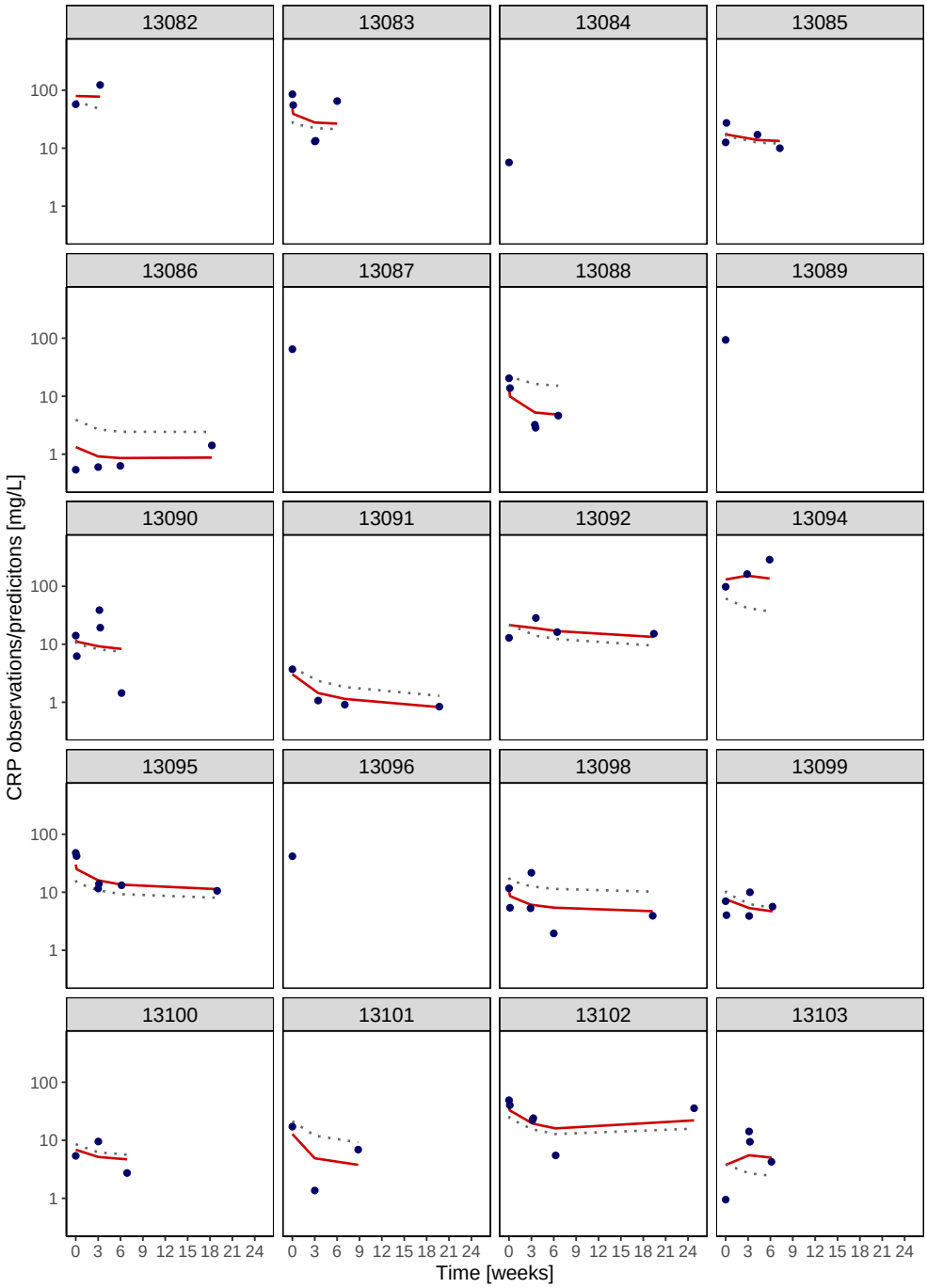

Figure B.8.1: *Cont.* Individual plots of CRP concentrations from the coupled tumour dynamics - CRP model showing the observations (blue dots), individual predictions (solid red line) and population predictions (dotted grey line) in the logarithmic scale. Each panel represents one patient indicated by the shaded number on top. CRP: C-reactive protein.
Note: results are of the coupled tumour dynamics - CRP model hence the deviation from the steady state predictions.

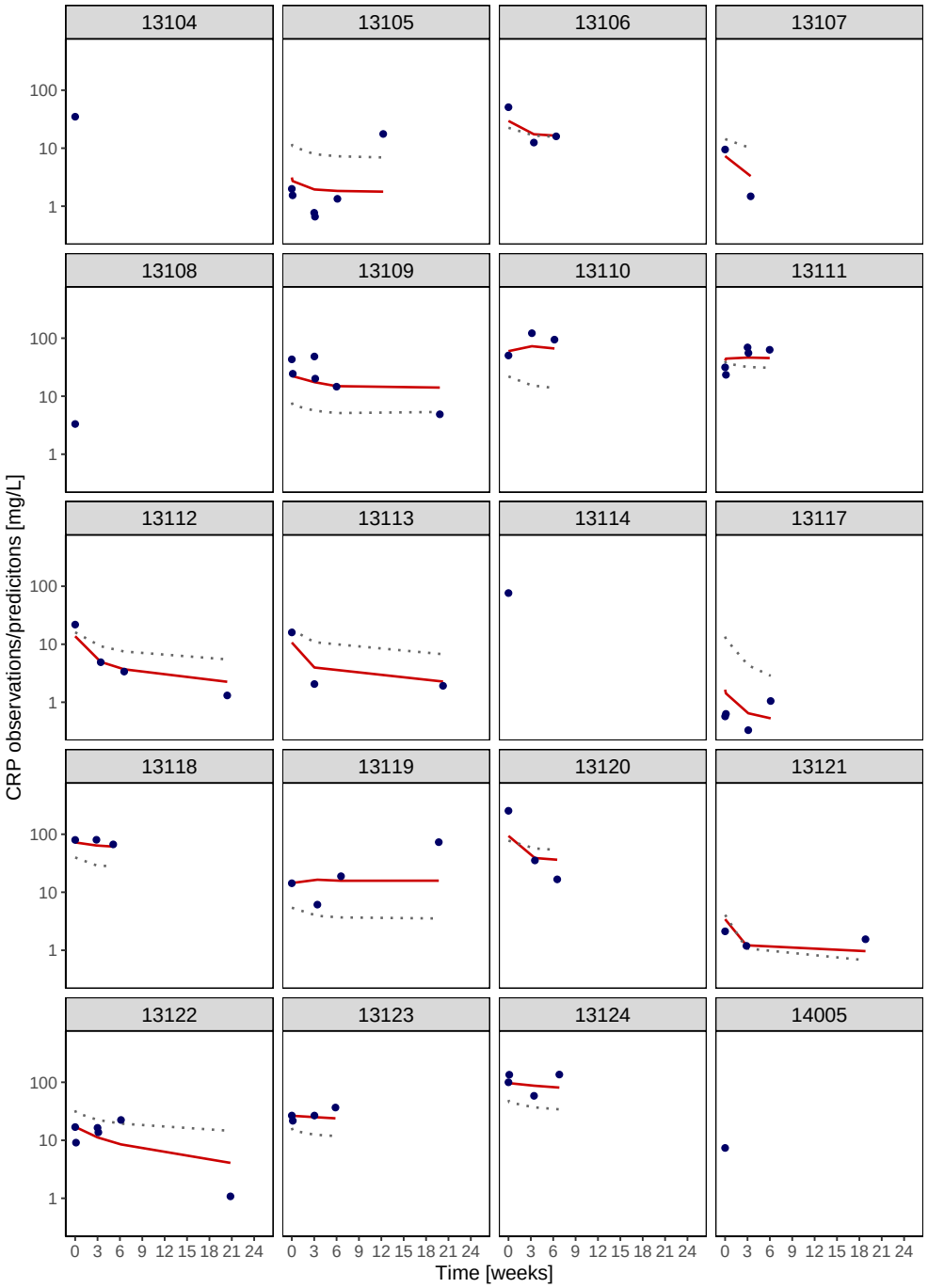

Figure B.8.1: *Cont.* Individual plots of CRP concentrations from the coupled tumour dynamics - CRP model showing the observations (blue dots), individual predictions (solid red line) and population predictions (dotted grey line) in the logarithmic scale. Each panel represents one patient indicated by the shaded number on top. CRP: C-reactive protein.
Note: results are of the coupled tumour dynamics - CRP model hence the deviation from the steady state predictions.

