Aus dem

Charité
Centrum Diagnostische und Präventive Labormedizin
 Institut für Pathologie

Direktor: Prof. Dr. Manfred Dietel

Habilitationsschrift

Quantitative Methoden in der Molekular- und Systempathologie

zur Erlangung der Venia Legendi für das Fach Experimentelle Pathologie

vorgelegt dem Fakultätsrat der Medizinischen Fakultät Charité Universitätsmedizin Berlin

von

Dr. Frederick Klauschen

Eingereicht: Oktober 2012

Dekanin: Prof. Dr. med. Annette Grüters-Kieslich

1. Gutachter: Prof. Dr. med. Hartmut Arps

2. Gutachter: Prof. Dr. med. Gunter Haroske

Inhaltsverzeichnis

Abkürzungen		
Einleit	ung	5
1.1	Quantitative Methoden in den	
	Wissenschaften	5
1.2 1.3	Die Relevanz der quantitativen Messbarkeit	6
	Zukunft der Histopathologie	8
1.4	Systembiologische Modellierung und	
	Simulation	9
1.5	Zielsetzung	10
Eigene	Arbeiten	11
2.1	Quantifizierung zellulärer Interaktionsdynamik in 3D Fluores-	
	zenzmikroskopiedaten	11
2.2	${\bf Sphingosin \hbox{-} 1\hbox{-} Phosphat\ mobilisiert\ Osteoklastenvorl\"{a}uferzellen\ und}$	
	reguliert Knochenhomöostase	20
2.3	Detektion und Segmentierung von Zellkernen in virtueller Mikro-	
	skopie:	
	Ein Minimum-Modell-Ansatz	27
2.4	Computerbasierte Rekonstruktion von Zellen und Gewebeober-	
	flächen für Modellierung und Datenanalyse	35
2.5	SAP-kontrollierte T-B Zellinteraktionen sind notwendig für die	
	Bildung von Keimzentren	44
2.6	Computerbasierte Modellierung von räumlich-dynamischen zel-	
	lulären Signalprozessen	52
2.7	Quantitative Analyse der diagnostischen	
	Richtlinien zur Bestimmung des Her2-Status	64
Diskus	sion	73
3.1	Bildanalyse	73

3.2	Modellierung und Simulation	76
3.3	Methodenevaluation	77
Zusammenfassung		
Literat	cur	80
Danksagung		
Erklär	ung	87

Abkürzungen

3D Drei Dimensionen bzw. drei-dimensional

Pixel Picture element, dt. Bildelement

RANKL Receptor Activator of NF- κ B Ligand

TNF– α Tumornekrosefaktor α

SAP Signaling lymphocyte activation molecular-associated protein

S1P Sphingosin-1-Phosphat

S1P₁ Sphingosin-1-Phosphat Rezeptor 1

Her2 Humaner epidermaler Wachstumsfaktorrezeptor 2

WSI Whole Slide Imaging, dt. digitalisierte Darstellung vollständiger histologischer Schnitte

 ${\sf Omics}$ Zusammenfassung für die auf -omics oder dt. -omik endenden Gebiete Proteomik, Genomik, etc.

 ${\sf FDA}\,$ Food and Drug Administration, US-amerikanische medizinische Zulassungsbehörde

Einleitung

1.1 Quantitative Methoden in den Wissenschaften

Während die Erfolge der modernen Physik und Ingenieurswissenschaften ohne parallele Entwicklungen in der Mathematik und deren Anwendung auf die Modellierung und Lösung physikalischer Probleme nicht möglich wären, ist die Biomedizin traditionell eine primär qualitative oder semi-quantitative, experimentelle Wissenschaft. Der Vorsprung der physikalischen Wissenschaften liegt einerseits daran, dass Mathematik schon im Altertum für naturphilosophische Betrachtungen für die Formulierung von Gesetzen verwendet wurde¹ und insbesondere in der Himmelsmechanik eine wichtige Rolle gespielt hat [1]. Außerdem haben sich die theoretischen und experimentellen physikalischen Naturwissenschaften parallel zueinander entwickelt, während die modernen molekularen Biowissenschaften erst viel später entstanden sind.

Ein anderer Grund liegt in der unterschiedlichen Komplexität der physikalischen gegenüber den Biowissenschaften: Naturgesetzte wie die Newton'schen Axiome lassen sich durch einfache mathematische Gleichungen formulieren, die trotz dieser Einfachheit in der makroskopischen Welt präzise Modelle ermöglichen [2], mit denen man z. B. den Flug einer Raumsonde zum Mars berechnen kann. Und sogar die Quanten- und Relativitätstheorie lassen sich durch relativ wenige, wenn auch komplexere mathematische Gleichungen beschreiben [3, 4]. Aber nicht nur physikalische Phänomene, sondern auch chemische Reaktionen lassen sich mathematisch, und zwar durch das Massenwirkungsgesetz, beschreiben [5]. Damit sind prinzipiell auch biologische Prozesse präzise mathematisch modellierbar, da sie sich aus biochemischen Reaktionen zwischen Proteinen, RNA, DNA und anderen Molekülen zusammensetzen. Allerdings ist die Funktion auch einfacher biologischer Vorgänge — ganz abgesehen vom Verhalten ganzer Zel-

¹Die erste quantitative Formulierung eines Natugesetzes wurde bereits vor 500 v.Chr. von der Pythagoreern durchgeführt, indem sie beschreiben, dass eine Harmonie der Töne erreicht wird, wenn sich die Saitenlängen wie kleine ganze Zahlen zueinander verhalten, S.61–63 in [1].

len oder eines Organs — abhängig von einer Vielzahl miteinander in komplexen Netzwerken wechselwirkender Komponenten. Zur sinnvollen quantitativen Analyse und Vorhersage der Funktion solcher Netzwerke sind die teilweise in der Physik möglichen Vereinfachungen nicht bzw. nur sehr begrenzt zulässig: Während z. B. zur Beschreibung der Bewegung eines Objektes (z. B. der o. g. Raumsonde) im Schwerefeld eines Planeten die Kenntnis der Schwerpunktkoordinaten, der Masse und der Anfangsgeschwindigkeit ausreichen, können Details über Form, Material oder innere Struktur weitgehend vernachlässigt werden [6]. Im Gegensatz dazu können auch kleinste biologische Details, z. B. das Vorhandensein eines Adapterproteins, das die Bindung eines Signalproteins an einen Rezeptor vermittelt, eine entscheidende Bedeutung für das Verhalten bzw. den Phänotyp einer Zelle haben, insbesondere weil ein Signalnetzwerk normalerweise viele solcher Details enthält bzw. die Details das Netzwerk erst definieren. Eine realistische, quantitativ-mathematische Beschreibung muss daher diese Details zumindest initial vollständig abbilden, um nicht Gefahr zu laufen, zwar ggf. vorhandene experimentelle Beobachtungen zu beschreiben, aber nicht die zugrunde liegenden Mechanismen zu erklären und somit keinen allgemeinen Vorhersagewert zu haben. Anders ausgedrückt: Auch wenn bio-molekulare Reaktionen fundamentalen physikalisch-chemischen Gesetzen folgen, reichen deren Kenntnis oder simplifizierte phänomenologische Ansätze nicht für das Verständnis und die Vorhersage biologischer Vorgänge aus, da deren funktionelle Komplexität aus unterschiedlichen Signalkompenten erwächst und es nicht die Ausnahme, sondern eher die Regel ist, dass "der Teufel" sozusagen im Detail steckt [7,8].

1.2 Die Relevanz der quantitativen Messbarkeit

Ein weiterer eng mit der Modellbildung verbundener, wesentlicher Aspekt der Verwendung quantitativer Verfahren in den Wissenschaften ist die Messbarkeit der relevanten Parameter. Modelle, deren Parameter (in der Physik sind das z. B. die Naturkonstanten) nicht durch eine Messung bestimmt und deren Vorhersagen nicht experimentell bzw. empirisch validiert werden können, mögen interessante Gedankenspiele sein, einen Beitrag zum besseren Verständnis der realen Welt leisten sie jedoch kaum² [9, 10].

Hierbei bewegen sich die Biowissenschaften zwischen den traditionell mathematisch geprägten Wissenschaften Physik (und auch Chemie) und den "mathematisierten" Wirtschafts- und Sozialwissenschaften. Während eine ganz fundamen-

²Bernhard Orth: "Das Messen ist sowohl historisch als auch methodisch gesehen eine der Grundlagen der Wissenschaft. Ohne die Durchführung exakter Messungen lässt sich die Entwicklung der empirischen Wissenschaften, insbesondere der Naturwissenschaften, nicht vorstellen."

tale Voraussetzung unseres physikalischen Weltbildes die universelle Gültigkeit der physikalischen Gesetze an verschiedenen Orten und zu verschiedenen Zeiten ist und die Tatsache, dass alle allgemein akzeptierten physikalischen Theorien umfassend experimentell validiert wurden³ [11], ist eine entsprechende Annahme in den Wirtschafts- und Sozialwissenschaften nicht bzw. nur sehr eingeschränkt zulässig. Quantitative makroökonomische "Gesetze" machen häufig abstrakte, unrealistische Annahmen bzw. Modellparameter gelten z. B. nicht notwendigerweise in allen Regionen und ebensowenig zu allen Zeiten und damit sind darauf aufbauende Vorhersage-/Modellierungsansätze problematisch⁴ [15, 16]. Auf ein weiteres Problem bei der Nutzung solcher Ansätze zur Testung des Einflusses einer Änderung makroökonomischer Parameter wies Robert Lucas 1976 hin, nämlich dass die gemessenen/geschätzten Modellparameter nicht invariant gegenüber den zu testenden Änderungen sind [17]. Spätestens jedoch bei der intertemporalen Variabilität und der Einbeziehung individueller Präferenzen und psychologischer Aspekte stößt die ökonomische Modellierung an ihre Grenzen, da hier die Grenze zwischen lediglich großem Aufwand der Messungen und der prinzipiellen Nicht-Messbarkeit erreicht ist.

