

Anwendung mathematischer Modelle zur Vorhersage des Therapieverlaufs von CML-Patienten

Dissertationsschrift
zur Erlangung des akademischen Grades
Doctor medicinae (Dr. med.)

vorgelegt

der Medizinischen Fakultät Carl Gustav Carus
der Technischen Universität Dresden

von

Tino Rothe

geboren in Dresden am 21.11.1991

angefertigt am:

Institut für Medizinische Informatik und Biometrie
in Dresden 2016

Betreuer: Prof. Dr. rer. med. Ingo Röder

Mitbetreuung: Dr. rer. med. Ingmar Glauche

1. Gutachter: Prof. Dr. rer. med. Ingo Röder

2. Gutachter: Prof. Dr. med. Thoralf Lange

Tag der mündlichen Prüfung: 05. Dezember 2017

Prof. Dr. rer. nat. habil. Henning Morawietz
Vorsitzender der Promotionskommission

Inhaltsverzeichnis

1. Einleitung	1
2. Medizinischer Hintergrund	3
2.1. Hämatopoese	3
2.2. Chronische myeloische Leukämie	3
2.2.1. Epidemiologie und Ätiologie	5
2.2.2. Pathogenese und Molekulargenetik	5
2.2.3. Verlauf	6
2.2.4. Diagnostik und Untersuchungsverfahren	7
2.3. Therapie der CML	9
2.3.1. Therapieziele	9
2.3.2. Medikamentöse Therapie	11
3. Systembiologischer Hintergrund	14
3.1. Statistische Beschreibung des Therapieverlaufs der CML	14
3.1.1. Messungen des Zeitverlaufes	14
3.1.2. Stückweise lineare Regression	14
3.1.3. Biexponentielle Beschreibung	17
3.2. Mathematische Modelle	18
3.2.1. Populationsbasierte Modelle	18
3.2.2. Einzelzellbasierte Modelle	19
3.3. Patientenindividuelle Anpassungsverfahren	20
4. Fragestellung und Zielsetzung dieser Arbeit	21
5. Material und Methoden	24
5.1. IRIS-Datensatz	24
5.1.1. Vorbereitung der Messwerte	25
5.1.2. Statistische Beschreibung der Patientenkohorte	25
5.1.3. Deskriptive Analyse des „medianen Patienten“	27
5.2. Einzelzellbasiertes Stammzellmodell	30

5.3.	Anpassungsmethode M0: Parameteridentifikation nach Horn et al. (2013) . . .	37
6.	Ergebnisse	39
6.1.	Alternativen in der Anpassung der biexponentiellen Beschreibung	39
6.1.1.	Nutzung des maximum likelihood-Verfahrens	39
6.1.2.	Vergleich der Anpassungsverfahren der biexponentiellen Beschreibung	40
6.2.	Variation der Modellparameter und deren Einfluss auf die statistische Therapiebeschreibung	42
6.3.	Konstruktion eines Simulationspools als Grundlage für weitere Untersuchungen	44
6.3.1.	Eingrenzung des Parameterraums	44
6.3.2.	Simulation des kompletten Modellparameterraumes	46
6.3.3.	Statistische Auswertung der Simulationen	48
6.4.	Entwicklung eines Gütemaßes zum Vergleich unterschiedlicher Therapieverläufe	53
6.5.	Anpassungsmethode M1: Berechnung der Modellparameter mittels „direkter“ Regression	53
6.6.	Anpassungsmethode M2: Parametervorhersage anhand abstandsminimierter Vergleiche von Therapieverläufen (lookup-table)	55
6.7.	Vergleich der verschiedenen patientenindividuellen Anpassungsverfahren . . .	56
6.8.	Analyse der vorhergesagten Verläufe der simulierten Tumorlast im Knochenmark	60
7.	Zusammenfassung und Diskussion	69
8.	Abstract	73
9.	Summary	76
A.	Appendix	78
A.1.	Genutzte Parameterkonfiguration des Stammzellmodells	78
A.2.	Abbildungen des Datenparameterraumes der 270.400 <i>in silico</i> Therapiesimulationen	79
A.3.	Abbildungen des Verfahrensvergleichs für alle Patienten	82

1. Einleitung

Die klinische Therapie von Krebserkrankungen entwickelt sich ständig weiter. Ausgehend von der klassischen Chemotherapie, welche die grundlegende Tumorbiologie nutzt (v. a. schnelle Proliferation), entwickelte man Medikamente, die gezielt ausgewählte Veränderungen der Tumorzellen beeinflussen (sogenannte *targeted therapies*) (Sawyers 2004). Um diese optimal anwenden zu können, müssen die Subgruppen der Tumoren identifiziert werden, die empfänglich für eine bestimmte Therapie sind. Eine weitere Diversifikation tritt durch die Entwicklung von Resistenzen ein, die ein bestimmtes Medikament in seiner Wirkung abschwächen können. Die große Anzahl der Tumorsubgruppen in Verbindung mit verschiedenen Therapiemöglichkeiten kann in präklinischen Studien nicht untersucht werden. An dieser Stelle erwächst das Potential systembiologischer Modellierung, die vielversprechenden Kombinationen vor Studien an Patienten identifizieren kann (Michor und Beal 2015). Zusätzlich kann man versuchen, den Schritt zur individualisierten Medizin zu gehen und über die Modellierung den Verlauf einer Behandlung vorherzusagen.

Ein gut geeignetes Beispiel ist die chronische myeloische Leukämie (CML). Diese ist eine Erkrankung des blutbildenden Systems, welche von Mutationen einzelner hämatopoetischer Stammzellen ausgeht. Sie erhalten dabei die Fähigkeit zu ungehinderter Entwicklung, der letztendlich zur Verdrängung der gesunden Blutbildung führt. Molekulargenetisch sind die Zellen, die dabei unreguliert proliferieren, durch ein pathologisches Fusionsgen gekennzeichnet. Dessen Genprodukt ist eine konstitutiv aktive Tyrosinkinase, die durch Tyrosinkinase-Inhibitoren (TKI) gezielt gehemmt werden kann.

Mit dieser Behandlung hat sich die Prognose für Patienten mit CML deutlich verbessert. Aufgrund der guten Wirksamkeit der verwendeten Medikamente fallen erhebliche Therapiekosten an. Außerdem sind die TKI mit unerwünschten Nebenwirkungen für den Patienten verbunden. Daraus ergibt sich das Bestreben, die Therapie bei beständigem Heilungserfolg abzusetzen. Eine Heilung ist jedoch schwierig zu beurteilen, da die Erkrankung vom leukämischen Stammzellen im Knochenmark ausgeht. Dieses ist im klinischen Alltag nur relativ aufwändig zu untersuchen. So kann bei immer weiter abnehmender Tumorlast das periphere Blut zwar frei von Tumorzellen werden, jedoch treten Rezidive nach Absetzen des TKI auf (Hehlmann und Hochhaus 2010). Es wird vermutet, dass die krankhaft veränderten Stammzellen im Knochenmark in einen ruhenden Zustand übergehen, in welchem TKI nicht

effektiv wirken können. Die leukämischen Stammzellen sind jedoch in der Lage, sich nach Ende der TKI-Gabe erneut zu aktivieren und damit ein Rezidiv der CML zu verursachen (Essers und Trumpp 2010).

Mathematische Vorgehen sind eine wichtige Strategie, um Aussagen über einen möglichen Absetzzeitpunkt zu erhalten, indem sie die Möglichkeit eröffnen, den Verlauf der Therapie auf Stammzellebene vorherzusagen. Die Grundlage dafür bilden Modelle, welche die Blutbildung abbilden und die Simulation der Entstehung bzw. Therapie der CML ermöglichen. Ein solches Modell ist neben anderen Vorgehensweisen (beschrieben im Kapitel 3.2.1) das einzelzellbasierte Stammzellmodell nach Roeder und Loeffler (2002) (detailliert beschrieben im Kapitel 5.2), welches die Grundlage der vorliegenden Arbeit bildet. Mit dieser Simulation ist es möglich die Therapie einer CML unter TKI zu simulieren. Ohne Weiteres können dabei Stufen der Hämatopoese analysiert werden, die klinisch nur schwer zu erreichen sind. Dies ist als Vorteil gegenüber der Beobachtung am realen Patienten zu sehen, bei welchem eine Analyse des Knochenmarks mit einer schmerzhaften Knochenmarkpunktion verbunden ist. Eine Anpassung an individuelle Patienten ist notwendig, um für diese den Abfall der leukämischen Stammzellen unter Therapie vorherzusagen. Damit besteht die Möglichkeit einen Zeitpunkt zu identifizieren, an welchem ein Therapieende höchstwahrscheinlich nicht zu einem Rezidiv führt.

Bereits von Horn et al. (2013) wurden Methoden zu Anpassung und Vorhersage entwickelt. Es zeigten sich aber noch Abweichungen bei der optimalen Beschreibung einzelner Patienten. Aus diesem Grund sollen in der vorliegenden Arbeit die Probleme der vorbeschriebenen Methoden identifiziert werden. Auf dieser Grundlage werden neue Herangehensweisen entwickelt und angewendet, welche genauere Vorhersagen ermöglichen sollen. Im letzten Teil erfolgt anschließend eine kritische Prüfung der neuen Methoden, die zeigen soll, ob die getroffenen Vorhersagen plausibel sind.

2. Medizinischer Hintergrund

2.1. Hämatopoese

Das blutbildende System im menschlichen Körper hat die Aufgabe, sämtliche Zellen des Blutes in adäquater Menge zu produzieren, bei Verlust zu erneuern und das Blut an sich ändernde Anforderungen (z. B. Infektionen, Aufenthalt in großen Höhen) anzupassen. Es wird aus hämatopoetischen Stammzellen (HSC) gebildet, deren wichtigste Eigenschaften nach Loeffler und Roeder (2002) die Folgenden sind: Sie sind undifferenzierte Zellen, die proliferieren, dabei ihre eigene Population aufrecht erhalten, aber auch eine große Menge an differenzierten Zellen produzieren. Zwischen all diesen Eigenschaften kann flexibel abgewogen werden, damit das System sich an die Anforderungen anpassen kann. Außerdem ist die Hämatopoese an geeignete Mikroumgebungen (auch Stammzellnischen genannt) gebunden. Die Blutbildung findet ausgehend von den HSC als irreversibel absteigende Hierarchie vor allem im Knochenmark, aber auch in Organen des lymphatischen Systems und der Milz statt. Man unterscheidet dabei zwei Entwicklungsreihen: die lymphatische Reihe, aus der die B- und T-Lymphozyten sowie die plasmazytären dendritischen Zellen hervorgehen und die myeloische Reihe, die für die Produktion aller anderen Blutzellen (Erythrozyten, Granulozyten, Thrombozyten und Makrophagen) verantwortlich ist. Die Entwicklung verläuft jeweils von den undifferenzierten Stammzellen über Progenitorzellen und Blasten zu den endgültig ausdifferenzierten Blutzellen (Wilson et al. 2008). Abbildung 2.1 gibt dazu einen schematischen Überblick. Sollte ein Mangel an Blutzellen vorhanden sein, so kann dieser durch gesteigerte Proliferation der jeweiligen Vorläuferzellen ausgeglichen werden (Uchida et al. 1993).

2.2. Chronische myeloische Leukämie

Die chronisch myeloische Leukämie (CML) ist eine klonale myeloproliferative Erkrankung. Sie geht von pluripotenten Stammzellen des Knochenmarks aus und ist durch das Fusionsgen BCR-ABL1 gekennzeichnet. Diese pathognomonische Gensequenz kann auch zur Unterscheidung von anderen myeloproliferativen und sonstigen Neoplasien genutzt werden (nach Hehlmann und Hochhaus (2010) sowie Leitner und Hehlmann (2011)).

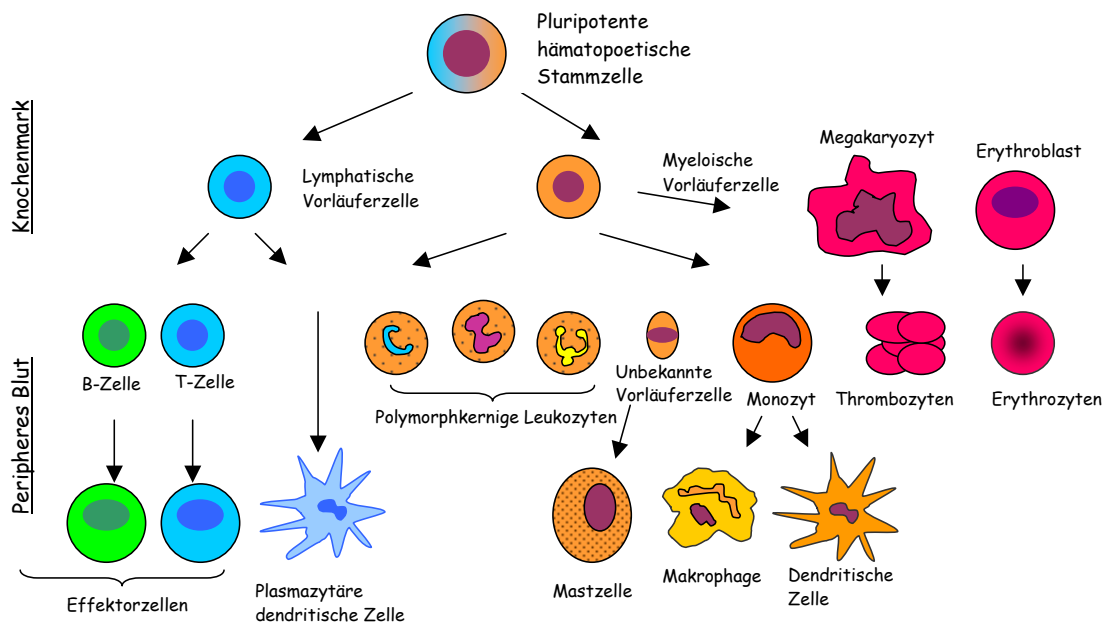


Abbildung 2.1.: Schematische Darstellung der Blutbildung. Ausgehend von hämatopoetischen Stammzellen, die auch ruhend ohne Teilungsaktivität vorkommen können, entwickeln sich über Vorläuferzellen (Progenitorzellen) und Blasten die unterschiedlichen differenzierten Blutzellen. Dabei gibt es zwei unterschiedliche Entwicklungsreihen: die lymphatische und die myeloische Reihe. Die lymphatische Reihe produziert B- und T-Lymphozyten, während aus der myeloischen Reihe die anderen Blutzellen hervorgehen. (Abbildung und Beschreibung nach Regn (2002)).

2.2.1. Epidemiologie und Ätiologie

Die CML hatte in Deutschland laut Robert-Koch-Institut (RKI) im Jahr 2011 eine Inzidenz von ca. 1-2/100.000, die mit dem Alter ansteigt. Die CML wird vom RKI in die Gruppe der Leukämien eingeordnet, die neben der CML noch von den akuten und chronischen lymphatischen Leukämien (ALL und CLL) sowie der akuten myeloischen Leukämie (AML) und sonstigen (unspezifischen oder nicht eindeutig klassifizierbaren) Formen gebildet wird. Mit einem Anteil von 10% an den genannten Leukämien ist die CML in Deutschland eine eher seltene Erkrankung. Lediglich der Anteil der ALL (akute lymphatische Leukämie) ist unter den Leukämien noch kleiner. Das mittlere Erkrankungsalter liegt bei ca. 65 Jahren. Eine familiäre Häufung oder geografische Variation ist nicht bekannt (Hehlmann und Hochhaus 2010; Kaatsch et al. 2013).

Als Risikofaktor für die Entwicklung einer CML ist lediglich ionisierende Strahlung bekannt, da ein Anstieg der Erkrankungshäufigkeit bei strahlentherapeutisch behandelten Patienten mit Morbus Bechterew und nach den Atombombenabwürfen in Hiroshima und Nagasaki beobachtet wurde. Beim überwiegenden Teil der CML-Erkrankten konnte jedoch keine erhöhte Exposition zu Strahlung identifiziert werden. Andere Gründe können deshalb nicht ausgeschlossen werden (Hehlmann und Hochhaus 2010).

2.2.2. Pathogenese und Molekulargenetik

Bei über 90% der Patienten, die unter einer CML leiden, finden sich komplexe Chromosomentranslokationen, die jedoch alle die Chromosomen 9 und 22 einbeziehen. Dabei findet eine reziproke Translokation der beiden langen Arme dieser Chromosomen statt und es entsteht das pathogenetisch relevante BCR-ABL1-Fusionsgen. Dieses findet sich auf dem verkürzten Chromosom 22, welches man auch als Philadelphia-(Ph)-Chromosom bezeichnet (siehe Abbildung 2.2). Diese Philadelphia-positiven (Ph^+) Zellen kommen bei erkrankten Patienten nicht nur im hämatopoetischen System sondern auch in Gefäßendothelien vor (Gunsilius et al. 2000; Hehlmann und Hochhaus 2010).

Das BCR-ABL1-Gen kodiert für eine Tyrosinkinase, die sich durch eine erhöhte, unregelte Aktivität auszeichnet. Dieses Protein war das Erste, für das die genaue onkogene Mutation beschrieben wurde (Shtivelman et al. 1985). Es wird als $\text{P210}^{\text{BCR-ABL1}}$ bezeichnet und hat dementsprechend ein Molekulargewicht von 210kD. Anders als das ABL1-Transkriptionsprodukt ($\text{P145}^{\text{ABL1}}$), welches überwiegend nukleär lokalisiert ist, befindet sich das $\text{P210}^{\text{BCR-ABL1}}$ -Protein zytoplasmatisch. Die konstitutiv aktivierte Tyrosinkinase erlaubt in Kombination mit der Struktur des BCR-ABL1-Proteins multiple Interaktionen mit der Signaltransduktion der Zelle. Dabei überträgt das Protein Phosphatgruppen von ATP auf Tyrosinreste der unterschiedlichsten Substrate. Über verschiedene Adapterproteine besteht

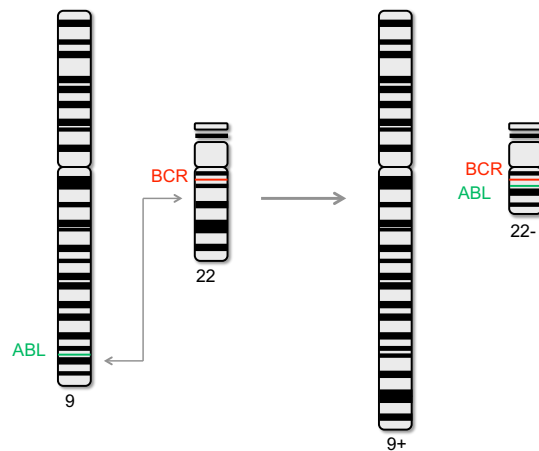


Abbildung 2.2.: Reziproke Translokation zwischen den langen Armen der Chromosomen 9 und 22. Dabei entsteht neben dem verlängerten Chromosom 9 (9q+) ein verkleinertes Chromosom 22 (22q-, Ph), welches das pathognomonische Fusionsgen BCR-ABL1 beinhaltet (Abbildung von Rick Proschmann zur Verfügung gestellt).

dabei z. B. ein Einfluss auf den RAS- sowie den JAK-STAT-Signalweg. Damit wird die Apoptose gehemmt und die Proliferation gesteigert. Außerdem besteht ein Einfluss auf das Zytoskelett, welcher zu verringerter Adhäsion der transformierten Zellen im Knochenmark führt. Damit entwickelt sich durch die Ausschwemmung der Tumorzellen das klinische Bild des „weißen Blutes“ (Leukämie) (Hehlmann und Hochhaus 2010).

2.2.3. Verlauf

Der Verlauf der CML lässt sich nach Hehlmann und Hochhaus (2010) in drei Stadien unterteilen:

- **Präklinisches Stadium**, welches meist nur zufällig durch Nachweis der Mutation (Philadelphia-Chromosom) erkennbar wird.
- **Chronische Phase**, in der die Myelopoese expandiert, aber eine gute therapeutische Einflussnahme möglich ist.
- **Terminale Phase**, die als Blastenkrise oder Akzeleration manifest wird. In dieser Phase ist die CML schwer zu kontrollieren und führt letztendlich zum Tod des Patienten.

Die Diagnose einer CML erfolgt meist im chronischen Stadium. Diese relativ stabile und symptomarme Phase kann 4 bis 10 Jahre dauern. Es treten überwiegend gut differenzierte

Vorläuferzellen der Granulopoese auf. Ist der Übergang in das terminale Stadium langsam progredient, so spricht man von einer Akzeleration. Diese steht im Gegensatz zur Blastenkrise, bei der dieser Übergang relativ schnell und dramatisch verläuft. Aufgrund des langsamen Übergangs kann eine Akzeleration oft erst im Verlauf diagnostiziert werden. Grundsätzlich können die Symptome, die dabei auffallen, vielfältig sein. Hinweisend sind jedoch meist zunehmender Medikationsbedarf bzw. diskrete Blutbildveränderungen wie steigende Leukozytenzahlen, zunehmende Anämie, Thrombopenie oder -zytose. Weiterhin zeigen sich neben dem Ph-Chromosom neue Chromosomenaberrationen, zunehmende Splenomegalie oder leukämische Infiltrate außerhalb des Knochenmarks. Auch Allgemeinsymptome können sich verschlechtern, Knochenschmerzen auftreten, unerklärliches Fieber und Gewichtsverlust einsetzen.

Eine Blastenkrise ist durch die Deutsche CML-Studiengruppe eindeutig zytologisch definiert und stellt das sich krisenhaft entwickelnde Analog zur Akzeleration dar. Folgende Punkte sind ausschlaggebend:

- Anteil der Blasten größer als 30 % der Leukozyten im peripheren Blut
- Anteil der Blasten größer als 50 % der kernhaltigen Zellen im Knochenmark oder
- extramedulläre blastäre Infiltrate, welche histologisch oder zytologisch gesichert werden müssen

Dabei ist zu beachten, dass nach der neuen WHO-Definition ein Grenzwert von 20 % Blasten im peripheren Blut festgelegt wurde. Für die aktuellen deutschen Leitlinien wird noch der alte Grenzwert verwendet. Begründet wird dies mit der signifikant besseren Prognose von Patienten, die einen Blastenanteil zwischen 20 % und 30 % haben, gegenüber Patienten mit einer Belastung von mehr als 30 % Blasten (Hochhaus et al. 2013).

2.2.4. Diagnostik und Untersuchungsverfahren

Zur Diagnosestellung und Verlaufsbeurteilung der CML werden hämatologische bzw. klinische, zyto- und molekulargenetische Kriterien einbezogen. Von klinischer Seite beurteilt man das Differentialblutbild, den histologischen Befund des Knochenmarks und die Milzgröße. Damit lässt sich auch der EUTOS-Prognopsescore berechnen, der aus dem Blastenanteil im peripheren Blut und der Ausdehnung der Milz unter dem Rippenbogen die Wahrscheinlichkeit einer kompletten zytogenetischen Remission vorhersagt (Hasford et al. 2011).

Zytogenetisch ist das Auftreten des Philadelphia-Chromosoms relevant. Da eine zytogenetische Remission (Verschwinden des Ph-Chromosoms im peripheren Blut) mit der Prognose korreliert, liefern wiederholte genetische Untersuchungen der Zellen des peripheren Blutes

wichtige Informationen über den Therapieverlauf. Außerdem können chromosomale Aberrationen aufgedeckt werden, die man als klonale Evolution bezeichnet und die auf ein Therapieversagen hinweisen können (Hasford et al. 1998; Leitner und Hehlmann 2011). Zur Verlaufsbeobachtung benötigt man eine Quantifizierung der Tumorlast. Diese kann über den Anteil der Ph^+ -Zellen an allen kernhaltigen Blutzellen ausgedrückt werden.

Molekulargenetisch bestimmt man mittels qPCR (quantitative Polymerasekettenreaktion, auch real-time PCR (RT-PCR) genannt) das $\frac{BCR-ABL1}{ABL1}$ -Transkriptverhältnis, welche eine Abschätzung des Verhältnisses $\frac{Ph^+-Zellen}{Ph^+-Zellen + Ph^- - Zellen}$ liefern soll. Die maximale Auflösung liegt im Bereich von 0,01%, das heißt, dass unter 10^5 Transkripten ein Transkript einer malignen Zelle identifiziert werden kann. Man verwendet Reaktanden, welche die Menge der DNA durch Fluoreszenz anzeigen. Diese Fluoreszenz wird parallel zur PCR-Reaktion überwacht und gemessen. Die kontinuierliche Messung führt zur Bezeichnung „real-time“. Ist die entsprechende Menge normaler Transkripte in der Probe kleiner, verändert sich die Schwelle der Detektierbarkeit entsprechend (Mensink et al. 1998; Gabert et al. 2003; Velden et al. 2003). Dazu sei folgendes Beispiel konstruiert: enthält eine Probe insgesamt 100 Transkripte, so ist der kleinste messbare Wert 1% (eines aus 100 Transkripten ist pathologisch). Besteht eine Probe aus den 10000 Transkripten, so kann minimal ein Wert von 0,01% gemessen werden (eines aus 10000 Transkripten wird als krank identifiziert). Weiterhin sei bereits an dieser Stelle erwähnt, dass große Tumorlasten mit dem $\frac{BCR-ABL1}{ABL1}$ -Verhältnis eher unterschätzt werden (Hanfstein et al. 2014). In der vorliegenden Arbeit ist jeweils das $\frac{BCR-ABL1}{ABL1}$ -Verhältnis als Tumorlast angegeben, welches durch einen zwischen den unterschiedlichen an Messungen beteiligten Laboren anhand eines Korrekturfaktors normiert wurde. Dadurch soll eine Vergleichbarkeit zwischen den unterschiedlichen Messungen ermöglicht werden. Durch diesen Korrekturfaktor kann es auch zu Messwerten oberhalb von einem $\frac{BCR-ABL1}{ABL1}$ -Verhältnis von 100% kommen. Weiterhin sei an dieser Stelle angemerkt, dass mit dem $\frac{BCR-ABL1}{ABL1}$ -Verhältnis die Menge von DNA-Transkripten und keine exakten Zellzahlen bestimmt werden. Weitere mögliche Messfehler kommen bei der Messung von Tumorlasten oberhalb von 10% vor, wenn das $\frac{BCR-ABL1}{ABL1}$ -Verhältnis bestimmt wird, da das ABL1-Gen sowohl im pathologischen Fusionsgen vorkommt und gleichzeitig als Maß für die gesunden Zellen dient. Nutzt man statt ABL1 ein anderes Referenzgen (z. B. GUS) kann die Messgenauigkeit verbessert werden (Hanfstein et al. 2014). Da mit der RT-PCR auch eine geringe Menge leukämischer DNA nachgewiesen werden kann, ist durch die Überwachung von Patienten in Remission frühzeitig eine Aussage über ein eventuelles Therapieversagen möglich. Hinweis darauf wäre eine wieder ansteigende Tumorlast (Hughes et al. 2006; Leitner und Hehlmann 2011).

Eine weitere Steigerung der Spezifität, um auch besonders kleine Tumorlasten identifizieren zu können, ist über eine sogenannte *nested PCR* möglich. Dabei wird mit dem Produkt des ersten kompletten PCR-Durchlaufs eine weitere Reaktion durch Zugabe von anderen

Reaktanden gestartet. Dadurch kann sehr spezifisch nur eine Sequenz vervielfältigt werden, da im zweiten Ansatz die Länge der vorliegenden DNA deutlich reduziert wurde. Somit können die Primer des zweiten Ansatzes mit deutlich höherer Sicherheit an die richtige Sequenz der Matrize binden und falsche Bindungen werden nahezu ausgeschlossen (Busch et al. 2010).

Bei Diagnosestellung im chronischen Stadium (siehe Kapitel 2.2.3) imponieren klinisch neben einem allgemeinen Krankheitsgefühl zumeist eine Spleno-, seltener auch eine Hepatomegalie. Das Differentialblutbild zeigt eine Leukozytose mit teilweise extremer Linksverschiebung. Nach bereits oben angeführter Definition muss in der chronischen Phase dabei der Anteil der Blasten an der im peripheren Blut gemessenen Leukozytenzahl unter 20 % sein. Im Knochenmark zeigt sich das Bild einer expandierten Granulo- und meist auch Megakaryopoese, wobei das Fettmark fast vollständig durch blutbildendes Knochenmark ersetzt wird. Der Anteil an retikulären Fasern, die bei einigen Patienten im Mark zu finden sind, steigt im Verlauf meist an (Hehlmann und Hochhaus 2010).

Differentialdiagnostisch muss man andere chronisch myeloproliferative, wie eine beginnende Osteomyelofibrose (OMF), und myelodysplastische Erkrankungen, vor allem in Form einer chronischen myelomonozytären Leukämie (CMML), von der CML abgrenzen. Dabei hilft im überwiegenden Teil der Patienten der molekulargenetische Nachweis des Philadelphia-Chromosoms (Hehlmann und Hochhaus 2010).

2.3. Therapie der CML

Die Therapie der CML hat im vergangenen Jahrhundert eine bedeutende Entwicklung genommen. Konnten frühere Therapiemöglichkeiten lediglich die Symptome mindern, stehen heute Therapien auf molekularer Ebene zur Verfügung, die eine deutliche Verbesserung der Prognose bewirken. Außerdem etablierte sich die Therapie mit Imatinib zum oft angeführten Modell der zielgerichteten onkologischen Behandlung (Leitner und Hehlmann 2011).

Problematisch bei allen bisher entwickelten Therapieverfahren ist das Auftreten von Rezidiven, die wie die ursprüngliche Erkrankung von den hämatopoetischen Stammzellen ausgehen. Damit verbietet sich eine Therapiebeendigung nach aktuellen Kenntnissen, da eine Messung der Tumorlast im Stammzellkompartiment bisher routinemäßig nicht möglich ist (Hehlmann und Hochhaus 2010).

2.3.1. Therapieziele

Die Therapie hat vorrangig das Ziel eine hämatologische Remission zu erreichen. Diese ist gekennzeichnet durch die Normalisierung des Blutbildes und damit einhergehend der Besserung aller Krankheitssymptome. Die hämatologische Remission bildet dann die Grundlage für die zytogenetische Remission, welche durch die Abnahme des Ph-Chromosoms im

peripheren Blut gekennzeichnet ist. Eine echte Heilung der CML ist nach aktueller Lehrmeinung lediglich mit einer allogenen Stammzelltransplantation möglich. Jedoch fanden klinische Studien statt, in denen Patienten auch nach Absetzen der Therapie in Remission blieben. Man spricht dabei aber noch nicht von Heilung, da ein erneutes Anwachsen der Tumorlast nicht ausgeschlossen werden kann. In die Festlegung der Therapie sollten neben den Empfehlungen des European LeukaemiaNet auch die Präferenzen und Wünsche des Patienten einfließen, da gewisse Therapieoptionen (z. B. Alkylantien oder die Therapie in Vorbereitung einer Stammzelltransplantation) mit Einschränkungen der Fertilität oder auch Teratogenität einhergehen (Baccarani et al. 2006; Hehlmann und Hochhaus 2010; Mahon 2015). Eine schnelle Zytoreduktion bei Hyperleukozytose und Leukostasesyndrom kann mittels Apherese erreicht werden, die sich auch zur Behandlung von Schwangeren eignet (Leitner und Hehlmann 2011).

Wichtige Definitionen des Therapieansprechens sind in Tabelle 2.1 zusammengefasst (aus Hochhaus et al. (2013) mit Material aus Baccarani et al. (2009) und Cross et al. (2012))

Tabelle 2.1.: Definitionen des Ansprechens auf eine Therapie mit TKI anhand hämatologischer, zytogenetischer und molekulargenetischer Kriterien (Hochhaus et al. 2013).

Method	Remission	Abkürzung	Parameter
Hämatologisch	komplett	CHR	Leukozyten $< 10 \cdot 10^9/l$
			Basophile $< 5\%$
			keine Myelozyten, Promyelozyten oder Myeloblasten im Differentialblutbild
			Thrombozyten $< 450 \cdot 10^9/l$
			Milz nicht tastbar $< 5\%$
zytogenetisch	komplett	CCyR	keine Ph ⁺ -Metaphasen
	partiell	PCyR	1 – 35 % Ph ⁺ -Metaphasen
	minor	mCyR	36 – 65 % Ph ⁺ -Metaphasen
	minimal	minCyR	66 – 95 % Ph ⁺ -Metaphasen
	keine	keine CyR	$> 95\%$ Ph ⁺ -Metaphasen
molekulargenetisch	tief	MR ^{4,5}	$\frac{BCR-ABL1}{ABL1}$ -Verhältnis $< 0,0032\%$
	tief	MR ⁴	$\frac{BCR-ABL1}{ABL1}$ -Verhältnis $< 0,01\%$
	major	MMR	$\frac{BCR-ABL1}{ABL1}$ -Verhältnis $< 0,1\%$

2.3.2. Medikamentöse Therapie

Im 19. Jahrhundert hatte man lediglich Arsen als rudimentäre Therapieoption. Man fand Dosierungen um Anämie, Leukozytose, Fieber und Milzgröße zu bessern. Später kombinierte man diese Therapie intermittierend mit externer Strahlentherapie. Auch dies führte zu keiner Verlängerung der Überlebenszeit der Patienten, konnte jedoch die Symptome lindern und damit die Lebensqualität verbessern. Forschung für die chemische Kriegsführung im Verlauf des 2. Weltkrieges bildete die Grundlage für viele im weiteren Verlauf genutzte Chemotherapeutika. Auf Basis von Senfgas (auch Stickstoff-Lost genannt) entwickelte man alkylierende Medikamente, die auch bei strahlentherapeutisch nicht mehr beherrschbarer Erkrankung die Symptome lindern konnten (Goldman und Daley 2007). Ein Beispiel dafür ist Busulfan, welches aber heute in der Primärtherapie aufgrund schwerer Nebenwirkungen wie Panzytopenie oder irreversibler Lungenfibrose nicht mehr indiziert ist (Hehlmann und Hochhaus 2010).

Hydroxyharnstoff

Mit Hydroxyharnstoff (Hydroxyurea) fand in den 70er Jahren des vergangenen Jahrhunderts erstmals ein Medikament Eingang in die Therapie, welches das Überleben der Patienten geringfügig verlängern konnte. Es handelt sich um einen Hemmstoff der Ribonukleotidreduktase, der auch noch heute zur schnellen initialen Zytoreduktion und in palliativen Situationen eingesetzt wird. Eine zytogenetische Remission ist aber auch mit diesem Medikament nicht zu erreichen bzw. aufrechtzuerhalten (Leitner und Hehlmann 2011).

Interferon- α

Interferon- α (IFN- α) hatte sich vor Entwicklung der Tyrosin-Kinase-Inhibitoren (TKI), namentlich Imatinib, als Standardtherapie etabliert. Es konnte stabile hämatologische und teilweise auch zytogenetische Remissionen induzieren, indem es eine myelosuppressive Wirkung entfaltet. Damit erlangten die behandelten Patienten einen Überlebensvorteil, der besonders ausgeprägt war, wenn innerhalb von drei Monaten eine komplette hämatologische Remission erreicht werden konnte. Trotzdem ist mittels RT-PCR eine Belastung mit BCR-ABL1-Transkripten, auch nach zytogenetischer Remission, nachweisbar (Hochhaus et al. 2000; Hehlmann und Hochhaus 2010).