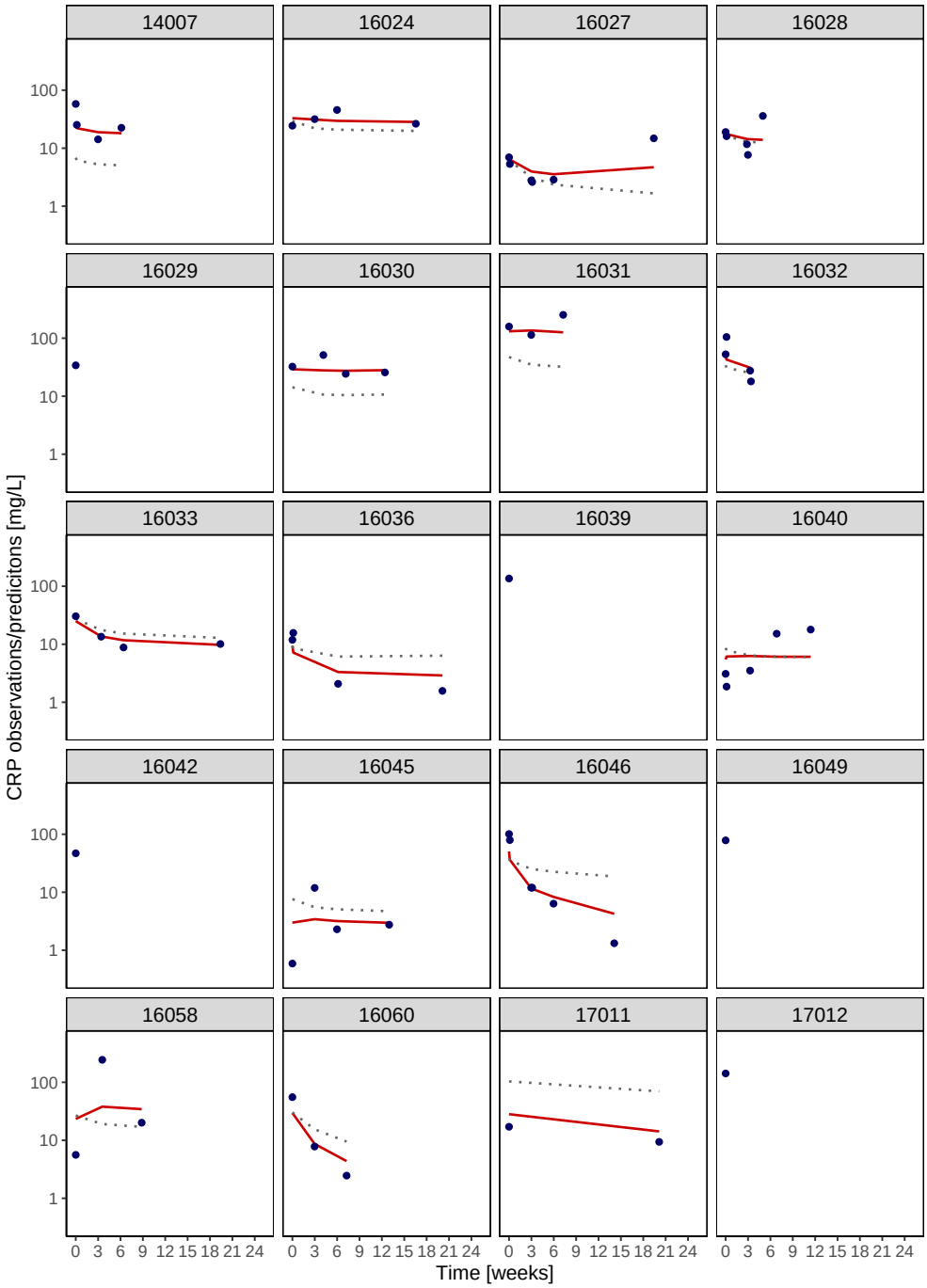

Figure B.8.1: *Cont.* Individual plots of CRP concentrations from the coupled tumour dynamics - CRP model showing the observations (blue dots), individual predictions (solid red line) and population predictions (dotted grey line) in the logarithmic scale. Each panel represents one patient indicated by the shaded number on top. CRP: C-reactive protein.
Note: results are of the coupled tumour dynamics - CRP model hence the deviation from the steady state predictions.

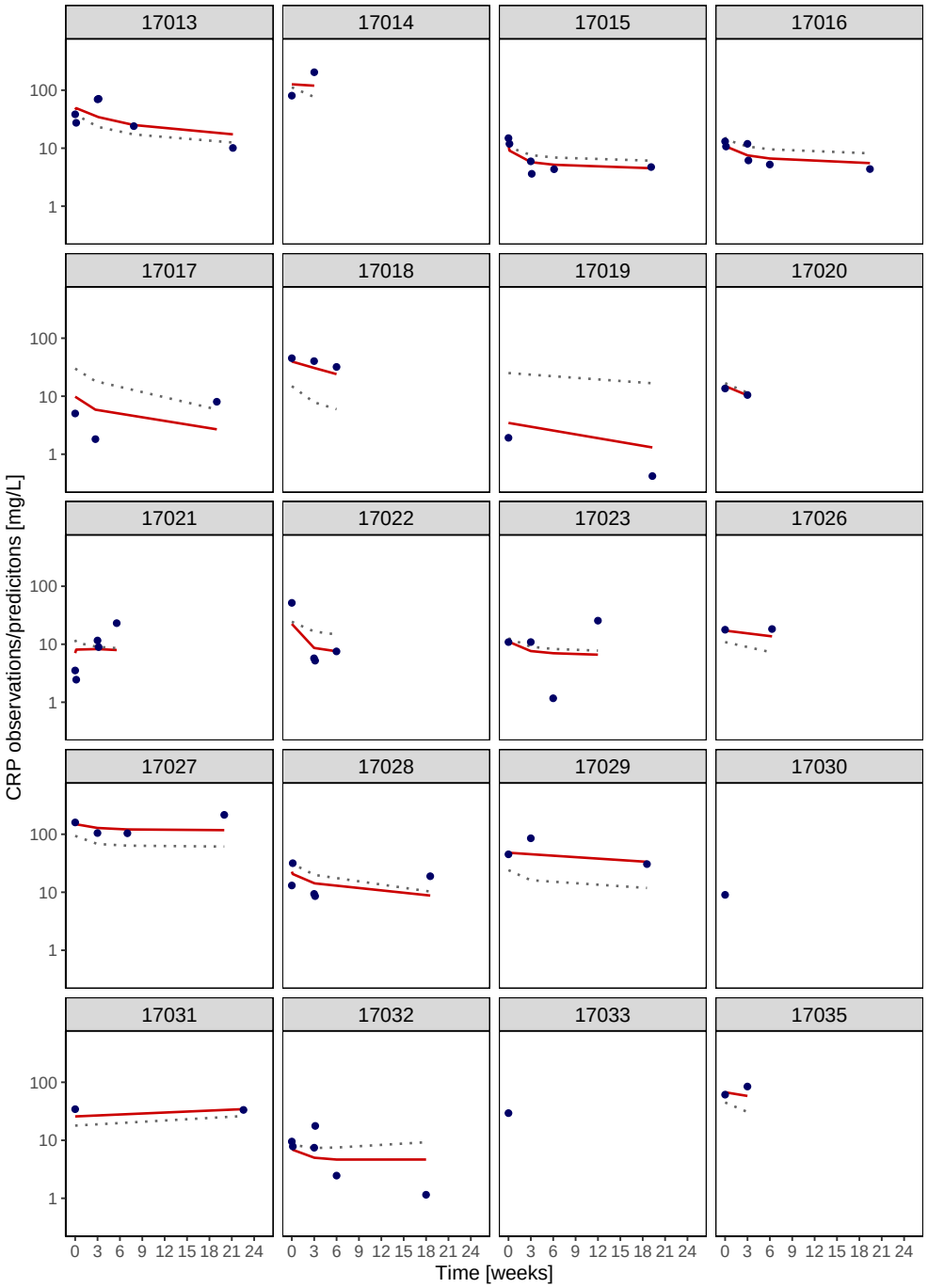

Figure B.8.1: *Cont.* Individual plots of CRP concentrations from the coupled tumour dynamics - CRP model showing the observations (blue dots), individual predictions (solid red line) and population predictions (dotted grey line) in the logarithmic scale. Each panel represents one patient indicated by the shaded number on top. CRP: C-reactive protein.
Note: results are of the coupled tumour dynamics - CRP model hence the deviation from the steady state predictions.