Diese Ausführungen aus anderen Fachbereichen veranschaulichen, welche konzeptionellen Probleme bei der Modellierung komplexer Prozesse allgemein bestehen können, und sollen dabei helfen, die Situation in den Biowissenschaften im Folgenden besser beurteilen zu können.

In den Biowissenschaften besteht zusätzlich der Nachteil im Vergleich zur Physik, dass nicht nur eine überschaubare Zahl an – wenn auch technisch hochanspruchsvollen – Messungen, sondern eine Vielzahl an Detailmessungen notwendig ist, um die Komplexität biologischer Prozesse angemessen zu erfassen. Die Situation ist allerdings nicht so hoffnungslos wie in der Ökonomie, da die Modellparameter wenigstens prinzipiell messbar und – da sie auf den physikalischen Grundgesetzen beruhen – räumlich und zeitlich konstant⁵ sind. Die Voraussetzung ist allerdings, dass die Modellierung sich an den detaillierten molekularen Mechanismen orientiert.

Dass die Biomedizin traditionell nicht als quantitative Wissenschaft betrachtet wird, ist dementsprechend nicht darauf zurückzuführen, dass die einzelnen Para-

 $^{^3}$ Unter den nicht im täglichen Leben erfahrbaren physikalischen Phänomene gilt Einsteins allgemeine Relativitätstheorie als eine der am besten experimentell validierten Theorien.

⁴Ein prominentes Beispiel ist der in der Philipps-Kurve beschriebene Zusammenhang zwischen Inflationsrate und Arbeitslosigkeit, der von William Philipps für die Jahre 1861 bis 1957 für Grobritannien nachgewiesen wurde [12] und für darauf aufbauende Arbeiten wurden mehrere Nobelpreise verliehen. Die volkswirtschaftlichen Daten der folgenden Jahre bis heute zeigen hingegen nicht mehr das gleiche Phänomen und die Gültigkeit der Philipps-Kurve musste stark eingeschränkt werden [13, 14].

⁵Gemeint sind die den biologischen Prozessen zugrunde liegenden biophysikalischen oder biochemischen Gesetze; räumlich/zeitliche Änderungen aufgrund biologischer Prozesse (z. B. Mutationen oder Änderungen der Expression von Signalkomponenten) sind wiederum auch messbar.

meter prinzipiell nicht messbar sind, sondern dass die experimentellen Techniken erst in den letzten Jahren begonnen haben, quantitative Messungen zu ermöglichen. Hier erlauben es insbesondere Hochdurchsatzverfahren aus Genomik und Proteomik sowie neue hochaufgelöste fluoreszenzmikroskopische Verfahren, detaillierte und oft quantitative Informationen über biologische Prozesse zu gewinnen [18–23]. Allerdings wurde in der Vergangenheit bei aller zu rechtfertigenden Begeisterung über die neuen Messmethoden oft der Aufwand unterschätzt, die resultierenden Datenmengen zu analysieren und bzgl. ihrer inhaltlichen Aussage zu interpretieren [24, 25]. Ein Teil der dafür erforderlichen, computergestützten mathematischen Analysemethoden wird im Folgenden vorgestellt.

1.3 Bedeutung der Bildanalyse für quantitative Messungen und die Zukunft der Histopathologie

Die ersten medizininformatischen Ansätze, mit Hilfe der automatischen Bildanalyse Zellmerkmale zu quantifizieren, wurden vor mehreren Jahrzehnten entwickelt [26, 27]. Dass das Thema heute trotzdem immer noch ein hochaktuelles Forschungsgebiet ist, hat einerseits technische Gründe: So werden die mikroskopischen Techniken immer besser und erlauben, z. B. mit immer mehr gleichzeitig durch Fluoreszenzfarbstoffe abbildbaren Signalmolekülen, räumlich- und zeitlich hochaufgelöste Messungen biologischer Prozesse in situ oder sogar in vivo durchzuführen. Im Fall der Hellfeld-Lichmikroskopie von Gewebe, wie sie in der Histopathologie benutzt wird, sind erst in den letzten Jahren sog. "Whole Slide Scanner" (WSI, engl. "Whole Slide Imaging") entwickelt worden, mit denen komplette histologische Schnitte digitalisiert werden können. Des Weiteren sind die für die Datenspeicherung und -analyse notwendigen Computer auch erst in den letzten Jahren so leistungsfähig geworden, dass Computeranalysen für den Labor- oder gar diagnostischen Routineeinsatz praktikabel wurden [28,29]. Letzteres gilt insbesondere für die WSI-Anwendungen, da hierbei Bilddaten bis zu mehreren Gigabyte pro Präparat entstehen können.

Außerdem gibt es – wie bereits im vorangegangen Abschnitt angedeutet – inhaltliche Gründe für eine quantitative Bildauswertung. Je präziser die experimentellen Methoden werden und je mehr Daten simultan akquiriert werden können, desto weniger zielführend ist eine konventionelle "manuelle", d. h. qualitative oder höchstens semi-quantitative Auswertung. Neben einer nicht zu vernachlässigenden subjektiven Komponente bei einer manuellen (Bild-)datenanalyse ermöglicht es erst eine standardisierte quantitative Auswertung, das volle Potenzial der Daten zu nutzen. Dabei ist allerdings die quantitative Datenanalyse

wiederum nur eine wichtige Komponente bzw. die Basis der inhaltlichen Interpretation. Hierfür wird eine weiteres Feld, die Datenmodellierung und -simulation von immer größerer Bedeutung für die Biowissenschaften, wie der folgende Abschnitt erläutert.

1.4 Systembiologische Modellierung und Simulation

Wie in den vorangegangenen Abschnitten beschrieben, erfordern die immer komplexer werdenden experimentellen Daten auch immer aufwändigere Datenanalysestrategien. So liegt z. B. der Aufwand, in den Rohdaten einer kompletten Genomsequenzierung (Next-Generation- / Whole-Genome-Sequencing) die Genesequenzen und Mutationen zu detektieren, in der gleichen Größenordnung wie die experimentelle Komponente [30]. Ähnlich sieht es in der Massenspektrometrie-basierten Proteomik aus, wo statisch-bioinformatische Methoden angewandt werden müssen, um komplexen Peptidgemischen zuverlässig die entsprechenden Proteine zuzuordnen [31]. Dieser Teil der Datenanalyse ist allerdings nur ein erster Schritt und ermöglicht selbst noch keine Interpretation der Ergebnisse. Hierfür kommen Modellierungs- und Simulationsverfahren als weitere Datenanalyseebene ins Spiel, die es oft mit Hilfe von topologischen Informationen über biologische Netzwerke (z. B. Signaltransduktions- oder Genregulationsnetzwerke) erlauben, die in der initialen Analyse unabhängig erscheinenden Gen- oder Proteindaten miteinander in funktionelle Verbindungen zu setzen. Als ein frühes erfolgreiches Beispiel dient das Hodgkin-Huxley-Modell zur Simulation neuronaler Aktionspotenitale [32]. Davon abgesehen hat die traditionelle theoretische Biologie jedoch kaum wesentliche wissenschaftliche Beiträge geleistet, was, wie eingangs erwähnt, vor allem an der hohen biologischen Komplexität und fehlenden quantitativen Parametermessungen gelegen hat und sich der Wandel hin zu realitätsnahen, biologisch relevanten Simulationsmodellen erst in den letzten Jahren, ausgelöst durch die beschriebenen neuen Messtechniken und leistungsfähigeren Computer, eingestellt hat [33]. Eine der im Folgenden vorgestellten Arbeiten leistet einen Beitrag zur realitätsgetreuen dynamisch-räumlichen Simulation von Zellen und Zellverbänden. Die Integration von komplexen experimentellen Ergebnissen mit solchen neuartigen, quantitativen Modellierungsansätzen wird aktuell unter dem Oberbegriff "Systembiologie" oder auch "Systempathologie" zusammengefasst, ein interdisziplinäres Feld, von dem sich die Wissenschaft neue Antworten auf komplexe biomedizinische Fragen insbesondere im Bereich der personalisierten Medizin verspricht [34–37].

1.5 Zielsetzung

In der vorliegenden Arbeit sollen Analyse- und Simulationsmethoden für die Biomedizin und insbesondere die Molekular-/Systempathologie entwickelt und angewendet werden. Dies hat zum Ziel, experimentelle Daten computergestützt quantitativ-standardisiert auszuwerten sowie die Ergebnisse in dynamische Simulationsmodelle von zellulären Prozessen einfliessen lassen zu können. Die vorgestellten Methoden sollen eine effiziente und objektive Datenanalyse, insbesondere im Bereich der mikroskopischen Bildgebung, ermöglichen sowie es erlauben, die Vereinbarkeit experimenteller Daten mit Annahmen über biologische Mechanismen mit Hilfe von Simulationsmodellen in in silico Experimenten zu überprüfen. Des Weiteren soll der Nutzen mathematischer Methoden für die Evaluierung diagnostischer Standards in der Molekularpathologie demonstriert werden.