Tyrosin-Kinase-Inhibitoren (TKI)

Da man das CML-auslösende Gen und sein Produkt identifizieren konnte (siehe Kapitel 2.2.2), unternahm man Versuche das deregulierte BCR-ABL1-Fusionsprotein gezielt zu

hemmen. Über das systematische Screening von verschiedenen Substanzen, welche die ATP-Bindungsstelle blockieren können, fand man CGP 57148. Es hemmt neben dem BCR-ABL1-Protein lediglich die Tyrosin-Kinase des Rezeptors des *Platelet-derived growth factor (PDGF)*. Die leukämischen Zellen werden damit nicht nur in ihrer unkontrollierten Proliferation gehemmt, sondern es kann über den programmierten Zelltod (Apoptose) ein Verschwinden der transformierten Zellen erreicht werden. Dieser Effekt konnte erstmals 1996 sowohl *in vitro* als auch *in vivo* an Mäusen gezeigt werden (Druker et al. 1996). Im Jahre 1998 begannen die klinischen Studien und es zeigte sich eine deutlich Überlegenheit der Substanz, die man später als Imatinib bezeichnete und die Novartis unter dem Handelsnamen Glivec vertreibt. Es konnte sowohl in IFN-refraktären als auch in allen fortgeschrittenen Stadien der CML ein überragender Wirksamkeitsnachweis in Form hämatologischer und zytogenetischer Remissionen erbracht werden. Auch eine Monotherapie mit Imatinib zeigte in einer Phase-III-Studie (IRIS, siehe Kapitel 5.1) diese Überlegenheit (Hehlmann und Hochhaus 2010).

Nach aktuellem Kenntnisstand muss Imatinib lebenslang verabreicht werden, da selbst bei molekular nicht mehr nachweisbarer Resterkrankung Rezidive nach Absetzen auftreten können. Außerdem zeigt die CML in fortgeschrittenen Stadien weitere chromosomale Aberrationen (siehe Kapitel 2.2.2), die eine Imatinibresistenz bewirken können. Dazu werden alternative Tyrosin-Kinase-Inhibitoren, wie Dasatinib und Nilotinib eingesetzt, die eine höhere Wirksamkeit als Imatinib besitzen und bis auf eine bestimmte Mutation (T315I) alle weiteren Chromosomenabberationen erkennen können. Neben der erhöhten Wirksamkeit zeigt Dasatinib in der DASISION-Studie auch neue unerwünschte Wirkungen wie Pleuraergüsse und eine erhöhte Rate an Tod durch Infektion, die unter Imatinib so nicht auftraten. Sonstige Nebenwirkungen waren vergleichbar mit denen von Imatinib (Kantarjian et al. 2012; Jabbour et al. 2014). Für Nilotinib hat sich das Nebenwirkungsspektrum von eher gastrointestinalen Problemen unter Imatinib zu vorwiegend dermatologischen Auffälligkeiten, wie Ausschlägen oder Juckreiz, und Kopfschmerzen verschoben. Auch die Leberfunktion ist unter Nilotinib häufiger beeinflusst. Dies zeigt sich an erhöhten Leberenzymen und gesteigerten Bilirubinwerten (Saglio et al. 2010).

Nach Dasatinib und Nilotinib als TKI der zweiten Generation wurden 2013 die Substanzen Bosutinib und Ponatinib als Drittgenerations-TKI in der EU zugelassen. Bosutinib konnte in der Standardtherapie keine Überlegenheit gegenüber Imatinib zeigen, hat aber ein anderes Nebenwirkungsprofil. Damit könnte es bei Imatinibunverträglichkeit eingesetzt werden (Cortes et al. 2012b). Ponatinib ist insofern interessant, als dass es auch gegen die Imatinib-resistente Mutation T315I wirksam ist (Cortes et al. 2012a). Trotzdem muss sein Einsatz kritisch abgewogen werden, da in Studien unter Ponatinib schwere Nebenwirkungen auftraten. Beispielhaft seien arterielle Thrombosen, die sich als Herzinfarkte oder sich verschlechternde koronare Herzkrankheit (KHK) äußerten, bzw. Leberversagen genannt (Price et al. 2012). Neben den

TKI der neueren Generationen werden Kombinationstherapien erprobt, um zum Beispiel die T-Zell-induzierende Wirkung von IFN- α mit der von Imatinib zu kombinieren und so die Therapie zu optimieren. Diese und weitere Kombinationen testet die deutsche CML-Studie IV (CML-IV; Hehlmann und Hochhaus (2010)).

Ein weiterer wichtiger Punkt, der gesundheitspolitische Relevanz hat, sind die extremen Therapiekosten einer Therapie mit TKI. Die Prinzipien der Marktwirtschaft von Angebot und Nachfrage haben den Preis für Imatinib nicht fallen, sondern steigen lassen, nachdem andere TKI auf den Markt kamen. Man geht für Imatinib von Jahrestherapiekosten von 92.000\$ (ca. 86.000€, Stand 2012) aus. Im Jahr der Zulassung (2001) lagen die Kosten bei 30.000\$ (ca. 28.000€). Die TKI der dritten Generation verursachen Jahrestherapiekosten im unteren sechsstelligen Bereich (Experts in Chronic Myeloid Leukemia 2013). Es bleibt abzuwarten, wie sich der Preis entwickelt, nachdem der Patentschutz für die ersten TKI abläuft. Auch wenn der Patentschutz für Imatinib in den USA bereits Mitte 2015 ausgelaufen ist, wurde durch spezielle Absprachen zwischen *Novartis* (Anbieter des originalen Imatinib-Präparates *Glivec*) und *Sun Pharmaceuticals Ltd.* der Beginn einer generischen Vermarktung bis Mitte 2016 verschoben. In Europa wird der Patentschutz für Imatinib im Laufe des Jahres 2016 auslaufen (Conti et al. 2015).

3. Systembiologischer Hintergrund

Wie bereits im Kapitel 2.2.4 beschrieben, kann die Tumorlast der CML sehr effizient und genau mittels PCR-Verfahren ermittelt werden. Damit ist die Erkrankung als prototypisches Beispiel für einen systembiologischen Zugang zur Tumorbilogie geeignet. Im Folgenden wird zunächst erläutert, wie der Therapieverlauf der CML unter Behandlung mit TKI durch statistische Parameter beschrieben werden kann. Nachfolgend erfolgt die Beschreibung der verschiedenen Ansätze der Übertragung der CML-Therapie in Modelle.

3.1. Statistische Beschreibung des Therapieverlaufs der CML

3.1.1. Messungen des Zeitverlaufes

Der Therapieverlauf der CML wird beschrieben durch eine zeitliche Abfolge von $\frac{BCR-ABL1}{ABL1}$ -Messwerten. In Abbildung 3.1a ist ein beispielhafter Therapieverlauf aus der IRIS-Studie (siehe Kapitel 5.1) dargestellt.

Es ist deutlich zu erkennen, dass die Messwerte nach relativ kurzer Behandlungsdauer in einen Bereich unterhalb von 1 % $\frac{BCR-ABL1}{ABL1}$ -Last abfallen. Damit findet der für die Aussage über Remission oder Langzeitverhalten interessante Therapieverlauf in einem visuell schwer erfassbaren Bereich des Diagramms statt. Um dieses Problem zu umgehen, werden die $\frac{BCR-ABL1}{ABL1}$ -Messwerte typischerweise auf einer log-Skala abgetragen (siehe Abbildung 3.1b). Daran lässt sich außerdem das Level der molekularen Remission ablesen. Beispielsweise wird eine Reduktion um vier log-Stufen als MR⁴ bezeichnet und bedeutet eine Reduktion um vier log-Stufen bezogen auf 100 % $\frac{BCR-ABL1}{ABL1}$ -Verhältnis. Diese MR-Level beziehen sich dabei immer auf einen standardisierten Ausgangswert, der zwischen den beteiligten Laboren abgestimmt wurde (Hughes et al. 2003).

In Abbildung 3.1b ist der Therapieverlauf aus 3.1a in einem Diagramm mit der eben beschriebenen log-transformierten y-Achse dargestellt.

3.1.2. Stückweise lineare Regression

In dieser Darstellung mit einer logarithmisch eingeteilten y-Achse, lässt sich die zeitliche Abfolge von $\frac{BCR-ABL1}{ABL1}$ -Messwerten durch zwei Geraden approximieren (stückweise linear).

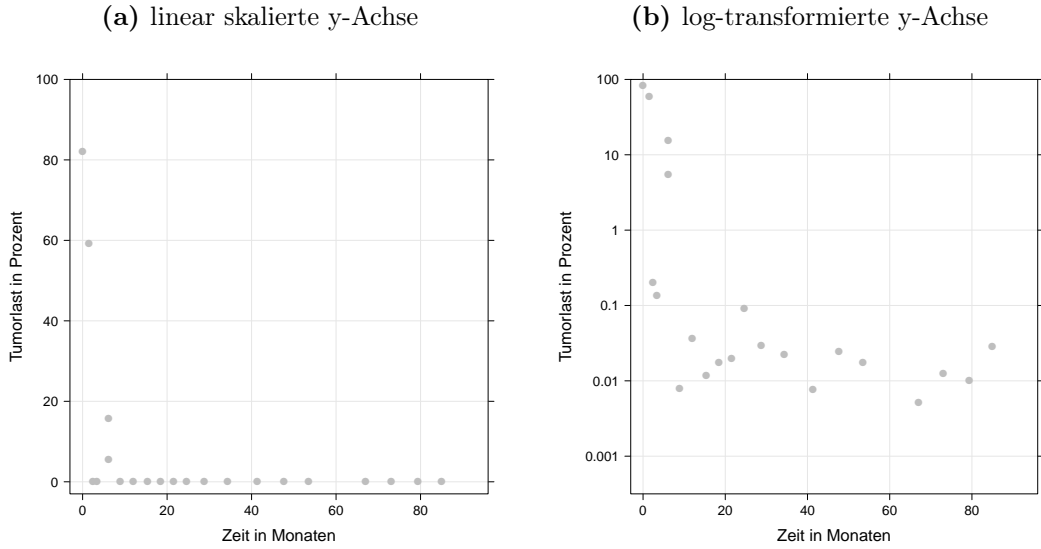


Abbildung 3.1.: Beispielhafter Therapieverlauf eines Patienten mit CML. In Abbildung 3.1a auf linear skalierten Achsen abgetragen. Zur besseren Auflösung wird üblicherweise eine log-transformierte y-Achse genutzt (Abbildung 3.1b)

Diese Herangehensweise wurde erstmals von Michor et al. (2005) beschrieben. Dabei gibt es eine steilere erste Gerade und einen flacheren zweiten Abfall. Der Schnittpunkt der beiden Geraden wird als Bruchpunkt bezeichnet. Die erste Gerade vom Zeitpunkt $t = 0$ bis zum Bruchpunkt folgt der Funktion

$$\log_{10} \left(\frac{BCR - ABL1}{ABL1} \right) (t) = A + \alpha \cdot t \quad (3.1)$$

Der Anstieg der ersten Gerade wird durch α charakterisiert. Die Variable A gibt den dekadischen Logarithmus der $\frac{BCR-ABL1}{ABL1}$ -Konzentration zum Zeitpunkt $t = 0$ an. Für $t = 0$ gilt demnach

$$A = \log_{10} \left(\frac{BCR - ABL1}{ABL1} \right) (t = 0) \quad (3.2)$$

Analog zu dieser mathematischen Beschreibung für den ersten Bereich, gibt es eine Funktion für den zweiten Abfall, der ab dem Bruchpunkt bis zum Ende der Messserie betrachtet wird:

$$\log_{10} \left(\frac{BCR - ABL1}{ABL1} \right) (t) = B + \beta \cdot t \quad (3.3)$$

Dabei gibt B den dekadischen Logarithmus der $\frac{BCR-ABL1}{ABL1}$ -Konzentration am Schnittpunkt der verlängerten Funktion mit der y-Achse an. Analog zu α beschreibt β den Anstieg der

Funktion im zweiten Teil. Für Patienten mit einem biphasischen Verlauf gilt jeweils für die Beträge der Anstiege $\alpha > \beta$. Zur Veranschaulichung dient hier Abbildung 3.2a.

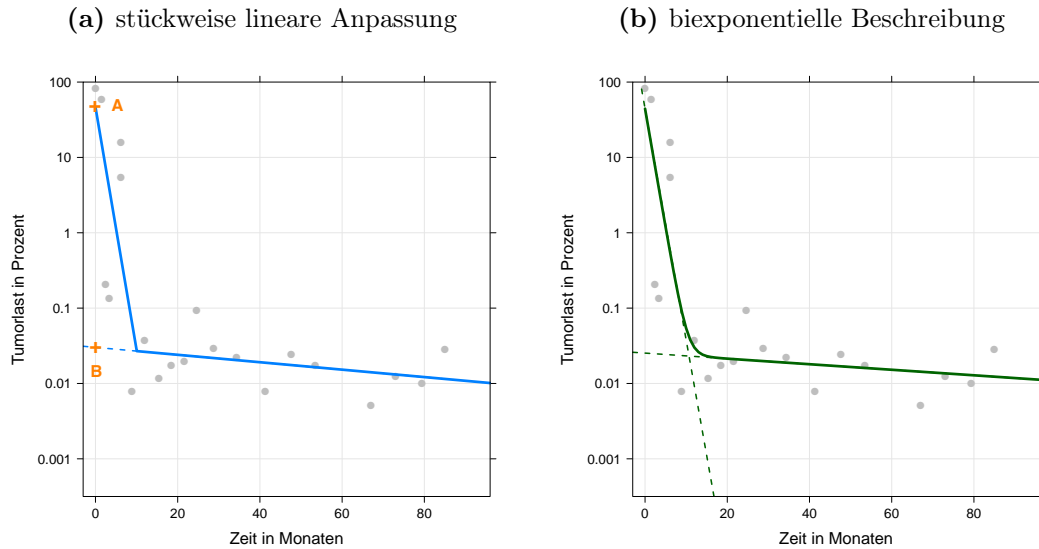


Abbildung 3.2.: Beispielhafter Therapieverlauf mit zusätzlich zur Abbildung 3.1b eingezeichneten Beschreibungen durch zwei verschiedene Anpassungsverfahren.

Abbildung 3.2a zeigt dabei die Anpassung mittels zweier Geraden, die durch lineare Regression ermittelt wurden (stückweise linear). Jede dieser Geraden wird durch eine eigene Funktion beschrieben. Der Übergang ist nicht kontinuierlich, da sich dort ein Bruchpunkt befindet. Für die nebenstehende Abbildung 3.2b wurde hingegen ein nichtlineares Regressionsverfahren genutzt, um eine biexponentielle, kontinuierliche Beschreibung über den gesamten Therapieverlauf zu erhalten. Beide Verfahren basieren auf der Minimierung des Abstandes der Beschreibung zu den Messpunkten.

Es wurde ein „Least-squares“-Ansatz verfolgt, um für die Beschreibung eines realen oder *in silico* simulierten Patienten die optimalen Parameter zu finden, welche die stückweise lineare Anpassungen repräsentieren. Dabei wird eine Beschreibung identifiziert, deren Abstand zu den einzelnen Messpunkten möglichst klein ist (Minimierung der Summe der quadratischen Abstände, weitere Erläuterungen siehe Fahrmeir et al. (2009)). Als Grundlage findet eine lineare Regression zur Beschreibung des gesamten Therapieverlaufs ohne Bruchpunkt statt. Ausgehend von dieser Regression wird getestet, ob ein Bruchpunkt das Anpassungsergebnis signifikant verbessern kann (gekennzeichnet durch kleinere Residuen durch Einfügen des Bruchpunktes im Vergleich zur Anpassung ohne Bruchpunkt). Dieser Test wird nach dem Erstbeschreiber als Davies-Test bezeichnet (Davies 1987). Ist das Ergebnis positiv, wird durch Verschieben dieses Bruchpunktes dessen optimale Lage bestimmt. Sie ist durch besonders kleine Residuen (Abstände zwischen Beschreibung und Messpunkten) gekennzeichnet (technisch

umgesetzt mit dem Paket *segmented* (Version 0.5-1.4) der Programmiersprache *R* Version 3.2.3).

3.1.3. Biexponentielle Beschreibung

Neben der Beschreibung des Therapieverlaufs mittels zweier Geraden auf der logarithmischen Skala gibt es ein biexponentielles Regressionsmodell, welches von Stein et al. (2013) erstmals auf die Daten der CML angewendet wurde. Dabei folgt die gesamte Kurve der Funktion

$$\log_{10} \left(\frac{BCR - ABL1}{ABL1} \right) (t) = \log_{10} (A \cdot e^{-\alpha t} + B \cdot e^{-\beta t}) \quad (3.4)$$

Der erste Summand ($A \cdot e^{-\alpha t}$) in der biexponentiellen Funktion hat dabei besonders großen Einfluss auf den vorderen steileren Abfall im Therapieverlauf, da α in der Regel betragsmäßig deutlich größer als β ist. Im flacheren Langzeitverlauf gewinnt der Einfluss des zweiten Summanden ($B \cdot e^{-\beta t}$) deutlich an Relevanz. Abbildung 3.2b dient hier zur Illustration. Die Verlängerung der beiden linearen Teilabschnitte des in Abbildung 3.2b gezeigten Therapieverlaufs bildet einen Schnittpunkt, der dem Bruchpunkt der stückweise linearen Anpassung entspricht (siehe Abbildung 3.2a).

Um analog zum Parameter A aus dem stückweise linearen Ansatz den Schnittpunkt mit der y-Achse zu berechnen, setzt man $t = 0$ und erhält

$$\log_{10} \left(\frac{BCR - ABL1}{ABL1} \right) (t = 0) = \log_{10} (A + B) \quad (3.5)$$

Da $B \ll A$ ist, gilt näherungsweise

$$\log_{10} \left(\frac{BCR - ABL1}{ABL1} \right) (t = 0) \approx \log_{10} (A) \quad (3.6)$$

Die biexponentielle Beschreibung weist gegenüber dem stückweise linearen Ansatz den Unterschied auf, dass der Bereich des Übergangs zwischen dem ersten und zweiten Therapieabschnitt kontinuierlich beschrieben wird. Auch eine komplette Differenzierbarkeit im mathematischen Sinne ist für diesen Kurvenverlauf gegeben. In Abbildung 3.2 wird die Ähnlichkeit zwischen dem stückweise linearen Ansatz und der biexponentiellen Anpassung deutlich.

Analog zur Beschreibung mittels stückweise linearem Ansatz muss auch hier eine Identifikation der besten Parameter zur Anpassung eines bestimmten Therapieverlaufs stattfinden. Dazu wurde die Funktion *nls* (Programmiersprache *R* Version 3.2.3) verwendet. Auch diese schätzt, ausgehend von Startparametern, Parameter der oben genannten Funktion so, dass die Residuen zwischen den Messwerten und der biexponentiellen Beschreibung im Sinne einer

nicht-linearen Regression minimiert werden (weitere Details Backhaus et al. (2015a) und Backhaus et al. (2015b)). Die Parameter A , B , α und β der biexponentiellen Beschreibung werden im Weiteren auch als Datenparameter bzw. als statistische Parameter der biexponentiellen Beschreibung bezeichnet.

3.2. Mathematische Modelle

Zur Modellierung des hämatopoetischen Systems gibt es populationsbasierte und einzelzellbasierte Modellierungsansätze. Für die erste Möglichkeit werden Zellen in ähnlichen Zuständen zusammengefasst und als Population betrachtet. Veränderungen, wie Zellzahl oder Konzentration, werden mittels Differentialgleichungen beschrieben, welche das mittlere Verhalten der Population charakterisieren. Eine Analyse einer einzelnen Zelle ist dabei nicht möglich.

Alternativ kann man jede Zelle als individuelles Element des Modells beschreiben. Jede Zelle hat ihre eigene Kombination an Eigenschaften und verhält sich als Individuum nach bestimmten Regeln. Auch die Interaktion zwischen verschiedenen Zellen ist möglich. Im Folgenden werden diese beiden Möglichkeiten anhand von Beispielen nochmals dargestellt, wobei das einzelzellbasierte Modell nach Roeder und Loeffler (2002) als Grundlage für die weiteren Ausführungen im Rahmen der vorliegenden Arbeit dient.

3.2.1. Populationsbasierte Modelle

Ein prototypisches Beispiel für ein differentialgleichungsbasiertes Leukämiemodell ist das von Michor et al. (2005) vorgestellte Verfahren. Darin werden die Differenzierungsstufen „Stammzelle“, „Progenitorzelle“, „differenzierende“ und „endgültig differenzierte“ Zelle abgebildet, welche die Zellen nacheinander durchlaufen. Für die Leukämie wird eine gesteigerte Proliferation der leukämischen Zellen angenommen, was dazu führt, dass sich das Verhältnis der Zellpopulationen zugunsten der kranken Zellen verschiebt. Die Therapie mit Imatinib wird durch eine geringere Anzahl an möglichen Zellteilungen simuliert. Der biphasische Abfall kommt durch folgende Effekte zustande: Zu Beginn fällt durch die geringere Teilungsaktivität der differenzierten leukämischen Zellen die Tumormast relativ schnell ab, da diese zahlenmäßig am größten vertreten sind. Nach einer gewissen Therapiedauer haben die differenzierten Zellen ein neues Gleichgewicht erreicht und es beginnt der zweite Abfall, in dem die Reduktion der Belastung mit BCR-ABL1-positiven Zellen von der geringeren Proliferation der Progenitorzellen getragen wird. Auch wenn im realen biologischen System die leukämischen Zellen nicht mehr komplett ausdifferenzieren (Aufreten von Blasten im peripheren Blut, die nicht mehr funktionstüchtig sind), werden in den modellhaften Beschreibungen auch von Leukämiezellen alle Differenzierungsstufen durchlaufen.

Eine relevante Depletion der leukämischen Stammzellen wurde nicht angenommen, da die zum Zeitpunkt der Veröffentlichung von Michor et al. (2005) und Abbott und Michor (2006) vorliegenden Daten zum Absetzen von Imatinib eine schnelle Rückkehr zu $\frac{BCR-ABL1}{ABL1}$ -Leveln, die vor Beginn der Therapie vorlagen, aufwiesen. Somit wurde eine Reduktion der Zellzahlen nur auf der Ebene der Progenitorzellen modelliert. In den ergänzenden Erläuterungen zu Michor et al. (2005) wird eine mögliche Apoptoseinduktion durch Imatinib als weitere Erklärung angeführt, die einen ähnlichen Verlauf des $\frac{BCR-ABL1}{ABL1}$ -Verhältnisses erzeugen könnte.

Ein komplettes Verschwinden mutierter Zellen lässt sich nicht simulieren, da unter Therapie mit Imatinib ein neues, durch geringe Tumorlasten gekennzeichnetes Gleichgewicht erreicht wird. Aus diesem heraus steigt die Tumorlast nach dem simulierten Absetzen von Imatinib wieder rapide an. Außerdem werden die leukämischen Stammzellen nicht aus dem System entfernt, lediglich ihre Differenzierung wird reversibel gehemmt. Durch Einführung einer dritten Zellklasse neben den gesunden und leukämischen Zellen, kann auch die Entwicklung von Resistenzen gegen Imatinib beschrieben werden (Michor et al. 2005).

3.2.2. Einzelzellbasierte Modelle

Im Gegensatz zu populationsbasierten Modellen erlauben einzelzellbasierte Modelle (auch agentenbasiert genannt) den Zugriff auf die Beschreibung einzelner simulierter Zellen. Als prototypische Herangehensweise sei an dieser Stelle das Modell nach Roeder und Loeffler (2002) genannt. Es wurde ursprünglich für die Simulation des hämatopoetischen Systems der Maus entwickelt und später von Roeder et al. (2006) und Horn et al. (2008) auf den Menschen angepasst. Das Modell betrachtet jede Zelle einzeln mit einer Kombination bestimmter Eigenschaften, die sich nach vorbestimmten Regeln in einem stochastischen Prozess ändern. Leukämische und gesunde Zellen unterscheiden sich durch unterschiedliche Parameter, die Einfluss auf ihre Entwicklung haben. Alle Zellen können sich in unterschiedlichen Kompartimenten aufhalten. Neben einem Kompartiment, welches im realen Patienten dem Knochenmark entspricht und in welchen die Stammzellen proliferieren oder ruhen können, gibt es das simulierte periphere Blut. Dort leben die Zellen eine vorbestimmte Zeit, bevor sie nach Ablauf der Lebensdauer aus dem System entfernt werden. Die beiden Zellsorten stehen im Wettbewerb um begrenzte Ressourcen (z. B. Größe bzw. maximale Zellanzahl im simulierten Knochenmark). Da die leukämischen Stammzellen einen Vorteil gegenüber den gesunden Zellen haben, verdrängen sie nach und nach die gesunde Blutbildung. Eine Therapie mit TKI wird durch die Änderung der Proliferationsparameter der leukämischen Zellen simuliert. Gleichzeitig wird pro Simulationszeitschritt ein bestimmter Anteil kranker Stammzellen aus dem System entfernt. Mit dieser Beschreibung ist es möglich Tumorlastverläufe im peripheren

Blut zu erzeugen, welche denen von realen Patienten entsprechen. Gleichzeitig besteht die einfache Möglichkeit Zugriff auf den Verlauf der Belastung mit Tumorzellen im simulierten Knochenmark zu bekommen (Horn et al. 2013; Horn 2014). Für eine detaillierte Beschreibung des Modells sei auf Kapitel 5.2 verwiesen.

3.3. Patientenindividuelle Anpassungsverfahren

Um Zugriff auf die Stammzellzahlen zu bekommen, muss die *in silico* Simulation optimal mit dem Therapieverlauf des betrachteten Patienten übereinstimmen. Dazu muss sie mit einem bestimmten Modellparametersatz gestartet werden. Zur Identifikation des optimalen Startparameterdatensatzes gibt es verschiedene Ansätze. In der Arbeit von Horn et al. (2013) wurde mittels eher einfacher Regression eine Beziehung zwischen den Modellparametern und Datenparametern hergestellt. Diese Beschreibung wird genutzt, um aus einer bestimmten Kombination aus statistischen Datenparametern die zur Modellierung nötigen, optimalen Modellparameter vorherzusagen. Die resultierende Simulation erlaubt schließlich die Rückschlüsse auf die Stammzellzahlen. Dieser Ansatz wird im Kapitel 5.3 nochmals vorgestellt und bildet den Ausgangspunkt für die weiteren Verfahren, die im Rahmen der vorliegenden Arbeit entwickelt werden. Ein Überblick über bestehenden und in der vorliegenden Arbeit neu entwickelte Verfahren ist in Abbildung 4.1 dargestellt. Es ist davon auszugehen, dass eine Simulation, deren Therapieverlauf im simulierten peripheren Blut nahe am korrespondierenden Therapieverlauf des Patienten liegt, eine gute Basis für weitere Vorhersagen ist. Leider ist diese Übereinstimmung nicht für alle analysierten Patienten gegeben (Horn 2014). Somit muss man annehmen, dass für diese Patienten eine sichere Aussage über den Verlauf der Belastung mit leukämischen Zellen im Knochenmark nur schwer möglich ist. Um eine gute Grundlage für die Beschreibung der Therapie im Knochenmark zu haben, muss die Anpassung für einen größeren Teil der Patienten besser werden.

4. Fragestellung und Zielsetzung dieser Arbeit

Aus den im letzten Kapitel 3.3 erläuterten Problemen folgen die anschließend genannten zentralen Fragen, die im Rahmen der Arbeit beantwortet werden sollen:

- Welche alternativen Verfahren sind geeignet, um die Qualität des Anpassungsverfahrens, welches in Horn et al. (2013) beschrieben wird, zu verbessern?
- Sind die vorhergesagten Therapieverläufe im Stammzellkompartiment geeignet, als Ausgangspunkt für Absetzvorhersagen zu dienen?

Um die genannten Fragen zu analysieren und umfassend zu beantworten, werden folgende Punkte umgesetzt:

- Erzeugung einer größeren Datenbasis, die als Grundlage für neue Anpassungsverfahren dienen soll
- Entwicklung eines Gütemaßes (basierend auf Abstandsberechnungen) mit dem sich unterschiedliche Simulationen untereinander und mit realen Therapieverläufen vergleichen lassen
- Entwicklung neuer Anpassungsverfahren
 - Entwicklung eines neuen Regressionsverfahrens zu Parameteridentifikation (M1, vgl. Abbildung 4.1) basierend auf dem Verfahren von Horn et al. (2013) (M0))
 - Entwicklung eines Anpassungsverfahrens, welches auf dem direkten Vergleich der Therapieverläufe basiert, und den Abstand als Auswahlkriterium nutzt (M2)
- Vergleich der neuen Anpassungsverfahren untereinander und mit dem Verfahren nach Horn et al. (2013) anhand des entwickelten Gütemaßes
- Prüfung der Robustheit der getroffenen Stammzellvorhersagen anhand ausgewählter Therapieverläufe im simulierten Knochenmark

Abbildung 4.1 gibt eine Übersicht über die verwendeten Verfahren, die basierend auf dem Therapieverlauf eines Patienten im peripheren Blut Aussagen über einen möglichen Absetzzeitpunkt treffen sollen. Ausgehend vom Therapieverlauf eines CML-Patienten werden entweder mittels stückweise linearer oder biexponentieller Anpassung die statistischen Parameter der Therapie ermittelt.

Die Methode M0 nach Horn et al. (2013) nutzt die zwei Anstiegsparameter der stückweise linearen Beschreibung (α und β), um anhand einer iterativen linearen Regression Rückschluss auf Modellparameter r_{deg} und f_ω zu ziehen. Diese Regression basiert auf einer Vorbetrachtung von 1600 Simulationsläufen, bei deren Erstellung die Modellparameter r_{deg} und f_ω variiert wurden und der Effekt auf α und β ermittelt wurde. Mit den ermittelten Werte für r_{deg} und f_ω muss eine erneute Modellsimulation gestartet werden, die dann der simulierten Therapie des Patienten entspricht.

Basierend auf der M0 erfolgte die Entwicklung der Methode M1. Dabei erfolgt die Beschreibung des Therapieverlaufs anhand einer biexponentiellen Funktion, welche von Stein et al. (2013) erstmals auf Daten der CML angewendet wurde. Anschließend werden alle vier Parameter dieser biexponentiellen Beschreibung (A , B , α und β) genutzt um anhand einer Regression, die auf Iteration verzichtet, Rückschluss auf fünf variierte Modellparameter (f_α , f_ω , r_{deg} , $trans$ und $init$) zu erhalten. Dieses Regressionsverfahren basiert auf einer Datenbasis von 270.400 Simulationen, welche durch die systematische Variation der genannten fünf Modellparameter erstellt wurde. Auch hier muss mit den ermittelten Modellparametern ein erneuter Simulationslauf gestartet werden, der die Therapiesimulation des Patienten zum Ergebnis hat.

Die Methode M2 nutzt die gleiche Datenbasis wie M1. Dabei erfolgt die Ermittlung der am besten zum Patienten passenden Therapiesimulation anhand des Abstandes zwischen den biexponentiellen Beschreibungen aller 270.400 Simulationen und der biexponentiellen Beschreibung der Patientendaten. Jede individuelle Simulation der Datenbasis wird mit Θ_j bezeichnet, wobei j eine natürliche Zahl zwischen 1 und 270.400 sein kann. Die Simulation mit dem geringsten Abstand wird sofort als Therapiesimulation Θ^* des Patienten identifiziert. Der erneute Simulationslauf entfällt.

Die mittels der Verfahren M0 bis M2 identifizierten Therapiesimulationen erlauben anschließend eine Vorhersage über den Verlauf der Tumorlast im Stammzellkompartiment. Anhand der sich unter Therapie reduzierenden Anzahl der leukämischen Stammzellen kann ein Absetzzeitpunkt vorhergesagt werden, nach welchem höchstwahrscheinlich kein Rezidiv mehr auftreten sollte. Die aufgeführten Methoden M0 bis M2 werden ausführlich in den Kapiteln 5.3 bzw. 6.5 und 6.6 erläutert.

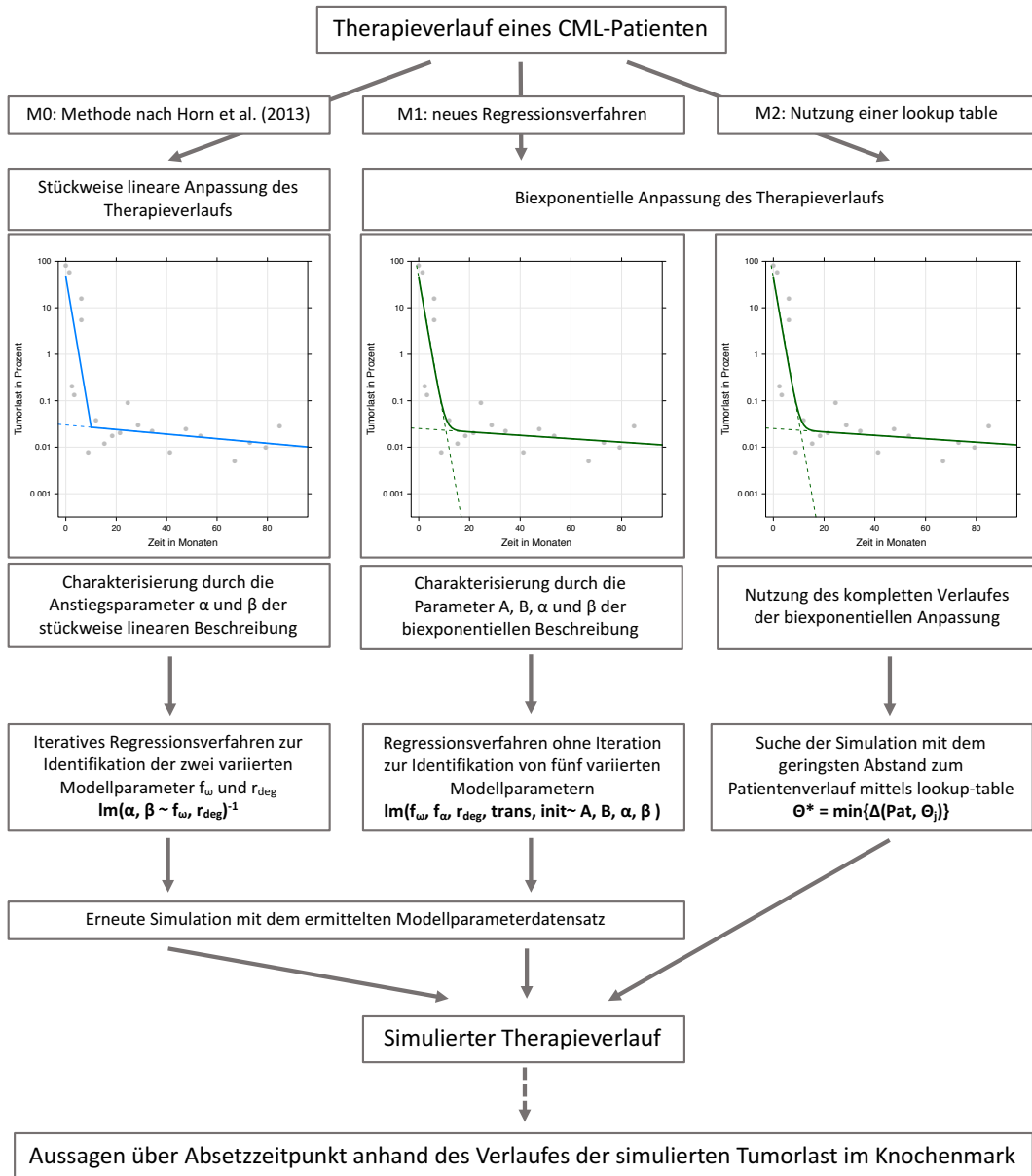


Abbildung 4.1.: Übersichtsabbildung der verwendeten Methoden. Diese zeigt wie anhand der statistischen Beschreibung des Therapieverlaufs eines Patienten (stückweise lineare oder biexponentielle Anpassung) über verschiedene Regressions- bzw. Suchverfahren mittels lookup-table ein simulierter Therapieverlauf identifiziert wird. Diese erlaubt Aussagen über einen höchstwahrscheinlich rezidivfreien Absetzzeitpunkt anhand des Verlaufes der Tumorlast auf Ebene der leukämischen Stammzellen. Eine ausführliche Erläuterung der Verfahren M0 bis M2 erfolgt in den Kapiteln 5.3 bzw. 6.5 und 6.6.