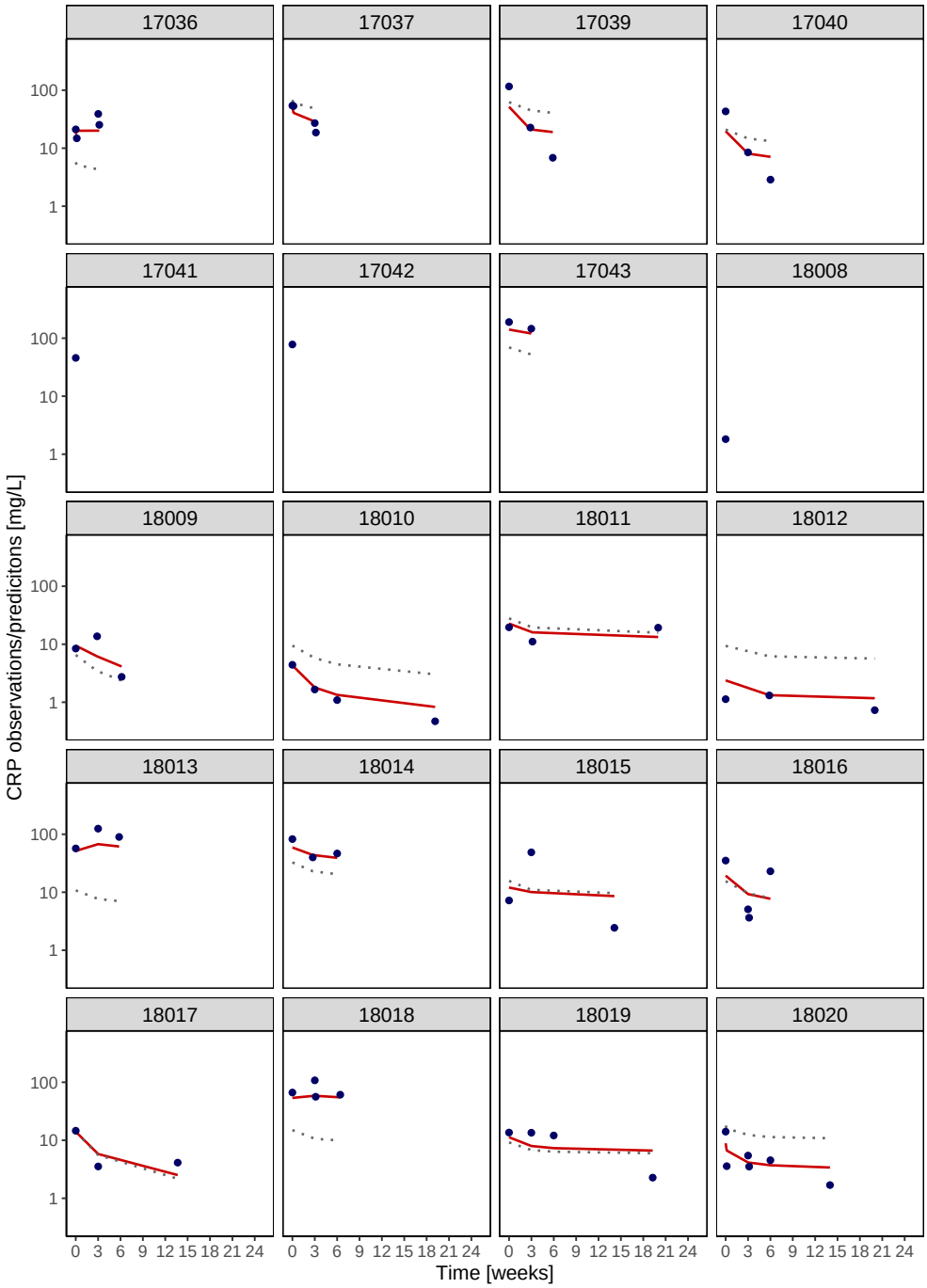

Figure B.8.1: *Cont.* Individual plots of CRP concentrations from the coupled tumour dynamics - CRP model showing the observations (blue dots), individual predictions (solid red line) and population predictions (dotted grey line) in the logarithmic scale. Each panel represents one patient indicated by the shaded number on top. CRP: C-reactive protein.
Note: results are of the coupled tumour dynamics - CRP model hence the deviation from the steady state predictions.

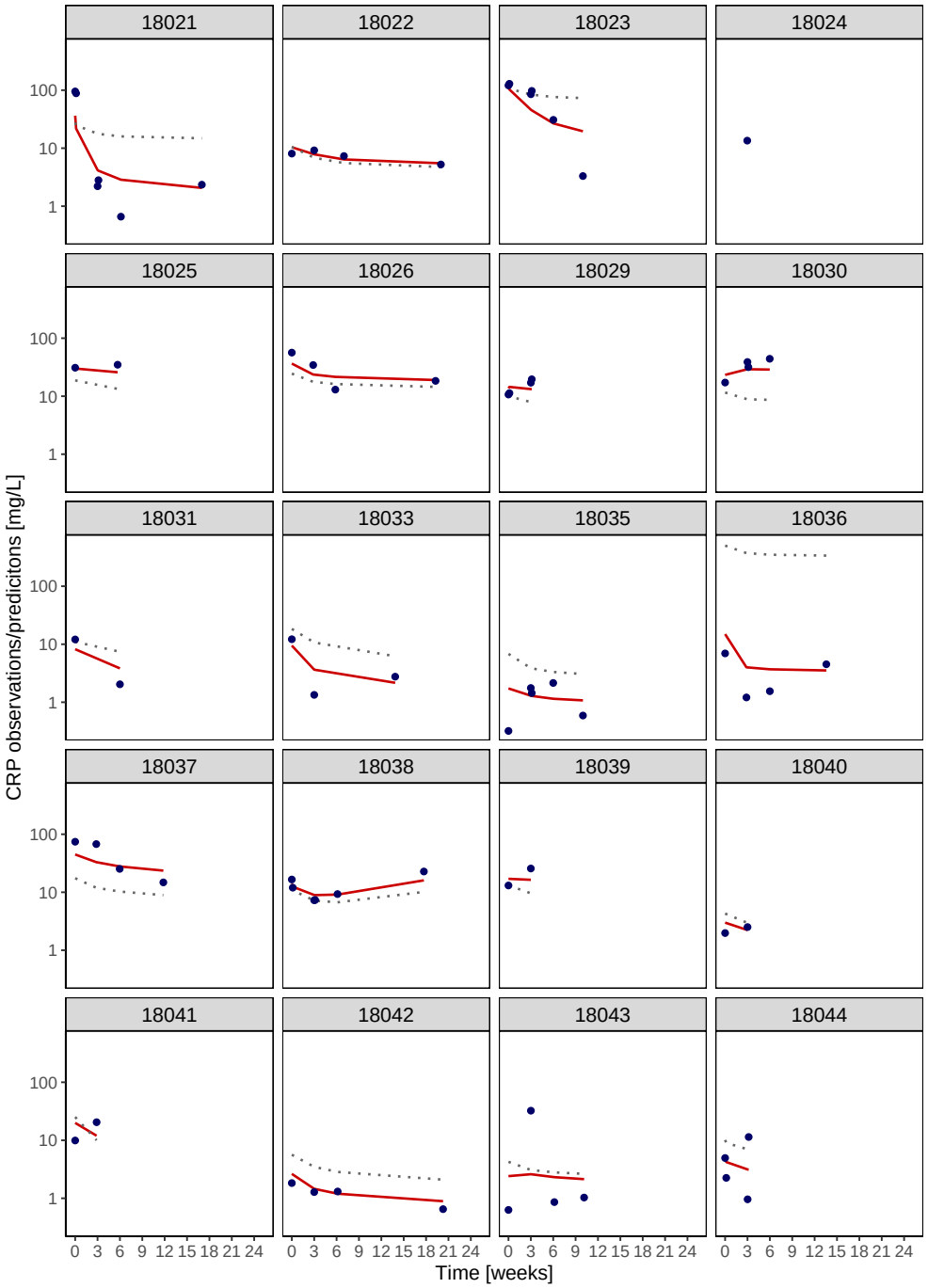

Figure B.8.1: *Cont.* Individual plots of CRP concentrations from the coupled tumour dynamics - CRP model showing the observations (blue dots), individual predictions (solid red line) and population predictions (dotted grey line) in the logarithmic scale. Each panel represents one patient indicated by the shaded number on top. CRP: C-reactive protein.
Note: results are of the coupled tumour dynamics - CRP model hence the deviation from the steady state predictions.

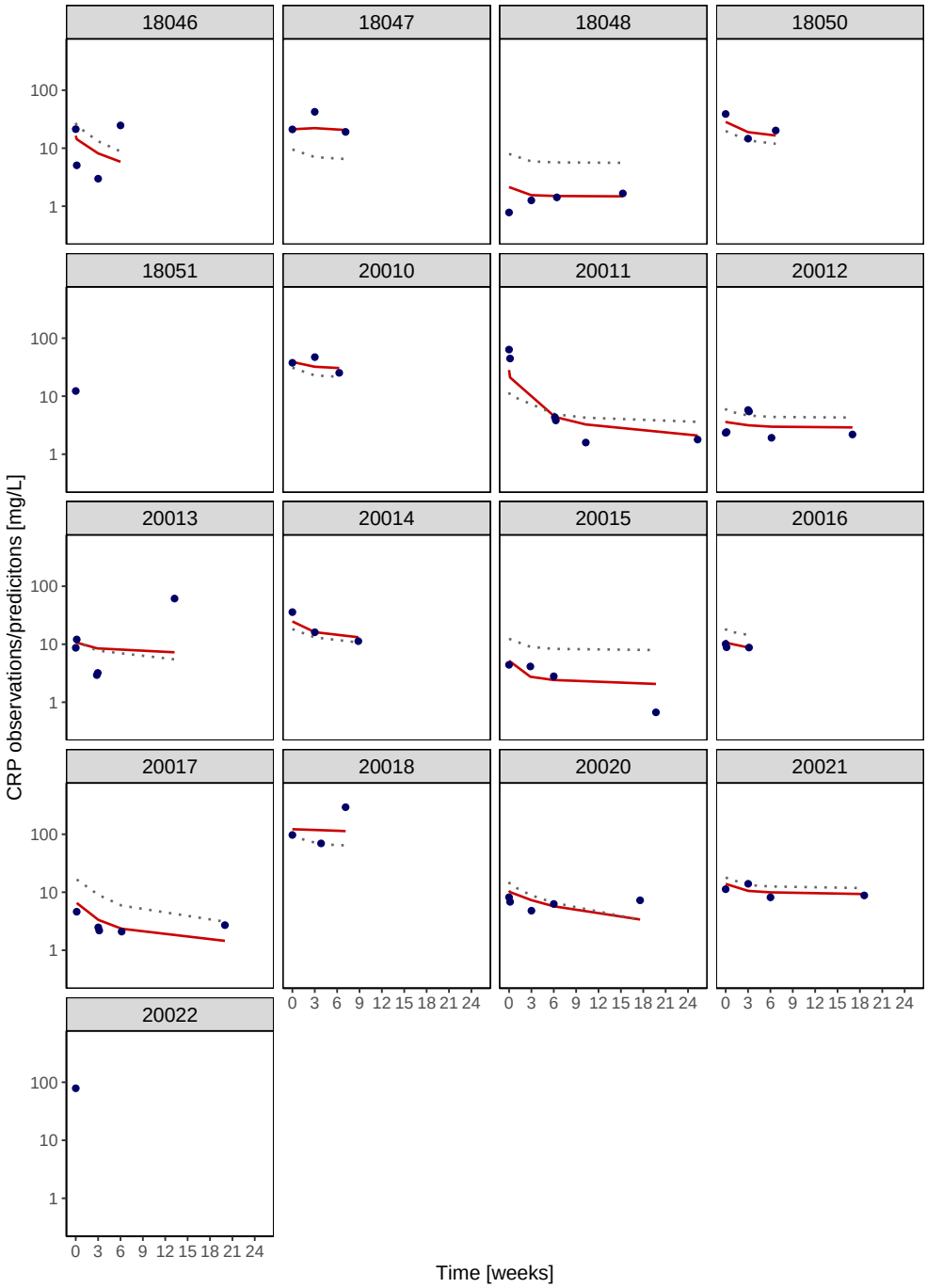

Figure B.8.1: *Cont.* Individual plots of CRP concentrations from the coupled tumour dynamics - CRP model showing the observations (blue dots), individual predictions (solid red line) and population predictions (dotted grey line) in the logarithmic scale. Each panel represents one patient indicated by the shaded number on top. CRP: C-reactive protein.
Note: results are of the coupled tumour dynamics - CRP model hence the deviation from the steady state predictions.