Eigene Arbeiten

2.1 Quantifizierung zellulärer Interaktionsdynamik in 3D Fluoreszenzmikroskopiedaten

Moderne Fluoreszenzmikroskopietechniken wie z. B. konfokale oder Multiphoton-Mikroskopie erlauben die Analyse biologischer Prozesse mit hoher räumlicher und zeitlicher Auflösung. Entsprechende Experimente liefern große und komplexe Datensätze über dynamisches Zellverhalten mit Hunderten bis Tausenden von Zellen unterschiedlichen Typs im Gewebekontext. Eine manuelle Auswertung ist daher nicht oder nur sehr eingeschränkt möglich, jedoch hat bisher die Entwicklung computergestützter Analysemethoden nicht Schritt gehalten mit den experimentellen Techniken.

Die hier vorgestellte Arbeit beschreibt, wie charakteristische Eigenschaften der Histogramme von Fluoreszenzmikroskopiedaten genutzt werden können, um – weitgehend unabhängig von der Bildqualität – Zellen und Gewebestrukturen zu segmentieren und zu detektieren. Insbesondere wird durch diesen Ansatz ermöglicht, automatisch Zelltrajektorien in zeitlich aufgelösten Daten zu berechnen und Zell-Zell-Interaktionen auszuwerten. Dabei können Interaktionen nicht nur detektiert und damit gezählt werden, sondern auch die Synapsengröße von Zell-Zell- oder Zell-Gewebe-Kontakten quantifiziert werden, was z. B. ein Maß für die Synapsenstabilität in T-B-Zellinteraktionen ist (vgl. 2.5) und ebenfalls in Form der Synapsenfläche zwischen Osteoklasten und Knochengewebe einen wichtigen Indikator im Knochenstoffwechsel darstellt, wie wir in der folgenden Studie anhand von 2-Photon-Fluoreszenzmikroskopiedaten zeigen.

Zusammengefasst stellt die Arbeit eine Histogramm-basierte computergestützte, sich automatisch an Bildqualitätsmerkmale anpassende Analysemethode zur Auswertung fluoreszenzmikroskopischer Daten vor, mit deren Hilfe gezeigt wird, dass Sphingosin-1-Phosphat (S1P) einen wesentlichen Einfluss auf den Knochenstoffwechsel über eine negative Regulation der Osteoklasten-Knochen-Interaktion ausübt (siehe auch 2.2).

Klauschen F, Ishii M, Qi H, Bajénoff M, Egen JG, Germain RN, Meier-Schellersheim M. Quantifying cellular interaction dynamics in 3D fluorescence microscopy data. *Nature Protoc*. 2009;4(9):1305-11.

S. 12-19

2.2 Sphingosin-1-Phosphat mobilisiert Osteoklastenvorläuferzellen und reguliert Knochenhomöostase

Osteoklasten sind als einzige Zelle in der Lage, Knochen zu resorbieren und sind somit eine wichtige Komponente im Mineralstoffwechsel. Dementsprechend sind Osteoklasten ebenfalls an der Pathogenese von Knochenerkrankungen wie Osteoporose oder rheumatoider Arthritis beteiligt. Während die Entstehung von Osteoklasten durch RANKL und TNF- α aus Monozyten relativ gut verstanden ist, untersucht die vorliegende Arbeit die Bedeutung des aus der Lymphoyztenmigration bekannten Lipids Sphingosin-1-Phosphat (S1P) für die Migration und Rekrutierung von Monozyten aus dem Blut und deren Fusion zu Osteoklasten an der Knochenoberfläche. S1P fördert die Rezirkulation von Lymphozyten aus Lymphknoten ins Blut.

Zur Analyse werden neben makroskopischen Knochendichtemessungen per Computertomographie 2-Photon-Mikroskopieexperimente durchgeführt, mit denen die Interaktion zwischen Fluoreszenzfarbstoff-markierten Monozyten / Osteoklasten und autofluoreszierendem Knochengewebe auf Einzelzellebene detailliert dargestellt werden kann. Da trotz gleicher Bedingungen jeder Datensatz in einem unabhängigen in vivo Experiment erzeugt wurde und dementsprechend variierende Bildqualitäten vorkommen, ist eine standardisierte Auswertung besonders wichtig. Hier kommt die im vorangegangenen Abschnitt 2.1 vorgestellte Analysemethode zum Einsatz, die es erlaubt, automatisiert die große Zahl der zur Hypothesentestung notwendigen Datensätze zu prozessieren und präzise quantitative Messwerte der Kontaktfläche zwischen Osteoklasten und Knochengewebe zu liefern, die wiederum als Maß für die Aktivität der Osteoklasten und damit den Knochenabbau herangezogen werden kann.

Die Ergebnisse zeigen, dass $S1P_1$ -(S1P-Rezeptor-1)-knock-out Mäuse osteoporotische Veränderungen aufweisen, die durch eine erhöhte Anlagerung der Osteoklasten an das Knochengewebe charakterisiert sind. Des Weiteren ergibt die Analyse, dass in $S1P_1$ -Wildtyp-Mäusen mit Ovarektomie-induzierter Osteoporose die Applikation des $S1P_1$ -Agonisten FTY720 (Fingolimod) die Osteoklastenanlagerung und damit auch die Osteoporose reduziert.

Die Arbeit zeigt beispielhaft, wie computergestützte Analysen standardisierte quantitative Messungen zellulärer Kontakte ermöglichen und damit einen wesentlichen Beitrag zur Überprüfung von Hypothesen über Wirkmechanismen von – potentiell klinisch relevanten – Medikamenten liefern können.

Ishii M, Egen JG, Klauschen F, Meier-Schellersheim M, Saeki Y, Vacher J, Proia RL, Germain RN. Sphingosine-1-phosphate mobilizes osteoclast precursors and regulates bone homeostasis. *Nature*. 2009 Mar 26;458(7237):524-8.

S. 21-26

2.3 Detektion und Segmentierung von Zellkernen in virtueller Mikroskopie: Ein Minimum-Modell-Ansatz

Während die in den beiden vorangegangenen Abschnitten vorgestellten Experimente und Methoden vor allem auf konfokaler und Multiphoton-Fluoreszenzmikroskopie basierten, die primär in der Grundlagenforschung angewendet werden, ist die konventionelle Hellfeldmikroskopie nach wie vor die dominante Methode in der diagnostischen Pathologie und in weiten Teilen der bildbasierten Biomarkerforschung.

Um in der quantitativen Bildanalyse nicht nur auf Fluoreszenzdaten beschränkt zu sein, wird in diesem Abschnitt eine Methode zur Segmentierung und Detektion von Zellen in konventionellen histologischen Schnittpräparaten vorgestellt. Bei der experimentell relativ aufwändigen (konfokalen/multiphoton) Fluoreszenzmikroskopie erleichtert die Tatsache, dass oft für die unterschiedlichen Zelltypen spezifische Marker eines umschriebenen Frequenzspektrums verwendet werden, die Bildanalyse im Vergleich mit konventioneller Mikroskopie. Bei der letzteren entsteht durch die Verwendung von Hellfeldbeleuchtung ein breites Spektrum an Farben bzw. Schattierungen im Rahmen der verwendeten Farbstoffe (i. d. R. Hämatoxylin und Eosin) und es ist daher nicht trivial, Zellen bzw. Zellkerne zu segmentieren, da diese nur schwer von unspezifischen Strukturen zu unterscheiden sind. Viele Methoden zur Zellerkennung machen daher Modellannahmen über die Form und andere Merkmale der Zellen und Zellkerne. Der Ansatz ist allerdings problematisch, weil diese Eigenschaften in Abhängigkeit von technischen Faktoren (Schnittdicken, Färbeintensität, Beleuchtung) und Gewebeeigenschaften sehr variabel sein und Modellannahmen daher ein Bias bzgl. der Detektion bewirken können. Insbesondere schwierig kann dies bei der Analyse von Tumorgewebe (einer naheliegenden Anwendung) werden, da Tumorzellen eine besonders hohe Form- und Größenvariabilität aufweisen können. Wir entwickeln daher einen Kontour-basierten Ansatz mit nur minimalen Modellannahmen und einer formunabhängigen Zellkernerkennung. Unsere Tests an einem manuell annotierten Goldstandarddatensatz aus mehr als 8000 Zellen verschiedenster Gewebe ergeben eine Präzision der Zellkerndetektion von über 90%.

Zusammengefasst beschreibt die Arbeit einen computergestützten Ansatz, der weitgehend unabhängig von zellulären Größen- und Formvariationen robuste Zellkernsegmentierungen ermöglicht und damit als Basis für automatische (immun-)histologische Analysen auch morphologisch vielgestaltiger Gewebe, wie z. B. maligner Tumoren, dienen kann.

Wienert S, Heim D, Saeger K, Stenzinger A, Beil M, Hufnagl P, Dietel M, Denkert C, Klauschen F. Detection and segmentation of cell nuclei in virtual microscopy images: a minimum-model approach. *Scientific Reports*. 2012;2:503.