5. Material und Methoden

5.1. IRIS-Datensatz

In der „International Randomized Trial of IFN- α vs. STI571 (Imatinib)“ (IRIS-Studie) wurde die Wirksamkeit einer Therapie mit Interferon- α (IFN- α) und niedrig dosiertem Cytarabin (Ara-C) gegen eine Monotherapie mit Imatinib (IM; Handelsname: *Glivec*, Novartis) untersucht. Es handelte sich dabei um eine prospektive zweiarmige randomisierte Phase-III-Studie. Dabei wurden weltweit 1106 Patienten (553 je Arm) in die Studie eingeschlossen. Primärer Endpunkt war die Progression der CML, wobei auch Tod jeglicher Ursache, Übergang in eine Blastenkrise, Akzeleration oder der Anstieg weißer Blutkörperchen auf über $20 \cdot 10^9/l$ und Remissionsverlust eingeschlossen wurden. Sekundär wurde das Erreichen hämatologischer und zytogenetischer Remissionen betrachtet. Die Patienten durften unter bestimmten Voraussetzungen (fehlende oder sich verschlechternde Wirkung, steigende Tumorlast unter Therapie, nicht beherrschbare Nebenwirkungen) zwischen den beiden Studienarmen wechseln. Bereits nach einer medianen Beobachtungszeit von 18 Monaten hatten über 300 Patienten des IFN- α -Armes diese Möglichkeit genutzt, in den deutlich wirksameren und nebenwirkungsärmeren IM-Arm zu wechseln (O'Brien et al. 2003). Somit entwickelte sich die IRIS-Studie zu einer fast reinen Beobachtungsstudie der Behandlung mit Imatinib. Es zeigte sich dabei die Überlegenheit und gute Verträglichkeit von IM in einer täglichen Dosierung von 400 mg (Hochhaus et al. 2009).

Für die vorliegende Arbeit wurden die Therapieverläufe der Patienten aus dem deutschen Arm der IRIS-Studie ($n = 69$) nach folgenden Kriterien selektiert:

- mindestens sieben Messungen des $\frac{BCR-ABL1}{ABL1}$ -Verhältnisses
- mindestens zwei Jahre Follow-up innerhalb der IRIS-Studie
- innerhalb der ersten beiden Jahre der Therapie mindestens einen $\frac{BCR-ABL1}{ABL1}$ -Messwert unterhalb von 10 %

Es wurde der deutsche Arm der Studie aufgrund der Vergleichbarkeit zu den Analysen von Horn et al. (2013) gewählt. Nur für diesen liegen Messungen anhand des Verfahrens der *nested PCR* vor. Insgesamt wurden nach Anwendung der Ausschlusskriterien 51 Patienten in die Analysen einbezogen.

5.1.1. Vorbereitung der Messwerte

Die Messung des $\frac{BCR-ABL1}{ABL1}$ -Verhältnisses wurde mittels qPCR (siehe Kapitel 2.2.4) durchgeführt. Parallel zu den Messungen wurde immer die Detektionsschwelle der qPCR ermittelt. Für den Fall, dass die qPCR ein negatives Ergebnis hervorbrachte, wurde eine *nested PCR* durchgeführt, die wie oben beschrieben eine deutlich höhere Sensitivität und Spezifität aufweist, jedoch keine quantifizierbare Aussage erlaubt. Diese beantwortet damit nur die Frage, ob überhaupt BCR-ABL1-Gene in der DNA-Probe vorliegen. War auch diese negativ, so wurde keine $\frac{BCR-ABL1}{ABL1}$ -Last in die Betrachtungen einbezogen. War die *nested PCR* positiv, so wurde die Detektionsschwelle der qPCR als konservative Annahme der $\frac{BCR-ABL1}{ABL1}$ -Tumorlast als Messwert registriert. Im Folgenden ist aufgeführt, welcher Wert bei einer bestimmten Ergebniskombination der verwendeten PCR-Verfahren als Messwert in die Analysen einbezogen wurde:

- die qPCR ermittelt ein $\frac{BCR-ABL1}{ABL1}$ -Verhältnis $> 0 \rightarrow$ Messwert entspricht dem konkreten Ergebnis der qPCR
- die qPCR ermittelt ein $\frac{BCR-ABL1}{ABL1}$ -Verhältnis von 0 und die *nested PCR* ist positiv \rightarrow die Detektionsschwelle der qPCR wird als Messwert registriert
- die qPCR ermittelt ein $\frac{BCR-ABL1}{ABL1}$ -Verhältnis von 0 und die *nested PCR* ist negativ \rightarrow es wird ein $\frac{BCR-ABL1}{ABL1}$ -Verhältnis von 0 als Messwert registriert

In aktuellen Arbeiten, deren Veröffentlichung nach Durchführung der vorliegenden Analysen erfolgte, wurde das Problem der Standardisierung zwischen den unterschiedlichen Laboren weiter analysiert. So wurden in der Publikation von Cross et al. (2015) Vorschläge zur Behandlung der Messwerte im Bereich sehr kleiner Tumorlasten gemacht. Auch die Standardisierung der Messwerte bei der Verwendung von Transkripten anderer Referenzgene als *ABL1* ist dort beleuchtet.

5.1.2. Statistische Beschreibung der Patientenkohorte

Zur statistischen Beschreibung der Patientenkohorte werden an dieser Stelle sogenannte „bean-plots“ eingeführt, die auch bei der Datenauswertung des Verfahrensvergleiches im Ergebnisteil verwendet werden. Bei diesen Diagrammen werden durch kleine waagerechte Striche (in der vorliegenden Arbeit jeweils in der Farbe grün gehalten) die genauen Werte der untersuchten Größe verdeutlicht. Haben mehrere Messungen den gleichen Wert, so sind die entsprechenden Linien verlängert dargestellt. Die sehr lange waagerechte Linie (in den Diagrammen in der Farbe schwarz) verdeutlicht den Mittelwert aller aufgetragenen Beobachtungen. Die farblich ausgefüllte Hüllkurve (jeweils in lila dargestellt) wird durch

eine senkrecht abgedruckte und mittig gespiegelte Dichtefunktion, welche die Verteilung der dargestellten Daten repräsentiert, dargestellt. Ist die Hüllkurve in einem Bereich besonders breit, so liegen viele Messwerte in diesem Datenbereich. Ist sie entsprechend schmal, sind nur wenige Messwerte in diesem Messbereich zu erwarten (Kampstra 2008).

Die Abbildungen 5.1 und 5.2 zeigen die Verteilungen unterschiedlicher Parameter der verwendeten Patientenkohorte. Das Alter bei Diagnosestellung (siehe Abbildung 5.1a) ist im Bereich von 20 bis 70 Jahre verteilt. Dabei liegt der Mittelwert bei 50 Jahren. Es bestätigen sich die in Kapitel 2.2.1 getroffenen Aussagen über die Epidemiologie der CML, die die Krankheit als solche des älteren Patienten beschreiben. Der initiale Wert der $\frac{BCR-ABL1}{ABL1}$ -Belastung im peripheren Blut (siehe Abbildung 5.1b) ist im Bereich von unter 1 % bis über 100 % weit gestreut. Die theoretisch unmöglichen Messungen von über 100 % treten methodenbedingt bei der Verwendung der *qPCR* (siehe Kapitel 2.2.4) auf. Sowohl das Maximum als auch die mittlere initiale Tumorlast der Patienten liegt im Bereich von 100 %.

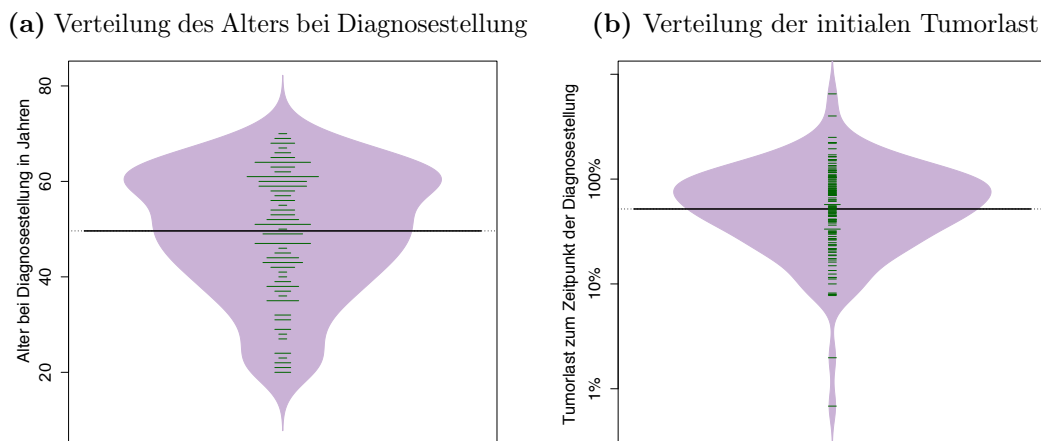


Abbildung 5.1.: Alter und Tumorlast bei Diagnosestellung jeweils für alle IRIS-Patienten des deutschen Armes aufgetragen. Dabei zeigt sich, dass das Alter der Patienten im Bereich von 20 bis 70 Jahren liegt. Das mittlere Alter bei Stellung der Diagnose CML beträgt ungefähr 50 Jahre (Abbildung 5.1a). Die initiale Tumorlast liegt gestreut im Bereich von unter 1 % bis weit über 100 % $\frac{BCR-ABL1}{ABL1}$ -Verhältnis verteilt. Das Maximum sowie der mittlere Wert liegen dabei bei 100 % (Abbildung 5.1b). Die logarithmische Einteilung der y-Achse ist zu beachten. (Werte von über 100 %, die theoretisch nicht möglich sind, treten methodenbedingt bei der Bestimmung der Tumorlast mittels *qPCR* auf.)

Bei der Analyse der Verteilung der Beobachtungsdauer (siehe Abbildung 5.2a) aller verwendeten Patienten der IRIS-Studie fällt auf, dass ein Großteil etwa sieben Jahre im Rahmen der Studie nachbeobachtet wurde. Patienten, die kürzer eingeschlossen waren, verteilen sich gleichmäßig auf den Bereich unterhalb von fünf Jahren. Die Anzahl der gemessenen Daten-

punkte (siehe Abbildung 5.2b) zeigt eine andere Verteilung mit einem Mittelwert von 15 Messungen pro Patient, sodass nicht davon ausgegangen werden kann, dass eine besonders lange Beobachtungsdauer mit vielen Messpunkten einhergeht. Die Verteilung der Anzahl von Messungen oberhalb von 10 % (siehe Abbildung 5.2c) ist insofern interessant, als dass die Patienten mit vielen Beobachtungen größer 10 % eher keinen Abfall in der Tumorlast unter Therapie zeigen. Somit entsprechen diese Patienten nicht dem typischen biexponentiellen Abfall und werden von Methoden, die im Rahmen dieser Arbeit genutzt wurden, nicht abgedeckt. Es sind keine Vorhersagen über einen Absetzzeitpunkt möglich und sinnvoll, da bei einer Therapie ohne Abfall der Tumorlast eher über ein Wechsel des Therapieregimes nachgedacht werden muss.

5.1.3. Deskriptive Analyse des „medianen Patienten“

Um die IRIS-Studie auszuwerten, wurde der Therapieverlauf eines theoretischen, „medianen“ Patienten unter Behandlung mit Imatinib berechnet. Dazu wurden in einem bestimmten Zeitfenster (3 Monate) die Messwerte aller Patienten betrachtet. Von diesen wurde der Median, die 25 %- und 75 %-Quantile bestimmt. Anschließend wurde das Zeitfenster um einen Monat verschoben, um erneut dieselben Betrachtungen durchzuführen. Dieses Vorgehen bezeichnet man als *moving average* (technisch exakt eher als *moving median*). Die resultierenden Mediane und Quantile wurden analog zu den Therapieverläufen der einzelnen Patienten aufgetragen (siehe Abbildung 5.3).

Betrachtet man den Therapieverlauf des medianen Patienten, so fällt auf, dass er bereits nach ca. 18 Monaten, noch im Bereich des ersten steilen Abfalls der Tumorlast, die MMR (major molecular response) erreicht. Dies entspricht einer Reduktion der Tumorlast auf 0,1 % bzw. ausgehend von 100 % um drei log-Stufen (Horn 2014). In diesem Zustand sind die Patienten in einer zytologischen Remission und klinische Symptome treten nicht mehr auf. Trotzdem muss die Therapie mit Imatinib fortgesetzt werden, da ein Absetzen zu einem sehr schnell einsetzenden Rezidiv führen würde (Michor et al. 2005). Im weiteren Therapieverlauf fällt die Tumorlast flacher ab, bevor sie im Median nach etwas mehr als 90 Monaten Behandlungsdauer noch 0,01 % beträgt. Dies entspricht einer Reduktion um vier log-Stufen. Nur an einem einzigen Messpunkt eines Patienten trat eine nicht messbare Tumorlast auf (es konnte in der Probe kein BCR-ABL1-Transkript identifiziert werden). Dies ist zum einen der hohen Sensitivität und Spezifität einer *nested PCR* zuzuschreiben, zum anderen entfernt Imatinib in der angewendeten Konzentration über den Beobachtungszeitraum nicht alle leukämischen Zellen aus den Körpern der Patienten der Studie.

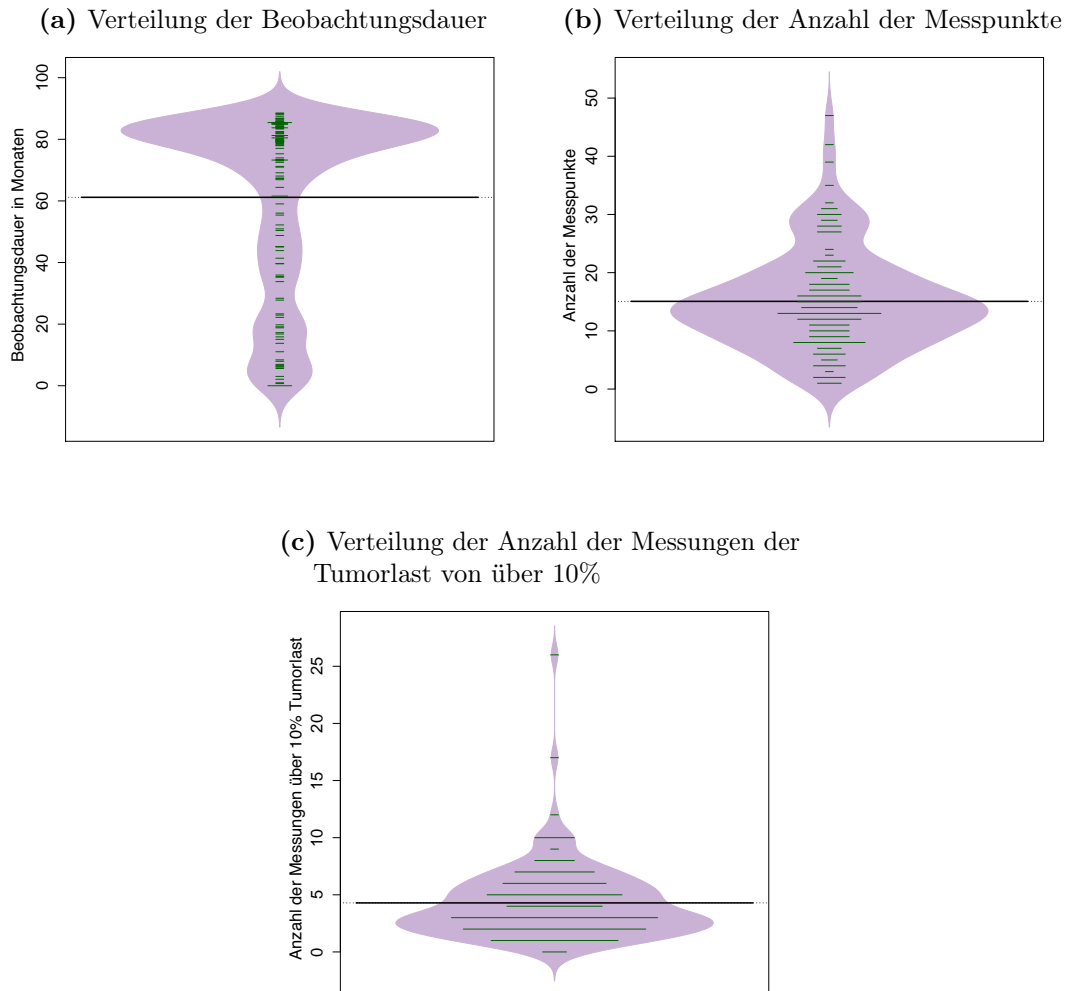


Abbildung 5.2.: Beobachtungsdauer, Anzahl der Messungen und Anzahl der Messungen über 10% jeweils für alle IRIS-Patienten des deutschen Armes aufgetragen. Der größte Teil der Patienten wurde circa sieben Jahre im Rahmen der Studie nachbeobachtet. Für den Rest liegt die Beobachtungsdauer gleichmäßig verteilt im Bereich unterhalb von fünf Jahren (Abbildung 5.2a). Anders sieht die Verteilung der Anzahl der Messpunkte aus. Dort zeigt sich, dass ein Großteil der Patienten zwischen zehn bis 20 Messungen im Therapieverlauf erhalten hat (Abbildung 5.2b). Messungen über 10% treten nur bei einem kleinen Teil der Patienten häufig auf, da anhaltend hohe Tumorlasten eher Hinweis auf schlechtes Therapieansprechen sind (Abbildung 5.2c).

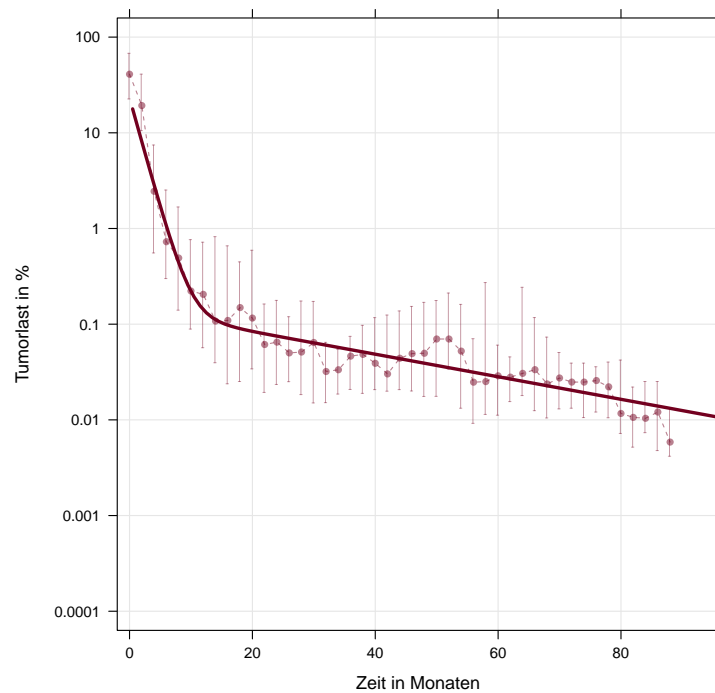


Abbildung 5.3.: Darstellung des medianen Patienten der IRIS-Studie unter Behandlung mit Imatinib. In einem Zeitfenster von 3 Monaten wurden jeweils der Median und die 25 %- und 75 %-Quantile bestimmt. Diese werden als Messpunkt mit den Antennen aufgetragen. Anschließend wird das Zeitfenster um einen Monat verschoben, um den nächsten Punkt zu bestimmen. Mithilfe der ermittelten Punkte ist es möglich, eine biexponentielle statistische Anpassung zu berechnen (rote Linie, siehe Kapitel 3.1.3).

5.2. Einzelzellbasiertes Stammzellmodell

Das von Roeder und Loeffler (2002) entwickelte Stammzellmodell dient als Grundlage für die angewendeten Simulationen des Therapieverlaufs von CML-Patienten unter TKI-Therapie. Es wurde mit Erweiterungen ausführlich unter anderem in folgenden Veröffentlichungen beschrieben: Roeder und Loeffler (2002), Loeffler und Roeder (2002), Roeder et al. (2006) und Glauche (2010) und Horn et al. (2013). Die nachfolgenden Erläuterungen beziehen sich explizit auf diese Quellen, die aufgrund der besseren Lesbarkeit innerhalb des gesamten Kapitels 5.2 nicht erneut genannt werden. Zitate aus anderen Quellen sind entsprechend kenntlich gemacht.

Für das Modell wird angenommen, dass Stammzellen einzeln und unabhängig voneinander auf Signale ihrer Umgebung reagieren können. Jede Zelle hat dabei einen bestimmten Zustand, der durch eine individuelle Kombination von Variablen beschrieben wird. Die Variablenkombinationen werden in jedem Simulationszeitschritt, der typischerweise einer Stunde entspricht, nach bestimmten Bedingungen aktualisiert. Das Stammzellkompartiment des Modells besteht aus zwei konzeptionell unterschiedlichen Unterkompartimenten A und Ω (siehe auch Abbildung 5.4). Neben der aktuellen Zuordnung zu einem bestimmten Kompartiment wird eine Zelle noch durch die weiter unten beschriebene Affinität a zur Umgebung A und einen Zellzykluszähler c definiert. Durch diese Variable erfolgt eine Zuordnung zur jeweiligen Position im Zellzyklus (G1, M, G2, S oder G0). Hat eine Zelle den Zellzyklus vervollständigt, werden aus der Mutterzelle zwei identische Tochterzellen erzeugt, deren Zellzykluszähler zurückgesetzt werden.

Das Kompartiment A charakterisiert dabei eine Umgebung, in der die Stammzelle in einen ruhenden Zustand, üblicherweise mit der Zellzyklusphase G0 assoziiert, übergeht. Damit wird ihre Ausreifung verhindert und die Fähigkeit der Regeneration bleibt erhalten bzw. wird verbessert. Dies sind die Eigenschaften, die im Allgemeinen mit einer Stammzellnische in Verbindung gebracht werden (Schofield 1978). Im Modell wird die Affinität zu dieser Nische über die bereits oben genannte Variable a abgebildet. Der Wert von a erhöht sich in jedem Simulationszeitschritt um den Faktor r (*Regenerationsrate*, $r \geq 1$), solange sich die Zelle im Kompartiment A befindet, bis zu einem $a_{\max} = 1$. Zellteilung findet in A nicht statt, dazu muss die Zelle erst ins Kompartiment Ω wechseln.

Im Gegensatz zu A charakterisiert das Kompartiment Ω einen Bereich, in dem die Stammzellen proliferieren. Dabei verlieren sie die Bindung zur jeweiligen Nische und auch die Fähigkeit zur Repopulation (Regeneration) nimmt ab. Im Modell wird ein pro Zeitschritt um den Faktor $\frac{1}{d}$ abnehmendes a verwendet. Dabei beschreibt d die sogenannte *Differenzierungsrate* ($d \geq 1$). Sinkt die Variable a unter ein bestimmtes a_{\min} , so kann die betroffene Zelle nicht mehr ins Kompartiment A zurückwechseln. Solche Zellen fungieren als sogenannte proliferierende

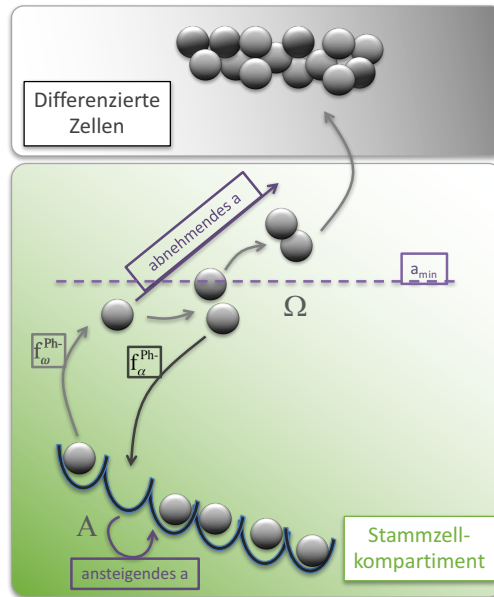


Abbildung 5.4.: Stammzellmodell mit normaler Hämatopoese. Die Stammzellen wechseln zwischen den Kompartimenten A und Ω im Stammzellkompartiment, welches üblicherweise dem Knochenmark entspricht, beschrieben durch die Übergangsfunktionen 5.1 und 5.2. Einfluss auf diese Funktionen haben $f_{\alpha}(N_A)$ und $f_{\omega}(N_{\Omega})$, zwei Faktoren, die jeweils von der Anzahl der Zellen im jeweiligen Kompartiment A oder Ω abhängig sind. Im Kompartiment A nimmt a um den Faktor r pro Simulationsschritt bis zu einem definierten a_{max} zu, im Kompartiment Ω um den Faktor $1/d$ bis zu a_{min} ab. Wird a_{min} unterschritten, so kann die betroffene Zelle nicht mehr ins Kompartiment A zurück wechseln. Sie kann sich als proliferierende Vorläuferzelle weiter teilen und wechselt später in den Bereich der differenzierten Zellen, welcher üblicherweise dem peripheren Blut entspricht. (Diese Abbildung wurde, genau wie Abbildung 5.6 bis 5.8 nach einem Entwurf von Ingmar Glauche erstellt.)

Vorläufer-Zellen und teilen sich demnach weiter. Nach vollständiger Ausreifung leben die Zellen als nicht proliferierende, ausdifferenzierte Einheiten noch bis zum Ende ihrer Lebensdauer in einem weiteren Kompartiment aus differenzierten Zellen, welches dem peripheren Blut entspricht. Nach Ablauf der Lebenszeit wird die Zelle aus dem System entfernt, was einem Absterben der Zelle entspricht.

Die Wahrscheinlichkeit $\omega(a, N_\Omega)$ für einen Übergang von Kompartiment A nach Ω wird durch

$$\omega(a, N_\Omega) = \frac{a_{min}}{a} \cdot f_\omega(N_\Omega) \quad (5.1)$$

beschrieben. Für den umgekehrten Wechsel errechnet sich die Wahrscheinlichkeit $\alpha(a, N_A)$

$$\alpha(a, N_A) = \frac{a}{a_{max}} \cdot f_\alpha(N_A) \quad (5.2)$$

Die Wahrscheinlichkeit für einen Wechsel des Kompartiments (ω, α) ist vom aktuellen Affinitätswert a der jeweiligen Zelle und dem Füllungsgrad N_A bzw. N_Ω der Zielkompartimente abhängig. Bei einem kleiner werdenden a steigt die Wahrscheinlichkeit, dass die Zelle nach Ω wechselt bzw. dort verbleibt. Hingegen drückt ein $a = a_{max}$ eine hohe Affinität zu A aus. Es errechnet sich somit eine große Wahrscheinlichkeit für den Wechsel nach A und eine kleine für den nach Ω .

Die sigmoidalen Funktionen $f_\alpha(N_A)$ und $f_\omega(N_\Omega)$ haben die Form

$$f(N) = \frac{1}{v_1 + v_2 \cdot e^{(v_3 \cdot \frac{N}{\tilde{N}})}} + v_4 \quad (5.3)$$

Die Faktoren v_1, v_2, v_3 und v_4 lassen sich nach Roeder (2003) in die leichter erfassbaren Werte $f(0), f(\frac{N}{2}), f(\tilde{N})$ und $f(\infty) = \lim_{N \rightarrow \infty} f(N)$ umrechnen. Dabei sind die Faktoren v_1 bis v_4 jeweils verschieden für gesunde (Ph^-) und leukämische (Ph^+) Stammzellen. Damit ergeben sich unterschiedliche Wahrscheinlichkeiten für den Wechsel zwischen den Kompartimenten. Diese sind so gestaltet, dass die Ph^+ -Zellen eine größere Wahrscheinlichkeit haben, vom ruhenden in den aktiven Zustand zu wechseln und umgekehrt. Damit haben sie einen Proliferationsvorteil gegenüber den gesunden Stammzellen. Die Form der Funktionen $f_\alpha(N)$ und $f_\omega(N)$ ist in Abbildung 5.5 dargestellt.

Da Stammzellen in A als ruhend angesehen werden, ändert sich pro Simulationszeitschritt nur die Affinität a wie oben beschrieben. Der Zell-Zyklus-Zähler c bleibt gleich.

Für die homöostatische Situation wird das System üblicherweise nur mit gesunden Zellen initialisiert. Zur Simulation der Pathogenese der CML fügt man nach Erreichen der Homöostase eine Ph^+ -Zelle zum System dazu und setzt die Simulation fort. Die Ph^+ -Zellen unterscheiden sich von den Ph^- -Zellen durch andere Übergangsfunktionen f_α und f_ω (siehe

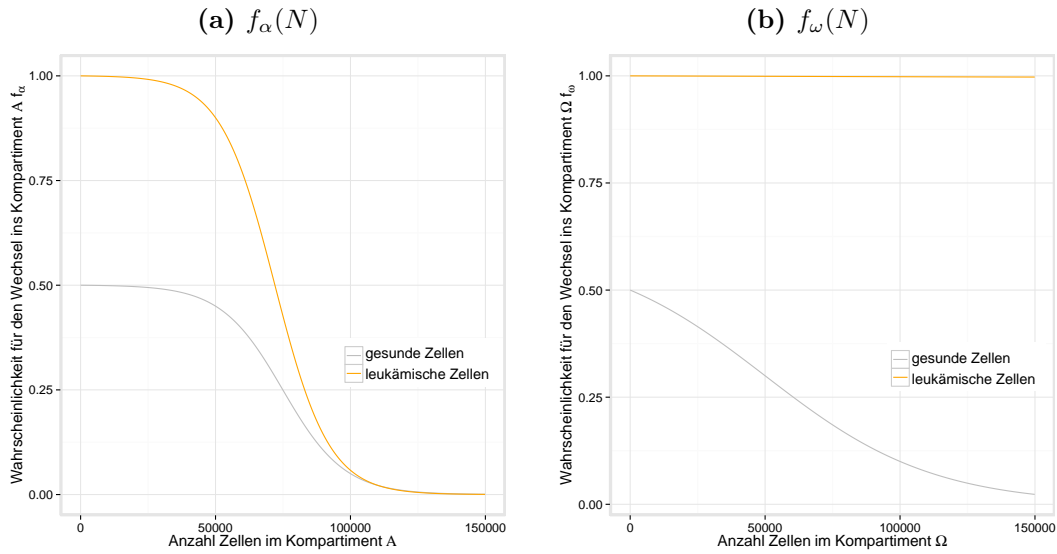


Abbildung 5.5.: Funktionen, die die Wahrscheinlichkeit des Wechsels ins Kompartiment A (Abbildung 5.5a) bzw. Ω (Abbildung 5.5b) pro Zeitintervall (eine Stunde) für gesunde und leukämische Stammzellen zeigen. Dabei wird deutlich, dass Ph^+ -Zellen immer eine höhere Wahrscheinlichkeit für den Wechsel ins jeweils andere Kompartiment haben als gesunde Zellen. Dies verschafft ihnen einen Proliferationsvorteil, der letztlich zur Verdrängung der gesunden Hämatopoese führt.

oben). Zur Verdeutlichung dient hier auch Abbildung 5.6.

Die Ph^+ -Stammzellen verdrängen die normalen Zellen der Hämatopoese im Stammzellkompartiment und sorgen für die Produktion von BCR-ABL1-positiven Zellen, die anschließend auch im peripheren Blut zu finden sind. Dort kann der Anteil leukämischer Zellen analog zur Messung im realen Patienten (siehe Kapitel 2.2.4) als $\frac{\text{BCR-ABL1}}{\text{ABL1}}$ -Verhältnis bestimmt werden kann. Da man davon ausgeht, dass pro Zelle das Gen ABL1 zweimal, das Fusionsgen BCR-ABL1 in einer leukämischen Zelle aber nur einmal vorkommt, nutzt man zur Berechnung der peripheren Tumorlast folgende Formel:

$$\frac{\text{BCR} - \text{ABL1}}{\text{ABL1}} = \frac{n_{\text{leuk. Zellen}}}{n_{\text{leuk. Zellen}} + 2 \cdot n_{\text{gesunde Zellen}}} \cdot 100\% \quad (5.4)$$

Diese Werte sind damit analog zu den durch die verschiedenen Zählverfahren bestimmten $\frac{\text{BCR-ABL1}}{\text{ABL1}}$ -Anteilen im realen Patienten und können deshalb in gleichen Koordinatensystemen dargestellt werden.

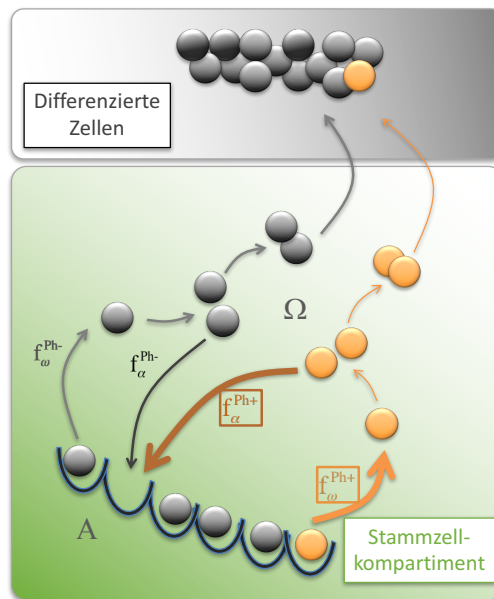


Abbildung 5.6.: Stammzellmodell mit leukämischen Stammzellen in der Entstehung einer CML. Diese Ph^+ -Zellen sind gegenüber den gesunden Zellen der Hämatopoese durch größere Übergangswahrscheinlichkeiten ($f_{\alpha}(N_A)$ und $f_{\omega}(N_{\Omega})$) zwischen den Kompartimenten gekennzeichnet (dickere Pfeile) und haben damit einen Proliferationsvorteil, der für die Verdrängung der gesunden Hämatopoese sorgen wird.

werden mittels eines Faktors auf einen Verlauf deutlich unterhalb der Funktionen der unbeeinflussten leukämischen Stammzellen und auch der gesunden Zellen eingestellt. Damit werden die Ph^+ -Stammzellen stärker in das ruhende Kompartiment A gedrückt. Desweiteren ist der Einfluss des TKI durch einen zytotoxischen Effekt auf einen gewissen Anteil TKI-beeinflusster Ph^+ -Zellen im Kompartiment Ω gekennzeichnet. Im Modell und in Abbildung 5.8 wird dieser Parameter als r_{deg} bezeichnet. Es wird pro Zeitschritt ein der Größe dieses Parameters entsprechender Anteil leukämischer Zellen aus dem Kompartiment Ω entfernt. Durch die genannten Veränderungen sind die Zellen der gesunden Hämatopoese im Vorteil, können sich wieder ausbreiten und die leukämischen Zellen verdrängen.