B.9 Correlation matrix between pre-selected covariates tested in the time-to-event models

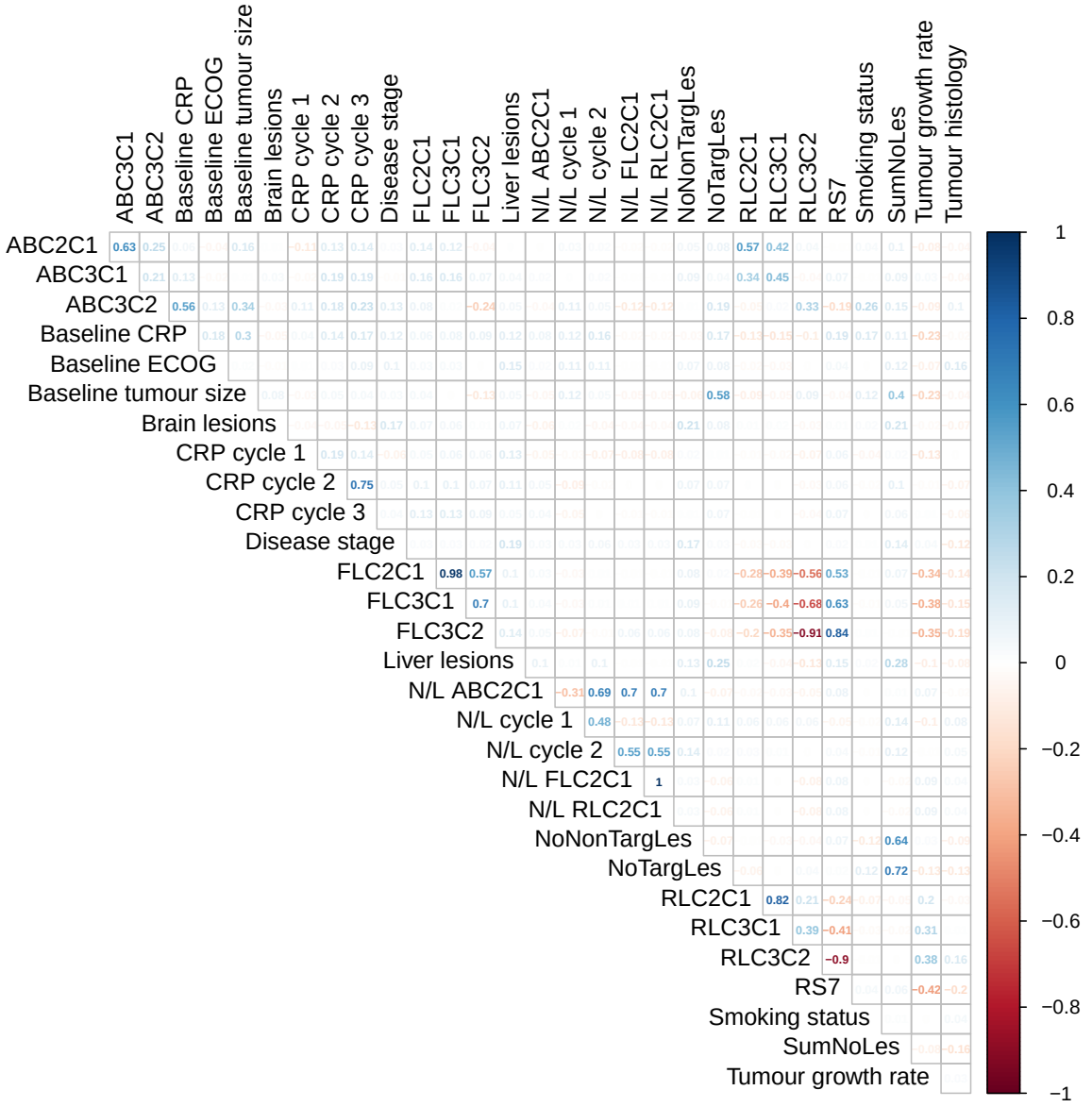

Figure B.9.1: Correlation matrix between pre-selected covariates (predictors) tested in the time-to-event models. Correlations are based on the full predictor information in the time-to-event dataset before applying the landmark time.
AB*: Absolute difference in CRP concentration; FL*: Fold change in CRP concentration; RL*: Relative change in CRP concentration; N/L*: Neutrophil-to-lymphocyte ratio; N/L AB*: Absolute difference in neutrophil-to-lymphocyte ratio; N/L FL*: Fold change in neutrophil-to-lymphocyte ratio; N/L RL*: Relative change in neutrophil-to-lymphocyte ratio; *C2C1: cycle 2 from cycle 1; *C3C1: cycle 3 from cycle 1; *C3C2: cycle 3 from cycle 2; CRP: C-reactive protein; ECOG: Eastern Cooperative Oncology Group; RS7: Tumour size at week 7 relative to baseline tumour size; NoTargLes: Number of target lesions; NoNonTargLes: Number of non-target lesions; SumNoLes: Sum of target and non-target lesions.
Note: Only strong positive (dark blue) or negative (dark orange) correlations (i.e. $r > \pm 0.6$) were of concern.