S. 28-34

2.4 Computerbasierte Rekonstruktion von Zellen und Gewebeoberflächen für Modellierung und Datenanalyse

Computergestützte Datenanalyse wird in den meisten wissenschaftlichen Disziplinen im Allgemeinen dafür verwendet, die Komplexität der Daten zu reduzieren und statistische bzw. quantitative Ergebnisse zu erzielen. Dies war auch der Fall bei den in den vorangegangenen Abschnitten vorgestellten Arbeiten, in denen Bildanalysemethoden zur standardisierten und quantitativen Auswertung mikroskopischer Daten verwendet wurden.

In Fachgebieten, die sich mit Modellierung und Simulation von komplexen Systemen beschäftigen, wie z. B. in der Kosmologie, den Wirtschaftswissenschaften und dem relativ jungen Feld der Systembiologie, dient die Datenanalyse hingegen zusätzlich auch der Erzeugung von Modellparametern. An dieser Schnittstelle der klassischen Datenanalyse zur Modellparametergenerierung liegt der Beitrag der hier vorgestellten Arbeit, die eine Methode zur dreidimensionalen Rekonstruktion biologischer Objekte wie Zellen oder Gewebestrukturen vorstellt. Die Methode ist primär für die Verwendung von konfokalen oder Multi-photon-Fluoreszenmikroskopiedaten entwickelt worden, kann jedoch flexibel auch mit anderen Bildmodalitäten verwendet werden. Der Ansatz basiert auf einem geometrischen Konzept, den sog. Voronoi (oder auch Dirichlet) Repräsentationen, die über eine iterative Optimierung Objektoberflächen mit definierbarer Präzision numerisch approximieren. Wir demonstrieren die Methode in zwei Anwendungen: Erstens anhand der Rekonstruktion von einzelnen Zellen aus hochaufgelösten konfokalen Mikroskopiedaten, mit der man detaillierte morphologische 3-D Modelle als Basis für die Durchführung dynamischer zellulärer Simulationen erzeugen kann (zur konkreten Anwendung in Simulationsmodellen s. 2.6). Im zweiten Beispiel wird gezeigt, wie Geweberegionen rekonstruiert werden können. Diese Anwendung kann zur Visualisierung und Parameterisierung (d. h. Vermessung) von Gewebestrukturen verwendet werden.

Zusammenfassend betrachtet ermöglicht die vorgestellte Methode, durch sich automatisch optimierende Voronoi-Gitter realistische dreidimensionale Modelle von Zellen oder Gewebearealen aus Mikroskopiedaten zu erzeugen, die zur Visualisierung, direkten Datenanalyse oder als Basis für dynamische Simulationsmodelle verwendet werden können.

In der im nächsten Abschnitt vorgestellten Arbeit wird diese Methode zur quantitativen Analyse des Migrationsverhaltens von Lymphozyten in Lymphknoten angewendet.

Klauschen F, Qi H, Egen JG, Germain RN, Meier-Schellersheim M. Computational reconstruction of cell and tissue surfaces for modeling and data analysis. *Nature Protoc*. 2009;4(7):1006-12.

S. 36-43

2.5 SAP-kontrollierte T-B Zellinteraktionen sind notwendig für die Bildung von Keimzentren

Die hier vorgestellte Arbeit zeigt die Anwendung der im vorangehenden Abschnitt beschriebenen Methode.

Keimzentren sind im Lymphknoten die Orte, an denen die langfristige adaptive Antikörper-vermittelte Immunität u. a. durch Antigen-spezifische Interaktion zwischen T- und B-Zellen erzeugt wird. Aus Krankheitsstudien und Mausmodellen ist bekannt, dass das T-Zell Adapterprotein SAP "signaling lymphocyte activation molecular-associated protein" für die Entstehung von Keimzentren notwendig ist, ohne dass jedoch bisher ein Mechanismus bekannt war. Die vorgestellte Arbeit zeigt mit Hilfe von intravitaler Multiphoton-Fluoreszenzmikroskopie, dass SAP-Defizienz selektiv die Stabilität der Interaktion von CD4-T-Zellen mit spezifischen B-Zellen reduziert, während Kontakte mit dentritischen Zellen unbeeinflusst bleiben. Dies führt zu einer ineffizienten Rekrutierung von T-Zellen in die Keimzentren und dadurch zu einer insuffizienten Keimzentrumsreaktion. Die experimentelle Untersuchung der Lymphozytenmigration in Bezug zu den Keimzentren ist möglich, da die Multiphoton-Fluoreszenzmikroskopie zeitlich aufgelöste dreidimensionale (d. h. "4-D") Bilddaten mit mehreren Farbkanälen für die unterschiedlichen Zelltypen (T-, B- und dendritische Zellen) liefert. Dies erlaubt eine Nachverfolgung individueller Zellen (sog. "cell-tracking" zur Extraktion der Zelltrajektorien) im Gewebekontext. Da Analyse neben der Komplexität der einzelnen Datensätze mehrere Experimentwiederholungen bewältigen muss, ist eine manuelle Auswertung extrem zeitaufwändig und fehleranfällig. Daher wurde eine automatisierte computergestützte Analysemethode auf Basis der im vorangehenden Abschnitt vorgestellten Methode entwickelt. Diese wurde verwendet, um die Keimzentrumsmorphologie zu parameterisieren, d. h. in ein morphologisches Modell zu überführen, das mit den aus der zeitlich aufgelösten Mikroskopie erhaltenen Zelltrajektorien eine quantitativ-statistische Analyse der Zellmigration in und aus den Keimzentren erlaubt und die oben genannten Ergebnisse ermöglichte.

Zusammengefasst zeigt die Arbeit unter Verwendung von aus Mikroskopiedaten rekonstruierten Modellen von Keimzentren (vgl. 2.4) und einer Analyse der Tund B-Zellmigration und -interaktion die Bedeutung des T-Zell-Adapterproteins SAP für die Bildung von Keimzentren in Lymphknoten.

Qi H, Cannons JL, Klauschen F, Schwartzberg PL, Germain RN. SAP-controlled T-B cell interactions underlie germinal centre formation. *Nature*. 2008 Oct 9;455(7214):764-9.

S. 45-51

2.6 Computerbasierte Modellierung von räumlichdynamischen zellulären Signalprozessen

Wie in der Einleitung beschrieben, dient Bildanalyse nicht ausschließlich zur Erzeugung von Ergebnissen aus experimentellen Daten, sondern kann auch eingesetzt werden, um Parameter für Simulationsmodelle zu generieren. Als Beispiel dafür wurde in Abschnitt 2.4 gezeigt, wie morphologische Zellmodelle aus konfokalen Mikroskopiedaten erzeugt werden können. In der im Folgenden vorgestellten Arbeit wird die letzte Stufe in der Sequenz (Bild-)Datenanalyse \rightarrow $Modellgenerierung \rightarrow Modellsimulation vorgestellt.$ Dabei handelt es sich um den ersten systembiologischen Ansatz, der es ermöglicht, realistische, biochemisch und morphologisch dynamische Simulationsmodelle von Signaltransduktionsprozessen zu erstellen. So können auch Biowissenschaftler ohne mathematische Kenntnisse komplexe Signaltransduktionsmodelle anhand von bimolekularen Wechselwirkungen definieren und die Simulationergebnisse aufgrund der morphologisch-biochemischen Detailtreue direkt mit den experimentellen Daten abgleichen. Aus experimentellen Mikroskopiedaten und dem Wissen bzw. Hypothesen über die zu untersuchenden Signalpfade werden morphologisch-biochemische (Zell-)Modelle erzeugt, die im nächsten Schritt über eine bestimmte Zeitspanne simuliert werden; schließlich erhält man simulierte Ergebnisbilder, die es z. B. ermöglichen, unmittelbar visuell die Validität der Hypothese anhand von Abweichungen bzw. Übereinstimmungen zwischen Experiment und Simulation zu überprüfen.

Zusammengefasst stellt die Arbeit die erste Methode zur sowohl biochemischals auch morphologisch realistischen und dynamischen Modellierung und Simulation von zellulären Signalprozessen vor. Solche Simulationsmodelle können in Zukunft im Rahmen eines molekular- und systempathologischen Ansatzes einen wichtigen Beitrag für das Verständnis von komplexen Pathomechanismen leisten, insbesondere mit Hinblick auf die Entstehung und Progression von Tumorzellen sowie die Vorhersage der Wirksamkeit von Kombinationstherapien (s. Diskussion).

Angermann BR, Klauschen F, Garcia AD, Prustel T, Zhang F, Germain RN, Meier-Schellersheim M. Computational modeling of cellular signaling processes embedded into dynamic spatial contexts. *Nature Methods*. 2012 Jan 29;9(3):283-9.

S. 53-64

2.7 Quantitative Analyse der diagnostischen Richtlinien zur Bestimmung des Her2-Status

In der letzten der hier vorgestellten Arbeiten wird ein weiterer Aspekt der quantitativen Datenanalyse im Zusammenhang mit Simulationsmethoden beschrieben. Während der Fokus der in den vorausgehenden Abschnitten auf der quantitativen Auswertung vor allem mikroskopischer experimenteller Daten sowie darauf aufbauender zellulärer Simulationsmodelle lag, wird hier Datenanalyse in Kombination mit sog. Monte-Carlo-Simulationen dafür verwendet, die Validität einer weit verbreiteten diagnostischen Methode zu untersuchen. Es handelt sich dabei um die in der täglichen Routinediagnostik verwendete in-situ-Hybridisierung zur Detektion einer Amplifikation des Her2-Gens in Magen- und Mammakarzinomen. Patienten mit einem Her2-positiven Tumor können von der Behandlung mit Trastuzumab profitieren, einem den Her2-Rezeptor blockierenden Antikörper. Die Bestimmung des Her2-Status wird in zwei Stufen durchgeführt: Fälle ohne oder mit nur sehr geringer membranärer Expression von Her2 (Score 0 oder 1+) in der Immunhistologie werden als Her2-negativ eingeordnet, hochgradige Her2-Signale ergeben einen Score 3+ entsprechend Her2-positiv. Bei mäßiggradiger Expression von Her2 wird per in-situ Hybridisierung die Kopiezahl der Her2-Gene ins Verhältnis zur Anzahl von Chromosom 17 gesetzt und beim Vorliegen einer (spezifischen) Amplifikation (Ratio Her2/Chr17 2,0) der Fall als Her2-positiv diagnostiziert.