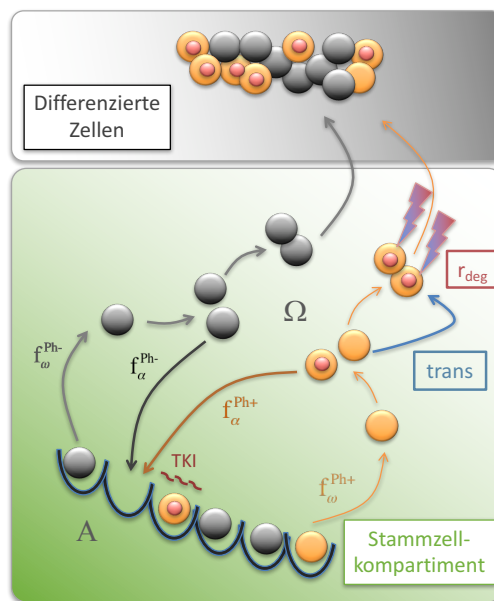


Abbildung 5.8.: Stammzellmodell bei der Simulation einer TKI-Therapie. Pro Simulationsschritt wird ein gewisser Anteil Ph^+ -Zellen im Kompartiment Ω in einen TKI-beeinflussten Zustand umgesetzt (*trans*; gekennzeichnet durch einen roten Punkt). Diese umgesetzten Zellen sind nun durch geringere $f_\alpha(N_A)$ und $f_\omega(N_\Omega)$ gekennzeichnet. Damit werden die Zellen stärker in das ruhende Kompartiment A gedrückt, teilen sich weniger und können sich einer Beeinflussung durch das TKI entziehen. Außerdem wird pro Zeitschritt ein bestimmter Anteil leukämischer Stammzellen im Kompartiment Ω entfernt (r_{deg}).

Die in den Untersuchungen konkret genutzte Konfiguration aller Modellparameter kann im Anhang (Appendix A.1) eingesehen werden.

Stammzellzahlen Während der Simulation der Entstehung und Therapie einer CML sind die Stammzellzahlen einfach zu analysieren. Dies ist als Vorteil zum realen Patienten zu

sehen, da dort die Analyse des Knochenmarks mit einer für den Patienten unangenehmen Prozedur der Knochenmarkpunktion verbunden ist. Außerdem kann dabei nicht das komplette blutbildende System, sondern nur ein kleiner Ausschnitt erfasst werden. Es ist vorstellbar, dass man wenige vorhandene leukämische Stammzellen (gerade im Verlauf der Langzeittherapie) mit einer Knochenmarkpunktion nicht findet. Bisher werden im Therapieverlauf nur bei Diagnosestellung und Progression Knochenmarkentnahmen vorgenommen (Hehlmann und Hochhaus 2010), sodass eine breite klinische Grundlage an Stammzellendaten nicht vorhanden ist.

5.3. Anpassungsmethode M0: Parameteridentifikation nach Horn et al. (2013)

In der Arbeit von Horn et al. (2013) wurden 1600 individuelle Kombinationen von r_{deg} und f_{ω} simuliert. Alle anderen Eingangsparameter der Simulation wurden konstant gehalten. Es erfolgte die Auswertung der resultierenden *in silico* Therapieverläufe nach der stückweise linearen Beschreibung. Zur Ermittlung der optimalen Startparameter wurde die rekursive Vorschrift 5.5 abgeleitet. Diese nutzt die Informationen über den Einfluss der Startparameter der Simulation auf die resultierenden Datenparameter der stückweise linearen Beschreibung. Über den umgekehrten Weg können aus den Datenparametern eines Patienten Modellparameter für seine Simulation vorhergesagt werden.

$$lm(\alpha, \beta \sim f_{\omega}, r_{deg})^{-1} \quad (5.5)$$

Auf dieser Basis konnten mithilfe einer nichtlinearen Regression die folgenden konkreten Beziehungen zwischen den Anstiegen des biphasischen Abfalls und den beiden variierten Startparametern aufgestellt werden (5.6 und 5.7). Dabei sind c_i die Koeffizienten der Regression (Anpassungsmethode *nls*, Programmiersprache *R* Version 3.2.3).

$$\ln(-\alpha) = c_1 \ln(r_{deg}) + c_2 f_{\omega} + c_3 \ln(r_{deg}) f_{\omega} + c_4 + \zeta \quad (5.6)$$

$$\ln(-\beta) = c_5 \ln(f_{\omega}) + c_6 r_{deg} + c_7 \ln(f_{\omega}) r_{deg} + c_8 + \zeta \quad (5.7)$$

ζ bezeichnet jeweils eine um 0 normalverteilte Zufallsgröße, die Messfehler und biologische Variabilität abbildet. Da mit den beiden Gleichungen 5.6 und 5.7 lediglich eine Vorhersage der Anstiege des Therapieverlaufs aus den Startparametern der Simulation möglich ist, mussten beide nach r_{deg} bzw. f_{ω} aufgelöst werden. Zur Lösung dieses numerischen Problems, wurde ein iterativer Ansatz (jeweils ohne ζ) von definierten Startwerten \hat{r}_{deg} und \hat{f}_{ω} formuliert. Dabei

erhält man die Gleichungen 5.8 und 5.9.

$$\hat{r}_{deg} = e^{\frac{\ln(-\alpha) - c_2 \hat{f}_\omega - c_4}{c_1 + c_3 \hat{f}_\omega}} \quad (5.8)$$

$$\hat{f}_\omega = e^{\frac{\ln(-\beta) - c_6 \hat{r}_{deg} - c_8}{c_5 + c_7 \hat{r}_{deg}}} \quad (5.9)$$

Ausgehend von $\hat{f}_\omega = 0,01$ werden anschließend mehrere Durchläufe der Gleichungen 5.8 und 5.9 gerechnet, deren Ergebnisse jeweils ineinander eingesetzt werden (Iteration). Damit ist die Vorhersage eines bestimmten Parameterpaares r_{deg} und f_ω möglich. Da hier eine Umformung der Gleichungen und Lösung mittels Iteration notwendig ist, wird das Verfahren im Rahmen der vorliegenden Arbeit als „indirekte“ Regression bezeichnet. Diese Formulierung steht im Gegensatz zum Verfahren der „direkten“ Regression, welches im Kapitel 6.5 erläutert wird. Mit den ermittelten Werten des Verfahrens nach Horn et al. (2013) wird eine erneute *in silico* Therapiesimulation gestartet. Die resultierenden Simulationen der Therapie für die genutzten Patienten der IRIS-Studie werden im Kapitel 6.7 mit den im Rahmen der vorliegenden Arbeit neu entwickelten Verfahren verglichen.

6. Ergebnisse

6.1. Alternativen in der Anpassung der biexponentiellen Beschreibung

Wie bereits oben beschrieben, können bei der Ermittlung der Tumorlast Messungenauigkeiten auftreten (siehe Kapitel 2.2.4). Diesen Messfehlern sollte, soweit sie bekannt sind, bei der Beschreibung der Patientendaten Rechnung getragen werden. Man könnte dazu die betroffenen Messwerte schlicht aus der Analyse ausschließen. Dies würde jedoch viel verwertbare Information aus den Therapieverläufen entfernen. Aus diesem Grund wurde eine alternative Möglichkeit der Anpassung der biexponentiellen Beschreibung gesucht, welche die Unsicherheit der möglicherweise fehlerbehafteten Datenpunkte einbezieht.

6.1.1. Nutzung des maximum likelihood-Verfahrens

Basierend auf der nichtlinearen Regression zur Anpassung der biexponentiellen Beschreibung (Abbildung 6.1a), die alle gemessenen Punkte als exakte Messwerte betrachtet, erfolgte eine Anpassung nach dem maximum likelihood-Ansatz (Abbildung 6.1b).

Dazu wurde ein Bereich definiert, in dem die Messungen mit Unsicherheit behaftet sind. Es ist jedoch sicher, dass die Messpunkte in diesem Bereich liegen. So flossen alle Messwerte zur Erstellung der Abbildung 6.1b, die oberhalb einer $\frac{BCR-ABL1}{ABL1}$ -Last von 10 % lagen, als Messungen ein, die in Realität an einem unbekanntem Punkt im Bereich oberhalb von 10 % und unterhalb von 100 % liegen. Das Verfahren bezeichnet man auch als Zensierung. Im Ergebnis gibt es eine fast gleiche Anpassung der biexponentiellen Funktion an die Therapiedaten des Patienten.

Dieses Verfahren funktioniert im Detail folgendermaßen: Jeder Therapieverlauf eines Patienten besteht aus einer zeitlichen Abfolge von Messungen des $\frac{BCR-ABL1}{ABL1}$ -Verhältnisses. Die Messpunkte sollen auch durch eine biexponentielle Anpassung beschrieben werden. Wie oben bereits erläutert (siehe Kapitel 3.1.3) beinhaltet die verwendete Funktion vier Parameter (α , β , A und B). Zusätzlich wird nun um die biexponentielle Funktion an jedem Zeitpunkt eines vorhandenen Messpunktes die Dichtefunktion einer Normalverteilung mit Maximum am konkreten Funktionswert angenommen. Die Breite dieser Dichtefunktion wird für alle

Messwerte gleich angenommen und durch den weiteren Parameter σ beschrieben. Für eine bestimmte Kombination der genannten fünf Parameter kann nun für jeden konkreten Messwert der Wert der Dichtefunktion an dieser Stelle ermittelt und für alle Messwerte aufsummiert werden. Das Ergebnis bezeichnet man als likelihood. Liegen die Messpunkte sehr nahe an der biexponentiellen Funktion, so ist die likelihood groß. Liegen die Messwerte eher weit entfernt, ergibt sich eine kleine likelihood. Werden die fünf Parameter der Funktion variiert, so verändert sich die likelihood entsprechend. Es kann die sogenannte likelihood-Funktion ermittelt werden, welche die likelihood in Abhängigkeit der Parameterkombination beschreibt. Man möchte den Parametersatz identifizieren für den die likelihood am größten ist, d. h. welche Parameterkombination am wahrscheinlichsten ist. Die Suche nach dem Maximum erfolgt mit einem Algorithmus nach Newton-Raphson. Für unsichere Messwerte wird nicht der konkrete Wert der Dichtefunktion in die Berechnung der likelihood einbezogen, sondern die Fläche unter der Dichtefunktion aus dem Bereich, in dem sich der zensierte Datenpunkt befindet. Im vorliegenden Beispiel demnach jeweils im Bereich von 10% bis 100% für alle $\frac{BCR-ABL1}{ABL1}$ -Verhältnisse oberhalb von 10% (nach Beal (2001), Casella und Berger (2002) und Henningsen und Toomet (2010)).

Für sehr kleine Messwerte (in der *RT-qPCR* nicht messbar, aber in der *nested PCR* positiv) wurde als konservative Annahme die Detektionsschwelle des ersten PCR-Verfahrens registriert (beschrieben im Kapitel 5.1, analog zur Vorarbeit von Horn et al. (2013)). Der Fall *nested PCR* negativ tritt in den verwendeten IRIS-Daten nur zu einem Zeitpunkt bei einem einzelnen Patienten auf. Dieser Messwert floss als 0% in die Betrachtungen ein. Zum darauf folgenden Messzeitpunkt diese Patienten wurde eine positive Tumorlast nachgewiesen.

Für die Anpassung der *in silico* simulierten Therapieverläufe ist eine Verwendung des maximum likelihood-Ansatzes nicht notwendig, da alle Zellzahlen exakt bekannt sind und damit jegliche $\frac{BCR-ABL1}{ABL1}$ -Verhältnisse ohne Unsicherheit angegeben werden können.

6.1.2. Vergleich der Anpassungsverfahren der biexponentiellen Beschreibung

Im Rahmen der vorliegenden Arbeit wurden zwei unterschiedliche Verfahren angewendet, um die biexponentielle Beschreibung auf Therapieverläufe von Patienten anzupassen: einerseits die Anpassung mittels nicht linearer Regression, in der alle Messpunkte ohne Messunsicherheit betrachtet werden (siehe Kapitel 3.1.3), und andererseits die Einbeziehung von zensierten Messbereichen und damit der Messunsicherheit mit Hilfe des maximum likelihood-Verfahrens (siehe oben). Anwendung fand diese Zensierung auf Werte oberhalb von 10% $\frac{BCR-ABL1}{ABL1}$, um damit unterschätzten Tumorlasten Rechnung zu tragen (Hanfstein et al. 2014). Um nun beide Verfahren miteinander zu vergleichen, wurden diese auf alle IRIS-Patienten des

- (a) Anpassung mittels nichtlinearer Regression
 (b) Anpassung unter Einbeziehung der Messunsicherheit oberhalb von 10% (maximum likelihood-Ansatz)

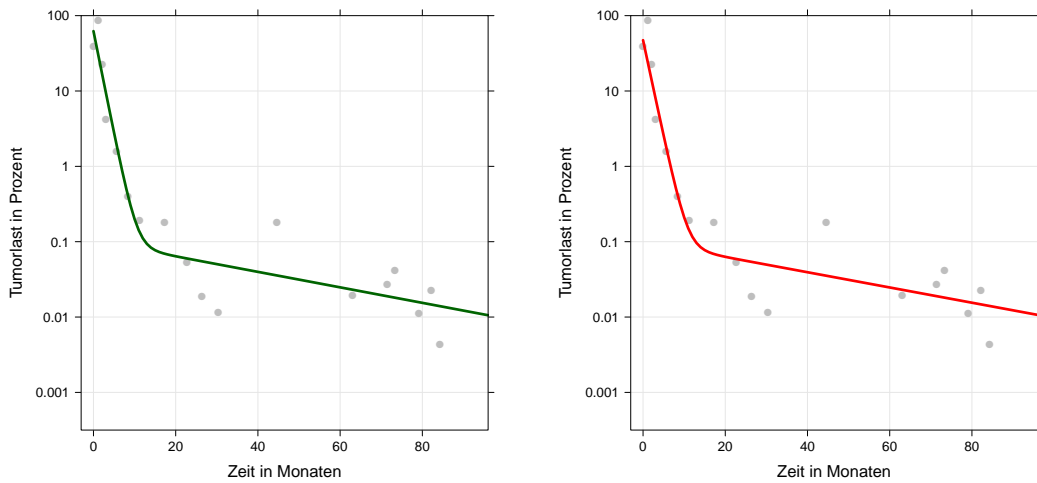


Abbildung 6.1.: Vergleich der Anpassungsverfahren der biexponentiellen Beschreibung. Beide Abbildungen verdeutlichen die unterschiedlichen Parameter der biexponentiellen Beschreibung bei der Verwendung unterschiedlicher Methoden der Anpassung. Beide Abbildungen basieren auf den selben Patientendaten. Für Abbildung 6.1a wurden alle Messwerte als exakt so gemessen angenommen. Wendet man zur Anpassung das maximum likelihood-Verfahren mit zensierten Datenpunkten an, um möglichen Unsicherheiten Rechnung zu tragen, so ergibt sich ähnliche Beschreibung des Therapieverlaufes (Abbildung 6.1b).

deutschen Armes ($n = 69$) angewendet und die resultierenden Parameter gegenübergestellt (siehe Abbildung 6.2).

Die Abbildungen zeigen eine annähernde Äquivalenz zwischen den Parametern der beiden unterschiedlichen Anpassungsverfahren. Die Abweichungen von der exakten Gleichheit können folgendermaßen erklärt werden: Die Parameter A und α , die den größten Einfluss auf die erste, steile Anpassung des Therapieverlaufs haben, zeigen größere Abweichungen, als die Parameter mit dem größeren Einfluss auf den Langzeitverlauf (B und β). Begründet wird dies durch den Umstand, dass bei der Anpassung nach der maximum likelihood-Methode die Messunsicherheit oberhalb von 10% in den Verlauf der angepassten biexponentiellen Beschreibung einfließt. Im Langzeitverlauf zeigen sich geringere Abweichungen, da durch die festgelegten Regeln im Umgang mit sehr kleinen Messwerten (siehe Kapitel 5.1) die Einbeziehung von Messunsicherheiten in der vorliegenden Arbeit nicht notwendig war. Eine Anpassung unter Berücksichtigung der Messunsicherheiten sehr kleiner Messwerte wurde bei der Analyse von Datensätzen nachfolgender Studien, welche so noch nicht publiziert sind, genutzt. Ob sich deutliche Abweichungen zur Anpassung ohne Zensierung ergeben, kann mit einer analogen Untersuchungen gezeigt werden.

6.2. Variation der Modellparameter und deren Einfluss auf die statistische Therapiebeschreibung

In der Arbeit von Horn et al. (2013) wurde der Einfluss einer Änderung der dynamischen Modellparameter r_{deg} und der Übergangsfunktion f_{ω} auf die statistischen Patientenparameter α (siehe Formel 3.1) und β (siehe Formel 3.3) untersucht (siehe Kapitel 5.3). Dabei wiesen einige resultierende Therapiesimulationen deutliche Abweichungen zu den realen Verläufen auf. In der vorliegenden Arbeit sollen zusätzlich zu den Variablen aus der Veröffentlichung von Horn et al. (2013) die Parameter der Übergangsfunktion f_{α} , Umsatzrate (*trans*) und initialer Pool (*init*) verändert und der Einfluss auf alle vier statistischen Patientenparameter, die aus der Auswertung mittels biexponentieller Beschreibung (siehe Kapitel 3.1) folgen, untersucht werden.

Um eine Aussage über den Einfluss von unterschiedlichen Eingangsparametern der Simulation (Modellparameter) auf die statistische Therapiebeschreibung zu erhalten, wurden Vorab-Simulationen durchgeführt. Dabei wurde jeweils nur ein Modellparameter verändert und die Auswirkung auf die Parameter der biexponentiellen Beschreibung untersucht. In Abbildung 6.3 ist auf den x-Achsen jeweils der veränderte Modellparameter und auf den y-Achsen die resultierenden statistischen Werte der biexponentiellen Beschreibung abgetragen. Folgende Zusammenhänge lassen sich ableiten:

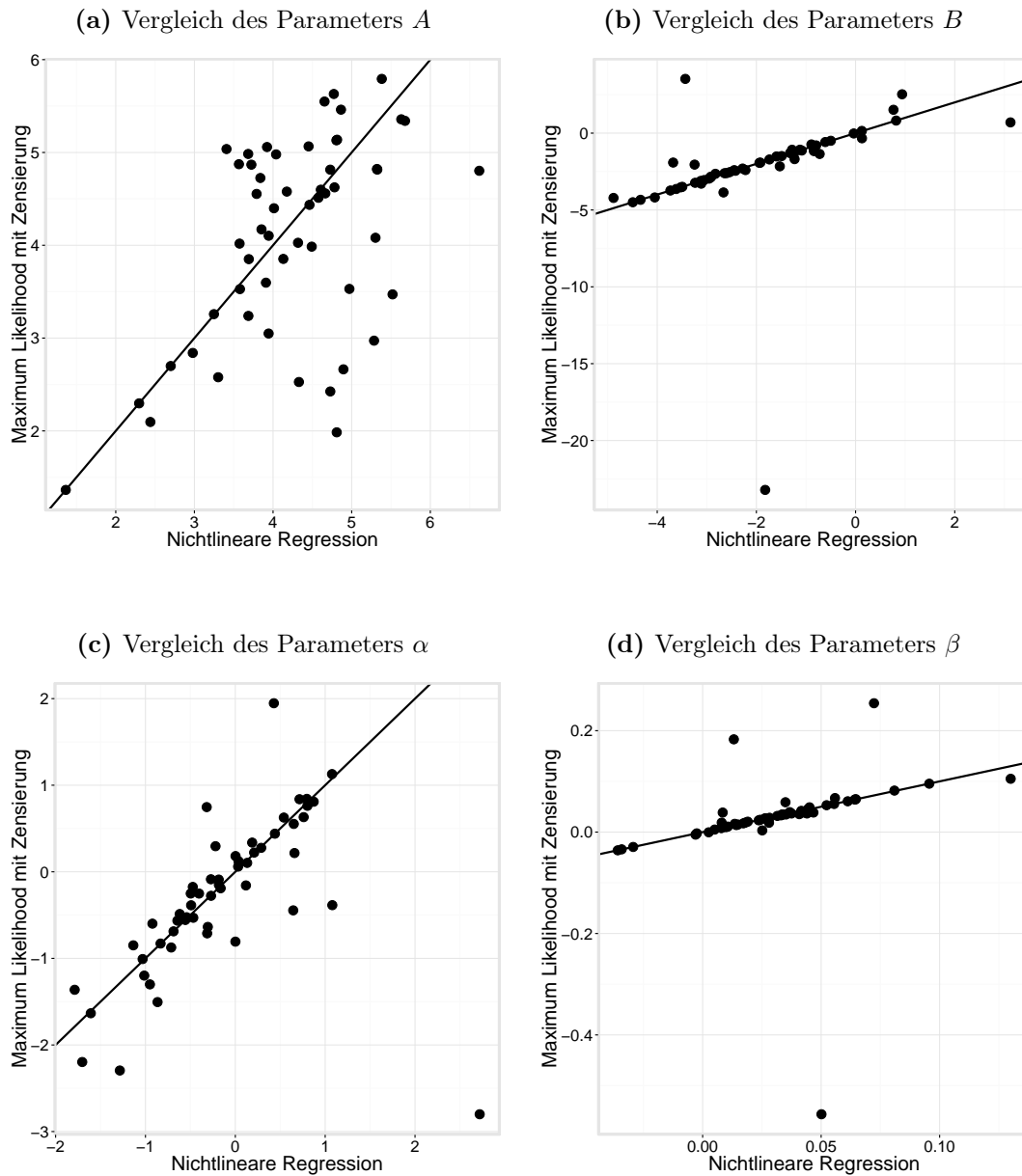


Abbildung 6.2.: Vergleich der Parameter der biexponentiellen Beschreibung aller IRIS-Patienten des deutschen Armes der Studie ($n = 69$) unter Nutzung unterschiedlicher Anpassungsmethoden (siehe Kapitel 3.1). An den Achsen sind dabei jeweils die in den Bildüberschriften genannten Parameter der biexponentiellen Beschreibung aufgetragen (für die x-Achsen ermittelt mittels nichtlinearer Regression ohne Einbeziehung zensierter Messbereiche, an den y-Achsen für den maximum likelihood-Ansatz, der Messunsicherheiten einbezieht). Die Abweichungen von der eingezeichneten Linie zeigen, dass die Parameter nicht exakt gleich sind. A und α zeigen die größten Abweichungen, da der erste steile Tumorlastabfall unter großem Einfluss der einbezogenen Messunsicherheit oberhalb von 10% steht.

- r_{deg} beeinflusst vor allem die initialen Werte der beiden Therapieabschnitte (A bzw. B)
- f_α und f_ω beeinflussen die Steilheit des langfristigen Therapieverlaufs (β)
- $trans$ beeinflusst die Parameter des ersten Therapieabschnitts (A und α)

Gerade der Einfluss von f_ω auf die Steilheit des zweiten Abfalls war bereits aus den Untersuchungen von Horn et al. (2013) bekannt, sodass dort eine größere Schrittweite für die Vorab-Untersuchungen gewählt werden konnte.

Die Parameter *Differenzierungsrate* (d), *Regenerationsrate* (r) und die Parameter, die den Zellzyklus beschreiben, werden in der in Tabelle A.1 genannten Konfiguration genutzt und nicht variiert. Sie sind Grundlage für ein stabiles System, welches ein realitätsnahes Verhalten zeigt (Horn et al. 2008). Exemplarische Untersuchungen haben gezeigt, dass eine Änderung dieser Parameter das simulierte System instabil machen. Nur in sehr engen Grenzen ist eine Variation möglich, ohne dass sich dabei eine systematische Änderung der resultierenden Simulationen zeigen würde.

6.3. Konstruktion eines Simulationspools als Grundlage für weitere Untersuchungen

6.3.1. Eingrenzung des Parameterraums

Die Therapieverläufe der Patienten, zu denen Aussagen getroffen werden sollen, müssen sich im Datenbestand der Therapiesimulationen wiederfinden. Damit benötigt man eine große Grundlage an *in silico* Therapieverläufen. Je größer die Datengrundlage ist, umso exakter stimmen die identifizierten Simulationen mit dem realen Patientenverlauf überein. Um den Parameterraum einzugrenzen, der diesen Datenbestand definiert, wurde zuerst eine Auswertung der Therapieverläufe der Patienten der IRIS-Studie (siehe Kapitel 5.1.3) vorgenommen. Anhand dieser wurden extreme Therapieverläufe identifiziert (z. B. besonders steile oder flache Verläufe in den beiden Therapieabschnitten) und diese anschließend nach einer manuellen Suche nach Startparametern für das Modell simuliert. Die dabei entstandenen Erfahrungen dienten als Grundlage für die manuelle Festlegung des Parameterraumes (Tab. 6.1). Dabei ist als obere Grenze des *initialen Pools* ein Wert von 99% angegeben, da ein System mit einer Tumorlast von 100% keine gesunden Zellen mehr enthalten würde und damit eine Therapiesimulation nicht möglich ist. Es ergibt sich ein fünfdimensionaler Raum der Modellparameter, der anhand der in der Tabelle 6.1 genannten Schrittweiten aus 270.400 individuellen Parameterkombinationen besteht.

Anschließend wurden alle Kombinationen der oberen und unteren Grenzen der unterschiedlichen Parameter („Ecken“) in einer Pilotstudie simuliert und nach der biexponentiellen

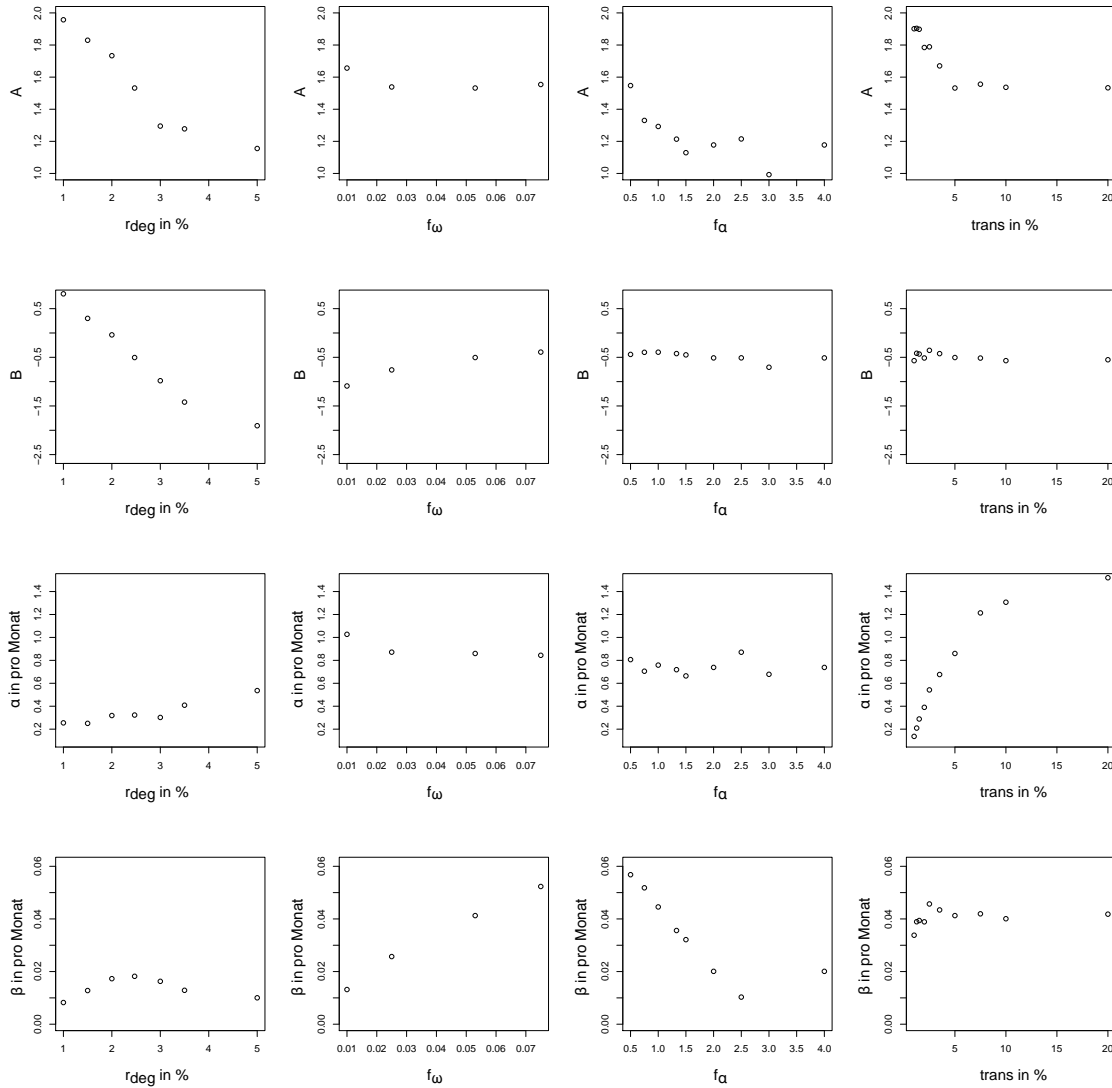


Abbildung 6.3.: Vorab-Untersuchungen zur Abschätzung des Einflusses unterschiedlicher dynamischer Simulationsparameter auf die statistischen Parameter. Auf den x-Achsen ist jeweils der veränderte Modellparameter und auf den y-Achsen die resultierenden statistischen Parameter abgetragen. r_{deg} hat vor allem Einfluss auf die initialen Werte der Therapieabschnitte (A bzw. B). A wird genauso wie α stark von der Umsatzrate $trans$ beeinflusst. Starken Einfluss auf die Steilheit des zweiten Abfalls haben die beiden Umsatzfunktionen f_α und f_ω .

Tabelle 6.1.: Grenzen des Parameterraumes der Vorab-Simulationen.

Parameter	untere Grenze	obere Grenze	Schrittweite	Anzahl der Schritte
<i>initialer Pool</i>	10 % $\frac{BCR-ABL1}{ABL1}$	99 % $\frac{BCR-ABL1}{ABL1}$	10 % $\frac{BCR-ABL1}{ABL1}$	10
<i>trans</i>	2 %	6,8 %	0,4 %	13
<i>r_{deg}</i>	1,5 %	7,5 %	0,4 %	16
<i>f_ω</i>	0,01	0,1	0,01	10
<i>f_α</i>	0,5	0,98	0,04	13

Beschreibung ausgewertet. Eine ausreichende Größe des Parameterraumes wird deutlich, wenn man nun die statistischen Datenparameter der IRIS-Patienten (gewonnen aus den biexponentiellen Beschreibungen der Therapieverläufe) neben denen der „Ecken“ des Parameterraums (Kombinationen der oberen und unteren Grenzen der einzelnen Modellparameter) darstellt (Abbildung 6.4). Dabei wurde jeweils in schwarz die Verteilung der Datenparameter der Patienten der IRIS-Studie ($n = 69$, ohne Anwendung der Selektionskriterien), in grau die Verteilung der analogen Parameter aus der Pilotstudie aufgetragen. So decken die Datenparameter der simulierten Ecken des Parameterraums jeweils die analogen Parameter der Patienten aus der IRIS-Studie in wesentlichen Bereichen ab. Die sichtbaren Abweichungen ergeben sich unter anderem durch die kleinere Datengrundlage der Pilotstudie im Vergleich zur IRIS-Studie. Besonders die deutlichen Abweichungen des Parameters A kommen dadurch zustande, dass am Patienten methodenbedingt Messwerte mit einem $\frac{BCR-ABL1}{ABL1}$ -Verhältnis von über 100% möglich sind (siehe Kapitel 2.2.4). Dies kann in der Simulation nicht auftreten, da die Tumorlast dort einzelzellbasiert exakt ausgerechnet werden kann (im realen Patienten erfolgt die Messung einer Transkriptzahl in der PCR). Die Hypothese, ob die Daten des kompletten Parameterraumes zu den Daten der IRIS-Studie passen, wird im Kapitel 6.3.3 überprüft.

6.3.2. Simulation des kompletten Modellparameterraumes

Um statistische Einflüsse auf den Simulationsverlauf zu minimieren, wurden die oben beschriebenen 270.400 individuellen Parameterkombinationen jeweils fünf Mal simuliert. Somit ergaben sich 1.352.000 Simulationsläufe. Ein ähnliches Verfahren wurde auch schon von Horn et al. (2013) verwendet, wobei dort nur 1600 Parameterkombinationen und damit 8000 Durchläufe zu simulieren waren. Diese extrem große Anzahl von Simulationen erforderte eine Parallelisierung, um den Zeitaufwand zu minimieren. Durch Pilotsimulationen auf einem Rechencluster des Zentrums für Informationsdienste und Hochleistungsrechnen (ZIH) der

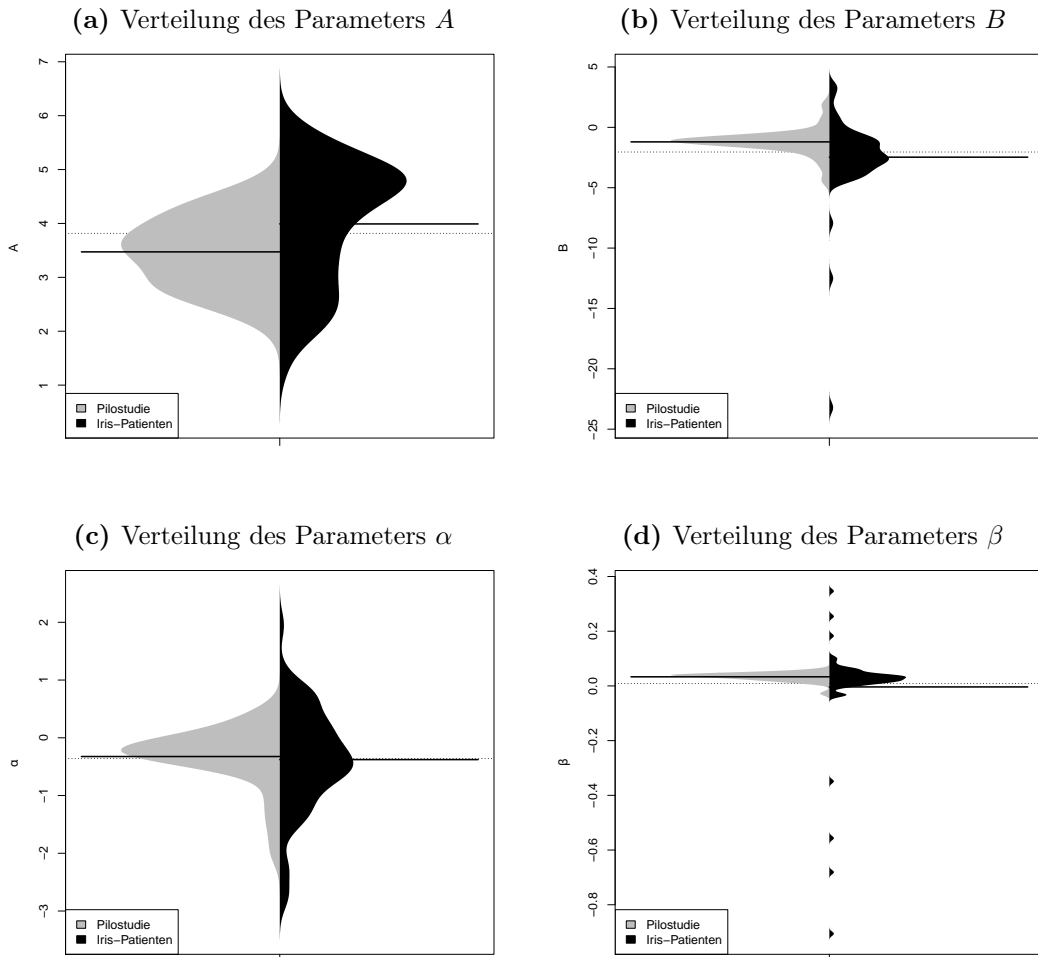


Abbildung 6.4.: Vergleich der Datenparameter (jeweils in der Bildüberschrift genannt) der Pilotsimulationen (Ecken des Parameterraumes) mit denen der IRIS-Patienten. Es zeigt sich eine weitgehende Übereinstimmung zwischen den Parametern der biexponentiellen Beschreibungen. Man kann davon ausgehen, dass der resultierende Parameterraum die IRIS-Patienten abdecken wird. Zur Erstellung der Abbildung wurden alle 69 Patientenverläufe aus dem Deutschen Arm der Studie verwendet, ohne die in Kapitel 5.1 genannten Selektionskriterien anzuwenden.