C | NONMEM codes

C.1 Project I: Data-generating (true) model for a strong drug effect based on a linear dose-response relationship

```
$SIZES     PD = 66

$PROBLEM

$INPUT     C
           ID       ; Unique subject identifier
           TIME     ; Time since first record in days
           TYPE     ; Type of record
           DV       ; Log-transformed dependent variable
           EVID     ; NONMEM event identification
           ARM      ; Treatment arm
           PLACEBO  ; Placebo
           DOSE     ; Dose
           LSTTIM   ; Censoring time
           TSFDD    ; Time since first dose
           SPLIT    ; Replicate number

$DATA      ../../data/datasetname.csv
           IGNORE(C.EQ.C)
           IGNORE(TYPE.EQ.0)
           WIDE

$ABBR  COMRES = 11
$SUBR  ADVAN13 TOL = 9
$MODEL NCOM = 3

$PK

IF(NEWIND.NE.2) THEN ; for each new individual
LLOQ = LOG(5)
ENDIF
```

```
;; == DOSE ==
IF(NEWIND.NE.2) THEN
AF_DOSE = 0
ENDIF

;; == progression ==
IF(NEWIND.NE.2) THEN
YCUR  = 0
LOWTM = 0
PDR   = 0
CENS  = 0
ENDIF
IF(ICALL.EQ.4.AND.TYPE.EQ.1) CENS = PDR

;; == DOSE ==
IF(NEWIND.NE.2) THEN
AF_DOSE = 0
ENDIF

IF(NEWIND.NE.2) THEN
RATIO = 0
RELCHG = 0
ENDIF

;================ Sum of longest diameter model
TVBASE =  THETA(1)              ; baseline SLD
IBASE  =  TVBASE*EXP(ETA(1))

TVKG   =  THETA(2)/7/1000       ; SLD growth rate constant
KG     =  TVKG*EXP(ETA(2))

TVLAM  =  THETA(3)/7/1000       ; rate constant for the resistance appearance
LAMBDA =  TVLAM*EXP(ETA(3))

TVDRU  =  THETA(4)/7/1000       ; drug effect rate constant
KDRUG  =  TVDRU*EXP(ETA(4))

;================ Non-target lesions

BL_NTR = THETA(5)

  PREV=0                              ; used for both NL and NTR
```

```
  IF(NEWIND.NE.2) THEN
    LAST   = PREV
    ISOBS  = 0
    TIM    = 0
    SNTR   = 0
    SNL    = 0
  ENDIF

ISOBS = 1
IF(TYPE.NE.2) ISOBS = 0

LAST = LAST*(1-ISOBS) + TIM*ISOBS   ; LAST is time of last observation record
 ; --> for obs records (ISOBS=1), LAST is TIME (current time)
 ; --> for dosing records (ISOBS=0), LAST is the time of last obs

;--------- Define DTIME --------------
DTIME = (TIME - LAST) ; DTIME is the difference between the current time
                      ;and the time of last obs record

;================ New-lesions

BL_NEWLES = THETA(6)

;============= Death before progression model

IF(NEWIND.NE.2) THEN
 DTHDV = 0
ENDIF

CONST_HZ = THETA(7)/1000

;====== Drop out

IF(NEWIND.NE.2) THEN
 DPDV = 0
ENDIF

HZLAM = THETA(8)/1000
HZALP = THETA(9)

LSTTIME = LSTTIM
;; THE SIMULATION PART FOR TTE SIMULATIONS ;;
IF (ICALL.EQ.4) THEN ; The event time sim $problem
IF (NEWIND.EQ.0) THEN ; Only for the first record
COM(6) = 1 ; Reset simulation ID counter
```

```
ENDIF
IF(NEWIND.NE.2) COM(4) = LSTTIME ; Set max time/censoring time
IF (NEWIND.EQ.1) THEN ; For every new ind except first in dataset
ICOUNT = COM(6) + 1 ; Update individual counter over simulations
COM(6) = ICOUNT
ENDIF

IF (NEWIND.NE.2) THEN ; For every new individual
CALL RANDOM(5,R)
COM(3) = -1  ; Variable for survival at event time
COM(2) = R   ; Store the random number
COM(1) = -1  ; Variable for the event time
COM(7) = 0   ; Individual event counter
ENDIF

IF (NEWIND.NE.2) THEN ; For every new individual
CALL RANDOM(4,R)
COM(10) = -1  ; Variable for survival at event time
COM(9)  = R   ; Store the random number
COM(8)  = -1  ; Variable for the event time
COM(11) = 0   ; Individual event counter
ENDIF
ENDIF

;---------MTIME for increasing precision in $DES --------
IF (NEWIND.NE.2) THEN
TEMP = 0
ENDIF
TEMP = TEMP+7
MTIME(1) = TEMP
MTDIFF = 1

;-----Compartment initialisation-----------
A_0(1) = IBASE    ; initialise with baseline SLD

$DES

RESIST = LAMBDA

E0 = THETA(13)
EFFECT = E0+((AF_DOSE)*KDRUG) ; Linear dose-response relationship

;=============== Sum of longest diameter model
DADT(1) =  KG*A(1)- EFFECT*EXP(-(RESIST*T))*A(1)
```

```
;============= Death model
DADT(2) = CONST_HZ

DESDTSUR = EXP(-A(2))

IF(COM(9).GT.DESDTSUR.AND.COM(8).EQ.-1) THEN
COM(8) = T
COM(10) = DESDTSUR
ENDIF

;============= Dropout model
DELX = 1E-16

DWEIB = HZLAM*HZALP*(HZLAM*(T+DELX))**(HZALP-1)

DADT(3) =  DWEIB

DESDPSUR = EXP(-A(3))

IF(COM(2).GT.DESDPSUR.AND.COM(1).EQ.-1) THEN
COM(1) = T
COM(3) = DESDPSUR
ENDIF

$ERROR
"FIRST
"@CHARACTER(LEN=100)::FMT ! Define FORMAT string for writing dataset

;Sim_start : add for simulation
; IF(TYPE.EQ.2) THEN
; F_FLAG=1
; ENDIF
;
; IF(TYPE.EQ.3) THEN
; F_FLAG=1
; ENDIF
;
; IF(TYPE.EQ.4) THEN
; F_FLAG=1
; ENDIF
;
; IF(TYPE.EQ.5) THEN
; F_FLAG=1
; ENDIF
;Sim_end
```

```
A1 = A(1)
IF(A1.LE.0) A1 = 1E-16

IPRED = LOG(A1)

LOQ = LLOQ
SDSIGMA = SQRT(SIGMA(1,1))

DUM = (LOQ-IPRED)/(SDSIGMA*EXP(ETA(5)))

CUMD = PHI(DUM)

; IF (SIMF.EQ.0.AND.TYPE.EQ.1.AND.DV.GT.LOQ) THEN
; F_FLAG=0
; Y      = IPRED+ EPS(1)*EXP(ETA(7))
; ENDIF
;
; IF (SIMF.EQ.0.AND.TYPE.EQ.1.AND.DV.LE.LOQ) THEN
; F_FLAG=1
; Y=CUMD
; ENDIF

IF (TYPE.EQ.1) THEN
Y = IPRED+ EPS(1)*EXP(ETA(5))
ENDIF

IF (TYPE.EQ.1) AF_DOSE = DOSE

; == Progression calculations ==

IF(TIME.EQ.0.AND.TYPE.EQ.1) YCUR  = EXP(Y)
IF(TIME.EQ.0.AND.TYPE.EQ.1) LOWTM = EXP(Y)

IF(TIME.GT.0.AND.EVID.EQ.0.AND.TYPE.EQ.1) YCUR = EXP(Y)

IF(PDR.EQ.0.AND.TIME.GT.0.AND.YCUR.LT.LOWTM.AND.TYPE.EQ.1) LOWTM = YCUR
RECIST = 1.2*LOWTM
FIVM   = 5+LOWTM
IF(TIME.GT.42.AND.YCUR.GT.RECIST.AND.YCUR.GT.FIVM.AND.TYPE.EQ.1) PDR = 1


; Change from baseline
IF(ICALL.EQ.4.AND.TIME.EQ.0) BASE = YCUR
```

```
IF(EVID.EQ.0.AND.TYPE.EQ.1)SCUR = EXP(DV)
IF(ICALL.EQ.4.AND.EVID.EQ.0.AND.TYPE.EQ.1) SCUR = EXP(Y)

IF(TYPE.EQ.1.AND.TIME.GT.0.AND.EVID.EQ.0) RATIO = 100*((SCUR-BASE)/BASE)
IF(TYPE.EQ.1.AND.TIME.GT.0.AND.EVID.EQ.0) RELCHG = ((SCUR-BASE)/BASE)

;=====================Non-target lesions

IF(TYPE.EQ.2) THEN
P1_NTR   =    BL_NTR+ (RELCHG*THETA(10)+ THETA(11)*LOG(BASE))
ENDIF

IF(TYPE.EQ.2) THEN
PROB_NTR =   EXP(P1_NTR)/(1+EXP(P1_NTR))
ENDIF

IF(TYPE.EQ.2.AND.EVID.EQ.0) THEN
PD_NTR   =   1-(1-PROB_NTR)**(DTIME) ; prob of les out between two tumour
                                     ; assessment
PND_NTR  =   (1-PROB_NTR)**(DTIME)   ; prob of not les out between two
                                     ; tumour assessment
ENDIF

;Sim_start : add for simulation
;    IF(TYPE.EQ.2.AND.DV.EQ.0) Y=PND_NTR
;    IF(TYPE.EQ.2.AND.DV.EQ.0) SNTR =0
;
;    IF(TYPE.EQ.2.AND.DV.EQ.1) Y=PD_NTR
;    IF(TYPE.EQ.2.AND.DV.EQ.1) SNTR = 1
;Sim_end

;=====================New-lesions
  BETA1 = THETA(7)

IF(TYPE.EQ.3) THEN
P1_NEWLES   =    BL_NEWLES + (RELCHG*BETA1+ THETA(12)*LOG(BASE))
ENDIF

IF(TYPE.EQ.3) THEN
PROB_NEWLES =  EXP(P1_NEWLES)/(1+EXP(P1_NEWLES))
ENDIF

IF(TYPE.EQ.3.AND.EVID.EQ.0) THEN
PD_NEWLES   =   1-(1-PROB_NEWLES)**(DTIME)   ; prob of new les out between two
                                             ; tumour assessment
```

```
PND_NEWLES  =   (1-PROB_NEWLES)**(DTIME)      ; prob of not new les out between
                                              ; two tumour assessment
ENDIF

;Sim_start : add for simulation
;    IF(TYPE.EQ.3.AND.DV.EQ.0) Y=PND_NEWLES
;    IF(TYPE.EQ.3.AND.DV.EQ.0) SNL =0
;
;    IF(TYPE.EQ.3.AND.DV.EQ.1) Y=PD_NEWLES
;    IF(TYPE.EQ.3.AND.DV.EQ.1) SNL = 1
;Sim_end

;----- time set for non-target lesions AND new-lesions
IF(TYPE.EQ.2) THEN
LCMT = 1
TIM = TIME
ENDIF
ISOBS = 1

;====================Death
CHZDTH = A(2)
SURDTH = EXP(-CHZDTH)

HAZNOWDTH  = CONST_HZ

;Sim_start : add for simulation
; IF(TYPE.EQ.4.AND.EVID.EQ.0.AND.DV.EQ.0) THEN ; PFS
; Y = SURDTH
; DTHDV = 0
; ENDIF
;
; IF(TYPE.EQ.4.AND.EVID.EQ.0.AND.DV.EQ.1) THEN ; CENS
; Y = SURDTH*HAZNOWDTH
; DTHDV = 1
; ENDIF
;Sim_end

;====================Dropout
CHZDP  = A(3)
SURDP  = EXP(-CHZDP)
DEL    = 1E-16

EWEIB    = HZLAM*HZALP*(HZLAM*(TIME+DEL))**(HZALP-1)
HAZNOWDP = EWEIB
```

```
;Sim_start : add for simulation
; IF(TYPE.EQ.5.AND.EVID.EQ.0.AND.DV.EQ.0) THEN ; PFS
; Y = SURDP
; DPDV = 0
; ENDIF
;
; IF(TYPE.EQ.5.AND.EVID.EQ.0.AND.DV.EQ.1) THEN ; CENS
; Y = SURDP*HAZNOWDP
; DPDV = 1
; ENDIF
; Sim_end

;; simulations
IF(ICALL.EQ.4) THEN
   IF(NEWIND.NE.2) THEN      ; for each new individual
     FLAG2 = 0               ; initialise event counter
     RTTE_NT = 0
   ENDIF

     CALL RANDOM (2,R)       ; call a random number
     TMPNTR = R              ; store the random number

   IF(TIME.GT.0.AND.TYPE.EQ.2.AND.TMPNTR.LE.PD_NTR.AND.FLAG2.EQ.0) THEN
   ; event
     DV      = 1
     SNTR    = 1
     FLAG2   = 1
     RTTE_NT = 1
   ENDIF
ENDIF

IF(ICALL.EQ.4) THEN
   IF(NEWIND.NE.2) THEN      ; for each new individual
     FLAG3 = 0               ; initialise event counter
   ENDIF

     CALL RANDOM (3,R)       ; call a random number
     TMPNL = R               ; store the random number

   IF(TIME.GT.0.AND.TYPE.EQ.3.AND.TMPNL.LE.PD_NEWLES.AND.FLAG3.EQ.0) THEN
   ; event
     DV    = 1
     SNL   = 1
     FLAG3 = 1
   ENDIF
```

```
ENDIF

IF(ICALL.EQ.4) THEN
   IF(NEWIND.NE.2) THEN      ; for each new individual
     FLAG4 = 0               ; initialise event counter
     CALL RANDOM (4,R)       ; call a random number
     TMPDTH = R              ; store the random number
   ENDIF

   IF(TIME.GT.0.AND.TYPE.EQ.1.AND.TMPDTH.GT.SURDTH.AND.FLAG4.EQ.0) THEN ; event
     DV     = 1
     DTHDV  = 1
     FLAG4  = 1
   ENDIF
ENDIF

IF(ICALL.EQ.4) THEN
   IF(NEWIND.NE.2) THEN      ; for each new individual
     FLAG5 = 0               ; initialise event counter
     CALL RANDOM (5,R)       ; call a random number
     TMPDP = R               ; store the random number
   ENDIF

   IF(TIME.GT.0.AND.TYPE.EQ.1.AND.TMPDP.GT.SURDP.AND.FLAG5.EQ.0) THEN ; event
     DV     = 1
     DPDV   = 1
     FLAG5  = 1
   ENDIF
ENDIF

 IF(NEWIND.NE.2) THEN
 PFS_RTTE = 0
 SPFS     = 0
 MPFS     = 0

 FLAGPFS = 0
 ENDIF

;Sim_start : add for simulation
; IF(TSFDD.EQ.PFSTMI) THEN
; IF(PFSCI.EQ.1)SPFS = 1
; PFS_RTTE = 1
; ENDIF
;Sim_end
```

```
IF(ICALL.EQ.4) THEN

IF(CENS.EQ.1.OR.SNTR.EQ.1.OR.SNL.EQ.1.OR.DTHDV.EQ.1) MPFS = 1
;CENS.EQ.1.OR.SNTR.EQ.1.OR.SNL.EQ.1.OR.DTHDV.EQ.1

IF(DPDV.EQ.0.AND.MPFS.EQ.1.AND.FLAGPFS.EQ.0) THEN
SPFS    = 1                         ;
PFS_RTTE = 1
FLAGPFS  = 1
ENDIF

IF(DPDV.EQ.1.AND.FLAGPFS.EQ.0) THEN
SPFS     = 0                         ;
PFS_RTTE = 1
FLAGPFS  = 1
ENDIF

IF(TSFDD.EQ.ENDDAYI.AND.FLAGPFS.EQ.0) THEN
SPFS     = 0                         ;
PFS_RTTE = 1
FLAGPFS  = 1
ENDIF
;
ENDIF


IF (COM(1).GT.COM(4)) THEN ;IF T > ENDTIME, T = ENDTIME
; Check survival again at endtime
IF (COM(2).GT.SURDP) THEN
COM(1) = COM(4)
ELSE
COM(1) = -1 ; Integrated too far, reset event
ENDIF
ENDIF


IF (COM(8).GT.COM(4)) THEN ;IF T > ENDTIME, T=ENDTIME
; Check survival again at endtime
IF (COM(9).GT.SURDTH) THEN
COM(8) = COM(4)
ELSE
COM(8) = -1 ; Integrated too far, reset event
ENDIF
ENDIF
```

```
IF(MPFS.EQ.1.OR.COM(8).NE.-1) SPFS = 1

EVT = COM(1) ; Save Event time
IF(SPFS.EQ.1.AND.COM(1).EQ.-1) EVT = TIME

RNM = COM(2)      ; Save random number, just for debugging
ENDTIME = COM(4) ; Endtime of study
TT = COM(5)       ; Analytic event time

IF (ICALL.EQ.4) THEN ; Initate DV to 0 (No event)
DV = 0
ENDIF
TMDV = 0
IF (EVID.GE.2) THEN ;Set MDV variable for output
TMDV = 1
ENDIF

ICOUNT = COM(6)+(IREP-1)*NINDR
ITER = IREP
; Define the format of the output file

" LAST
" FMT='(E13.7,9(1XE13.7))' ! The output FORMAT
" ! Write all events
" IF (NEWIND.EQ.0) THEN !Open file at first record
" OPEN (99, FILE = 'simtab.dat', POSITION='APPEND')
" IF (IREP.EQ.1) THEN !Write header for 1st subproblem
" WRITE (99,'(A,9(1XA))') 'ID','DV','TIME','RTTE','SURDP','SPFS','ITER',
" 'RAND','ARM','MPFS'
" ENDIF
" ENDIF
" IF (EVT.NE.-1) THEN !If an EVENT
" DV=1
" RTTE=1
" TMDV=0
" IF(MPFS.EQ.1) SPFS = 1
" ! Write SIM specific output
" WRITE (99,FMT) ID,DV,EVT,RTTE,COM(3),SPFS,ITER,COM(2),ARM,MPFS
" COM(1) = -1 !Reset Event time variable
" COM(2) = 0 !Reset Random variable
" COM(3) = -1 !Reset survival variable
" COM(7) = COM(7) + 1 !Update Event counter
" ELSE IF (LIREC.EQ.NDREC.AND.COM(7).EQ.0) THEN !Right Censoring
" (if no previous events)
```

```
" DV=0
" TMDV=0
" RTTE=1
" TMP=COM(4)
" SPFS = 0
" WRITE (99,FMT) ID,DV,TMP,RTTE,SURDP,SPFS,ITER,COM(2),ARM,MPFS
" ENDIF
" IF (NDREC.EQ.LIREC.AND.NIREC.EQ.NINDR) THEN ! Last record for last individual
" CLOSE(99) ! Close File pointer
" ENDIF

REP = IREP

;----------- Sum of longest diameter model
$THETA
(0,51.5831) ; 1 IBASE in mm
(0,1.04043) ; 2 KG*1000 in weeks
(0,391.146) ; 3 LAMBDA*1000 in weeks
(0,1.8)     ; 4  KDRUG*1000 in weeks

;-----------Non-target lesions
-8.19696    ; 5 NTR

;-----------New-lesions
-8.8932 FIX ; 6 NL intercept

;-----------Death
0.082753 FIX ; 7 DT constant hazard death

;----------- Dropout model
(0,1.02292) ; 8 DP lambda
(0,1)       ; 9 DP alpha

;-----------Non-target lesions covariate effect
2.59292     ; 10 Rel change on NTR
0.47028     ; 11 effect of BASE on NTR

;-----------New-lesions covariate effect
0.635 FIX   ; 12 NL beta baseline SLD

;----------- Baseline drug effect
0 FIX       ; 13 E0
```

```
$OMEGA
0.432048 ; 1 IBASE
0.60443  ; 2 KG
1.19485  ; 3 LAMBDA
0.110422 ; 4 KDRUG
0.347521 ; 5 RUV

$SIGMA 0.0105136

$SIM (1955199112 NEW) (678910 UNIFORM) (2076053753 UNIFORM)
          (12344 UNIFORM) (726447632 UNIFORM) ONLYSIM
          NSUBPROBLEMS=1
```