Während sich verschiedene Studien mit dem Einfluss der Laborprozessierung auf die Diagnostik beschäftigen, wird hier die erste rigorose quantitative Evaluation der diagnostischen Richtlinien zur Her2-Auswertung präsentiert. Dabei untersuchen wir mit Hilfe von Monte-Carlo-Simulationen die Abhängigkeit der diagnostizierten Her2-Chr17-Ratios von 1) der Stichprobengröße (die Anzahl der ausgezählten Zellen), 2) die statistischen Verteilungen von Gen- und Chromosomzahlen sowie 3) die für die Quotientenkalkulation verwendete Methode. Mit Hilfe von aus einer repräsentativen Zahl realer Daten gewonnener statistischer Verteilungen der Her2- und Chr17-Signale wird der Diagnosevorgang im Computer simuliert. Dadurch kann der Einfluss der Parameter auf die Diagnose systematisch evaluiert werden. Die Ergebnisse zeigen, dass die aktuellen Empfehlungen zur Her2-Diagnostik zu Variationen der Her2-Chr17-Ratios von bis zu 0,94 Punkten und damit zu fehlerhaften Her2-Status-Diagnosen führen kann, wenn man den oben genannten, von der amerikanischen Zulassungsbehörde FDA vorgegebenen Grenzwert von 2,0 verwendet.

Zusammenfassend sprechen unsere Ergebnisse für eine Revision der aktuellen Empfehlungen hin zu mehr ausgezählten Zellen und zeigen, dass die vorgestellte Methode von allgemeiner Bedeutung für die Evaluation von quantitativer Diagnostik ist.

Stenzinger A, von Winterfeld M, Aulmann S, Warth A, Weichert W, Denkert C, Rüschoff J, Dietel M, Klauschen F. Quantitative analysis of diagnostic guidelines for HER2-status assessment. *Journal of Molecular Diagnostics*. 2012 May-Jun;14(3):199-205.

S. 65-72

Diskussion

Quantitative Analysen werden nicht nur in vielen wissenschaftlichen Gebieten immer wichtiger, sie sind auch präsent in vielen gesellschaftspolitischen Diskussionen [38]. Das kann einerseits damit zusammenhängen, dass mathematischstatistische Analysen allgemein als standardisiert und damit objektiv angesehen werden, aber auch damit, dass die Basis unserer technologischen Errungenschaften und damit der wirtschaftlichen Prosperität weitgehend in den Erkenntnissen der durch die Mathematik geprägten Wissenschaften (Physik, Ingenieurswissenschaften, Chemie, Informatik) liegt. Es besteht jedoch auch die Gefahr, dass ein unkritischer Umgang mit mathematisch-quantitativen Methoden zu vermeintlich objektiven, aber dennoch falschen Schlussfolgerungen führt, was insbesondere in den Wirtschaftswissenschaften, aber auch der Medizin problematisch sein kann, wenn die Modellannahmen nicht der Realität entsprechen, jedoch die Schlussfolgerungen auf die Realität verallgemeinert werden oder wenn eine Analysemethode nicht die Präzision erlaubt, die für die letztendlich getroffene Aussage erforderlich wäre [16]. Trotz allem ist insbesondere in Medizin und Biowissenschaften ein intensiver Einsatz quantitativ-mathematischer Verfahren essentiell, um z. B. den Inhalt der im Zeitalter der sog. "Omics"-Verfahren gewonnenen Datenmengen zu verarbeiten und zu verstehen und um in der Medizin objektive, evidenz-basierte Aussagen zu neuen Diagnose- und Therapieverfahren treffen zu können [39–41].

Die in dieser Arbeit vorgestellten Studien beschäfitgen sich alle mit der Anwendung quantitativer Methoden in Medizin und Biologie, und zwar im Wesentlichen auf drei Themenfeldern, die im Folgenden diskutiert werden.

3.1 Bildanalyse

Im zweiten größeren Gebiet werden Arbeiten vorgestellt, die die Entwicklung neuer Methoden zur Bilddatenanalyse sowie deren Anwendung auf konkrete biologische Fragestellungen zeigen. Das Spektrum reicht hier von Grundlagenan-

wendungen aus dem Bereich der Fluoreszenzmikroskopie (konfokal und multi-photon) bis zu histopathologischer Zellkernsegmentierung mit dem Ziel des Einsatzes in der Routinediagnostik. Trotz der zahlreichen Beiträge über das Thema (z. B. [42–46]), gibt es auch heute noch keine umfassend einsetzbare, praxistaugliche Methode und die wissenschaftliche Diskussion über den besten methodischen Ansatz dauert an. Dass dieses Gebiet erst in den letzten Jahren auch verstärkt Interesse auf Seiten der Medizin erweckt hat, liegt, wie eingangs erklärt, zum einen an den wachsenden Anforderungen an standardisierte quantitative diagnostische Analysen, aber auch an neuen technischen Möglichkeiten, Bildanalyse überhaupt in der Praxis effizient einsetzen zu können. Die Gründe dafür, dass histologische Bildanalyse trotz intensiver Bemühungen bisher noch nicht in der Routinediagnostik Einzug gehalten hat, sind aber primär in den biologischen Gewebeeigenschaften und den histolabortechnischen Methoden zu finden. Die meisten Arbeiten benutzen nur relativ kleine Datensätze mit relativen homogenen Zell- und Gewebseigenschaften, da die neuentwickelte mathematische Methode im Vordergrund steht und die erfolgreiche beispielhafte Anwendung auf relativ wenige Bilddaten oft als ausreichend angesehen wird. Das Problem dabei ist aber die extrem hohe morphologische Variabilität schon normaler Gewebe und erst recht pathologischer Veränderungen. So wird die Malignität von Tumoren u. a. durch die Größen- und Formvariation der Zellkerne abgeschätzt. Daher kann es bei automatischen Zellerkennungsmethoden, die (wie es oft der Fall ist) Annahmen über die Zellgröße und -form machen, zu Fehlern kommen. Soll aber z. B. ein automatisches "Grading" (zur Malignitätsbestimmung) des Tumors durchgeführt werden [47], ist es essentiell, dass alle Varianten der Tumorzellen erkannt werden. Das von uns vorgestellte Zellkernsegmentierungsverfahren minimiert daher die Modellannahmen, so dass unterschiedlichste Zellformen zuverlässig erkannt werden. Wir greifen dabei das bei den Validierungsstudien angesprochene Thema der Gold-Standard-Daten wieder auf, indem wir knapp 10,000 von Pathologen annotierte Zellen verschiedenster Gewebe-/Tumotypen zur Validierung verwenden. Eine Erweiterung dieses Datensatzes mit über 100,000 Zellen ist bereits in Arbeit und kann in Zukunft zur Validierung verschiedenster Bildanalysemethoden verwendet werden.

Die Mehrzahl der Bildanalyseverfahren im Bereich der Histologie basiert auf der Zellsegmentierung, d. h. im Bild enthaltene Objekte werden erst voneinander getrennt ("Segmentierung"). Erst danach wird entschieden, ob es sich dabei um Zellen/Zellkerne oder andere Objekte handelt ("Klassifikation"). Dabei ist die Segmentierung ein kritischer und vergleichsweise fehleranfälliger Schritt und die Kritik an dem Ansatz lautet, dass die eigentliche Aufgabe, nämlich die Zellklassifikation, erst über den Umweg der Segmentierung erreicht wird. Als Alternative wird zur Zeit intensiv an der sog. "direkten Klassifikation" geforscht, bei der

unter Verwendung von neuen Algorithmen aus dem Bereich der Künstlichen Intelligenz und des Maschinellen Lernens eine Zelldetektion ohne vorausgehende Segmentierung möglich ist; allerdings ist hier die Qualität maßgeblich von der Qualität der notwendigen Trainingsdaten abhängig [48–50]. Außerdem lassen vorhandene Verfahren keinen Rückschluss auf die genaue Lokalisation der klassifizierten Objekte in einer gewissen Region zu. In einem aktuellen Projekt verwenden wir einen mehr als 100,000 Mammakarzinomzellen und Normalgewebe umfassenden, manuell annotierten Datensatz und entwickeln einen Klassifikationansatz, der es erlaubt, die Vorteile von Segmentierung und direkter Klassifikation zu verbinden und dadurch die Lokalisation der Tumorzellen auf Pixelebene zu berechnen.