TU Dresden (Megware PC-Farm (*Atlas*)) konnte eine Simulationsdauer von ca. 2 CPU-Stunden und eine genutzte Festplattenkapazität von ca. 250 KB pro Simulationsdurchlauf ermittelt werden. Hochgerechnet auf alle Simulationen bedeutet dies 3 Mio. CPU-Stunden und 500 GB Festplattenspeicher. Die vollständigen Simulationen konnten durch die großzügige Unterstützung des ZIH auf den 5888 Rechenkernen der Megware PC-Farm (*Atlas*) innerhalb eines Monats durchgeführt werden. Ein herkömmlicher Computer mit nur einer CPU wäre mit dieser Aufgabe annähernd 340 Jahre beschäftigt gewesen.

6.3.3. Statistische Auswertung der Simulationen

Die 1.352.000 Simulationsläufe wurden für jede individuelle Parameterkombination gemittelt. Dazu erfolgte die Berechnung des arithmetischen Mittels der simulierten Tumorlast an jedem Zeitpunkt der jeweils fünf Simulationen, die mit dem gleichen Modellparametersatz gestartet wurden. Somit ergab sich insgesamt eine Zahl von 270.400 simulierten Therapien. Diese Daten werden im Folgenden auch als Vorsimulationen bezeichnet. Jede individuelle Kombination wird weiterhin durch Θ_j gekennzeichnet, wobei j eine natürliche Zahl zwischen 1 und 270.400 ist. Anschließend erfolge eine Anwendung der biexponentiellen Anpassung des Therapieverlaufes, um die Datenparameter der Therapiesimulationen zu erhalten (siehe Kapitel 3.1). Für jeden Datenparameter (A , B , α bzw. β) ergibt sich in Abhängigkeit von allen dynamischen Parametern jeweils eine sechs-dimensionale Landschaft (bestehend aus jeweils einem Datenparameter und den fünf variierten Eingangsparametern der Simulation).

Beschreibung der Parameterlandschaft

Um herauszufinden, ob die Vorsimulationen geeignet sind, als Grundlage für die weiteren Verfahren der Modellparameteridentifikation zu sein, muss zuerst die Form der Parameterlandschaft ausgewertet werden. Da die Darstellung sechs-dimensionaler Räume (jeweils ein Datenparameter A , B , α oder β in Beziehung zu den fünf Eingangsparametern der Simulation) sehr komplex ist, wurden die einzelnen Parameterdimensionen des Parameterraumes mit Hilfe einer bidirektionalen Übertragung in drei Dimensionen überführt. Es entstand die Abbildung 6.5. Die Abbildungen für die weiteren Datenparameter sind im Anhang A.2 abgedruckt.

Beispielhaft sei der Vorgang der Transformation der Dimensionen anhand eines vierdimensionalen Parameterraumes beschrieben. Diesen kann man sich als Würfel vorstellen, wobei Breite, Tiefe und Höhe jeweils einen Parameter abbilden. Der vierte Parameter sei durch eine Farbe charakterisiert. Schneidet man diesen Würfel für jeden Wert des Parameters der senkrechten Achse in Scheiben und klappt diese jeweils nebeneinander in eine Ebene, so erhält man ein zweidimensionales Koordinatensystem. Eine Achse enthält dann nur einen Parameter, während die zweite Achse zwei Parameter darstellt. Die vierte Dimension wird

nach wie vor durch die Farbe charakterisiert.

In den vorliegenden Abbildungen enthält die waagerechte Achse drei Eingangsparameter der Simulation (*trans*, *initialerPool* und f_ω) und die senkrechte Achse die zwei übrigen Parameter (r_{deg} und $f\alpha$). Die Farbe charakterisiert jeweils den Parameter der biexponentiellen Beschreibung, der in der Überschrift der Abbildung genannt ist. Insgesamt sind die Abbildungen so gestaltet, dass sich zwischen zwei benachbarten Simulationen Θ_x und Θ_y jeweils nur ein Modellparameter um exakt einen diskreten Schritt ändert.

Zur besseren Veranschaulichung dient Abbildung 6.6, welche einen Ausschnitt aus Abbildung A.1 darstellt. Diejenigen Modellparameter, die in diesem Ausschnitt konstant sind, werden unter dem Diagramm aufgeführt, während die sich ändernden Parameter an den Achsen abgetragen sind. Auch hier ändert sich zwischen zwei benachbarten Simulationen Θ_x und Θ_y nur genau ein Modellparameter um einen diskreten Schritt.

Ziel bei der Erstellung der Abbildungen sollte sein, größere Sprünge in den Datenparametern optisch leicht erkennbar zu machen. Diese wären als Hinweis auf große Änderungen resultierenden Parameter der biexponentiellen Beschreibung bei kleinen Schritten der Modellparameter zu werten. Solche Sprungstellen wären von den gewählten Modellparameterkombinationen nicht abgedeckt und eine Aussage über das Verhalten der Simulation in solchen Bereichen wäre nicht möglich. Kleine Abstufungen sind unvermeidbar, da die Untersuchung des Parameterraums an diskreten Punkten und nicht durch kontinuierliche Veränderung der Modellparameter erfolgte.

Zusammenfassend kann festgestellt werden, dass größere Sprünge oder Ausreißer nicht erkennbar sind. Die gewählten Schrittweiten für die Startparameter sind demnach ausreichend klein und es ist festzuhalten, dass die Datengrundlage groß genug ist, um die Vorhersagefähigkeit der entwickelten Verfahren zu ermöglichen.

Validierung des Parameterraumes

Analog zum Vergleich zwischen den Pilotsimulationen zur Festlegung der Ecken des Parameterraumes mit den Datenparametern der IRIS-Patienten (siehe Kapitel 6.3.1) wurde auch die Datenparameter der Simulationen des komplette Parameterraum nochmals mit den Datenparametern aller IRIS-Patienten ($n = 69$) verglichen. Damit soll überprüft werden, ob der abgedeckte Parameterraum weiterhin den Raum Therapieverläufe der IRIS-Patienten abdeckt (siehe Abbildung 6.7).

Vergleicht man die Abbildung 6.7 und 6.4 fällt auf, dass vor allem bezogen auf den Parameter B der biexponentiellen Beschreibung die Abdeckung besser ist, als die Pilotstudien suggeriert hatten. Auch die anderen Parameter zeigen eine ausreichende Abdeckung. Patienten, die nicht vom kompletten Datensatz umfasst werden, zeigen entweder einen Wiederanstieg im

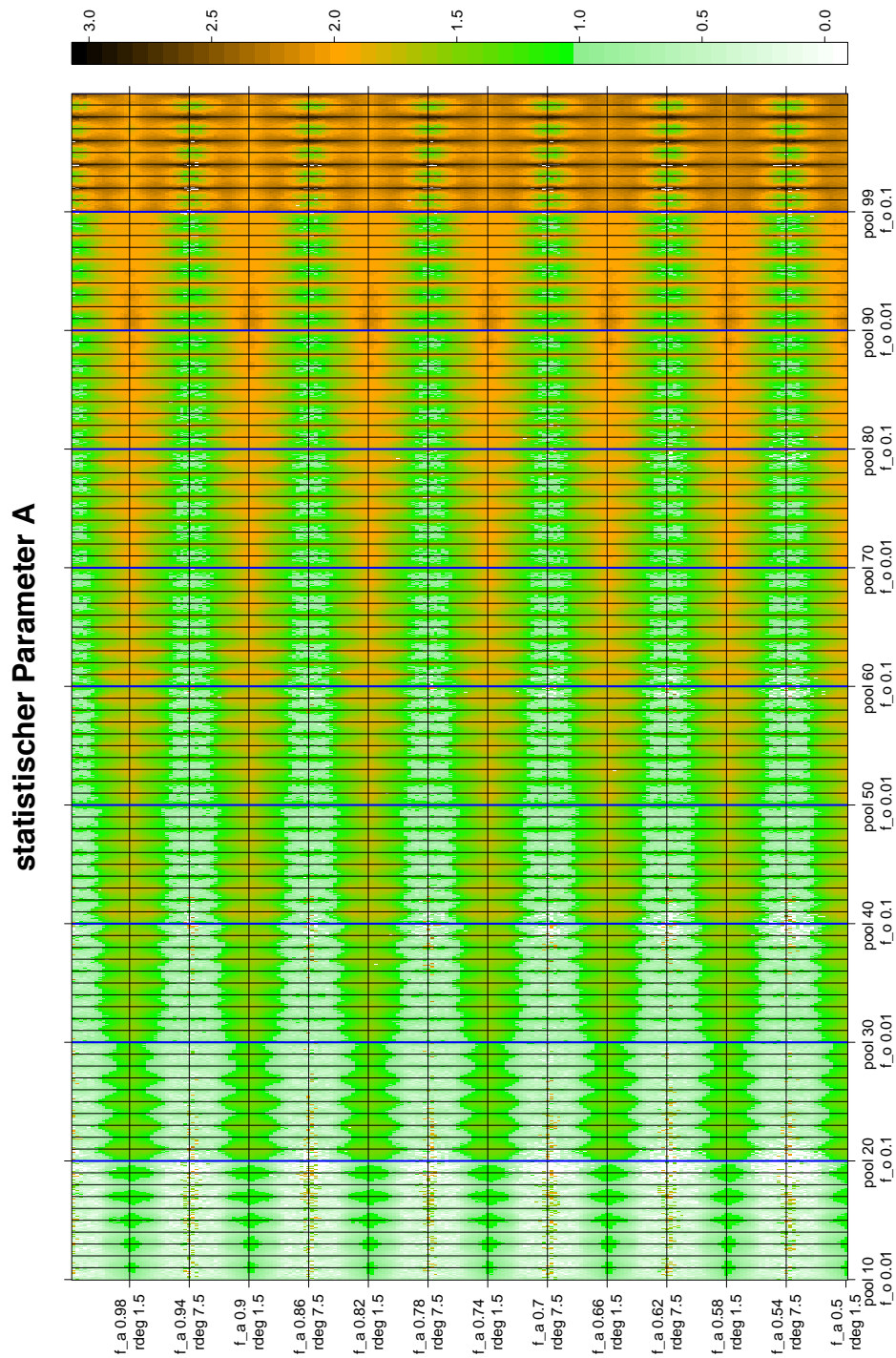


Abbildung 6.5.: Darstellung der Parameterlandschaft für den statistischen Parameter A. Farbkodiert dazu jeweils der Wert für A und an den beiden Achsen die sich ändernden Modellparameter der Simulationen. Zwischen zwei Feldern der Darstellung ändert sich jeweils genau ein Modellparameter um einen diskreten Schritt. Es besteht somit zwischen zwei benachbarten Simulationen Θ_x und Θ_y ein Unterschied von genau einem Schritt in einem Eingangsparameter der Simulation. Große Sprünge oder Ausreißer sind nicht zu erkennen. Die sichtbaren Absätze kommen durch die nicht-kontinuierliche Veränderung der Modellparameter zustande.

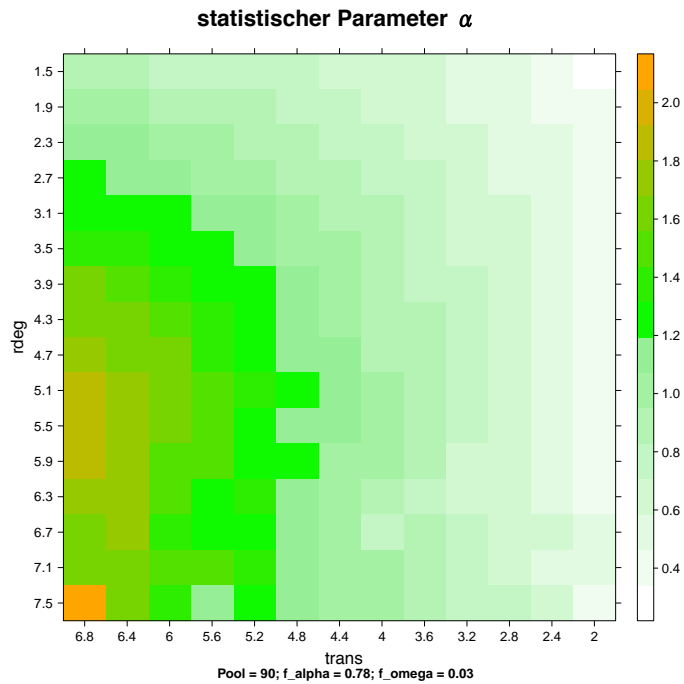


Abbildung 6.6.: Ausschnitt des linken unteren Feldes der Abbildung A.1 zur besseren Veranschaulichung. Farbkodiert hier der Wert von α . Die konstanten Startparameter sind unter dem Koordinatensystem genannt. Zwischen zwei Kästchen der Abbildung ändert sich der an der Achse genannte Modellparameter um genau einen diskreten Schritt. Auch hier gilt wieder die bereits in Abbildung 6.5 genannte Schrittweite von einer Änderung eines Modellparameters zwischen zwei benachbarten Simulationen Θ_x und Θ_y . Es zeigt sich, dass bei ansteigendem r_{deg} bzw. $trans$ der Wert für α immer größer wird. Das heißt, dass der erste Therapieabschnitt steiler verläuft. Erklärung dafür ist, dass bei einem größeren Wert für $trans$ pro Zeitschritt mehr maligne Zellen vom TKI beeinflusst werden. Ein größeres r_{deg} entfernt pro Zeitschritt einen größeren Anteil an Tumorzellen aus dem System. Beides führt zu einer schnell abfallenden Tumorlast im peripheren Blut.

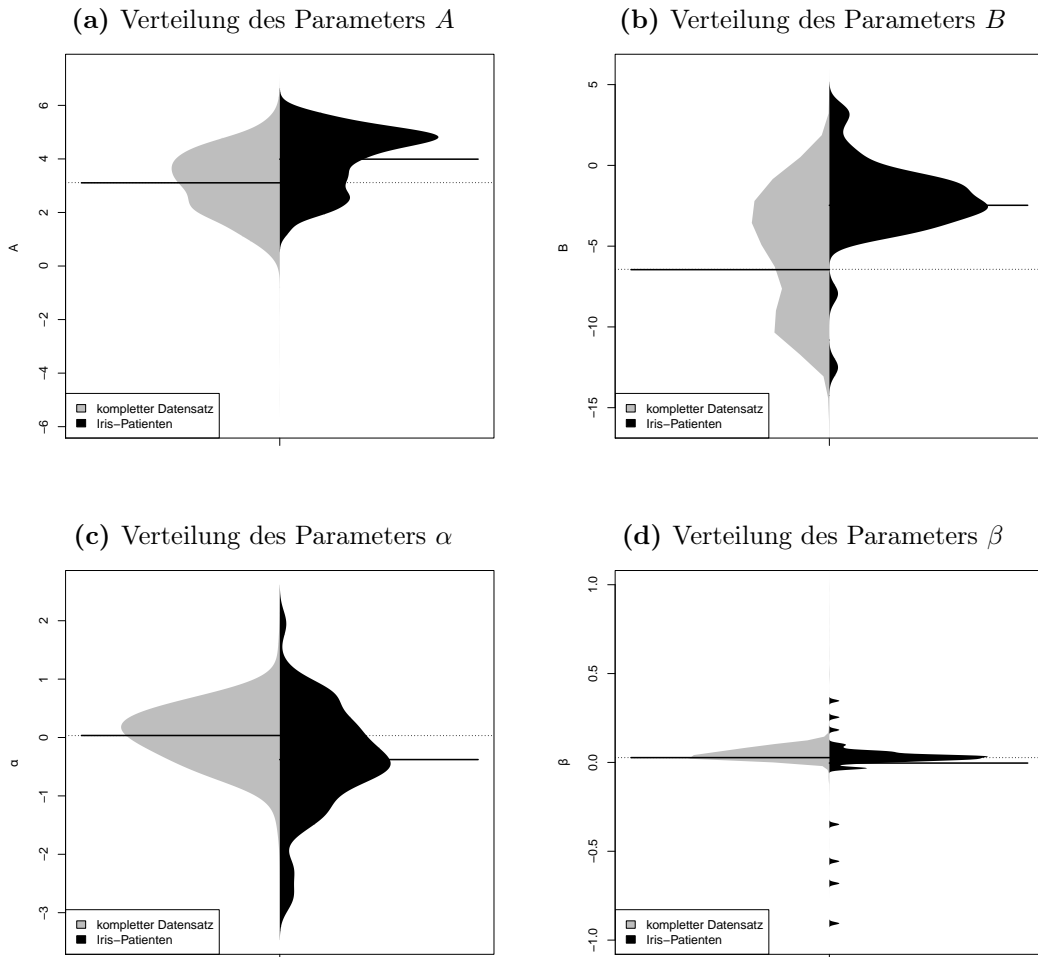


Abbildung 6.7.: Vergleich der Datenparameter nach biexponentieller Auswertung aller Simulationen des Parameterraumes mit den Datenparametern der IRIS-Patienten. Die Patienten, deren Werte in Abbildung 6.7d deutlich entfernt von den anderen Werten liegen, sind entweder solche mit Wiederanstieg unter Therapie (negatives β , werden als Therapieversager identifiziert) oder haben einen monophasischen Verlauf bei dem α und β ähnlich sind. Beide Fälle sollen von den beschriebenen Methoden nicht abgedeckt werden. Sind sie nicht im Parameterraum enthalten, gehen keine Informationen verloren.

langfristigen Therapieverlauf (negatives β), oder haben einen monophasischen Verlauf (α und β gleich bzw. ähnlich). Beide Varianten sollen von den beschriebenen Methoden nicht erfasst werden und schmälern damit nicht die Qualität der Datengrundlage.

6.4. Entwicklung eines Gütemaßes zum Vergleich unterschiedlicher Therapieverläufe

Um unterschiedliche Therapiesimulationen untereinander und mit dem zugrunde liegenden realen Therapieverlauf eines Patienten vergleichen zu können, ist ein Vergleichsmaß notwendig. Im Rahmen der vorliegenden Arbeit wurde ein solches entwickelt und angewendet. Es basiert auf Abstandsvergleichen zwischen den unterschiedlichen biexponentiellen Beschreibungen der Therapieverläufe. Es muss ein Vergleichszeitraum definiert werden. Für jeden Zeitpunkt i dieses Vergleichszeitraumes (es wurden jeweils monatliche Zeitpunkte gewählt) kann nun der absolute Abstand d zwischen den Werten verschiedener biexponentieller Beschreibungen f_P bzw. f_Θ der beiden Therapieverläufe (P und Θ) ermittelt werden. Diese Abstände werden für alle Zeitpunkte aufaddiert und anschließend auf die Anzahl der Vergleichszeitpunkte (i_{max}) normiert. Formalisiert ergibt sich folgende Vorschrift:

$$\overline{\Delta(P, \Theta)} = \frac{\sum_{i=1}^{i_{max}} d(P, \Theta)_i}{i_{max}} = \frac{\sum_{i=1}^{i_{max}} |f_{P_i} - f_{\Theta_i}|}{i_{max}} \quad (6.1)$$

Es gilt zu beachten, dass das errechnete Abstandsmaß auf der log-Skala angegeben ist, da die Darstellung des Therapieverlaufs und auch die biexponentielle Anpassung jeweils logarithmiert angegeben ist. Es ist dabei nicht unbedingt notwendig, dass ein Patientenverlauf P mit einer Simulation Θ verglichen wird. Ein Vergleich zwischen zwei unterschiedlichen Patientenverläufen oder auch von Simulationen untereinander ist problemlos möglich.

6.5. Anpassungsmethode M1: Berechnung der Modellparameter mittels „direkter“ Regression

Das Regressionsverfahren nach Horn et al. (2013) nutzt die Beziehung zwischen den Startparametern der Simulation und den resultierenden Datenparametern der biexponentiellen Beschreibung als Grundlage. Da in der praktischen Anwendung aber der umgekehrte Weg (aus den Datenparametern sollen die Startparameter der Simulation ermittelt werden) notwendig ist, erfolgt eine Transformierung der Regression und die Anwendung eines Iterationsverfahrens. Aus diesem Grund wird das Verfahren nach Horn et al. (2013) als „indirekte“ Regression bezeichnet. Im Gegensatz dazu erfolgt hier die Entwicklung eines Regressionsverfahrens,

welches die Nutzung der Iteration umgeht. Dies führt im Rahmen der vorliegenden Arbeit zur Bezeichnung „direkt“. Zur Berechnung der Modellparameter mittels dieser direkten Regression werden die 270.400 simulierten Therapieverläufe nach dem biexponentiellen Ansatz ausgewertet. Anschließend wurden die notwendigen Beziehungen einer nichtlinearen Regression aus der mehrdimensionalen Parameterlandschaft (siehe Kapitel 6.3.3) ermittelt. Dieses Vorgehen folgt der Beziehung

$$lm(f_\omega, f_\alpha, r_{deg}, trans, init \sim A, B, \alpha, \beta) \quad (6.2)$$

In der konkreten Anwendung der genannten Vorschrift 6.2 entstehen folgende Gleichungen mit den Koeffizienten c_i (Anpassungsmethode *nls*, Programmiersprache *R* Version 3.2.3):

$$trans = e^{c_1 \cdot A + c_2 \cdot \alpha + c_3 \cdot B + c_4 \cdot \beta} \quad (6.3)$$

$$r_{deg} = e^{c_5 \cdot A + c_6 \cdot \alpha + c_7 \cdot B + c_8 \cdot \beta} \quad (6.4)$$

$$f_\alpha = e^{c_9 \cdot A + c_{10} \cdot \alpha + c_{11} \cdot B + c_{12} \cdot \beta} \quad (6.5)$$

$$f_\omega = e^{c_{13} \cdot A + c_{14} \cdot \alpha + c_{15} \cdot B + c_{16} \cdot \beta} \quad (6.6)$$

$$pool = e^{c_{17} \cdot A + c_{18} \cdot \alpha + c_{19} \cdot B + c_{20} \cdot \beta} \quad (6.7)$$

Die Nutzung der Exponentialfunktion in Gleichung 6.3 bis 6.7 hat wie bei Horn et al. (2013) den Vorteil, negative Werte für die Startparameter der Simulation zu verhindern. Außerdem wurden keine Interaktionen oder Funktionen höherer Ordnung erlaubt, da diese die Gleichungen bei der großen Anzahl an Koeffizienten deutlich verkomplizieren. Alle erhaltenen Koeffizienten c_i der Vorschriften 6.3 bis 6.7 erlauben nun durch Einsetzen der statistischen Parameter des Therapieverlaufs eines Patienten die Berechnung des Modellparametersatzes für den Start der Simulationen.

6.6. Anpassungsmethode M2: Parametervorhersage anhand abstandsminimierter Vergleiche von Therapieverläufen (lookup-table)

Die Grundlage für die Auswahl des am besten passenden vorsimulierten Therapieverlaufs basiert wieder auf der Auswertung der *in silico* Therapievorsimulationen und des Therapieverlaufs des Patienten nach dem biexponentiellen Ansatz. Damit erhält man wie im Kapitel 3.1 beschrieben eine geglättete Beschreibung des Verlaufs der $\frac{BCR-ABL1}{ABL1}$ -Verhältnisse unter realer oder simulierter Therapie. Im Unterschied zu den beiden oben aufgeführten Verfahren, die eine Regression zur Parameteridentifikation nutzen, wird hier direkt aus dem Bestand der vorsimulierten Therapieverläufe ausgewählt.

Diese Auswahl erfolgt anhand einer Rangliste, in die alle 270.400 Therapiesimulationen einsortiert werden. Eine individuelle Kombination an Eingangsparametern für die Simulation sei für die folgende Formalisierung mit Θ bezeichnet. Die Rangliste wird im Rahmen der vorliegenden Arbeit als „lookup-table“ bezeichnet. Die Rangfolge wird durch den Abstand zwischen der biexponentiellen Anpassung des Patientenverlaufes und der Therapiesimulationen bestimmt. Zur Berechnung des Abstandes wird die gleiche Formel, die auch zur Ermittlung des Gütemaßes Anwendung findet, genutzt (Formel 6.1). Theoretisch wäre auch ein quadratischer oder eine andersartig definierte Differenz möglich gewesen. Es fiel die Entscheidung für den absoluten Abstand, um kleine oder große Werte nicht besonders zu werten. Als beste Simulation Θ^* wird diejenige mit der geringsten Differenz zum Patientenverlauf identifiziert. Formal wird folgende Vorschrift genutzt, die auf Formel 6.1 aufbaut und eine mittlere Differenz zwischen zwei verschiedenen biexponentiellen Beschreibungen berechnet:

$$\Theta^* = \min(\overline{\Delta(P, \Theta)}) = \min\left(\frac{\sum_{i=1}^{i_{max}} |f_{P_i} - f_{\Theta_i}|}{i_{max}}\right) \quad (6.8)$$

Ein Beispiel für eine solche Rangliste ist in Abbildung 6.8 abgedruckt.

Zwar können mit dem eben beschriebenen Verfahren nur die diskreten Schritte der Modellparameter genutzt werden, jedoch ist keine erneute Simulation des Patienten mit den durch Regression ermittelten Modellparametern notwendig. Selbst auf leistungsfähiger Hardware benötigt man für diesen Schritt der konkreten Patientensimulation ca. 10 CPU-Stunden (was bei Parallelsimulation der jeweils fünf Simulationsläufe zur Minimierung der statistischen Effekte einer Rechenzeit von 2 Stunden entspricht). Da der Bestand an Vorsimulationen jedoch sehr umfangreich ist und den Wertebereich aller betrachteten Patienten sehr gut abdeckt (siehe Kapitel 6.3.1 und 6.3.3), kann der Nachteil begrenzter Parameterkombinationen entkräftet werden.

Rang	Distanz	A	a	B	b	pool	transrate	rdeg	f_alpha	f_omega
Θ^* 1	0.8420774	4.682827	-1.1235729	0.69323794	0.07820963	90	2.0	1.9	0.50	0.09
2	1.0411667	4.646896	-1.1291675	0.81738619	0.07914418	90	2.0	1.9	0.58	0.10
3	1.1961300	4.676218	-1.1194011	0.85698280	0.08327692	90	2.0	1.9	0.54	0.10
4	1.4606732	5.189041	-1.0885082	0.76289823	0.08043196	99	2.8	2.3	0.70	0.10
5	1.4676458	5.223245	-1.0738732	0.59458099	0.07771111	99	2.8	2.3	0.62	0.09
6	1.4881109	4.997822	-0.9702667	0.64516736	0.07777007	99	3.6	2.3	0.74	0.10
7	1.5099214	5.009830	-1.0071478	1.19762793	0.08550934	99	4.4	1.9	0.54	0.10
8	1.5169844	5.259719	-1.0542638	0.50000122	0.07568306	99	2.8	2.3	0.54	0.08
9	1.5173683	5.042035	-1.0471487	1.14472565	0.08366570	99	4.0	1.9	0.54	0.10
10	1.5180981	4.979970	-0.9865166	1.18069101	0.08540528	99	4.8	1.9	0.54	0.10
11	1.5719880	5.172185	-1.0922885	0.72639219	0.07759148	99	2.8	2.3	0.74	0.10
12	1.5774386	5.071945	-1.0271765	0.61922149	0.07579473	99	3.2	2.3	0.74	0.10
13	1.5890074	4.946853	-0.9598673	1.15711628	0.08533464	99	5.2	1.9	0.54	0.10
14	1.5992920	5.123150	-1.0080822	0.71389321	0.08037966	99	3.2	2.3	0.70	0.10

Abbildung 6.8.: Beispielhafte Rangliste der Simulationen Θ_j zur Ermittlung der Simulation mit dem geringsten Abstand zum Therapieverlauf eines Patienten. Die Sortierung der Rangfolge erfolgt dabei nach ansteigender Distanz. Die Simulation mit dem geringsten Abstand wird als die am besten passende Simulation identifiziert und als Θ^* bezeichnet. Weiterhin sind die Daten- und Modellparameter der jeweiligen Therapiesimulation notiert.

Man hat die Möglichkeit, die Anpassung auf verschiedene Zeitfenster zu optimieren, indem man die Vergleichspunkte nur aus diesen auswählt. So könnten mittels der lookup-table Simulationen identifiziert werden, deren Verlauf besonders gut auf den ersten Teil des Therapieverlaufs des Patienten passen. Analog wäre eine Anpassung auf den hinteren Langzeitverlauf der Therapie möglich. Dazu müssen aber auch Messpunkte aus diesem Teil der Therapie vorhanden sein. In der vorliegenden Arbeit wurde Wert auf eine besonders gute Anpassung des gesamten Therapieverlaufs der Patienten gelegt, um ausgehend von dieser Vorhersagen treffen zu können. Die resultierenden Verläufe der Tumorlast im Stammzellkompartiment werden im Kapitel 6.8 analysiert.

6.7. Vergleich der verschiedenen patientenindividuellen Anpassungsverfahren

Wie oben beschrieben wurden im Rahmen dieser Dissertation zwei neue Verfahren für die Identifikation eines optimalen Startparameterdatensatzes für die Modellsimulationen vorgestellt (M1 und M2). Diese erweitern das von Horn et al. (2013) vorgestellte Verfahren (M0). An dieser Stelle soll der Vergleich der unterschiedlichen Verfahren erfolgen. Abbildung

6.9a zeigt dazu einen Beispielpatienten der IRIS-Studie im Vergleich zu den jeweiligen *in silico* Therapiesimulationen, die unter Verwendung der beschriebenen Identifikationsverfahren ermittelt wurden. Man erkennt deutlich, dass die Verwendung der lookup-table (M2) die am besten passende Therapiesimulation liefert, da zur Identifikation der Abstand der biexponentiellen Beschreibungen von Therapiesimulation und realer Therapie genutzt wurde. Die Verwendung der direkten Regression (M1) führt zu einem Modellparameterdatensatz, dessen Simulation eine deutliche Abweichung zeigt. Ergänzend wurden die Ergebnisse des Verfahrens nach Horn et al. (2013) in die Abbildung integriert (M0). Auch diese verfehlt den realen Therapieverlauf des Patienten deutlich, es ist zumindest eine Parallelität im Langzeitverlauf der Therapie gegeben. Auch bei Analysen der weiteren Patienten bestätigen sich diese Beobachtungen (siehe Anhang A.3).

(a) Beispielpatient mit eingezeichneten Therapiesimulationen (b) Verteilungen der Abstände zwischen realer und simulierter Therapie

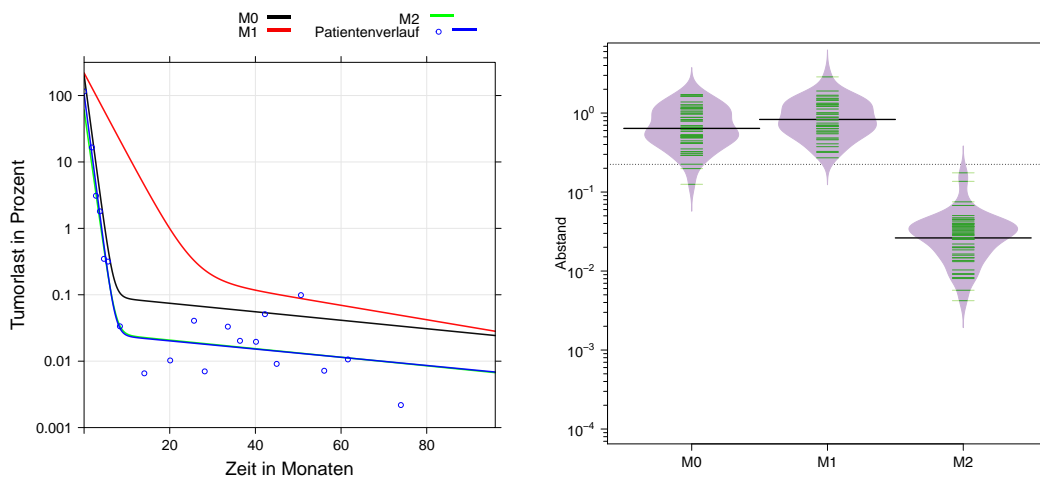


Abbildung 6.9.: Vergleich der unterschiedlichen Anpassungsverfahren der IRIS-Patienten. Man erkennt beispielhaft in Abbildung 6.9a, dass die Anpassung des realen Patientenverlaufes und der mittels lookup-table (M2) identifizierten Simulation fast übereinander liegen. Die Verwendung der direkten Regression (M1) ergibt eine Simulation, die extreme Abweichungen zum Therapieverlauf des Patienten zeigt. Das Verfahren nach Horn et al. (2013) (M0) zeigt einen simulierten Therapieverlauf, dessen Verlauf im ersten Abschnitt noch nahe an der Realität liegt. Im Langzeitverlauf zeigen sich jedoch deutliche Abweichungen. Abbildung 6.9b zeigt die Verteilungen der Abstände aller 51 verwendeten IRIS-Patienten zu den unterschiedlichen Simulationen. Dabei ist die y-Achse logarithmisch eingeteilt. Als Vergleichszeitraum wurden die Monate 0 – 60 der Therapie gewählt.

Für jeden verwendeten IRIS-Patienten ($n = 51$, Auswahl anhand der von Horn et al. (2013)

genutzten und im Kapitel 5.1 genannten Selektionskriterien) wurde nach dem beschriebenen Verfahren der Abstand zwischen der jeweiligen Simulation und dem realen Therapieverlauf als Gütekriterium für die Therapiesimulation ermittelt. Abbildung 6.9b zeigt die Verteilung dieser Abstände als Beanplots aufgetragen (Erklärung der Plots siehe Kapitel 5.1.2). Dabei wurde zunächst der komplette Therapieverlauf der Monate 0 – 60 in die Berechnung des Abstandes und damit in den Vergleich einbezogen.