C.2 Project I: Candidate model based on a linear dose-response relationship

```
$SIZES     PD = 65

$PROBLEM

$INPUT     C
           ID       ; Unique subject identifier
           TIME     ; Time since first record in days
           TYPE     ; Type of record
           DV       ; Log-transformed dependent variable
           ARM      ; Treatment arm
           LSTTIM   ; Censoring time
           TSFDD    ; Time since first dose
           DOSE     ; Dose
           EVID     ; NONMEM event identification
           PLACEBO  ; Placebo

$DATA     ../data/datasetname.csv
           IGNORE(C.EQ.C)
           WIDE

$SUBR      ADVAN13 TOL=6
$MODEL     NCOM=2

$PK
;; == DOSE ==
IF(NEWIND.NE.2) THEN
AF_DOSE = 0
ENDIF

;=============== Sum of longest diameter model
TVBASE =   THETA(1)              ; baseline SLD
IBASE  =   TVBASE*EXP(ETA(1))

TVKG   =   THETA(2)/7/1000       ; SLD growth rate constant
KG     =   TVKG*EXP(ETA(2))

TVLAM  =   THETA(3)/7/1000       ; rate constant for the resistance appearance
LAMBDA =   TVLAM

TVDRU  =   THETA(4)/7/1000       ; drug effect rate constant
KDRUG  =   TVDRU
```

```
;====== Drop out
IF(NEWIND.NE.2) THEN
 DPDV = 0
ENDIF

HZLAM = THETA(5)/1000
HZALP = THETA(6)

;-----Compartment initialisation-----------
A_0(1) =  IBASE          ; initialise with baseline SLD

$DES
RESIST = LAMBDA
E0     = THETA(7)
EFFECT = E0+((AF_DOSE)*KDRUG) ; Linear dose-response relationship

;================ Sum of longest diameter model

DADT(1) =  KG*A(1)- EFFECT*EXP(-(RESIST*T))*A(1)

;============= Dropout model
DELX = 1E-16

DWEIB   = HZLAM*HZALP*(HZLAM*(T+DELX))**(HZALP-1)
DADT(2) =  DWEIB

$ERROR
IF(TYPE.EQ.1) THEN
F_FLAG = 0
ELSE
F_FLAG = 1
ENDIF

A1 = A(1)
IF(A1.LE.0) A1 = 1E-16

IPRED = LOG(A1)

IF(TYPE.EQ.1)THEN
  F_FLAG = 0
  Y      = IPRED+ EPS(1)
ENDIF

IF(TYPE.EQ.1) AF_DOSE = DOSE
```

```
;=====================Dropout
CHZ  = A(2)
SUR  = EXP(-CHZ)
DEL  = 1E-16

EWEIB   = HZLAM*HZALP*(HZLAM*(TIME+DEL))**(HZALP-1)
HAZNOW  = EWEIB

IF(TYPE.EQ.2.AND.EVID.EQ.0.AND.DV.EQ.0) THEN ; CENS
F_FLAG = 1
Y      = SUR
DPDV   = 0
ENDIF

IF(TYPE.EQ.2.AND.EVID.EQ.0.AND.DV.EQ.1) THEN ; PFS
F_FLAG = 1
Y      = SUR*HAZNOW
DPDV   = 1
ENDIF

;----------- Sum of longest diameter model
$THETA
(0,50)         ; 1  IBASE in mm
(0,1.08)       ; 2  KG*1000 in weeks
(0,390)        ; 3  LAMBDA*1000 in weeks
(0,0.01)       ; 4  KDRUG*1000 in weeks

;----------- Dropout model
(0,0.7)        ; 5 hz_LAMBDA
1 FIX          ; 6 hz_ALPHA

;----------- Baseline drug effect
0 FIX        ; 7 E0

$OMEGA
0.432     ; 1 IBASE
1.5       ; 2 KG

$SIGMA 0.105

$ESTIM METHOD=1 INTER LAPLACE NUMERICAL PRINT=1 NOABORT NSIG=2 SIGL=9
RANMETHOD=P MCETA=10
```