Eine im Vergleich zum Einsatz in der Hellfeld-Lichtmikroskopie relativ junge Anwendung der computergestützten Bildanalyse ist die dreidimensionale konfokale oder Multi-Photon-Mikroskopie. Ein Vorteil aus bildanalytischer Sicht ergibt sich aus der Möglichkeit, durch Fluoreszenzfarbstoffe Zelltypen oder Zellkompartimente spezifisch zu markieren, so das ein Teil der Klassifikation bereits implizit in den Daten enthalten ist. In histologischen Hellfeldbildern sind zwar z. B. die Farbstoffe Hämatoxylin und Eosin auch spezifisch für Zellkerne und Zytoplasma, jedoch sind die Farbspektren vergleichsweise breit und stark überlappend. Dafür sind die Fluoreszenzsignale oft stärker verrauscht und man beobachtet Intensitätsinhomogenitäten innerhalb eines Bildes sowie insbesondere entlang der z-Achse bei Multi-Photon-Mikroskopie [51]. Ebenfalls unterscheiden sich die unterschiedlichen Fluoreszenzfarbstoffe, so dass die Bildqualität der unterschiedlichen Kanäle in gleichen Bildregionen variiert. Die hier vorgestellte Methode zur Analyse von 3-D Fluoreszenzmikroskopie nutzt deshalb Histogrammeigenschaften, passt die Segmentierung adaptiv an die Veränderungen innerhalb des Datensatzes an und gleicht automatisch Unterschiede zwischen verschiedenen Kanälen aus, so dass z. B. Interaktionen unterschiedlich markierter Zelltypen standardisiert bewertet werden können.

Wir zeigen unter anderem (s. 2.2), wie man die beschriebene Bildanalysemethode einsetzen kann, um die Wirkung eines potentiellen Medikamentes auf den Knochenstoffwechsel standardisiert zu messen. Der Wirkstoff FTY720 führt dabei zu einer Reduktion der Anlagerung von Osteoklasten an Knochen und damit im Mausmodell zu einer Reduktion der Osteoporose. Die Bildanalyse erlaubt es erstmals, automatisiert die Kontaktfläche zwischen Osteoklasten und Knochen zu quantifizieren und ist der konventionellen Methode der manuellen Morphometrie [52] bzgl. Standardisierung und Geschwindigkeit klar überlegen.

Als Fazit kann zur Bildanalyse festgehalten werden, dass es bei allen Fortschritten der letzten Jahre noch keine allgemein akzeptierte, praxistaugliche Methode z. B. zur vollautomatisierten Detektion von Tumorzellen und der immun-

histologischen Markerquantifizierung gibt. Insbesondere bei der Erkennung von Tumorzellen in einem komplexen histologischen Hintergrund ist zu erwarten, dass auch kurz- bis mittelfristig das menschliche Gehirn computerbasierten Ansätzen überlegen bleiben wird. Allerdings zeigen die diagnostische Praxis und systematische Studien, dass auch Experten i. d. R. nur sehr eingeschränkt eine robuste und objektive Quantifizierung (immun-)histologischer Merkmale bewältigen können [53,54]. An dieser Stelle können bereits jetzt und in naher Zukunft Computerassistenzsysteme einen Beitrag leisten, in vom Pathologen als Tumor identifiziertem Gewebe z. B. immunhistologische Marker objektiv und standardisiert zu quantifizieren.

3.2 Modellierung und Simulation

Das dritte Gebiet schlägt einen Bogen von der Daten- und Bildanalyse zu dynamischen Simulationsmodellen von zellulären Signalprozessen. Hierbei knüpft die Modellbildung direkt an die oben beschriebenen Arbeiten an, da für die Konstruktion realistischer räumlicher Zellmodelle direkt Mikroskopiedaten verwendet werden können, die zuvor bildanalytisch verarbeitet werden (siehe 2.4, 2.1). Der hier vorgestellte Ansatz zeichnet sich aus durch die einfache, auch für Biowissenschaftler ohne mathematisch-physikalisches Wissen handhabbare Modellgenerierung, ohne dass dadurch die Biologie nicht angemessen repräsentierende Vereinfachungen notwendig sind. Während andere Methoden zwar auch räumlich aufgelöste, aber morphologisch statische oder zwar morphologisch dynamische, aber biochemisch stark vereinfachte Modellierung ermöglichen [55,56], ist die hier beschriebene Methode die erste, die sowohl biochemisch als auch morphologisch dynamische hochrealistische Simulationsmodelle von Zellen oder Zellgruppen erlaubt. Dabei werden sämtliche, für die computerbasierte Simulation notwendigen mathematischen Gleichungen zur Abbildung der Signalnetzwerke automatisch erzeugt. Dies ist umso wichtiger, da selbst kleinere Signalnetzwerke aufgrund der intermolekularen Komplexbildung i. d. R. viel mehr Reaktionsgleichungen erzeugen als Molekülarten vorhanden sind [57-59]. Dadurch ist eine manuelle Erstellung der Gleichungen nicht nur sehr aufwändig, sondern auch sehr fehleranfällig, da z. B. bei der Einführung von "knock-out" Mutanten oder der in silico Testung von spezifischen Inhibitoren einer zielgerichteten Therapie nicht nur die Eigenschaften eines Proteins geändert werden müssen, sondern auch sämtliche das Protein enthaltenden Komplexe.

Dieser systempathologische Ansatz ist nicht nur für die Grundlagenforschung relevant, sondern kann auch einen wesentlichen Beitrag zur personalisierten Medizin an der Schnittstelle von Diagnostik und Therapie leisten: Sein volles Po-

tential entfaltet die Simulationsmodellierung dann allerdings erst in Kombination mit Ergebnissen experimenteller (Hochdurchsatz-) Verfahren, wie z. B. der Genomik, quantitativer Fluoreszenzmikroskopie und vor allem der (Phospho-)-Proteomik zusammen mit Informationen über Signalnetzwerkkonnektivität und Protein-Protein-Wechselwirkungen ("Interactomics") [60-63]. Insbesondere die Phosphoproteomik und die damit möglichen simultanen Messungen des Aktivitätszustandes einer großen Zahl von Proteinen sind geeignet, die für die Simulationsmodelle benötigten quantitativen Parameter zu liefern [64,65]. Während solche experimentellen Daten allein bzw. auch mit Hilfe von konventionellen statistisch-bioinformatischen Analysemethoden nur sehr eingeschränkt funktionelle Einsichten erlauben, ermöglicht die hier vorgestellte Methode die Integration mit in silico Experimenten und dadurch eine detaillierte funktionelle Analyse pathologischer Veränderungen von untersuchtem Tumorgewebe und wiederum damit die systematische Suche nach prädiktiven Biomarkern bzw. Ansatzpunkten zielgerichteter Kombinationstherapien (vgl. [66–68]). Wir verfolgen diesen Ansatz in einem aktuellen Projekt, in dem mit Massenspektrometrie-basierter Phosphoproteomik die Regulation der Signalnetzwerke in Abhängigkeit von aktivierenden EGFR-Mutationen in Tumoren untersucht wird.

3.3 Methodenevaluation

Die Arbeit zur Analyse der auf in-situ-Hybridisierung basierenden Her2-Genamplifikationsdiagnostik kann der Kategorie "Evaluation von Analysemethoden" zugeordnet werden. Dabei handelt es sich um ein Gebiet, das traditionell eher unterentwickelt ist, was wahrscheinlich u. a. damit zusammenhängt, dass auf Seiten der Endanwender oft keine Detailkenntnis der Analysemethoden vorhanden ist und somit wenig Anreiz besteht, eine einmal allgemein akzeptierte Methode in Frage zu stellen.

Daher ist es das Ziel der Studie über die Her2-Status-Bestimmung bei Magenund Mammakarzinomen, zu analysieren, inwieweit die Anwendung der durch die Richtlinien vorgegebenen diagnostisch-methodischen Kriterien zuverlässige und adäquate Diagnosen garantiert. Bei der in diesem Test verwendeten Genamplifikationsanalyse durch in-situ-Hybridisierung muss eine bestimmte Anzahl von Tumorzellen bzgl. der in den Zellkernen vorhandenen Sondensignale durch den Pathologen ausgezählt werden. Allerdings wurden die genauen Zahlenwerte in der diagnostischen Richtlinie von einer Expertenrunde ohne eine vorangehende quantitative Evaluierung der Kriterien festgelegt [69,70]. Die vorliegende Arbeit zeigt, dass die aktuellen Richtlinien in Abhängigkeit von der Variation der Genamplifikation innerhalb des Tumorgewebes auch bei sorgfältiger Auszählung zufällig zu falsch als Her2-positiv oder -negativ (oder umgekehrt) diagnostizierten Fällen führen können. Dadurch, dass der beschriebene Ansatz eine Abschätzung des zu erwartenden Fehlers in Abhängigkeit vom Auszählmodus erlaubt, können damit konkrete, statistisch fundierte Empfehlungen zur Verbesserung der Diagnostik gemacht werden.

Zur Umsetzung wurden hierbei simulierte Datensätze erzeugt, die auf einem breiten Spektrum aus realen Daten gewonnener typischer Verteilungen der Her2und Chromosom-17-Signale basieren. Aus diesen Verteilungen wurden dann zahlreiche Zufallsstichproben gezogen, die jeweils einen diagnostischen Vorgang simulieren und so eine Aussage darüber erlauben, wie wahrscheinlich unter Verwendung der Richtlinien eine falsche Bewertung stattfindet.

Diese auch unter dem Begriff der Monte-Carlo-Simulationen zusammengefassten Methoden erlauben es allgemein, simulierte, aber realistische Gold-Standard-Daten zu erzeugen, für die die wahren Eigenschaften bekannt sind und die somit zur Methodenevaluierung herangezogen werden können. Die Erkenntnisse können daher über die Her2-Evaluierung hinaus als beispielhaft für die Definition und Evaluierung von diagnostischen Richtlinien histologischer bzw. molekularer Tests angesehen werden.