Um einschätzen zu können, ob sich eines der verwendeten Verfahren zur Identifikation der besten Modellparameter eher für den anfänglichen oder den langfristigen Verlauf eignet, wurden die Monate 0 bis 12 bzw. 24 bis 60 nochmals einzeln betrachtet (Abbildung 6.10). Der Zeitraum von 12 bis 24 Monate wurde dabei aus der Betrachtung ausgeschlossen, um eine Verfälschung durch den Übergangsbereich vom ersten steileren zum zweiten flacheren Abfall im Therapieverlauf zu verhindern.

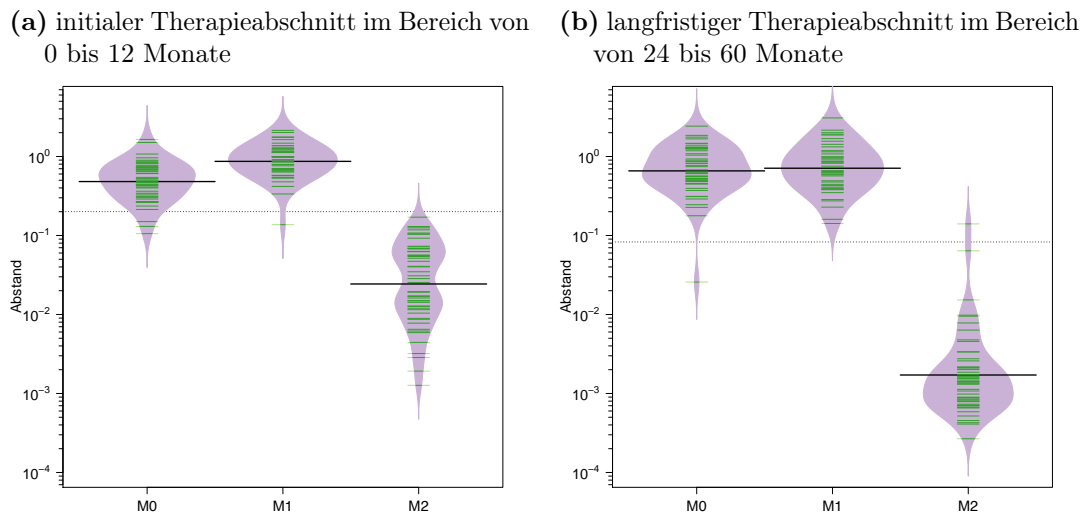


Abbildung 6.10.: Vergleich der unterschiedlichen Anpassungsverfahren für beide Therapieabschnitte. Werden die zwei Therapieabschnitte einzeln in die Analyse einbezogen (erster Bereich im Bereich von 0 bis 12 Monaten, Abbildungen **6.10a**, bzw. zweiter Langzeitverlauf im Bereich von 24 bis 60 Monaten, Abbildung **6.10b**), so bestätigt sich die Überlegenheit der lookup-table. Sie zeigt jeweils signifikant besser passende Simulationsläufe als die beiden Vergleichsverfahren (wie auch Abbildung 6.9). Der Zeitraum zwischen 12 und 24 Monaten wurde aus den Vergleichen ausgeschlossen, um eine Verfälschung der Ergebnisse im Übergangsbereich der Therapieabschnitte zu vermeiden.

Bereits Abbildung 6.9b zeigt die deutliche Überlegenheit der Identifikation der Modellparameter mittels lookup-table (M2). So hat sich der mittlere Abstand zum realen Therapieverlauf der Patienten von 0,639 beim Verfahren nach Horn et al. (2013) (M0) auf 0,026 deutlich

Tabelle 6.2.: Mittelwerte der mittleren Abstände zwischen Patientendaten und den Simulationen, die mittels unterschiedlicher Verfahren ermittelt wurden.

betrachtete Monate	Mittlere Abstände zwischen den durch die angegebenen Verfahren ermittelten Simulationen und dem Therapieverlauf des Patienten		
	0 bis 60	0 bis 12	24 bis 60
M0	0,639	0,481	0,657
M1	0,829	0,867	0,71
M2	0,026	0,024	0,002

erkennbar reduziert. Die Nutzung der direkten Regression (M1) zeigt hingegen eine Verschlechterung. So stieg hier im Median der Abstand auf 0,829. Dies entspricht einer nicht akzeptablen medianen Abweichung von fast einer log-Stufe.

Analog sind die Resultate, wenn unterschiedliche Zeitabschnitte des Therapieverlaufs betrachtet werden. Im Bereich von 0 bis 12 Monaten ist der Effekt der Verschlechterung durch Nutzung von M1 im Vergleich M0 noch deutlicher ausgeprägt (medianer Abstand 0,481 bzw. 0,867). Das Verfahren M2 ist mit einem medianen Abstand von 0,024 weiterhin deutlich überlegen.

Im Langzeitverlauf im Bereich von 24 bis 60 Monaten liegen die Verfahren M0 und M1 auf ähnlichem Niveau (medianer Abstand 0,657 bzw. 0,710). Die Nutzung von M2 hingegen ergibt Therapiesimulationen, die im Langzeitverlauf nahezu perfekt zum realen Therapieverlauf des Patienten passen. So beträgt der mediane Abstand 0,002 und unterscheidet sich damit erheblich von den beiden anderen Verfahren.

In Tabelle 6.2 sind die aus der Abbildung 6.10 ersichtlichen Mittelwerte zusätzlich numerisch angegeben.

Es ist zu beachten, dass der Abstand zwischen den unterschiedlichen Therapieverläufen auf den statistischen Beschreibungen derselben basiert und damit selbst mit einer gewissen Unsicherheit behaftet ist. Diese zu untersuchen, gegebenenfalls zu quantifizieren und zu reduzieren kann Gegenstand zukünftiger Arbeiten sein. So ist vorstellbar, die Analyse der Abstände direkt auf Ebene der Messwerte durchzuführen ohne vorher eine biexponentielle Anpassung durchzuführen.

Es kann zusammenfassend festgestellt werden, dass die Nutzung der lookup-table zur Identifikation des besten *in silico* Therapieverlaufs für einen Patienten zu einer deutlichen Verbesserung der Anpassung geführt hat. Die Nutzung einer direkten Regression hat im Vergleich zu Horn et al. (2013) keinen Vorteil gezeigt. Die hochkomplexen Parameterlandschaften können nicht ausreichend genau durch lineare Regressionen beschrieben werden. Die Identifikation der passenden Therapiesimulation mithilfe der lookup-table kann möglicherweise

weiter verbessert werden, wenn im Bereich der Modellparameter der passendsten Simulation mit kleineren Schrittweiten erneut Therapiesimulationen erfolgen. Damit erhöht sich zwar der Rechenzeit nach Identifikation der besten Simulation. Dies geschieht jedoch nur im Rahmen einer kleinen Anzahl von Simulationen.

Außerdem bringt die Anwendung der lookup-table einen weiteren Vorteil mit sich: So liegen die Therapiesimulationen bereits in einer Datenbank vor und werden anhand des Algorithmus ausgewählt. Der zeit- und rechenintensive Simulationsschritt mit berechneten Modellparametern entfällt. Weit vorausgedacht könnte man sich mobile Rechner (Smartphones oder Tablets), die am Bett des Patienten zur Dateneingabe verwendet werden, zunutze machen. Die Rechenleistung solcher Geräte wäre ausreichend, um die Datenparameter der biexponentiellen Beschreibung des Patienten zu ermitteln. Anhand dieser biexponentiellen Beschreibung kann dann leicht aus dem großen Datenbestand der vorsimulierten möglichen Therapieverläufe ausgewählt werden. Da dieser Schritt dennoch einigen Rechenaufwand fordert, sollten diese Informationen von einem am Netzwerk angeschlossenen, leistungsfähigen Server vorgehalten werden. Dieser wäre in der Lage innerhalb von wenigen Sekunden den am besten passenden *in silico* Therapieverlauf zu finden und an das mobile Gerät auszuliefern. Basierend auf den ebenfalls vorsimulierten Verläufen der Tumorlast im Stammzellkompartiment könnten dann Aussagen über den zukünftigen Therapieverlauf getroffen werden. Ein weiterer Vorteil in der Nutzung der lookup-table in Kombination mit der biexponentiellen Beschreibung ist darin zu sehen, dass die Therapiesimulationen, nachdem sie durchgeführt wurden, mittels weniger Parameter beschrieben werden können. So wird der Therapieverlauf im peripheren Blut durch vier Parameter und der Verlauf der Tumorlast im Stammzellkompartiment durch lediglich zwei Parameter charakterisiert. Somit reduziert sich auch der benötigte Speicherplatz ausgehend von ca. 500 GB für alle 270.400 Therapiesimulationen deutlich.

6.8. Analyse der vorhergesagten Verläufe der simulierten Tumorlast im Knochenmark

Der vorhergesagte Verlauf der simulierten Tumorlast im Stammzellkompartiment lässt sich durch eine Gerade in der logarithmischen Darstellung beschreiben. Anhand dieser Funktion kann man errechnen, wann die Anzahl der leukämischen Stammzellen im Knochenmark unter ein definiertes Limit fällt. Dies soll Ausgangspunkt für Vorhersagen über den Zeitpunkt einer Heilung der CML unter Behandlung mit TKI sein. Alternativ könnte man basierend auf dem Modell ein sicheres Absetzen vorhersagen. In der Arbeit von Horn et al. (2013) wurden zur Ermittlung der Stammzellzahl, unter der ein Therapieende möglich erscheint, Simulationen des Verhaltens der CML nach Absetzen der Therapie mit TKI vorgenommen. Dabei zeigten die Simulationen ein ähnliches Verhalten, wie Studien an realen Patienten mit einer bis dahin

erfolgten Nachbeobachtungszeit von etwa zwei Jahren (Ross et al. 2010; Mahon et al. 2010). Simulierte Patienten, die in dieser Zeit keinen Rückfall in messbare Tumorlasten hatten, wiesen zum Absetzzeitpunkt weniger als 200 leukämische Stammzellen auf (Horn et al. 2013).

Es fällt bei der großen Datenbasis *in silico* simulierter Therapieverläufe auf, dass im Falle eines Beispielpatienten aus der IRIS-Studie nicht nur die Simulation mit dem geringsten Abstand, sondern auch nachrangige Simulationen relativ nah am realen Therapieverlauf liegen (siehe Abbildung 6.11).

In Abbildung 6.12 sind die Verläufe der Tumorlast im Stammzellkompartiment für die ersten 1000 Ränge für den Beispielpatienten aus Abbildung 6.11 aufgetragen. Man erkennt die deutliche Streuung, welche die Verläufe der Tumorlast im Stammzellkompartiment aufweisen, obwohl der Verlauf im peripheren Blut einen eher geringen Abstand zum Therapieverlauf des Patienten hat. Geht man von der von Horn et al. (2013) postulierten Schwelle von 200 leukämischen Stammzellen, unter der ein sicheres Absetzen der Therapie sehr wahrscheinlich ist, aus, so ist der Zeitpunkt, zu dem die Simulationen diesen Wert erreichen, weit gestreut (siehe Abbildung 6.13). Weiterhin zeigt sich keine Tendenz innerhalb der Rangliste, da als nachrangig identifizierte Simulationen sowohl steilere als auch flachere Tumorlastverläufe auf Stammzellebene als die „beste“ Simulation aufweisen. Damit werden sowohl frühere als auch spätere Zeitpunkte zum sicheren Absetzen der Therapie vorausgesagt. Die geringe Abweichung der Verläufe im peripheren Blut deutet darauf hin, dass erheblich unterschiedliche Stammzellkonfigurationen ähnliche Tumorlastverläufe im peripheren Blut zur Folge haben können. Dies lässt die Annahme zu, dass mit dem Ansatz der lookup-table zwar eine peripher sehr gut passende Therapiesimulation gefunden werden kann. Der Tumorlastverlauf im Stammzellkompartiment ist jedoch sehr unsicher. Sollte man beispielsweise ein anderes Verfahren der Anpassung der biexponentiellen Beschreibung (siehe 3.1, z. B. maximum likelihood mit Zensierung) genutzt haben, welches leicht abweichende statistische Parameter der Therapie als Ergebnis hat, so kann die identifizierte Therapiesimulation einen stark abweichenden Stammzellverlauf zur Folge haben und damit große zeitliche Abweichungen bei der Vorhersage des sicheren Therapieendes. Weitere Analysen und Verbesserungen sind notwendig um dieses Problem einzugrenzen und Lösungsmöglichkeiten zu entwickeln.

Modellparameter mit starkem Einfluss auf den Therapieverlauf im Stammzellkompartiment Analysiert man die 100 am besten passenden Therapiesimulationen weiter, so fällt auf, dass es gewisse Gruppierungen in den Verläufen gibt. Diese kommen durch die gewählten Schrittweiten bei der Produktion der Datenbasis aus 270.400 Therapieverläufen zustande. Färbt man dazu die angepassten Therapieverläufe im simulierten Stammzellkompartiment entsprechend ihrer unterschiedlichen Modellparameter an, so ergeben sich die in Abbildung 6.14 gezeigten Grafiken.

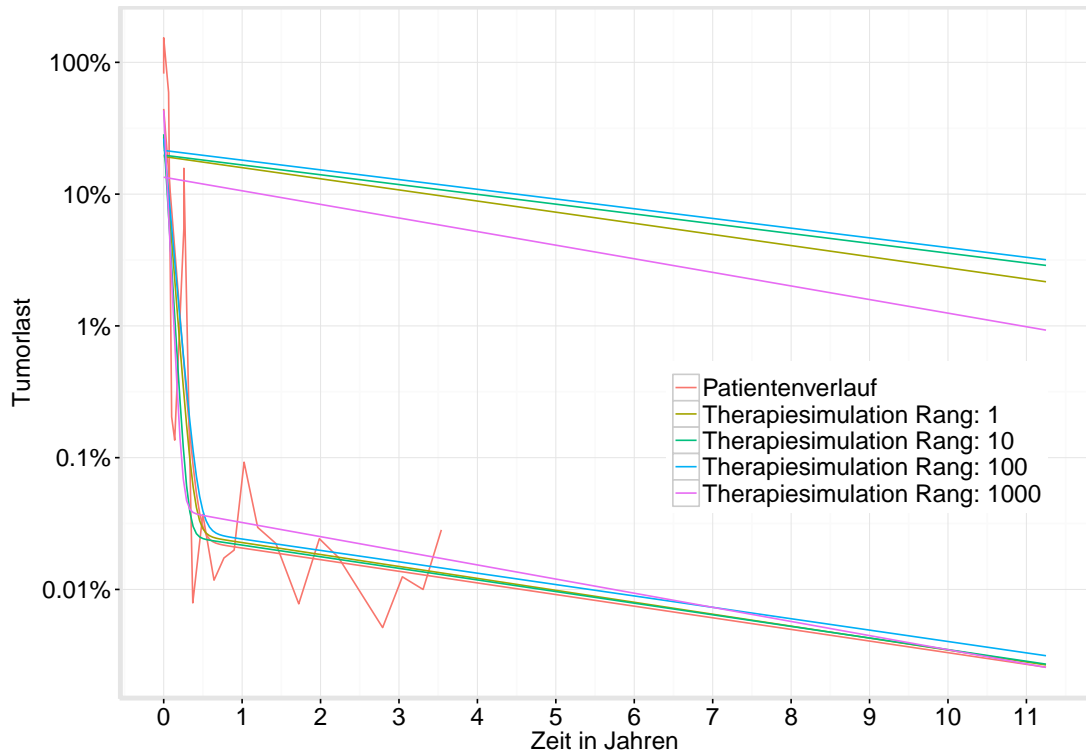


Abbildung 6.11.: Beispielpatient mit dazu passenden *in silico* Therapiesimulationen.

Diese wurden anhand der Rangfolge des Abstands aus den 270.400 vorsimulierten Therapien ausgewählt. Man erkennt, dass nicht nur die am nächsten an der Realität liegende Simulation (Rang 1) sondern auch die an Rang 10, 100 oder gar 1000 der nach Abstand geordneten 270.400 Therapiesimulationen gut den realen Therapieverlauf nachbildet. Parallel dazu ist der Verlauf der Tumorlast im Stammzellkompartiment (Geraden) aufgetragen. Die Stammzellverläufe der 1000 *in silico* Therapiesimulationen, die am nächsten am Therapieverlauf des Patienten liegen, sind in Abbildung 6.12 nochmals aufgetragen. Bereits hier wird deutlich, dass der Therapieverlauf im Stammzellkompartiment deutliche Abweichungen zeigt, obwohl der periphere Tumorlastverlauf relativ ähnlich ist. Außerdem ist festzustellen, dass die als nachrangig identifizierten Therapieverläufe sowohl steilere als auch flachere Verläufe der Tumorlast im Stammzellkompartiment zeigen. Dies bedeutet somit eine Vorhersage eines zeitigeren oder späteren Erreichens des Zeitpunktes von 200 Tumorzellen im Knochenmark, ohne eine Tendenz in der Rangfolge zu erkennen.

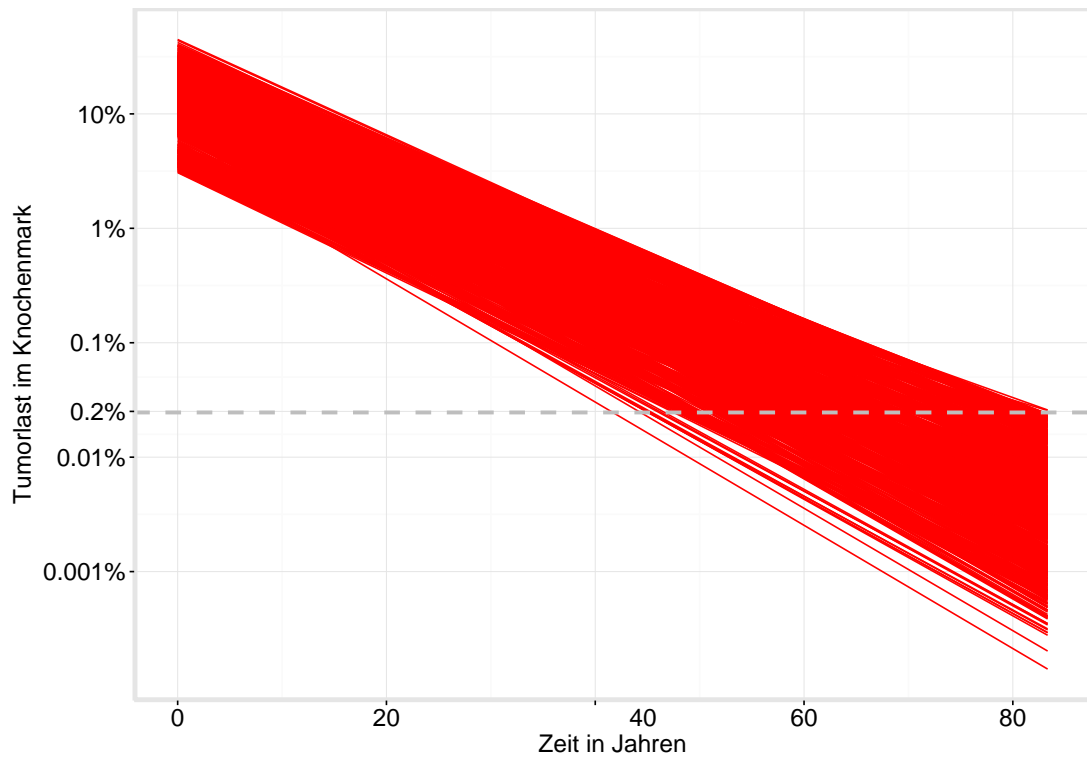


Abbildung 6.12.: Verläufe der Tumorlast im Stammzellkompartiment für die 1000 *in silico* Therapiesimulationen, deren Verlauf der Tumorlast im peripheren Blut am nächsten am Therapieverlauf des Beispielpatienten aus Abbildung 6.11 liegen. Man erkennt die deutliche Streuung der Verläufe der Tumorlast im Stammzellkompartiment für genannten Simulationen, obwohl diese einen ähnlichen Verlauf der Tumorlast im peripheren Blut zur Folge haben. Als grau gestrichelte Linie ist die im Text genannte Grenze von 200 leukämischen Stammzellen markiert, die nach Horn et al. (2013) die Grenze bildet, unter der von einer Heilung bzw. einer nicht wieder ansteigenden Tumorlast auszugehen ist.

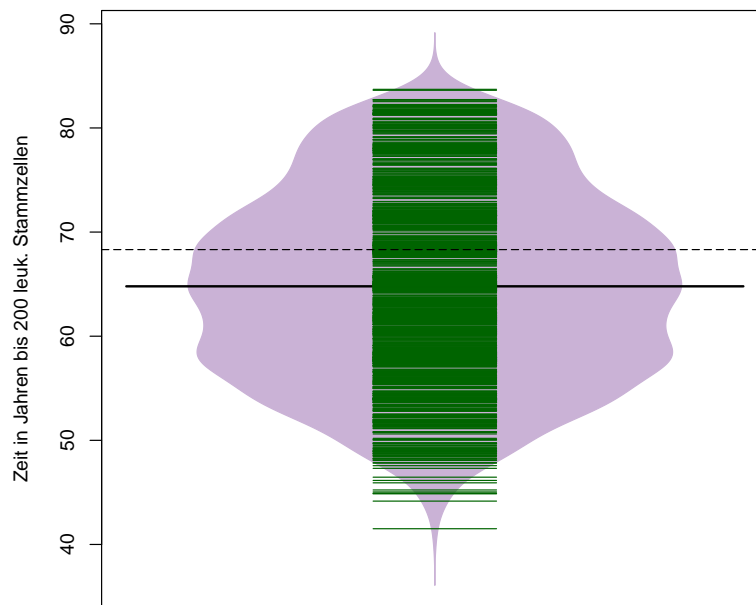


Abbildung 6.13.: Verteilung des simulierten Therapieendes (entspricht 200 leukämischen Stammzellen) für die 1000 *in silico* Therapieverläufe, deren Verlauf der Tumorlast im peripheren Blut am nächsten am Therapieverlauf des Beispielpatienten aus **Abbildung 6.11** liegen. Die in **Abbildung 6.12** dargestellte Streuung der Tumolastverläufe auf Stammzelebene wird erneut deutlich. Weiterhin eingezeichnet sind der Mittelwert (durchgezogene Linie) und der Zeitpunkt, zu dem die am nächsten am realen Patientenverlauf liegende Therapiesimulation nur noch 200 leukämische Stammzellen vorhersagt (gestrichelte Linie).

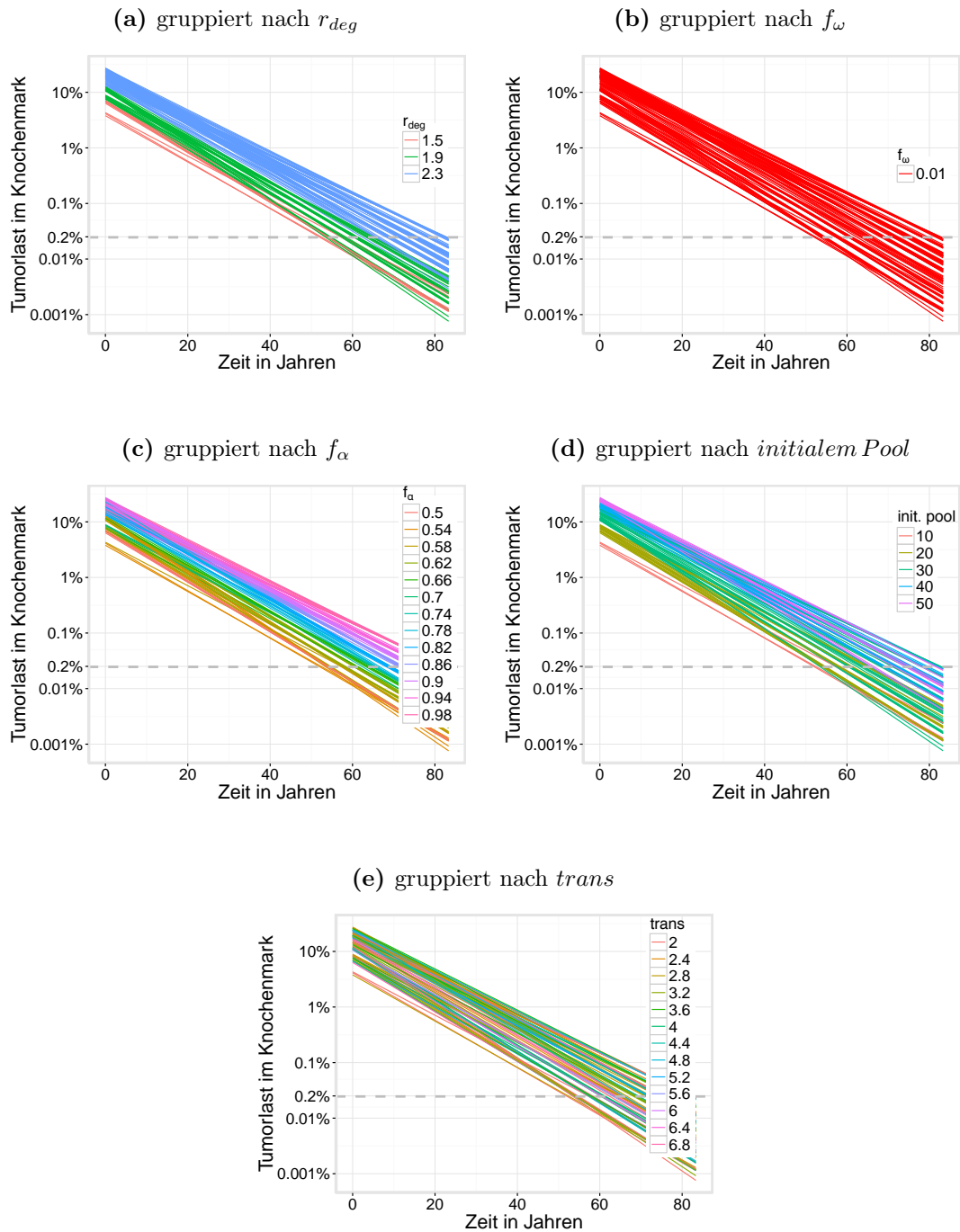


Abbildung 6.14.: Die 100 am nächsten am Patienten aus Abildung 6.11 liegenden Tumorlastverläufe auf Stammzelebene nach den konkreten Werten der unterschiedlichen Modellparameter eingefärbt. Therapiesimulationen, die mit einer hohen initialen Tumorlast starten, benötigen ein kleines r_{deg} um ähnliche Therapieverläufe zu erzeugen. Für die Übergangsfunktion f_{ω} gibt es nur einen Wert. Sie beschreibt relativ eindeutig den Therapieverlauf. Hingegen zeigt f_{α} eine weite Streuung. Minimale Veränderungen im Tumorlastverlauf fordern also abweichende Werte für diesen Parameter. Die Umsatzrate *trans* hat offensichtlich keinen systematischen Einfluss auf den Therapieverlauf, da sich keine Gruppierung zeigt.

So wird in Abbildung 6.14a deutlich, dass die 100 passendsten Therapiesimulationen durch drei verschiedene Werte für den Modellparameter r_{deg} beschrieben werden. Für Therapiesimulationen, die mit einer hohen initialen Tumorlast starten (siehe Abbildung 6.14d), reicht ein kleineres r_{deg} als bei niedriger initialer Tumorbelastung aus, um einen zum Patienten passenden Therapieverlauf im peripheren Blut hervorzubringen. Weiterhin ist die simulierte initiale Tumorlast für einen Patienten weit gestreut, obwohl der Verlauf im peripheren Blut relativ nah am realen Therapieverlauf liegt (siehe Abbildung 6.11). Der Modellparameter f_{ω} (siehe Abbildung 6.14b) beschreibt relativ eindeutig den Patientenverlauf, da nur ein Wert in allen 100 Therapiesimulationen vorkommt. Eine zweite Erklärung für dieses Phänomen könnte eine zu große Schrittweite für diesen Parameter sein. f_{α} (siehe Abbildung 6.14c) hingegen kommt mit vielen Werten vor. Es ist davon auszugehen, dass der Einfluss dieses Parameters so groß ist, dass minimal verschiedene Therapieverläufe andere Werte fordern. Die Umsatzrate $trans$ scheint keinen systematischen Einfluss auf den Therapieverlauf zu haben, da keine Gruppierung sichtbar wird (siehe Abbildung 6.14e). Insgesamt muss bei der Interpretation dieser Zusammenhänge jedoch beachtet werden, dass die Modellparameter abhängig voneinander den simulierten Therapieverlauf beschreiben. So kann beispielsweise ein um einen Schritt größerer Parameter durch einen anderen Modellparameter ausgeglichen werden (siehe oben die Beziehung zwischen initialer Tumorlast und r_{deg} um einen ähnlichen Therapieverlauf im peripheren Blut zu erzeugen). Weitere Analysen dieser Beziehungen sind in folgenden Arbeiten unbedingt notwendig, um die Qualität zu verbessern.

Bestimmt man für alle IRIS-Patienten anhand der am nächsten am realen Therapieverlauf gelegene Simulation den Zeitpunkt des Erreichens von 200 Tumorstammzellen, so ergibt sich die in Abbildung 6.15 gezeigte Kurve. In dieser zeigt sich, dass bis zu einer Therapiedauer von unter vier Jahren kein Therapieende vorhergesagt wird. Eine gewisse Therapiedauer vergeht also für jeden Patienten, bis überhaupt eine deutliche Zurückdrängung der leukämischen Blutbildung stattfindet. Fast die Hälfte der Patienten erreicht das Stadium des vorhergesagten sicheren Therapieendes im Zeitraum zwischen fünf und sieben Jahren. Danach flacht der Verlauf der Kurve deutlich ab, bis nach ca. 19 Jahren ungefähr 90% der Patienten eine Tumorlast von weniger als 200 leukämischen Stammzellen haben. Für die restlichen 10% der Patienten kann das Erreichen dieses Zustandes nicht vorhergesagt werden. Die Gründe dafür liegen in der fehlenden Anpassbarkeit des Therapieverlaufs. Weiterhin zeigt ein Teil dieser Patienten kein Ansprechen auf die Therapie, sodass überhaupt kein oder kein biphasischer Abfall der Tumorlast im peripheren Blut zu beobachten ist. Sollten sich durch Abweichungen zum beschriebenen Vorgehen andere Therapiesimulationen ergeben, die im peripheren Blut ähnlich zum realen Therapieverlauf sind, könnten sich deutlich abweichende Tumorlastverläufe im Stammzellkompartiment ergeben. Weiterhin wurde die Grenze des sicheren Therapieendes, die von Horn et al. (2013) postuliert wurde, in der vorliegenden Arbeit nicht überprüft. Sollten

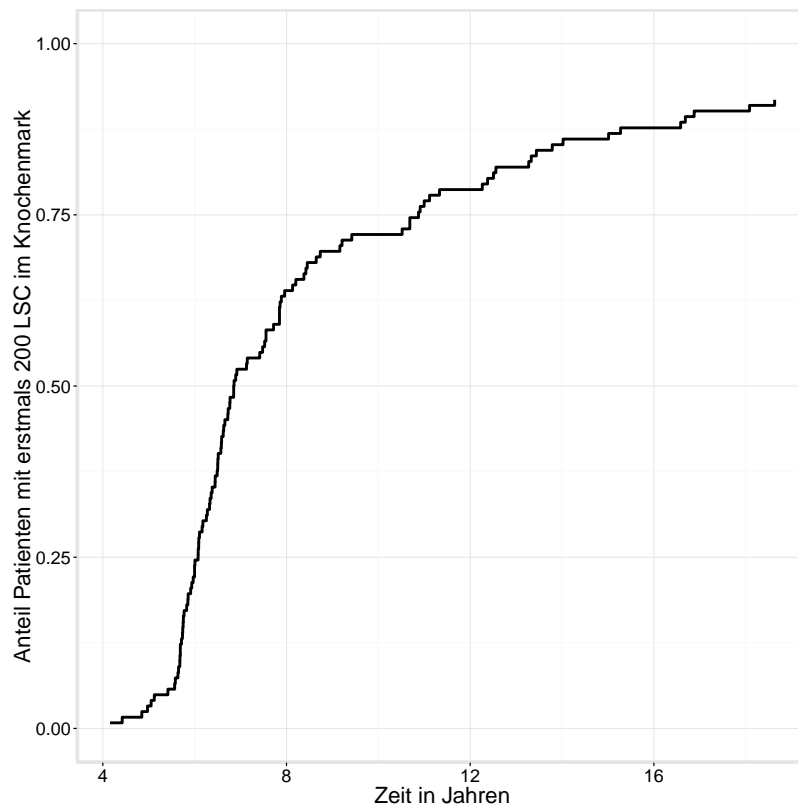


Abbildung 6.15.: Kurve, die den Anteil aller IRIS-Patienten des deutschen Armes ($n = 69$) zeigt, deren Simulation zu einem bestimmten Zeitpunkt die Zahl von 200 leukämischen Stammzellen erreicht haben. Bereits nach etwas weniger als sieben Jahren wird für die Hälfte der Patienten eine Tumorlast von weniger als 200 Zellen im Knochenmark vorhergesagt, was nach Horn et al. (2013) dem Zeitpunkt des gefahrlosen Therapieendes ohne Rezidiv entspricht. Nach circa 19 Jahren kann für alle Patienten, die einen biphasischen Abfall zeigen, ein Therapieende vorhergesagt werden. Der restliche Anteil von ungefähr 10% der Patienten kann durch die verwendeten Methoden nicht vorhergesagt werden, da überhaupt kein Anprechen oder kein anpassbarer Tumorlastverlauf unter Therapie vorliegen.

sich in weitergehenden Analysen andere Grenzen zeigen, so verschiebt sich die Vorhersage des Absetzzeitpunktes entsprechend.

Zusätzlich wird im Rahmen der vorliegenden Arbeit der Verlauf der Tumorlast im Stammzellkompartiment als Gerade in der bekannten logarithmischen Darstellung vom verwendeten Simulationsmodell beschrieben. Erste Analysen der leukämischen Stammzellen an realen Patienten lassen Zweifel an dieser vereinfachten Betrachtung aufkommen (Guevara et al. 2012; Mustjoki et al. 2013). Man müsste die vorhergesagten Absetzzeitpunkte im Rahmen von Studien in ihrer Anwendung prüfen, um zu entscheiden, ob eine andere Form der Beschreibung des Therapieverlaufs im Stammzellkompartiment besser geeignet ist. Als Vorstufe dazu könnten detailliertere Analysen der Tumorlast im Knochenmark von realen CML-Patienten die Genauigkeit der Modelle und der getroffenen Vorhersagen verbessern. Denkbar ist auch, dass anhand einer Messung der initialen Tumorlast im Knochenmark des Patienten die Anzahl der Simulationen, die als passend zu einem realen Therapieverlauf beschrieben werden, weiter eingegrenzt werden kann.