C.3 Project II: Tumour dynamics model

```
$SIZES PD = -100

$PROBLEM Paclitaxel AUC-driven drug-induced tumour decay

$INPUT ID       ; Patient number
       OCC      ; Treatment cycle
       DAY      ; Day within cycle
       TIME     ; Time of tumour size measurement
       DV       ; Tumour size (log-transformed)
       BLSD     ; Baseline tumour size
       EVID     ; Event identifier
       MDV      ; Missing dependent variable
       FLGIG    ; Flag for tumour size records
       AUC      ; Paclitaxel AUC per cycle
       SimNo    ; Simulation number (idnetifier with respect to multiple
                ; imputation)

$DATA data1.csv IGNORE=I

IGNORE(ID.EQ.16037)
IGNORE(FLGIG.EQ.1)  ; exclude non-tumour size records
IGNORE(TIME.GT.5040)

$SUBR ADVAN6 TOL=5

$MODEL COMP=(SD)            ;1 Sum of diameter
       COMP=(SIZE8)         ;2 Tumour size at week 8

$PK
" FIRST
" COMMON/PRCOMG/IDUM1,IDUM2,IMAX,IDUM4,IDUM5
" INTEGER IDUM1,IDUM2,IMAX,IDUM4,IDUM5
" IMAX=10000000


; Sum of diameters
 GR     = THETA(1) * EXP(ETA(1))          ; Growth rate
 BETA   = THETA(2) * EXP(ETA(2))          ; Drug induced decay
 BASE   = EXP(LOG(BLSD)+THETA(3)*ETA(3))  ; Baseline tumour size (B2 method)
 LAMBDA = THETA(4)*EXP(ETA(4))            ; Decline in drug effect in a cycle
```

```
 A_0(1) = BASE
 A_0(2) = BASE

$DES

; Sum of diameters
 EFF = BETA * AUC*EXP(-LAMBDA*TIME)
 DADT(1) = GR - EFF * A(1)
 CT = A(1)

;Output of week 8 size
 IF(T.LT.1344) SLOP2 = GR - EFF * A(1)       ; 8*7*24=1344
 IF(T.GE.1344) SLOP2 = 0                     ; 8*7*24=1344
 DADT(2) = SLOP2

$ERROR

;Output of amounts
 AA1 = A(1)
 AA2 = A(2)

;Relatve change of size at week 8
TSIZE = AA2
RS8    = 100/BLSD*TSIZE

;Residual error
 IPRED=0.0001
 IF(A(1).GT.0) IPRED=LOG(A(1))
 W = THETA(3)
 Y = IPRED + W *EPS(1)

 IRES  = DV-IPRED
 IWRES = IRES/W

$THETA
(0, 0.002)   ;1. GR
(0, 0.0003)  ;2. BETA
(0, 0.12)    ;3. ERR
(0, 0.0004)  ;4. LAMBDA

$OMEGA
 0.4      ;1. IIV GR
 0.6      ;2. IIV BETA
 1 FIX    ;3. IIV BLSD, residual
 0.5      ;4. IIV LAMBDA, residual
```

```
$SIGMA
 1 FIX    ;1. RES

$ESTIM PRINT=5 MAXEVAL=99999 METHOD=1 INTER NOABORT POSTHOC SIGDIG=3
MSFO=MSF001 NOTHETABOUNDTEST NOOMEGABOUNDTEST NOSIGMABOUNDTEST

$COV PRINT=E
```

C.4 Project II: Coupled tumour dynamics - CRP model

```
$SIZES    PD = -70
$PROBLEM  Coupled tumour-CRP turnover model

$INPUT ID       ; Patient number
       OCC      ; Treatment cycle
       DAY      ; Day within cycle
       TIME     ; Time of CRP measurement
       EVID     ; Event identifier
       MDV      ; Missing dependent variable
       DV       ; CRP (log-transformed)
       CMT      ; Compartment identifier (CRP, Tumour)
       FLAGCRP  ; Flag for BLQ and ULOQ value for CRP
       BLIL6    ; Baseline interleukin-6
       NEWBLIL6 ; Baseline interleukin-6, missing value imputed by median
       SMOK     ; Smoking status
       STAGE    ; Disease stage
       BLSD     ; Baseline tumour size
       IGR      ; Individual tumour growth rate
       IBETA    ; Individual drug effect
       ITBASE   ; Individual estimated baseline tumour size
       ILAMBDA  ; Individual drug resistance rate constant
       SEGR     ; Individual standard error of tumour growth rate
       SEBETA   ; Individual standard error of drug effect
       SELAMBDA ; Individual standard error of resistance
       SEERR    ; Individual standard error of baseline tumour size
       MAUC     ; Paclitaxel AUC per cycle

$DATA 2-CRP_TGI_Model.csv IGNORE=@

     IGNORE=(FLAGCRP.EQ.2)      ;Exclude flagged items (BLQ)
     IGNORE=(ID.EQ.13028)       ;Patient with no CRP (all are BLQ)

$ABBR PROTECT
$SUBR ADVAN13 TOL9

$MODEL     NCOMP=3
           COMP=(TS)            ;1 Tumour size
           COMP=(CRP)           ;2 CRP
           COMP=(SIZE8)         ;3 Tumour size at week 8
```

```
$PK
;------ CRP model parameters -------------------------------

;Covariate function: sum of diameters-exponential
KINSUMDIA = EXP(THETA(5)*(BLSD-8.25))

;Covariate function: IL6-linear
KINIL6 = ( 1 + THETA(4)*(NEWBLIL6-2.57))

;Covariate function: disease stage
IF(STAGE.EQ.1) KINSTAGE = 1 ; Most common
IF(STAGE.EQ.0) KINSTAGE = ( 1 + THETA(6))

;Covariate function: smoking status
IF(SMOK.EQ.1) KINSMOK = 1 ; Most common
IF(SMOK.EQ.2) KINSMOK = ( 1 + THETA(7))
IF(SMOK.EQ.3) KINSMOK = ( 1 + THETA(8))

; Production constant: Kin, Zero-order
TVKIN = THETA(1)*KINIL6*KINSUMDIA*KINSTAGE*KINSMOK
KIN   = TVKIN*EXP(ETA(1))

; Basal production constant
KINB = THETA(9)

KINT = KINB+KIN

; Degradation constant: Kout, 1st-order
TVKOUT = THETA(2)
KOUT   = TVKOUT*EXP(ETA(2))

; Baseline
BASE = KINT/KOUT

; Relationship between tumour size and Kin
TVSLP = THETA(10)
SLP   = TVSLP*EXP(ETA(7))

;------ Tumour growth inhibition model parameters ------------------------------
GR     = IGR*EXP(ETA(3)*SEGR)
BETA   = IBETA*EXP(ETA(4)*SEBETA)
LAMBDA = ILAMBDA*EXP(ETA(5)*SELAMBDA)
TBASE  = BLSD*EXP(ETA(6)*SEERR)
```

```
; Intialise compartments
A_0(1)    = TBASE  ; Baseline tumour size
A_0(2)    = BASE   ; CRP (mg/L)
A_0(3)    = TBASE  ; Baseline tumour size

$DES
; TGI model
EFF     = BETA * MAUC*EXP(-LAMBDA*T)
DADT(1) = GR - EFF * A(1)

; CRP
KINP    = KIN*SLP*(A(1)/TBASE)
DADT(2) = (KINB+KINP)-KOUT*A(2)     ; Time-course of CRP

;Output of tumour size week 8
 IF(T.LT.1344) SLOP2 = GR - EFF * A(1)       ; 8*7*24=1344
 IF(T.GE.1344) SLOP2 = 0                     ; 8*7*24=1344
 DADT(3) = SLOP2

$ERROR
;Output of amounts
TSIZE = A(1)

CRP   = A(2)
IPRED = 0
IF(CRP.NE.0) IPRED = LOG(CRP)

RWSIZE = A(3)

;Relatve change of size at week 8
RS8 = (RWSIZE/BLSD)*100

; CRP RUV model: additive
W     = THETA(3)                    ; Sigma fixed to 1
IRES  = DV-IPRED
IWRES = IRES/W
Y     = IPRED+W*EPS(1)

$THETA
   (0,0.4)              ;1. KIN
   0.036473684 FIX      ;2. KOUT
   (0,0.8)              ;3. SIGMA SD
   (0.3)                ;4. COV IL6 on Kin
   0 FIX                ;5. COV Sumdia on Kin
   (-0.4)               ;6. Stage 0: IIIB
```

```
   (0.6)                ;7. Smok 2: former
   (1.3)                ;8. Smok 3: current
   0.01094211 FIX       ;9. KINB set to equivalent to LLOQ
   (0.8)                ;10. TS:linear

$OMEGA 0.8                ;1. IIV-KIN
$OMEGA 0 FIX              ;2. IIV-KOUT
$OMEGA 1 FIX              ;3. IGR
$OMEGA 1 FIX              ;4. IBETA
$OMEGA 1 FIX              ;5. ILAMBDA
$OMEGA 1 FIX              ;6. IIV BSL, residual
$OMEGA 0.36               ;7. IIV slope

$SIGMA 1 FIX              ; CRP ERROR

$ESTIM METHOD=1 INTER MAX=9999 PRINT=1 NOABORT

$COV
```