Zusammenfassung

Nicht nur in der biomedizinischen Forschung, auch in der klinischen Medizin steigen die zu verarbeitenden Datenmengen rasant und es existiert ein allgemeiner Trend hin zu quantitativen Messwerten und Aussagen. Das Spektrum reicht dabei von histopathologischen Tests, in denen über quantitative Auswertungen (z. B. Her2-Statusbestimmung) prädiktive, d. h. therapieentscheidende Aussagen getroffen werden, bis zu systembiologischen Fragestellungen mit Tausenden von Simultanmessungen aus Hochdurchsatzverfahren der Genomik und Proteomik oder neuer mikroskopischer Techniken. Die Validität solcher quantitativer Diagnostik zu gewährleisten und die anfallenden Datenmengen biologisch interpretieren zu können, erfordert geeignete mathematisch-analytische Verfahren, deren Entwicklung gegenüber den technisch-experimentellen Verfahren zurück liegt. Die in dieser Schrift vorgestellten Arbeiten leisten hier einen Beitrag im Wesentlichen in drei teilweise miteinander überlappenden Feldern.

Als Hauptteil der Arbeit beschreiben wir die Entwicklung von Bildanalysemethoden für Fluoreszenz- und Hellfeldmikroskopie und wie deren Anwendung effiziente Analysen mit objektiven Aussagen zur Bedeutung des T-Zell Adapterproteins SAP für die Bildung von Keimzentren sowie zum Einfluss von S1P auf die Osteoklastenfunktion ermöglicht. Neben dieser primär analytischen Auswertung experimenteller Daten rückt im Rahmen systembiologischer Ansätze immer mehr die Modellierung und Simulation von biologischen Prozessen ins wissenschaftliche Interesse. Die in diesem Kontext hier vorgestellte Arbeit beschreibt erstmals die einfache Erstellung realistischer, biochemisch und morphologisch dynamischer Simulationsmodelle auch komplexer biologischer Signalnetzwerke. Die Methode kann in Zukunft einen wichtigen Beitrag für die personalisierte Medizin leisten, da sie die Integration von Ergebnissen aus experimentellen Genomik- oder vor allem Proteomikstudien über pathologisch veränderte Prozesse bei Tumoren mit funktionellen in silico Experimenten erlaubt und so die Mechanismen der Wirkung zielgerichteter (Kombinations-)Therapien systematisch evaluiert werden können.

Im letzten Teil wird mit der Methodenevaluation ein weiterer wichtiger Anwendungsbereich quantitativer Analyse- und Simulationsmethoden anhand der routinediagnostischen Bestimmung des Her2-Status beschrieben. Es wird gezeigt, dass erhebliche Ergebnisunterschiede zwar abhängig von Art und Qualität der Daten, jedoch primär durch Eigenschaften der Analysemethode zustande kommen können. Dass die Ergebnisfluktuationen im Fall der Her2-Diagnostik therapieentscheidend sein können, verdeutlicht, wie wichtig eine rigorose Validierung insbesondere quantitativer Methoden ist.

Insgesamt zeigen die hier vorgestellten Arbeiten den bereits aktuell wichtigen Beitrag quantitativ-analytischer Ansätze zur standardisierten Datenauswertung und Ergebnisinterpretation und lassen deren weiter zunehmende Bedeutung für die Zukunft der Biomedizin erkennen.

Literaturverzeichnis

- [1] K. Simonyi. Kulturgeschichte der Physik. Verlag Harri Deutsch, 2001.
- [2] I. Newton. *Philosophiae naturalis principia mathematica*. Tomus Primus, 1726.
- [3] A. Einstein. Die Grundlage der allgemeinen Relativitaetstheorie. *Annalen der Physik*, 49:769–822, 1916.
- [4] E. Schroedinger. Quantisierung als Eigenwertproblem. *Annalen der Physik*, 79:361, 489, 734, 1926.
- [5] K.-H. Lautenschlaeger, W. Schroeter, and A. Wanninger. Taschenbuch der Chemie. Verlag Harri Deutsch, 2005.
- [6] L. D. Landau and E. M. Lifschitz. Lehrbuch der theoretischen Physik: Mechanik. Verlag Harri Deutsch, 14. Auflage 1997.
- [7] F. Klauschen, B. R. Angermann, and M. Meier-Schellersheim. Understanding diseases by mouse click: the promise and potential of computational approaches in systems biology. *Clinical and Experimental Immunology*, 149(3):424–429, 2007.
- [8] M. Meier-Schellersheim, F. Klauschen, and B. Angermann. Computational modeling of signaling networks for eukaryotic chemosensing. *Methods Molecular Biology*, 571:507–526, 2009.
- [9] D.J. Hand. How measurement defines our world. Significance, 2:81–83, 2005.
- [10] B. Orth. Einfuehrung in die Theorie des Messens. Kohlhammer, 1974, S.9.
- [11] R. Feynman. The Character of Physical Law. M.I.T. Press, 1967.
- [12] A.W. Philipps. The relationship between unemployment and the rate of change of money wages in the united kingdom 1861-1957. *Economica*, 25 (100):283–299, 1958.

- [13] J. Rudd and K. Whelan. New tests of the new-keynesian philips curve. Journal of Monetary Economics, 52(6):1167–81, 2005.
- [14] J. Gali, M. Gertler, and J. D. Lopez-Salido. European inflation dynamics. *European Economic Review*, 45(7):1237–70, 2001.
- [15] N. G. Mankiw. The macroeconomist as scientist and engineer. *The Journal of Economic Perspectives*, 20(4):29–46, 2006.
- [16] C.-P. Ortlieb. Methodische Probleme und methodische Fehler der mathematischen Modellierung in der Volkswirtschaftslehre. Hamburger Beitraege zur Modellierung und Simulation, 18:1–24, 2004.
- [17] R. Lucas. Econometric policy evaluation: A critique. Carnegie-Rochester Conference Series on Public Policy, 1:19–46, 1976.
- [18] J. Bell. Predicting disease using genomics. Nature, 429(6990):453–456, 2004.
- [19] E. S. Lander. Initial impact of the sequencing of the human genome. *Nature*, 470(7333):187–197, 2011.
- [20] E. Phizicky, P. Bastiaens, H. Zhu, M. Snyder, and S. Fields. Protein analysis on a proteomic scale. *Nature*, 422(6928):208–215, 2003.
- [21] B.F. Cravatt, G.M. Simon, and J.R. Yates 3rd. The biological impact of mass-spectrometry-based proteomics. *Nature*, 450(7172):991–1000, 2007.
- [22] S. Stoll, J. Delon, T. M. Brotz, and R. N. Germain. Dynamic imaging of t cell-dendritic cell interactions in lymph nodes. *Science*, 296(5574):1873– 1876, 2002.
- [23] R. N. Germain, E. A. Robey, and M. D. Cahalan. A decade of imaging cellular motility and interaction dynamics in the immune system. *Science*, 336(6089):1676–1681, 2012.
- [24] M.S. Boguski and M.W. McIntosh. Biomedical informatics for proteomics. $Nature,\ 422(6928):233-237,\ 2003.$
- [25] L. Chin, W.C. Hahn, F. Getz, and M. Meyerson. Making sense of cancer genomic data. *Genes Dev*, 25(6):534–555, 2011.
- [26] M. Bibbo, P.H. Bartels, H.E. Dytch, and G.L. Wied. Computed cell image information. *Monogr Clin Cytol*, 9:62–100, 1984.
- [27] E. Bengtsson. The measuring of cell features. *Anal Quant Cytol Histol*, 9:212–217, 1987.

- [28] L. Pantanowitz, P. N. Valenstein, A. J. Evans, K. J. Kaplan, J. D. Pfeifer, D. C. Wilbur, L. C. Collins, and T. J. Colgan. Review of the current state of whole slide imaging in pathology. *J Pathol Inform*, 2:36, 2011.
- [29] T. C. Cornish, R. E. Swapp, and K. J. Kaplan. Whole-slide imaging: routine pathologic diagnosis. *Adv Anat Pathol*, 19(3):152–159, May 2012.
- [30] L. D. Stein. An introduction to the informatics of next-generation sequencing. *Curr Protoc Bioinformatics*, Chapter 11:Unit 11.1., 2011.
- [31] W. S. Noble and M. J. MacCoss. Computational and statistical analysis of protein mass spectrometry data. *PLoS Comput Biol*, 8(1):e1002296, 2012.
- [32] A.L. Hodgkin and A.F. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *Journal* of Physiology, 117(4):500–544, 1952.
- [33] S. Ghosh, Y. Matsuoka, Y. Asai, K.-Y. Hsin, and H. Kitano. Software for systems biology: from tools to integrated platforms. *Nature Reviews Genetics*, 12(12):821–832, 2011.
- [34] H. Kitano. Systems biology: a brief overview. Science, 295(5560):1662– 1664, 2002.
- [35] H. Kitano. Computational systems biology. Nature, 420(6912):206–210, 2002.
- [36] D. Faratian, R. G. Clyde, J. W. Crawford, and D. J. Harrison. Systems pathology-taking molecular pathology into a new dimension. *Nat Rev Clin Oncol*, 6(8):455–464, 2009.
- [37] D. Faratian. Systems pathology. Breast Cancer Res, 12 Suppl 4:S4, 2010.
- [38] D. B. Collier, E. Henry, and J. M. Box-Steffensmeier, editors. The Oxford Handbook of Political Methodology. Oxford Handbooks Online, 2009.
- [39] R. Kandpal, B. Saviola, and J. Felton. The era of 'omics unlimited. *Biotechniques*, 46(5):351–2, 354–5, 2009.
- [40] A. Stenzinger, F. Klauschen, D. Wittschieber, W. Weichert, C. Denkert, M. Dietel, and C. Roller. Would virchow be a systems biologist? a discourse on the philosophy of science with implications for pathological research. *Virchows Archive*, 456(6):599–607, 2010.
- [41] G. Damia, M. Broggini, S. Marsoni, S. Venturini, and D. Generali. New omics information for clinical trial utility in the primary setting. J Natl Cancer Inst Monogr, 2011(43):128–133, 2011.