7. Zusammenfassung und Diskussion

Im Rahmen der vorliegenden Arbeit wurden modellbasierte Simulationen zur Beschreibung der Behandlung von Patienten mit chronischer myeloischer Leukämie (CML) optimiert. Insbesondere wurde ein Verfahren entwickelt, um simulierte *in silico* Therapieverläufe effizient mit den tatsächlichen Verläufen einzelner Patienten zu vergleichen. Dieses Verfahren erlaubt es, Simulationen mit sehr hoher Übereinstimmung zu identifizieren. Eine Korrektheit des zugrundeliegenden Modells vorausgesetzt, ermöglichen diese Simulationen einen Einblick in klinisch schwer zu erfassende Laborwerte. Insbesondere können Tumorlastverläufe der hämatopoetischen Stammzellen im Knochenmark abgeschätzt werden.

Stammzellen: *in vivo* vs. *in silico* Stammzellen sind undifferenzierte Zellen, die in der Lage sind, sich asymmetrisch zu teilen. Damit können sie ihre eigene Population aufrecht erhalten, aber auch Zellen produzieren, die sich immer weiter ausdifferenzieren (Potten und Loeffler 1990; Loeffler und Roeder 2002). Bei den betrachteten hämatopoetischen Stammzellen (HSCs) handelt es sich um adulte Stammzellen, da sie nicht wie embryonale Stammzellen in der Lage sind, jedes Gewebe zu bilden, sondern bereits auf die Aufrechterhaltung der Hämatopoese spezialisiert sind. Die HSCs sind eine inhomogene Population, die sich auf funktionaler Ebene unterscheiden. Neben ruhenden HSCs mit geringer Teilungsaktivität (im Mausmodell im Schnitt nur eine Teilung in mehreren Monaten), aber hohem Repopulationspotential, gibt es auch sogenannte multipotente Progenitorzellen, die sich häufiger teilen, aber nur noch wenig Potential zur Repopulierung aufweisen. Die normale Hämatopoese wird vor allem von multipotenten Progenitorzellen aufrechterhalten. Erst der Einfluss von Stress (bspw. Blutung, Infektion, o. ä.) führt zur gesteigerten Aktivierung der ruhenden HSCs, die dann eine endogene Repopulation der Hämatopoese durchführen (Höfer und Rodewald 2016).

Die CML als myeloproliferative Erkrankung geht von mutierten HSCs aus. Diese unterhalten als leukämische Stammzellen (LSCs) die CML und verdrängen die gesunde Hämatopoese. Um die Hämatopoese zu beurteilen, sind Knochenmarkpunktionen geeignet. Damit können am Patienten Aussagen über die leukämische Last nicht nur im peripheren Blut, sondern auch auf Ebene der Stammzellen getroffen werden. Die oben angeführte Inhomogenität innerhalb der HSCs hat den Nachteil, dass nicht klar ist, wie hoch die Tumorbelastung auf den einzelnen Ebenen der Hämatopoese ist, da bei einer Knochenmarkanalyse alle kernhaltigen Zellen des

Punktats untersucht werden. Außerdem ist eine Knochenmarkpunktion für den Patienten unangenehm und mit nicht zu vernachlässigenden Risiken verbunden.

Im Modell nach Roeder und Loeffler (2002) sind im Stammzellkompartiment (bestehend aus den Unterkompartimenten der aktiv proliferierenden Stammzellen A und der ruhenden Stammzellen Ω) unterschiedliche Differenzierungsstufen enthalten. In Ω befinden sich darüber hinaus sogar Zellen, welche die Fähigkeit der Regenerierung verloren haben und damit nicht mehr in den ruhenden Zustand übergehen können (siehe Kapitel 6.8). Die Entstehung und Behandlung der CML wird im Modell als klonale Konkurrenz zwischen HSCs und LSCs in den genannten Kompartimenten beschrieben. Die Bestimmung der Tumorlast auf Ebene der Stammzellen erfasst im Modell alle Zellen in diesen Abschnitten.

Auch das simulierte periphere Blut ist im Modell einfach zu erfassen. Dort können die Daten leicht mit Messungen am Patienten verglichen werden, da regelmäßige Messungen der Tumorlast im peripheren Patientenblut etabliert zur Therapieüberwachung sind. Grundidee bei der Anwendung der Simulation ist, dass bei Übereinstimmung zwischen dem Tumorlastverlauf peripheren Blut von Patient und Simulation, rekursiv Zugriff auf das simulierte Stammzellkompartiment gegeben ist. Anhand dieser Daten könnte dann ein sicheres Absetzen der Therapie mit TKIs vorhergesagt werden.

Vom peripheren Blut zum Knochenmark In der vorliegenden Arbeit wurde anhand klinischer Daten der IRIS-Studie eine effiziente Möglichkeit der Abbildung von Patientendaten aus dem peripheren Blut im Modell gefunden. Über Simulationen konnte anschließend Zugriff auf das simulierte Stammzellkompartiment ermöglicht werden. Der vom Modell postulierte Verlauf der Tumorlast im Stammzellkompartiment ist eine Gerade im logarithmischen Koordinatensystem, d. h. exponentiell abfallend. Bei der beispielhaften Analyse eines einzelnen Patienten im Kapitel 6.8 zeigte sich, dass sehr ähnliche Therapieverläufe im peripheren Blut durch viele verschiedene Stammzellverläufe beschrieben werden können. Eine ein-eindeutige Zuordnung zwischen peripherer Kinetik und Stammzellverlauf ist somit nicht möglich.

Die mathematische Definition einer Geraden ist über zwei Punkte gegeben. Aus theoretischer Sicht und unter Annahme der Korrektheit des Modells reichen zwei Messpunkte, um eine exakte Vorhersage der Stammzellkinetik zu ermöglichen. Allerdings ist bisher die Form der Geraden nicht validiert. Es ist denkbar, dass man anhand häufigerer Knochenmarkpunktionen am Patienten eine andere Form der Funktion erhält und somit das in der jetzigen Form Modell widerlegt wäre.

Um als Grundlage für diese Möglichkeit der Validierung die Vergleichbarkeit zwischen Stammzelldaten aus dem Modell mit Daten aus Knochenmarkpunktionen zu gewährleisten, ist die Vorsortierung der im Punktat enthaltenen Zelltypen (z. B. aktive und ruhende HSCs sowie Progenitorzellen) anhand von Oberflächenmarkern sinnvoll (Mustjoki et al. 2013). Dieser

Ansatz wurde weiterhin bereits verwendet, um die mit etablierten Verfahren nicht mehr messbare Tumorlast in tiefer molekularer Remission zu quantifizieren (Kooten Niekerk et al. 2013). Definiert man im Modell die genau gleichen im Stammzellkompartiment enthaltenen Zelltypen kann der Vergleich so auf solide Füße gestellt werden.

Angenommen der exponentielle Abfall der leukämischen Stammzellen im Knochenmark ist korrekt, wären Messfehler nicht außer Acht zu lassen, da bei einer zweimaligen Messung bereits geringe Messungenauigkeiten an einer Stelle zu erheblichen Abweichungen im Langzeitverlauf führen könnten. Dies würde erheblich abweichende Absetzzeitpunkte ergeben.

Aus der einfachen Geradendefinition auf logarithmischer Skala könnte geschlussfolgert werden, dass das Modell für die Vorhersage eines Absetzzeitpunktes nicht notwendig ist. Dagegen steht jedoch die Frage, ob die Messung in Knochenmarkpunktionen aus einem einzelnen Knochen gerade bei geringer Tumorbelastung überhaupt repräsentative Ergebnisse für die gesamte Hämatopoese liefert.

Zusammengefasst scheint die Anwendung des noch zu validierenden Modells sinnvoll, da sie verhindert, dass übermäßig häufig Knochenmarkpunktionen vorgenommen werden müssen.

Am oben genannten Beispielpatienten (siehe Kapitel 6.8) zeigte sich, dass die simulierten Stammzellkinetiken passend zu ähnlichen peripheren Therapieverläufen häufig parallel zueinander verlaufen. Somit könnte zumindest eine initiale Messung der Tumorlast im Knochenmark die Anzahl der in Betracht zu ziehenden Stammzellverläufe deutlich einschränken.

Von der Stammzellsimulation zu Vorhersagen Aufgrund des Mangels an Daten über die Tumorbelastung im Knochenmark ist eine alternative Validierungsstrategie die Vorhersage und Durchführung von Absetzuntersuchungen. Ein Ansatz auf Populationsebene wurde von Horn et al. (2013) auf Basis desselben Modells von Roeder und Loeffler (2002) beschrieben. Dazu wurde als Grenze in der Simulation die Anzahl von 200 leukämischen Stammzellen (LSCs) identifiziert, unter welche die Patienten fallen müssen, bevor ein sicheres Absetzen der TKI-Therapie sinnvoll ist. Diese Grenze wurde anhand einer einzelnen LSC-Kinetik für jeden Patienten bestimmt, jedoch legt die vorliegende Arbeit die nicht eindeutige Zuordnung einer Kinetik im peripheren Blut zur Kinetik im Stammzellkompartiment nahe.

Zwei Annahmen wurden von Horn et al. (2013) zur Simulation des Absetzens getroffen. Einerseits wurde der Wert für den zytotoxischen Effekt r_{deg} abgeschaltet. Andererseits blieb der Wert für die Übergangsfunktion f_{ω} , die den Übergang der Stammzellen aus einem aktiven Zustand (Ω) in einen ruhenden Zustand im Kompartiment A beschreibt, auf dem Wert wie unter der Therapie. Basierend auf diesen Annahmen konnte für die analysierte Population eine Übereinstimmung zu den damals publizierten Absetzbeobachtungen gezeigt werden. So wiesen ca. 60% der Patienten einen Rückfall innerhalb des ersten Jahres auf und ca. 40% blieben bis zum Ende der Nachbeobachtungszeit von zwei Jahren rückfallfrei (Mahon et al.

2010; Ross et al. 2010; Horn et al. 2013). Aktuell liegen Beobachtungen mit einer längeren Nachbeobachtungszeit vor. Diese legen nahe, dass Patienten, die nach einem Jahr noch ohne erneutes Anwachsen der Tumorlast geblieben sind, über viele Jahre rückfallfrei bleiben (Saussele et al. 2016). Wendet man jedoch die getroffenen Annahmen in der Simulation an, so zeigt sich eine weitere Zunahme der Rückfälle im weiteren Zeitverlauf.

Diese Diskrepanz gilt es in der Modellvorhersage auszuräumen. Dazu geeignet erscheint die Etablierung einer weiteren Zellpopulation im Modell, die dem des Immunsystems entspricht. Analysen von Clapp et al. (2016) zeigen, dass mithilfe eines Immunsystems der Therapieverlauf einer CML und das Absetzen des TKIs möglicherweise besser beschreibbar sind. Grundlage für diese Annahme ist die Beobachtung, dass die Patienten, die über 12 Monate rückfallfrei bleiben, trotz *BCR-ABL1*-Positivität auch ohne laufende Therapie mit TKI über viele Jahre keinen Rückfall zeigen. Der Einfluss des Immunsystems kann jedoch erst bei kleiner Tumorbelastung relevant werden, da vorher das Immunsystem in seiner Funktion supprimiert ist. Sinkt die leukämische Last unter ein bestimmtes Level kann die Abwehrleistung die residualen LSCs beherrschen, bis so wenig LSCs vorhanden sind, dass das Immunsystem herunterreguliert werden kann. Daraufhin können die LSCs wieder etwas ungestörter proliferieren und kommen damit erneut in den Fokus. Insgesamt führt dies zu einer oszillierenden, aber geringen Tumorlast (Ilander et al. 2014; Clapp et al. 2016). Aktuelle Studien untersuchen den Einfluss von Immunzellen auf die CML.

In der vorliegenden Arbeit wurde die Schwelle von 200 LSCs aus der Publikation von Horn et al. (2013) übernommen und Vorhersagen über den Zeitpunkt des sicheren Absetzens auf dieser Basis getroffen. So zeigte sich in Abbildung 6.15 für alle analysierten Patienten eine Zeit von ungefähr vier Jahren, die behandelt werden muss, bevor überhaupt eine relevante LSC-Depletion bei den ersten Patienten erreicht werden kann. Nach Ablauf von sieben Therapiejahren wird für ungefähr die Hälfte der Patienten ein sicheres Absetzen vorhergesagt. Für 10% der Patienten wird auch nach einem Zeitraum von über 19 Jahren die Schwelle von 200 Zellen nicht erreicht. Diese Patienten konnten aufgrund ihres monophasischen Therapieverlaufs bzw. Wiederanstieges unter Therapie vom biexponentiellen Modell nicht beschrieben werden und waren somit nicht in der Simulation abzubilden. Absetzvorsagen auf Basis einzelner Patienten wurden nicht vorgenommen, sind jedoch essentiell für die klinische Anwendung in der individualisierten Medizin.

Zusammenfassend kann festgestellt werden, dass das einleitend postulierte Verbesserungspotential der Anpassungsroutine für den Therapieverlauf im peripheren Blut genutzt werden konnte. Die Suche nach optimal passenden Verläufen über die lookup-table ist schnell und effizient. Der einfache Zugriff auf die Stammzellkinetiken ist über das Modell zwar gegeben, dennoch zeigen sich verbesserungswürdige Punkte, welche die Übertragung auf die klinische Anwendung aktuell noch verhindern.

8. Abstract

Hintergrund Die chronische myeloische Leukämie (CML) ist eine myeloproliferative Erkrankung, die aufgrund ihres Modellcharakters unter der Behandlung mit Tyrosin-Kinase-Inhibitoren (TKI) gut für eine Beschreibung mittels computerbasierter Modelle geeignet ist. Grundlage für die Entstehung einer CML ist die Bildung eines Philadelphia-Chromosoms durch eine Translokation der Chromosomen 9 und 22. Es resultiert das Onkogen *BCR-ABL1*, welches für eine konstitutiv aktive Tyrosinkinase codiert. Diese führt zu unregelmäßiger Proliferation der betroffenen Zellen und zur Verdrängung der gesunden Blutbildung.

Das überaktivierte Protein kann durch TKIs gezielt gehemmt werden. Damit ist es möglich, die Tumorlast erheblich zu senken und das Fortschreiten der Erkrankung aufzuhalten. Aktuell werden in der klinischen Anwendung außerhalb von Studien TKIs für die gesamte Lebensdauer der Patienten eingesetzt. Absetzstudien zeigten, dass circa 50% der Patienten nach einer über zwei Jahren nicht nachweisbaren *BCR-ABL1*-Last nach Behandlungsstopp kein erneutes Anwachsen der Tumorlast aufwiesen.

Die Anwendung von computergestützten Modellsimulationen hilft, Zugriff auf die klinisch nur schwer zu messenden leukämischen Stammzellen zu bekommen und darüber Vorhersagen über den weiteren Therapieverlauf zu treffen.

Aufgabenstellung Im Rahmen der vorliegenden Arbeit sollen Möglichkeiten der Übertragung von Patientendaten auf das etablierte Modell nach Roeder und Loeffler (2002) verbessert werden. Die vom Modell vorhergesagten Stammzellkinetiken sollen abschließend auf Praxistauglichkeit geprüft werden.

Material und Methoden Aufgrund der Vergleichbarkeit zu früheren Untersuchungen erfolgte die Auswahl von 51 Patienten des deutschen Armes der IRIS-Studie. Deren Therapieverläufe wurden analysiert und können über eine biphasische exponentielle (biexponentielle) bzw. über eine stückweise lineare Funktion beschrieben werden. Als Erweiterung der Arbeiten von Horn et al. (2013) wurden alle Parameter der biexponentiellen Funktion in die Entwicklung neuer Methoden einbezogen. Zusätzlich wurde untersucht, ob die Einbeziehung von zensierten Messpunkte die Form der biexponentiellen Funktion verändert.

Basierend auf den Therapiedaten der IRIS-Patienten erfolgte die Ermittlung eines Para-

meterraumes für Eingangsparameter der Modellsimulation (Modellparameter), welcher in 270.400 individuelle Parameterkombinationen unterteilt wurde. Es folgten anschließend die Simulation und Auswertung nach der biexponentiellen Beschreibung. Auf Basis dieser erheblich größeren Datengrundlage konnten zwei neue Verfahren der Modellparameteridentifikation für individuelle Patienten entwickelt werden. Einerseits wurde in Anlehnung an die Arbeit von Horn et al. (2013) ein Verfahren unter Nutzung der Regression vorgestellt. Andererseits konnte über den Vergleich der Abstände zwischen simulierten und realen Therapieverläufen eine Suche (lookup-table) etabliert werden. Die Berechnung des Abstandes zwischen Therapieverläufen ermöglicht gleichzeitig den Vergleich der verschiedenen Verfahren und damit eine Aussage über deren Anpassungsgüte.

Zum Schluss wurde beispielhaft für einen Patienten das Verfahren der lookup-table angewendet und die resultierende Stammzellkinetik weiter analysiert.

Ergebnisse Einführend erfolgte die Analyse der resultierenden biexponentiellen Funktion mit und ohne Einbeziehung von Messunsicherheiten. Es zeigte sich, dass der Verlauf dieser Funktion besonders in Bereichen, die von einbezogenen Messunsicherheiten betroffen sind, abweichend ist. Die Beschreibung des Langzeitverlaufs erfolgt jedoch annähernd gleich.

Anschließend erfolgte die Validierung der Größe des vorsimulierten Datenpool anhand eines Vergleichs der statistischen Parameter von Patienten und Simulationen. Dieser zeigte sich dabei für die weiteren Untersuchungen geeignet.

Die Nutzung der lookup-table zur Identifikation der am besten zu einem Patienten passenden Therapiesimulation ist überlegen sowohl gegenüber von der Horn et al. (2013) beschriebenen als auch in dieser Arbeit neu entwickelten Regressionsverfahren. Diese ergeben deutliche Abweichungen zwischen Patientendaten und Simulation.

Eine Analyse des vorhergesagten Therapieverlaufes im Stammzellkompartiment ergibt jedoch, dass ähnliche Therapieverläufe im peripheren Blut durch stark unterschiedliche Stammzellkonfigurationen beschrieben werden können. Es resultiert eine starke Streuung der vorhergesagten Zeitpunkte eines möglichen Therapieendes.

Schlussfolgerungen Die Nutzung der lookup-table zu Identifikation einer passenden Therapiesimulation ist hoch effektiv und anderen Verfahren, die auf Regression basieren, überlegen. Die etablierte Computersimulation nach Roeder und Loeffler (2002) bietet Zugriff auf die Therapie in der Ebene der Stammzellen.

Die in weiteren Analysen gezeigten Streuungen der vorhergesagten Therapieverläufe im Stammzellkompartiment lassen den Schluss zu, dass Methoden zur Eingrenzung der Stammzellverläufe entwickelt werden müssen, um die Vorhersagen klinisch nutzbar zu machen. Weiterhin muss anhand von Messungen an Knochenmarkproben von realen Patienten geprüft werden,

ob die von der Simulation postulierten Verläufe der Tumorlast im Stammzellkompartiment der realen Behandlung entsprechen.

Ausblick Die in aktuellen Arbeiten beschriebene Rolle des Immunsystems im Therapieverlauf der CML (Saussele et al. 2016; Clapp et al. 2016) sollte in eine Verbesserung des Stammzellmodells nach Roeder und Loeffler (2002) einfließen. Weiterhin kann die Validierung der im Rahmen der Individualmedizin zu treffenden Absetzvorsagen letztendlich nur über klinische Absetzuntersuchungen ermöglicht werden.

9. Summary

Background Chronic myeloic leukaemia (CML) is a myeloproliferative disease, which is well suited for modelling approaches. It is characterized by the oncogenic *BCR-ABL1* fusion gene originating from an inverse translocation of the chromosomes 9 and 22 leading to the *Philadelphia chromosome*. The result is a constitutively activated tyrosine-kinase. This is followed by an extensive proliferation of leukaemic stem cells leading to a displacement of normal haematopoiesis.

The molecular specificity of CML forms the basis of a highly efficient, targeted therapy by tyrosine kinase inhibitors (TKIs). TKIs can decrease the tumour burden and slow down or eventually stop progressing of the disease. Currently, in clinical applications drugs are administered for the remaining life span. Interestingly, in recent treatment cessation trials patients were stopped after two years of non-detectable tumour burden and about 50% remained without relapse.

The application of computer-based modelling helps to gain access to stem cell counts being difficult to measure clinically. This forms the basis for predictions of long-term therapy outcomes.

Aim of this work This work aims on identifying a suitable algorithm to efficiently identify model simulations that optimally describe individual patient kinetics. Furthermore, the clinical usability of the new methods was investigated.

Material and methods The analysed group of patients was chosen out of the German cohort of the IRIS trial to ensure comparability to former investigations. It consists of 51 individuals. The course of leukaemic burden, i. e. leukaemic vs. non-leukaemic cells on a single patient level can be described as a biphasic exponential (bi-exponential) or a piecewise linear function. As an extension to former methods described by Horn et al. (2013) all parameters are included into further method development. Additionally, an investigation was conducted whether censored data points change the functional behaviour of a bi-exponential fit based on patients' data.

According to therapy data of all patients an input parameter space for the model simulation was delimited, such that all observed patient kinetics can be mimicked by the model. This parameter space was uniformly divided into 270.400 discrete parameter combinations. The

therapy simulation of each combination was conducted and described by a bi-exponential function likewise to the patients' fit. With the help of these huge variety of *in silico* therapies two new methods of model parameter identification for individual patients were developed. The first one is an advanced approach based on a regression model proposed by Horn et al. (2013). The second one by comparing distances between the patients' and the models' bi-exponential functions (lookup table). The comparison of the distances between different therapy courses (either simulated or patients' data) was also used to compare the quality of different methods.

As an example, for one patient the stem cell kinetics from the model were analysed in more detail and checked for robustness. Such a strategy, which might build the basis for clinical applications.

Results A comparison between the different bi-exponential functions with and without censored data points revealed differences especially in the area in which censoring was performed. However, for the long-term tumour burden censored data had no influence.

Secondly, an investigation was performed showing the sufficiency of the pre-simulated therapy courses for the new methods, i. e. lookup-table and regression models. The lookup-table turns out to be superior to identify a therapy simulation for a unique patient, since the complexity of linear regression models lead to increased deviations between patients' therapy courses and the simulations.

Unfortunately, distinct stem cell configurations lead to similar therapy descriptions in peripheral blood, assuming the correctness of the model. As a result, the prediction of a safe treatment cessation is often widely spread.

Conclusions The new developed lookup-table to identify model simulations suitable for an individual patient is highly effective and superior to other methods using regression models. The simulation of the TKI treatment using the agent-based model of Roeder und Loeffler (2002) gives easy access to therapy courses on the level of leukaemic stem cells. Unfortunately, the finding of a well fitting simulation within the peripheral blood is not enough to provide a point of safe treatment cessation, since different stem cell configurations can lead to similar therapy courses. Additionally, it is necessary to check which of the assumed therapy courses on the stem cell level is appropriate. This could be done by gathering more information from bone-marrow punctures during the course of treatment.

Outlook Investigations of new data showed the important role of the immune system in CML treatment (Saussele et al. 2016; Clapp et al. 2016). This should be taken into account by improving the model of Roeder und Loeffler (2002). Additionally, data from cessation trials can be used to validate the model assumptions.

A. Appendix

A.1. Genutzte Parameterkonfiguration des Stammzellmodells

Tabelle A.1.: Genutzte Parameterkonfiguration des Stammzellmodells. Zu jedem der im Simulationsmodell möglichen Parameter ist der jeweilige Wert getrennt für gesunde und leukämische Zellen aufgetragen. Dabei haben die Parameter folgende Bedeutungen: a_{min} , a_{max} = Grenzen des Parameters a , der die Affinität zum ruhenden Kompartiment A beschreibt; d = Differenzierungskoeffizient; r = Regenerationskoeffizient; τ_c = Dauer des Zellzyklus von Stammzellen; τ_S , $\tau_{G_2/M}$ = Dauer der S - bzw. G_2/M -Phase; $\tilde{\tau}_c$ = mittlere Generationszeit der proliferierenden ausdifferenzierten Zellen; λ_p = Zeitdauer der Teilungsfähigkeit; λ_m = Lebensdauer der nicht mehr proliferierenden ausgereiften Zellen; f_α , f_ω = Übergangsfunktion für den Wechsel von Kompartiment Ω nach A bzw. aus A nach Ω mit Funktionswerten, die durch die Faktoren \tilde{N}_A bzw. \tilde{N}_Ω beeinflusst werden.

Parameter	gesunde Zellen	leukämische Zellen
a_{min}	0,002	0,002
a_{max}	1,0	1,0
d	1,05	1,05
r	1,1	1,1
τ_c	48 h	48 h
τ_S	8 h	8 h
$\tau_{G_2/M}$	8 h	8 h
$\tilde{\tau}_c$	24 h	24 h
λ_p	20 d	20 d
λ_m	8 d	8 d
$f_\alpha(0)$	0,5	1,0
$f_\alpha(\frac{\tilde{N}_A}{2})$	0,45	0,9
$f_\alpha(\tilde{N}_A)$	0,05	0,058
$f_\alpha(\infty)$	0	0
\tilde{N}_A	10^5	10^5
$f_\omega(0)$	0,5	1
$f_\omega(\frac{\tilde{N}_\Omega}{2})$	0,3	0,999
$f_\omega(\tilde{N}_\Omega)$	0,1	0,998
$f_\omega(\infty)$	0,0	0,996
\tilde{N}_Ω	10^5	10^5

A.2. Abbildungen des Datenparameterraumes der 270.400 *in silico* Therapiesimulationen

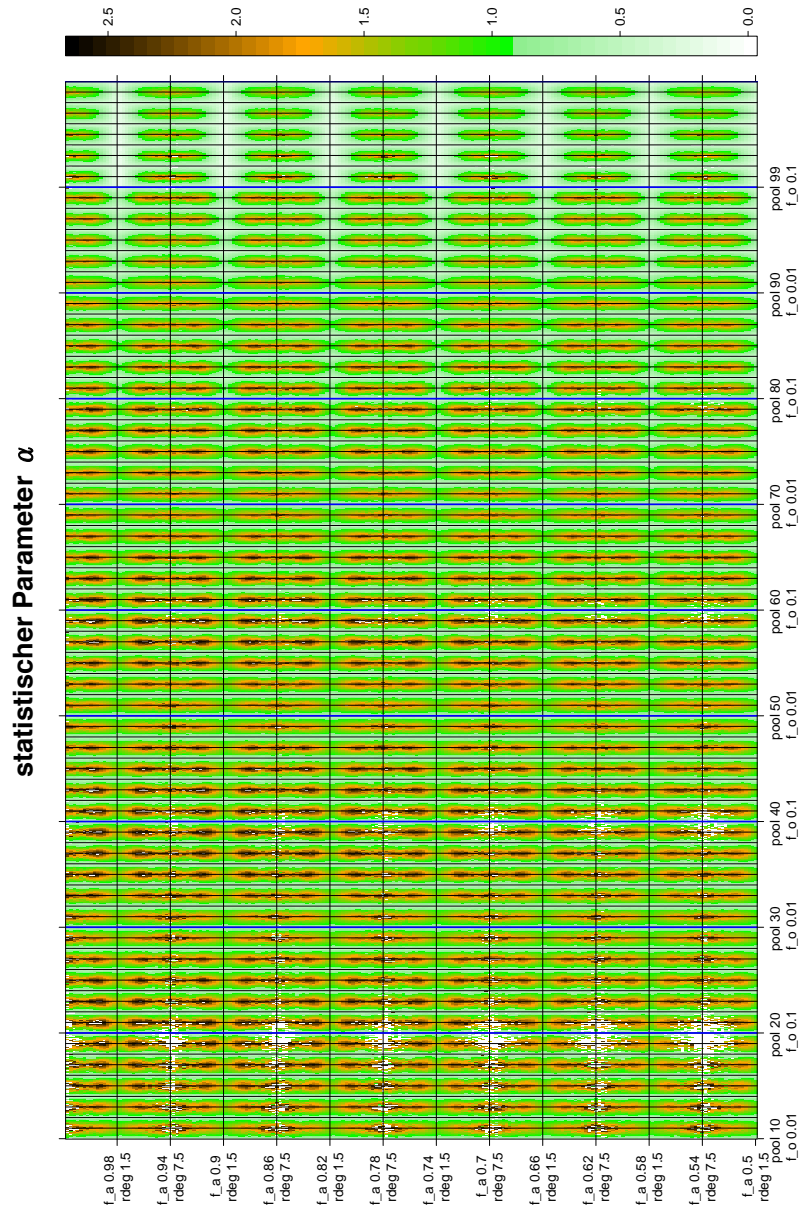


Abbildung A.1.: Darstellung der Parameterlandschaft für den statistischen Parameter α . Farbkodiert dazu jeweils der Wert für α und an den beiden Achsen die jeweils sich ändernden Modellparameter für die Simulationen.

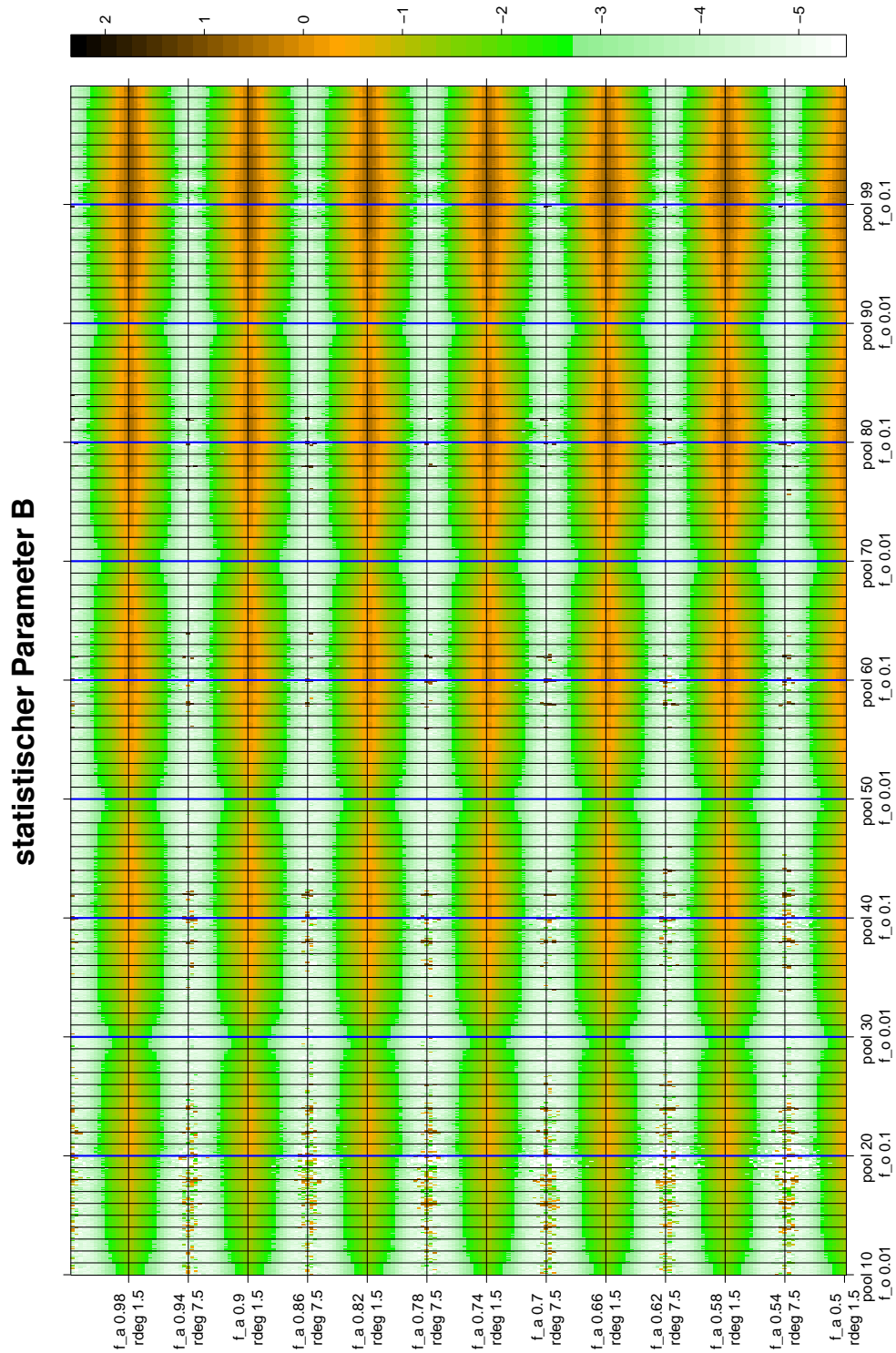


Abbildung A.2.: Darstellung der Parameterlandschaft für den statistischen Parameter B. Farbkodiert dazu jeweils der Wert für B und an den beiden Achsen die jeweils sich ändernden Modellparameter für die Simulationen.

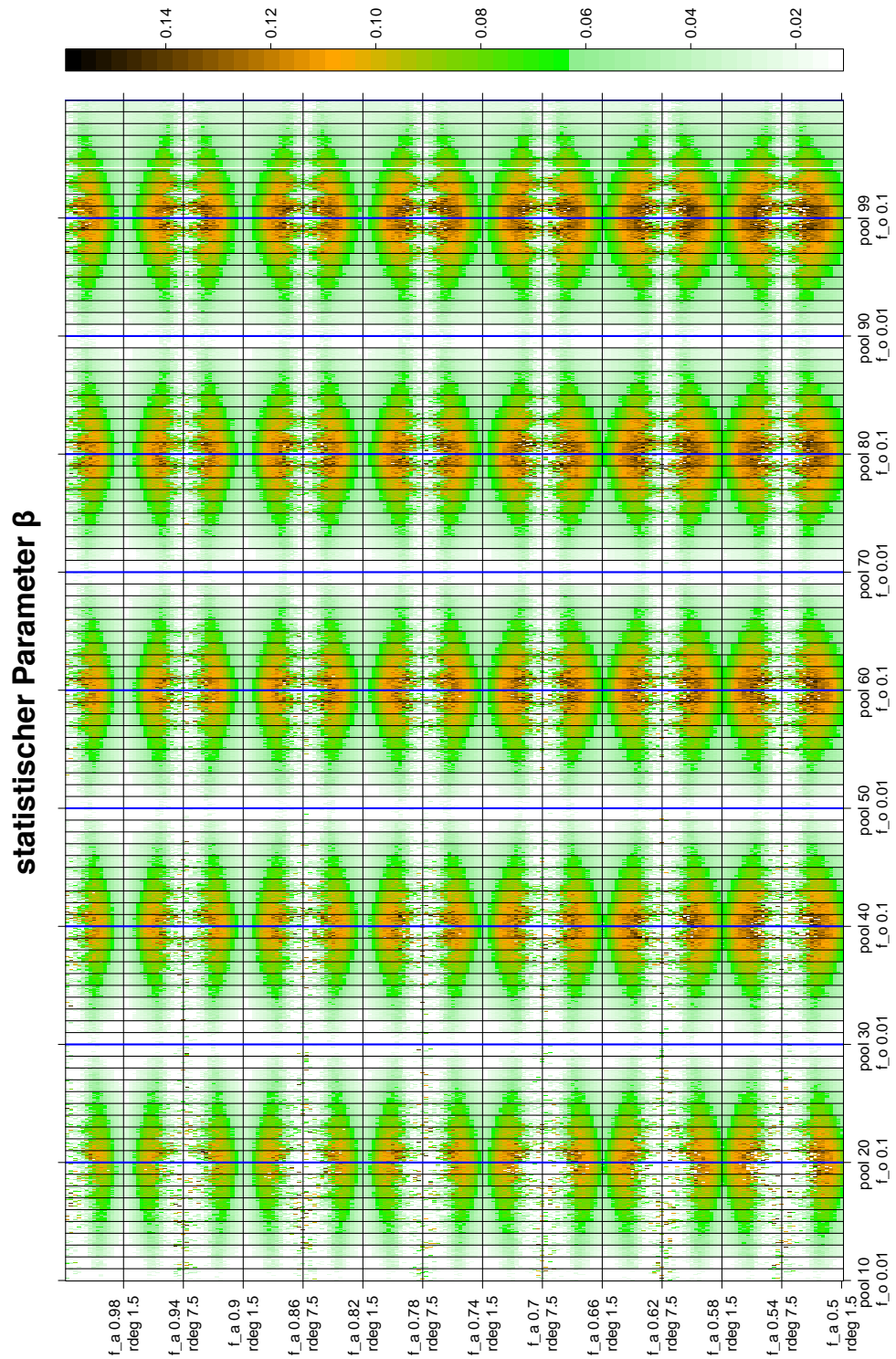


Abbildung A.3.: Darstellung der Parameterlandschaft für den statistischen Parameter β . Farbkodiert dazu jeweils der Wert für β und an den beiden Achsen die jeweils sich ändernden Modellparameter für die Simulationen.