C.5 Project II: Time-to-event model of progression-free survival—Final covariate model

```
$PROBLEM

$INPUT     ID     ; Patient number
           TIME   ; Time of event
           EVID   ; Even tidentifier
           DV     ; Progression/death/Censored event
           SIM    ; Flag for simulation records
           C3D1   ; CRP concnetration at cycle 3
           ABC3C2 ; Difference in CRP concnetration between cycle 3 and 2
           LM     ; Flag for landmark time (day 42)

$DATA      5-CRP_PFS_COV_LM.csv IGNORE=@
           IGNORE=(ID.EQ.13028)       ; Patient with no CRP (i.e. all are BLQ)
           IGNORE=(LM.EQ.1)           ; Patients with events before cycle 3

;Sim_start : Add/remove for simulation
      IGNORE=(SIM.EQ.1)
;Sim_end

$ABBR PROTECT
$SUBR ADVAN = 13 TOL = 9
$MODEL COMP = (HAZARD)

$PK
;*************Covariates*************

;[1] CRP cycle 3
HZC3D1 = (1 + THETA(3)*(C3D1))

;[2] Difference in CRP concentration between cycle 3 and cycle 2
HZABC3C2 = ((ABC3C2)**THETA(4))
;----------------------
HZCOV = HZC3D1*HZABC3C2
;----------------------

;*************Hazard function: lognormal************

SDTTP = THETA(1)*EXP(ETA(1)) ; standard deviation of lognormal hazard model
MUTTP = THETA(2)             ; mean of lognormal hazard fucntion
PI   = 3.14159265
```

```
$DES
DEL   = 1E-16            ; Small number to avoid LOG(0)
TIM   = T+DEL
LTIM  = LOG(TIM)

X1X     = (LTIM-MUTTP)/SDTTP
PDF1X   = EXP(-(1/2)*((X1X)**2))/SQRT(2*PI)
LOGPFX  = ((1/(TIM*SDTTP))*PDF1X/(1-PHI(X1X))) ; lognormal baseline hazard
DADT(1) = LOGPFX*HZCOV

$ERROR
DELX = 1E-8 ; to avoid value zero of time

;----------TTE Model------------------------------
  IF(NEWIND.NE.2) OLDCHZ=0   ; Reset the cumulative hazard
  CHZ     = A(1)-OLDCHZ      ; Cumulative hazard
  OLDCHZ = A(1)              ; Rename old cumulative hazard
  SUR     = EXP(-CHZ)        ; Survival probability

TIMX    = TIME+DELX
LTIMX   = LOG(TIMX)

X1      = (LTIMX-MUTTP)/SDTTP
PDF1    = EXP(-(1/2)*(X1**2))/SQRT(2*PI)
LOGPFS  = ((1/(TIMX*SDTTP))*PDF1/(1-PHI(X1)))

HAZNOW = LOGPFS*HZCOV   ; hazard for PFS

  PDF = SUR*HAZNOW        ; Probability density function

  IF(DV.EQ.0) Y = SUR        ; Censored event (prob of not porgressing)
  IF(DV.NE.0) Y = SUR*HAZNOW ; Probability density function of event

  IF(ICALL.EQ.4) THEN ; For simulation
   CALL RANDOM (2,R)
   DV = 0              ; Event censeored or not
   RTTE = 0          ; Flag to tell whether there is an event or not
   IF(TIME.EQ.20832) RTTE = 1 ; for the censored observation at 124 weeks
   IF(R.GT.SUR) THEN
      DV = 1
      RTTE = 1
   ENDIF
  ENDIF
```

```
$THETA (0,0.9)    ; 1. Standard deviation of lognormal hazard model
$THETA (0,9)      ; 2. Mean of lognormal hazard fucntion
$THETA (0,0.1)    ; 3. Covariate: CRP cycle 3
$THETA (-0.26)    ; 4. Covariate: Difference in CRP concentration between
                  ;    cycle 3 and cycle 2

$OMEGA 0  FIX  ; 1. IIV standard deviation of lognormal hazard model

;Sim_start : add/remove for simulation
;$SIMULATION (5988566) (39978 UNIFORM) ONLYSIM NOPREDICTION SUB=100

$ESTIM MAXEVAL=9990 METHOD=0 LIKE PRINT=1 MSFO=msfb1 SIGL=9 NSIG=3
$COV PRINT=E
;Sim_end
```

C.6 Project II: Time-to-event model of overall survival—Final covariate model

```
$PROBLEM

$INPUT     ID    ; Patient number
           TIME  ; Time of event
           EVID  ; Event identifier
           DV    ; Death/Censored event
           SIM   ; Flag for simulation records
           BLSD  ; Baseline tumour size
           C3D1  ; CRP concentration at cycle 3
           ABC3C2 ; Difference in CRP concentration between cycle 3 and 2
           LIVERLES ; Liver lesions: Presence/absence
           LM    ; Flag for landmark time (day 42)

$DATA 4-CRP_SUR_COV_LM.csv IGNORE=@
       IGNORE=(ID.EQ.13028)       ;Patient with no CRP (all are BLQ)
       IGNORE=(LM.EQ.1)           ;Patients with events before cycle 3

;Sim_start : Add/remove for simulation
     IGNORE=(SIM.EQ.1)
;Sim_end

$ABBR PROTECT
$SUBR ADVAN = 13 TOL = 9
$MODEL COMP = (HAZARD)

$PK
;**************Covariates*************
;[1] CRP cycle 3
HZC3D1 = ((C3D1)**THETA(3))

;[2] Liver lesions
IF(LIVERLES.EQ.0) HZLIVERLES = 1 ; Most common
IF(LIVERLES.EQ.1) HZLIVERLES = ( 1 + THETA(4))

;[3] Difference in CRP concentration between cycle 3 and cycle 2
HZABC3C2 = ((ABC3C2)**THETA(5))

;[4] Baseline tumour size
HZBLSD = ((BLSD)**THETA(6))
;-------------------------------
HZCOV = HZC3D1*HZLIVERLES*HZABC3C2*HZBLSD
```

```
;*************Hazard function: Weibull************

HZLAM  = THETA(1)*EXP(ETA(1)) ; Scale factor
HZALPH = THETA(2)*EXP(ETA(2)) ; Shape factor

$DES
DEL     = 1E-8 ; To avoid value zero of time
DADT(1) = (HZLAM*HZALPH*(HZLAM*(T+DEL))**(HZALPH-1))*HZCOV ; Weibull

$ERROR
DELX = 1E-8 ; to avoid value zero of time

;----------TTE Model------------------------------
  IF(NEWIND.NE.2) OLDCHZ=0    ; Reset the cumulative hazard
  CHZ     = A(1)-OLDCHZ       ; Cumulative hazard
  OLDCHZ = A(1)               ; Rename old cumulative hazard
  SUR     = EXP(-CHZ)         ; Survival probability

 HAZNOW = (HZLAM*HZALPH*(HZLAM*(TIME+DELX))**(HZALPH-1))*HZCOV ;weibull
                          ; Rate of event each time point

  PDF = SUR*HAZNOW        ; Probability density function

  IF(DV.EQ.0) Y = SUR         ; Censored event (prob of survival)
  IF(DV.NE.0) Y = SUR*HAZNOW ; Probability density function of event

  IF(ICALL.EQ.4) THEN ; For simulation
   CALL RANDOM (2,R)
   DV = 0               ; Event censeored or not
   RTTE = 0           ; Flag to tell whether there is an event or not
   IF(TIME.EQ.23520) RTTE = 1 ; For the censored observation at 140 weeks
   IF(R.GT.SUR) THEN
      DV = 1
      RTTE = 1
   ENDIF
  ENDIF

$THETA (0.0000000001,0.00009)  ; 1. LAM: Scale factor
$THETA (0, 1.1)                ; 2. ALPHA: Shape factor
$THETA (0,0.4)                 ; 3. Covariate: CRP cycle 3
$THETA (0,0.5)                 ; 4. Covariate: Liver lesions
$THETA (-0.3)                  ; 5. Covariate: Difference in CRP concentration
                               ;     between cycle 3 and cycle 2
$THETA (0, 0.4)                ; 6. Covariate: Baseline tumour size
```

```
$OMEGA 0  FIX      ; 1. IIV LAM
$OMEGA 0  FIX      ; 2. IIV ALPHA

;Sim_start : add/remove for simulation
;$SIMULATION (5988566) (39978 UNIFORM) ONLYSIM NOPREDICTION SUB=100

$ESTIM MAXEVAL=9990 METHOD=0 LIKE PRINT=1 MSFO=msfb1 SIGL=9 NSIG=3
$COV PRINT=E
;Sim_end
```