- [42] D. J. Foran and R. A. Berg. A method for quantitative image assessment based on redundant feature measurements and statistical reasoning. *Comput Methods Programs Biomed*, 45(4):291–305, 1994.
- [43] L. Latson, B. Sebek, and K. A. Powell. Automated cell nuclear segmentation in color images of hematoxylin and eosin-stained breast biopsy. *Anal Quant Cytol Histol*, 25(6):321–331, 2003.
- [44] B. Ko, M. Seo, and J.-Y. Nam. Microscopic cell nuclei segmentation based on adaptive attention window. J Digit Imaging, 22(3):259–274, 2009.
- [45] F. Bunyak, A. Hafiane, and K. Palaniappan. Histopathology tissue segmentation by combining fuzzy clustering with multiphase vector level sets. Adv Exp Med Biol. 696:413–424, 2011.
- [46] R. Riber-Hansen, B. Vainer, and T. Steiniche. Digital image analysis: a review of reproducibility, stability and basic requirements for optimal results. APMIS, 120(4):276–289, 2012.
- [47] J. C. Underwood. Nuclear morphology and grading in tumours. *Curr Top Pathol*, 82:1–15, 1990.
- [48] T. Amaral, S.J. McKenna, K. Robertson, and A. Thompson. Classification of breasttissue microarray spots using colour and local invariants. *ISBI*, pages 999–1002, 2008.
- [49] A. Cruz-Roa, J. C. Caicedo, and F. A. Gonzalez. Visual pattern analysis in histopathology images using bag of features. CIARP, pages 521–528, 2009.
- [50] D. Baehrens, T. Schroeter, S. Harmeling, M. Kawanabe, K. Hansen, and Mueller K.-R. How to explain individual classification decisions. *Journal* of Machine Learning Research, 11:1803–1831, 2010.
- [51] A. Bullen, R.S. Friedman, and M.F. Krummel. Two-photon imaging of the immune system: a custom technology platform for high-speed, multicolor tissue imaging of immune responses. *Curr Top Microbiol Immunol*, 334:1– 29, 2009.
- [52] A.M. Parfitt. Bone histomorphometry: standardization of nomenclature, symbols, and units. Report of the ASBMR Histomorphometry Nomenclature Committee. J Bone Miner Res, 2:595–610, 1987.
- [53] M. Mengel, R. von Wasielewski, B. Wiese, T. Ruediger, H.-K. Mueller-Hermelink, and H. Kreipe. Inter-laboratory and inter-observer reproducibility of immunohistochemical assessment of the Ki-67 labelling index in a large multi-centre trial. *J Pathol*, 198(3):292–299, 2002.

- [54] Z. Varga, J. Diebold, C. Dommann-Scherrer, H. Frick, D. Kaup, A. Noske, E. Obermann, C. Ohlschlegel, B. Padberg, C. Rakozy, O. S. Sancho, S. Schobinger-Clement, H. Schreiber-Facklam, G. Singer, C. Tapia, U. Wagner, M. G. Mastropasqua, G. Viale, and H.-A. Lehr. How reliable is Ki-67 immunohistochemistry in grade 2 breast carcinomas? A QA study of the Swiss Working Group of Breast- and Gynecopathologists. *PLoS One*, 7(5):e37379, 2012.
- [55] B. M. Slepchenko and L. M. Loew. Use of virtual cell in studies of cellular dynamics. Int Rev Cell Mol Biol, 283:1–56, 2010.
- [56] A.E. Cowan, I.I. Moraru, J.C. Schaff, B.M. Slepchenko, and L.M. Loew. Spatial modeling of cell signaling networks. *Methods Cell Biol*, 110:195–221, 2012.
- [57] E. J. Deeds, J. Krivine, J. Feret, V. Danos, and W. Fontana. Combinatorial complexity and compositional drift in protein interaction networks. *PLoS One*, 7(3):e32032, 2012.
- [58] N.M. Borisov, N.I. Markevich, J.B. Hoek, and B.N. Kholodenko. Signaling through receptors and scaffolds: independent interactions reduce combinatorial complexity. *Biophys J*, 89(2):951–966, 2005.
- [59] K. W. Kohn, M. I. Aladjem, Kim S., J. N. Weinstein, and Y. Pommier. Depicting combinatorial complexity with the molecular interaction map notation. *Mol Syst Biol*, 2:51, 2006.
- [60] H. Blankenburg, F. Ramirez, J. Bach, and M. Albrecht. Dasmiweb: online integration, analysis and assessment of distributed protein interaction data. *Nucleic Acids Res*, 37(Web Server issue):W122–W128, 2009.
- [61] A. Brueckner, C.L. Polge, C. Polge, N. Lentze, D. Auerbach, and U. Schlattner. Yeast two-hybrid, a powerful tool for systems biology. *Int J Mol Sci*, 10(6):2763–2788, 2009.
- [62] V. Collura and G. Boissy. From protein-protein complexes to interactomics. Subcell Biochem, 43:135–183, 2007.
- [63] N.C. Hubner and M. Mann. Extracting gene function from proteinprotein interactions using quantitative bac interactomics (qubic). *Methods*, 53(4):453–459, 2011.
- [64] B. Blagoev, S.-E. Ong, I. Kratchmarova, and M. Mann. Temporal analysis of phosphotyrosine-dependent signaling networks by quantitative proteomics. *Nat Biotechnol*, 22(9):1139–1145, 2004.

- [65] A. Derouiche, C. Cousin, and I. Mijakovic. Protein phosphorylation from the perspective of systems biology. Curr Opin Biotechnol, 23(4):585–590, 2012.
- [66] C.R. Cho, M. Labow, M. Reinhardt, J. van Oostrum, and M.C. Peitsch. The application of systems biology to drug discovery. Curr Opin Chem Biol, 10(4):294–302, 2006.
- [67] J. D. Feala, J. Cortes, P.M. Duxbury, C. Piermarocchi, A.D. McCulloch, and G. Paternostro. Systems approaches and algorithms for discovery of combinatorial therapies. Wiley Interdiscip Rev Syst Biol Med, 2(2):181– 193, 2010.
- [68] X. W. Zhang, Y. Yap, L. Li, S. Guo, and F. Chen. Mass spectrometry-based Omics technologies in cancer diagnostics. Mass Spectrometry Reviews, 26(3):403–31, 2007.
- [69] A. C. Wolff, M. H. Hammond, J. N. Schwartz, K. L. Hagerty, D. C. Allred, R. J. Cote, M. Dowsett, P. L. Fitzgibbons, W. M. Hanna, A. Langer, L. M. McShane, S. Paik, M. D. Pegram, E. A. Perez, M. F. Press, A. Rhodes, C. Sturgeon, S.E. Taube, R. Tubbs, G. H. Vance, M. van de Vijver, T. M. Wheeler, D. F. Hayes, and American Society of Clinical Oncology & College of American Pathologists. American society of clinical oncology/college of american pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer. J Clin Oncol, 25(1):118–145, 2007.
- [70] E. Burandt and G. Sauter. [HER2 ASCO guidelines. The answer to everything?]. *Pathologe*, 31 Suppl 2:285–291, 2010.

Danksagung

Ich danke Herrn Prof. Dr. Manfred Dietel für die Betreuung und Unterstützung bei meiner Habilitation, Herrn Dr. Ronald Germain und Herrn Dr. Martin Meier—Schellersheim für die wissenschaftliche Ausbildung während meiner Postdoktorandenzeit an den National Institutes of Health (USA) sowie Herrn Prof. Dr. Gerhard Mack vom II. Institut für theoretische Physik der Universität Hamburg für die Förderung und Offenheit für meine interdisziplinäre Ausbildung. Des Weiteren danke ich meiner Familie und insbesondere meiner Frau Dr. Amrei Dilling für die Unterstützung meiner wissenschaftlichen Laufbahn.

ERKLÄRUNG

§ 4 Abs. 3 (k) der HabOMed der Charité

Hiermit erkläre ich, dass

- · weder früher noch gleichzeitig ein Habilitationsverfahren durchgeführt oder angemeldet wurde,
- · die vorgelegte Habilitationsschrift ohne fremde Hilfe verfasst, die beschriebenen Ergebnisse selbst gewonnen sowie die verwendeten Hilfsmittel, die Zusammenarbeit mit anderen Wissenschaftlern/Wissenschaftlerinnen und mit technischen Hilfskräften sowie die verwendete Literatur vollständig in der Habilitationsschrift angegeben wurden,

· mir die geltende Habilitationsordnung bek	annt ist.
Datum	Unterschrift