A.3. Abbildungen des Verfahrensvergleichs für alle Patienten

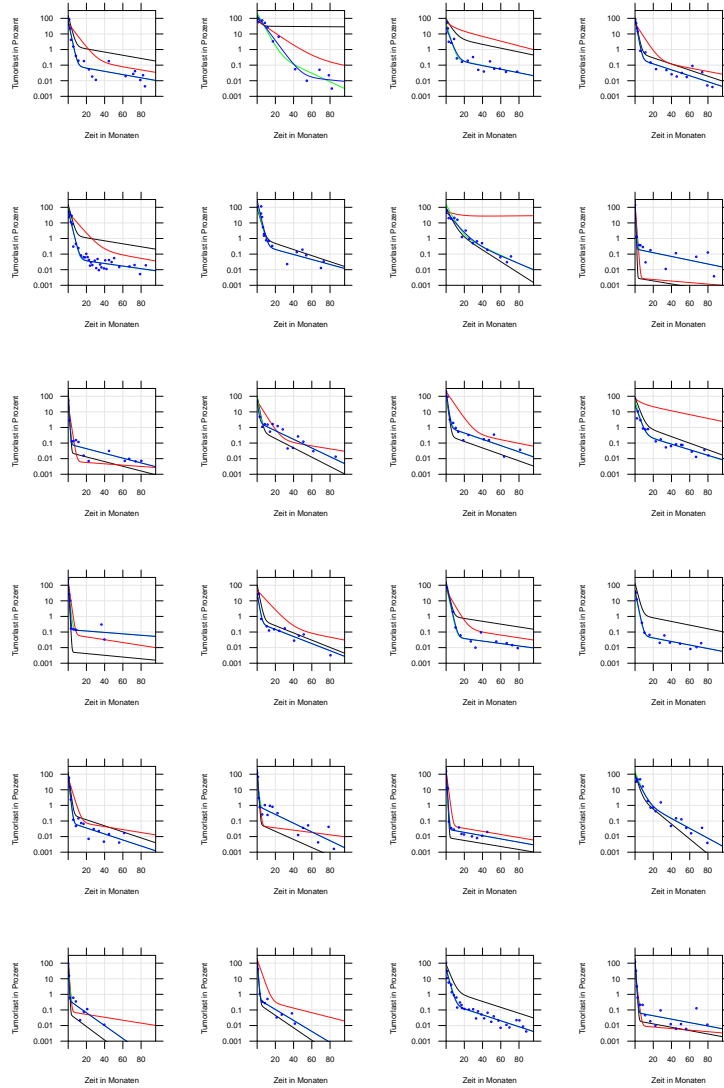


Abbildung A.4.: Verfahrensvergleich für alle verwendeten Patienten der IRIS-Studie.

Auf der y-Achse ist jeweils die Tumorlast, auf der x-Achse die Therapiedauer in Monaten aufgetragen. Blau charakterisiert den realen Therapieverlauf der Patienten (Punkte: Messwerte der Tumorlast im peripheren Blut, Linien: biexponentielle Beschreibung der realen Therapie). Die andersfarbigen Linien verdeutlichen jeweils die biexponentiellen Funktionen der *in silico* simulierten Therapieverläufe, die man erhält, wenn man die Simulation mit dem angegebenen Verfahren ermittelt (schwarz: indirekte Regression nach Horn et al. (2013), rot: direkte Regression, grün: Nutzung der lookup-table).

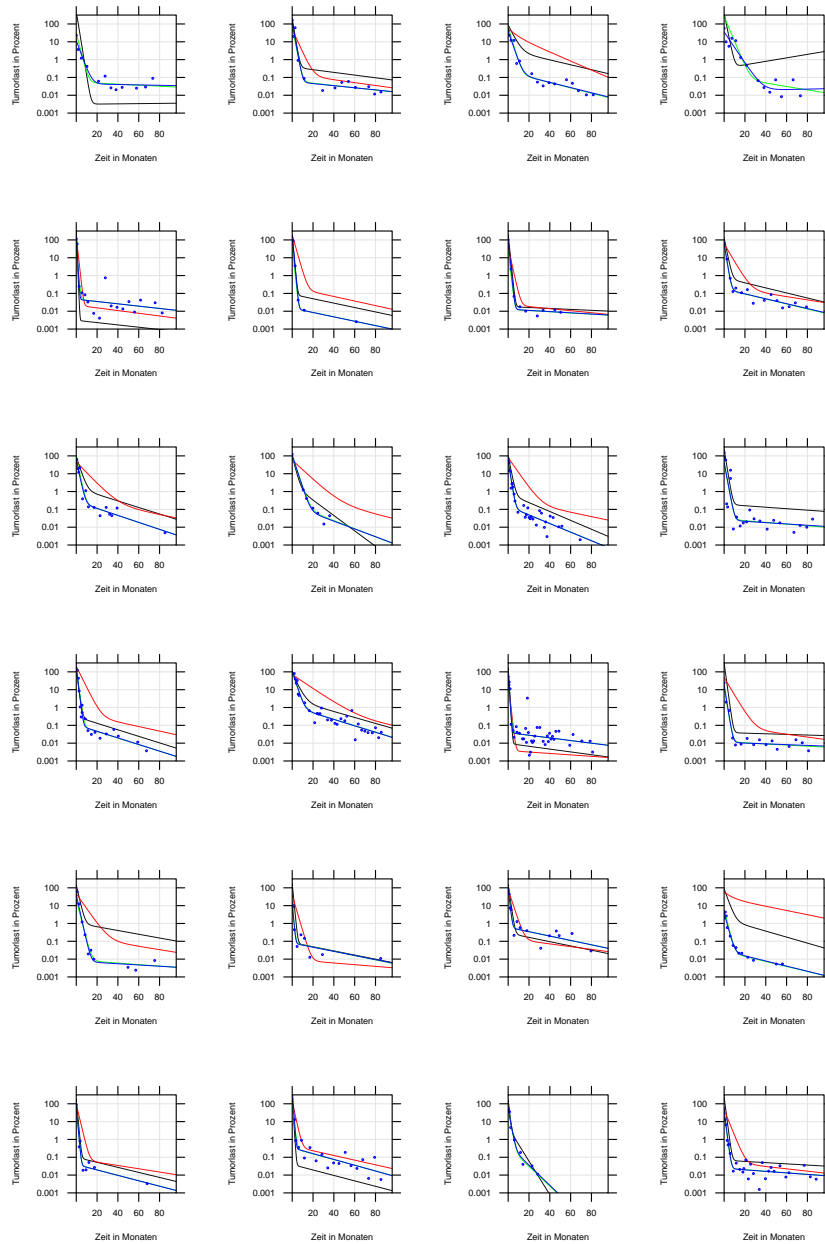


Abbildung A.5.: Fortsetzung der Abbildung zum Verfahrensvergleich für alle verwendeten Patienten der IRIS-Studie. Auf der y-Achse ist jeweils die Tumorklast, auf der x-Achse die Therapiedauer in Monaten aufgetragen. Blau charakterisiert den realen Therapieverlauf der Patienten (Punkte: Messwerte der Tumorklast im peripheren Blut, Linien: biexponentielle Beschreibung der realen Therapie). Die andersfarbigen Linien verdeutlichen jeweils die biexponentiellen Funktionen der *in silico* simulierten Therapieverläufe, die man erhält, wenn man die Simulation mit dem angegebenen Verfahren ermittelt (schwarz: indirekte Regression nach Horn et al. (2013), rot: direkte Regression, grün: Nutzung der lookup-table).

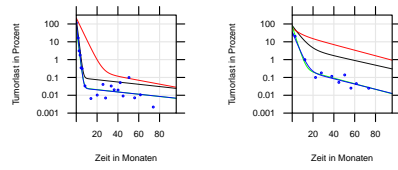


Abbildung A.6.: Fortsetzung der Abbildung zum Verfahrensvergleich für alle verwendeten Patienten der IRIS-Studie. Auf der y-Achse ist jeweils die Tumorlast, auf der x-Achse die Therapiedauer in Monaten aufgetragen. Blau charakterisiert den realen Therapieverlauf der Patienten (Punkte: Messwerte der Tumorlast im peripheren Blut, Linien: biexponentielle Beschreibung der realen Therapie). Die andersfarbigen Linien verdeutlichen jeweils die biexponentiellen Funktionen der *in silico* simulierten Therapieverläufe, die man erhält, wenn man die Simulation mit dem angegebenen Verfahren ermittelt (schwarz: indirekte Regression nach Horn et al. (2013), rot: direkte Regression, grün: Nutzung der lookup-table). Der 51. eingeschlossene Patient ist in Abbildung 6.9a gezeigt.

Literaturverzeichnis

- Abbott LH, Michor F. 2006. Mathematical models of targeted cancer therapy. *Br J Cancer* 95(9):1136–1141 (siehe S. 19).
- Baccarani M, Cortes J, Pane F, Niederwieser D, Saglio G, Apperley J, Cervantes FF, Deininger M, Gratwohl A, Guilhot F, Hochhaus A, Horowitz M, Hughes T, Kantarjian H, Larson R, Radich J, Simonsson B, Silver RT, Goldman JM, Hehlmann R. 2009. Chronic myeloid leukemia: an update of concepts and management recommendations of European LeukemiaNet. *J Clin Oncol* 27(35):6041–6051 (siehe S. 10).
- Baccarani M, Saglio G, Goldman JM, Hochhaus A, Simonsson B, Appelbaum F, Apperley J, Cervantes FF, Cortes J, Deininger M, Gratwohl A, Guilhot F, Horowitz M, Hughes T, Kantarjian H, Larson R, Niederwieser D, Silver R, Hehlmann R. 2006. Evolving concepts in the management of chronic myeloid leukemia: recommendations from an expert panel on behalf of the European LeukemiaNet. *Blood* 108(6):1809–1820 (siehe S. 10).
- Backhaus K, Erichson B, Weiber R. 2015a. *Fortgeschrittene multivariate Analysemethoden*. 3. Auflage. Eine anwendungsorientierte Einführung. Springer, Berlin, Heidelberg (siehe S. 18).
- Backhaus K, Erichson B, Weiber R, Plinke W. 2015b. Nichtlineare Regression. In: *Multivariate Analysemethoden*. Springer, Berlin, Heidelberg, S. 573–580 (siehe S. 18).
- Beal SL. 2001. Ways to fit a PK model with some data below the quantification limit. *J Pharmacokinet Pharmacodyn* 28(5):481–504 (siehe S. 40).
- Bokemeyer C, Schmitz N. 2014. Jahrestagung der Deutschen, Österreichischen und Schweizerischen Gesellschaften für Hämatologie und Medizinische Onkologie. In: Hallek M (Hrsg). *Oncology Reserach and Treatment*. Bd. 37. Supplement 5 (siehe S. 100, 102).
- Busch U, Becker B, Demmel A, Ehrmann M, Geisen R, Hahn A, Huber I, Hubner P, Konrad R, Kostmann D, Made D, Messelhauser U, Pavlovic M, Popping B, Renger S, Schalch B, Schulze M, Stachel C, Unterberger C, Wagner M, Waiblinger HU, Zeltz P.

2010. *Molekularbiologische Methoden in der Lebensmittelanalytik*. Busch U (Hrsg). Springer, Berlin, Heidelberg (siehe S. 9).
- Casella G, Berger RL. 2002. *Statistical Inference*. 2nd Edition. Duxbury Press, Pacific Grove, CA USA (siehe S. 40).
- Clapp GD, Lepoutre T, Nicolini FE, Levy D, Clapp GD, Lepoutre T, Levy D. 2016. BCR-ABL transcript variations in chronic phase chronic myelogenous leukemia patients on imatinib first-line: Possible role of the autologous immune system. *OncImmunity* 5(5):1–3 (siehe S. 72, 75, 77).
- Conti RM, Padula WV, Larson RA. 2015. Changing the cost of care for chronic myeloid leukemia: the availability of generic imatinib in the USA and the EU. *Ann Hematol* 94(2):249–257 (siehe S. 13).
- Cortes JE, Kantarjian H, Shah NP, Bixby D, Mauro MJ, Flinn I, O’Hare T, Hu S, Narasimhan NI, Rivera VM, Clackson T, Turner CD, Haluska FG, Druker BJ, Deininger MWN, Talpaz M. 2012a. Ponatinib in refractory Philadelphia chromosome-positive leukemias. *N Engl J Med* 367(22):2075–2088 (siehe S. 12).
- Cortes JE, Kim DW, Kantarjian HM, Brümmendorf TH, Dyagil I, Griskevicius L, Malhotra H, Powell C, Gogat K, Countouriotis AM, Gambacorti-Passerini C. 2012b. Bosutinib versus imatinib in newly diagnosed chronic-phase chronic myeloid leukemia: results from the BELA trial. *J Clin Oncol* 30(28):3486–3492 (siehe S. 12).
- Cross NCP, White HE, Colomer D, Ehrencrona H, Foroni L, Gottardi E, Lange T, Lion T, Polakova KM, Dulucq S, Martinelli G, Leibundgut EO, Pallisgaard N, Barbany G, Sacha T, Talmaci R, Izzo B, Saglio G, Pane F, Iler MCMu, Hochhaus A. 2015. Laboratory recommendations for scoring deep molecular responses following treatment for chronic myeloid leukemia. *Leukemia* 29(5):999–1003 (siehe S. 25).
- Cross NCP, White HE, Müller MC, Saglio G, Hochhaus A. 2012. Standardized definitions of molecular response in chronic myeloid leukemia. *Leukemia* 26(10):2172–2175 (siehe S. 10).
- Davies RB. 1987. Hypothesis testing when a nuisance parameter is present only under the alternative. *Biometrika* 74(1):33–43 (siehe S. 16).

- Druker BJ, Tamura S, Buchdunger E, Ohno S, Segal GM, Fanning S, Zimmermann J, Lydon NB. 1996. Effects of a selective inhibitor of the Abl tyrosine kinase on the growth of Bcr–Abl positive cells. *Nat Med* 2(5):561–566 (siehe S. 12).
- Essers MAG, Trumpp A. 2010. Targeting leukemic stem cells by breaking their dormancy. *Mol Oncol* 4(5):443–450 (siehe S. 2).
- Experts in Chronic Myeloid Leukemia. 2013. The price of drugs for chronic myeloid leukemia (CML) is a reflection of the unsustainable prices of cancer drugs: from the perspective of a large group of CML experts. *Blood* 121(22):4439–4442 (siehe S. 13).
- Fahrmeir L, Kneib T, Lang S. 2009. Regressionsmodelle. In: Dette H, Härdle W (Hrsg). *Regression (Statistik und ihre Anwendungen)*. Springer, Berlin, Heidelberg, S. 19–58 (siehe S. 16).
- Gabert J, Beillard E, Velden VHJ van der, Bi W, Grimwade D, Pallisgaard N, Barbany G, Cazzaniga G, Cayuela JM, Cavé H, Pane F, Aerts JLE, De Micheli D, Thirion X, Pradel V, González M, Viehmann S, Malec M, Saglio G, Dongen JJM van. 2003. Standardization and quality control studies of 'real-time' quantitative reverse transcriptase polymerase chain reaction of fusion gene transcripts for residual disease detection in leukemia - a Europe Against Cancer program. *Leukemia* 17(12):2318–2357 (siehe S. 8).
- Glauche I. 2010. Theoretical studies on the lineage specification of hematopoietic stem cells. Dissertation. Universität Leipzig (siehe S. 30).
- Glauche I, Baldow C, Fröhlich S, Schulze P, Roy A, Subar M, Wang X, Roeder I. 2014. Model-Based Characterization of the Molecular Response Dynamics of Tyrosine Kinase Inhibitor (TKI)-Treated CML Patients – a Comparison of Imatinib and Dasatinib First-Line Therapy. *Blood* 124(21):4562–4562 (siehe S. 100, 102).
- Goldman JM, Daley GQ. 2007. Chronic Myeloid Leukemia — A Brief History. In: Melo JV, Goldman JM (Hrsg). *Myeloproliferative Disorders*. Springer Berlin Heidelberg, Berlin, Heidelberg, S. 1–13 (siehe S. 11).
- Guevara G, González JA, Lopera DE, González M, Saavedra JD, Lobaton JF, Duque JE. 2012. Follow-up of the tumor load in patients with de novo chronic myeloid leukemia and in complete cytogenetic remission treated with imatinib in Colombia. *Colomb. Med.* 43(4):267–272 (siehe S. 68).

- Gunsilius E, Duba HC, Petzer AL, Kähler CM, Grünewald K, Stockhammer G, Gabl C, Dirnhofer S, Clausen J, Gastl G. 2000. Evidence from a leukaemia model for maintenance of vascular endothelium by bone-marrow-derived endothelial cells. *Lancet* 355(9216):1688–1691 (siehe S. 5).
- Hanfstein B, Shlyakhto V, Lauseker M, Hehlmann R, Saussele S, Dietz C, Erben P, Fabarius A, Proetel U, Schnittger S, Krause SW, Schubert J, Einsele H, Hänel M, Dengler J, Falge C, Kanz L, Neubauer A, Kneba M, Stegelmann F, Pfreundschuh M, Waller CF, Spiekermann K, Baerlocher GM, Pffirmann M, Hasford J, Hofmann WK, Hochhaus A, Müller MC. 2014. Velocity of early BCR-ABL transcript elimination as an optimized predictor of outcome in chronic myeloid leukemia (CML) patients in chronic phase on treatment with imatinib. *Leukemia* 28(10):1988–1992 (siehe S. 8, 40).
- Hasford J, Baccarani M, Hoffmann V, Guilhot J, Saussele S, Rosti G, Guilhot F, Porkka K, Ossenkoppele G, Lindoerfer D, Simonsson B, Pffirmann M, Hehlmann R. 2011. Predicting complete cytogenetic response and subsequent progression-free survival in 2060 patients with CML on imatinib treatment: the EUTOS score. *Blood* 118(3):686–692 (siehe S. 7).
- Hasford J, Pffirmann M, Hehlmann R, Allan NC, Baccarani M, Kluin-Nelemans JC, Ali-mena G, Steegmann JL, Ansari H. 1998. A new prognostic score for survival of patients with chronic myeloid leukemia treated with interferon alfa. Writing Committee for the Collaborative CML Prognostic Factors Project Group. *J Natl Cancer Inst* 90(11):850–858 (siehe S. 8).
- Hehlmann R, Hochhaus A. 2010. Chronische myeloische Leukämie. In: Hiddemann W, Bartram C (Hrsg). *Die Onkologie*. Springer, Berlin, Heidelberg, S. 1703–1721 (siehe S. 1, 3, 5, 6, 9–13, 37).
- Henningsen A, Toomet O. 2010. maxLik: A package for maximum likelihood estimation in R. *Comput Stat* 26(3):443–458 (siehe S. 40).
- Hochhaus A, O’Brien SG, Guilhot F, Druker BJ, Branford S, Foroni L, Goldman JM, Müller MC, Radich JP, Rudoltz M, Mone M, Gathmann I, Hughes TP, Larson RA. 2009. Six-year follow-up of patients receiving imatinib for the first-line treatment of chronic myeloid leukemia. *Leukemia* 23(6):1054–1061 (siehe S. 24).
- Hochhaus A, Reiter A, Saussele S, Reichert A, Emig M, Kaeda J, Schultheis B, Berger U, Shepherd PC, Allan NC, Hehlmann R, Goldman JM, Cross NC. 2000. Molecular heterogeneity in complete cytogenetic responders after interferon-alpha therapy for chronic

- myelogenous leukemia: low levels of minimal residual disease are associated with continuing remission. *Blood* 95(1):62–66 (siehe S. 11).
- Hochhaus A, Baerlocher GM, Brümmendorf TH, Chalandon Y, Dölken G, Thiede C, Wolf D, Coutre P le. 2013. *Chronische Myeloische Leukämie (CML)*. URL: <https://www.onkopedia.com/de/onkopedia/guidelines/chronische-myeloische-leukaemie-cml/@@view/html/index.html> (Aufruf am: 31.05.2016) (siehe S. 7, 10).
- Höfer T, Rodewald HR. 2016. ScienceDirect Output without input: the lifelong productivity of hematopoietic stem cells. *Current Opinion in Cell Biology* 43:69–77 (siehe S. 69).
- Horn M. 2014. Optimierung der Therapie von chronischer myeloischer Leukämie mit Hilfe eines dynamischen Modells normaler und leukämischer Stammzellorganisation. Dissertation. Universität Leipzig, Medizinische Fakultät (siehe S. 20, 27).
- Horn M, Glauche I, Müller MC, Hehlmann R, Hochhaus A, Loeffler M, Roeder I. 2013. Model-based decision rules reduce the risk of molecular relapse after cessation of tyrosine kinase inhibitor therapy in chronic myeloid leukemia. *Blood* 121(2):378–384 (siehe S. 2, 20–22, 24, 30, 37, 38, 40, 42, 44, 46, 53, 54, 56–61, 63, 66, 67, 71–74, 76, 77, 82–84).
- Horn M, Loeffler M, Roeder I. 2008. Mathematical modeling of genesis and treatment of chronic myeloid leukemia. *Cells Tissues Organs* 188(1-2):236–247 (siehe S. 19, 44).
- Hughes TP, Kaeda J, Branford S, Rudzki Z, Hochhaus A, Hensley ML, Gathmann I, Bolton AE, Hoomissen IC van, Goldman JM, Radich JP. 2003. Frequency of Major Molecular Responses to Imatinib or Interferon Alfa plus Cytarabine in Newly Diagnosed Chronic Myeloid Leukemia. *N Engl J Med* 349(15):1423–1432 (siehe S. 14).
- Hughes T, Deininger M, Hochhaus A, Branford S, Radich J, Kaeda J, Baccarani M, Cortes J, Cross NCP, Druker BJ, Gabert J, Grimwade D, Hehlmann R, Kamel-Reid S, Lipton JH, Longtine J, Martinelli G, Saglio G, Soverini S, Stock W, Goldman JM. 2006. Monitoring CML patients responding to treatment with tyrosine kinase inhibitors: review and recommendations for harmonizing current methodology for detecting BCR-ABL transcripts and kinase domain mutations and for expressing results. *Blood* 108(1):28–37 (siehe S. 8).
- Ilander M, Hekim C, Mustjoki S. 2014. Immunology and Immunotherapy of Chronic Myeloid Leukemia. *Curr Hematol Malig Rep* 9(1):17–23 (siehe S. 72).

- Jabbour E, Kantarjian HM, Saglio G, Steegmann JL, Shah NP, Boqué C, Chuah C, Pavlovsky C, Mayer J, Cortes J, Baccarani M, Kim DW, Bradley-Garelik MB, Mohamed H, Wildgust M, Hochhaus A. 2014. Early response with dasatinib or imatinib in chronic myeloid leukemia: 3-year follow-up from a randomized phase 3 trial (DASISION). *Blood* 123(4):494–500 (siehe S. 12).
- Kaatsch P, Spix C, Hentschel S, Katalinic A, Luttmann S, Stegmaier C, Caspritz S, Cernaj J, Ernst A, Folkerts J, Hansmann J, Kranzhöfer K, Krieghoff-Henning E, Kunz B, Penzkofer A, Treml K, Wittenberg K, Baras N, Barnes B, Bertz J, Buttman-Schweiger N, Dahm S, Franke M, Haberland J, Kraywinkel K, Wienecke A, Wolf U. 2013. *Krebs in Deutschland 2009/2010*. Robert Koch-Institut, Gesellschaft der epidemiologischen Krebsregister in Deutschland e.V. (Hrsg). 9. Ausgabe. Berlin, Lübeck (siehe S. 5).
- Kampstra P. 2008. Beanplot: A Boxplot Alternative for Visual Comparison of Distributions. *J Stat Softw* 28(Code Snippet 1):1–9 (siehe S. 26).
- Kantarjian HM, Shah NP, Cortes JE, Baccarani M, Agarwal MB, Undurraga MS, Wang J, Ipiña JJK, Kim DW, Ogura M, Pavlovsky C, Junghanss C, Milone JH, Nicolini FE, Robak T, Van Droogenbroeck J, Vellenga E, Bradley-Garelik MB, Zhu C, Hochhaus A. 2012. Dasatinib or imatinib in newly diagnosed chronic-phase chronic myeloid leukemia: 2-year follow-up from a randomized phase 3 trial (DASISION). *Blood* 119(5):1123–1129 (siehe S. 12).
- Kooten Niekerk PB van, Petersen CC, Nyvold CG, Ommen HB, Roug AS, Nederby L, Hokland P, Kjeldsen E. 2013. Cell sorting enables interphase fluorescence in situ hybridization detection of low BCR-ABL1 producing stem cells in chronic myeloid leukaemia patients beyond deep molecular remission. *Br J Haematol* 164(1):53–60 (siehe S. 71).
- Leitner AA, Hehlmann R. 2011. Modern therapy of chronic myeloid leukemia. *Internist* 52(2):209–217 (siehe S. 3, 8–11).
- Loeffler M, Roeder I. 2002. Tissue stem cells: definition, plasticity, heterogeneity, self-organization and models – a conceptual approach. *Cells Tissues Organs* 171(1):8–26 (siehe S. 3, 30, 69).
- Mahon FX. 2015. Discontinuation of tyrosine kinase therapy in CML. *Ann Hematol* 94(Suppl 2):187–193 (siehe S. 10).

- Mahon FX, Réa D, Guilhot J, Guilhot F, Huguet F, Nicolini F, Legros L, Charbonnier A, Guerci A, Varet B, Etienne G, Reiffers J, Rousselot P. 2010. Discontinuation of imatinib in patients with chronic myeloid leukaemia who have maintained complete molecular remission for at least 2 years: the prospective, multicentre Stop Imatinib (STIM) trial. *Lancet Oncology* 11(11):1029–1035 (siehe S. 61, 71).
- Mensink E, Locht A van de, Schattenberg A, Linders E, Schaap N, Geurts van Kessel A, De Witte T. 1998. Quantitation of minimal residual disease in Philadelphia chromosome positive chronic myeloid leukaemia patients using real-time quantitative RT-PCR. *Br J Haematol* 102(3):768–774 (siehe S. 8).
- Michor F, Beal K. 2015. Improving Cancer Treatment via Mathematical Modeling: Surmounting the Challenges Is Worth the Effort. *Cell* 163(5):1059–1063 (siehe S. 1).
- Michor F, Hughes TP, Iwasa Y, Branford S, Shah NP, Sawyers CL, Nowak MA. 2005. Dynamics of chronic myeloid leukaemia. *Nature* 435(7046):1267–1270 (siehe S. 15, 18, 19, 27).
- Mustjoki S, Richter J, Barbany G, Ehrencrona H, Fioretos T, Gedde-Dahl T, Gjertsen BT, Hovland R, Hernesniemi S, Josefsen D, Koskenvesa P, Dybedal I, Markevarn B, Olofsson T, Olsson-Stromberg U, Rapakko K, Thunberg S, Stenke L, Simonsson B, Porkka K, Hjorth-Hansen H. 2013. Impact of malignant stem cell burden on therapy outcome in newly diagnosed chronic myeloid leukemia patients. *Leukemia* 27(7):1520–1526 (siehe S. 68, 70).
- O’Brien SG, Guilhot F, Larson RA, Gathmann I, Baccarani M, Cervantes FF, Cornelissen JJ, Fischer T, Hochhaus A, Hughes T, Lechner K, Nielsen JL, Rousselot P, Reiffers J, Saglio G, Shepherd J, Simonsson B, Gratwohl A, Goldman JM, Kantarjian H, Taylor K, Verhoef G, Bolton AE, Capdeville R, Druker BJ. 2003. Imatinib compared with interferon and low-dose cytarabine for newly diagnosed chronic-phase chronic myeloid leukemia. *N Engl J Med* 348(11):994–1004 (siehe S. 24).
- Potten CS, Loeffler M. 1990. Stem cells: attributes, cycles, spirals, pitfalls and uncertainties. Lessons for and from the crypt. *Development* 110(4):1001–1020 (siehe S. 69).
- Price KE, Saleem N, Lee G, Steinberg M. 2012. Potential of ponatinib to treat chronic myeloid leukemia and acute lymphoblastic leukemia. *Onco Targets Ther* 6:1111–1118 (siehe S. 12).

- Proschmann R, Baldow C, Rothe T, Suttorp M, Thiede C, Tauer JT, Müller MC, Hochhaus A, Roeder I, Glauche I. 2016. Response dynamics of pediatric patients with chronic myeloid leukemia on imatinib therapy. *Haematologica* (siehe S. 100, 102).
- Regn S. 2002. Simultane Expansion zytotoxischer T-Lymphozyten gerichtet gegen Adeno- und Epstein-Barr-Virus-Epitope, zur Therapie von opportunistischen Infektionen nach allogener hämatopoetischer Stammzelltransplantation (HSCT). Dissertation. Technische Universität München (siehe S. 4).
- Roeder I. 2003. Dynamic Modeling of Hemato-poietic Stem Cell Organization. Dissertation. Universität Leipzig, Medizinische Fakultät (siehe S. 32).
- Roeder I, Horn M, Glauche I, Hochhaus A, Mueller MC, Loeffler M. 2006. Dynamic modeling of imatinib-treated chronic myeloid leukemia: functional insights and clinical implications. *Nat Med* 12(10):1181–1184 (siehe S. 19, 30).
- Roeder I, Loeffler M. 2002. A novel dynamic model of hematopoietic stem cell organization based on the concept of within-tissue plasticity. *Exp. Hematol.* 30(8):853–861 (siehe S. 2, 18, 19, 30, 70, 71, 73–75, 77).
- Ross DM, Branford S, Seymour JF, Schwarzer AP, Arthur C, Bartley PA, Slader C, Field C, Dang P, Filshie RJ, Mills AK, Grigg AP, Melo JV, Hughes TP. 2010. Patients with chronic myeloid leukemia who maintain a complete molecular response after stopping imatinib treatment have evidence of persistent leukemia by DNA PCR. *Leukemia* 24(10):1719–1724 (siehe S. 61, 72).
- Saglio G, Kim DW, Issaragrisil S, Coutre P le, Etienne G, Lobo C, Pasquini R, Clark RE, Hochhaus A, Hughes TP, Gallagher N, Hoenekopp A, Dong M, Haque A, Larson RA, Kantarjian HM. 2010. Nilotinib versus imatinib for newly diagnosed chronic myeloid leukemia. *N Engl J Med* 362(24):2251–2259 (siehe S. 12).
- Saussele S, Richter J, Hochhaus A, Mahon FX. 2016. The concept of treatment-free remission in chronic myeloid leukemia. *Leukemia* 30(8):1638–1647 (siehe S. 72, 75, 77).
- Sawyers C. 2004. Targeted cancer therapy. *Nature* 432(7015):294–297 (siehe S. 1).
- Schofield R. 1978. The relationship between the spleen colony-forming cell and the haemopoietic stem cell. *Blood Cells* 4(1-2):7–25 (siehe S. 30).

- Shtivelman E, Lifshitz B, Gale RP, Canaani E. 1985. Fused transcript of abl and bcr genes in chronic myelogenous leukaemia. *Nature* 315(6020):550–554 (siehe S. 5).
- Stein AM, Martinelli G, Hughes TP, Müller MC, Beppu L, Gottardi E, Branford S, Soverini S, Woodman RC, Hochhaus A, Kim DW, Saglio G, Radich JP. 2013. Rapid initial decline in BCR-ABL1 is associated with superior responses to second-line nilotinib in patients with chronic-phase chronic myeloid leukemia. *BMC Cancer* 13(1):173 (siehe S. 17, 22).
- Uchida N, Fleming WH, Alpern EJ, Weissman IL. 1993. Heterogeneity of hematopoietic stem cells. *Curr Opin Immunol* 5(2):177–184 (siehe S. 3).
- Velden VHJ van der, Hochhaus A, Cazzaniga G, Szczepanski T, Gabert J, Dongen JJM van. 2003. Detection of minimal residual disease in hematologic malignancies by real-time quantitative PCR: principles, approaches, and laboratory aspects. *Leukemia* 17(6):1013–1034 (siehe S. 8).
- Wilson A, Laurenti E, Oser G, Wath RC van der, Blanco-Bose W, Jaworski M, Offner S, Dunant CF, Eshkind L, Bockamp E, Lió P, Macdonald HR, Trumpp A. 2008. Hematopoietic stem cells reversibly switch from dormancy to self-renewal during homeostasis and repair. *Cell* 135(6):1118–1129 (siehe S. 3).

Abbildungsverzeichnis

2.1. Schematische Darstellung der Blutbildung	4
2.2. Reziproke Translokation zwischen den langen Armen der Chromosomen 9 und 22	6
3.1. Beispielhafter Therapieverlauf eines Patienten mit CML	15
3.2. Beispielhafter Therapieverlauf mit zusätzlich zur Abbildung 3.1b eingezeichneten Beschreibungen durch zwei verschiedene Anpassungsverfahren	16
4.1. Übersichtsabbildung der verwendeten Methoden	23
5.1. Alter und Tumorlast der IRIS-Patienten bei Diagnosestellung	26
5.2. Beobachtungsdauer, Anzahl der Messungen und Anzahl der Messungen über 10%	28
5.3. Darstellung des medianen Patienten der IRIS-Studie unter Behandlung mit Imatinib	29
5.4. Stammzellmodell mit normaler Hämatopoese	31
5.5. Übergangsfunktionen f_α und f_ω	33
5.6. Stammzellmodell mit leukämischen Stammzellen in der Entstehung einer CML	34
5.7. Stammzellmodell mit fast verdängter gesunder Hämatopoese	35
5.8. Stammzellmodell bei der Simulation einer TKI-Therapie	36
6.1. Vergleich der Anpassungsverfahren der biexponentiellen Beschreibung	41
6.2. Vergleich der Parameter der biexponentiellen Beschreibung aller IRIS-Patienten des deutschen Armes der Studie unter Nutzung unterschiedlicher Anpassungsverfahren	43
6.3. Vorab-Untersuchungen zur Abschätzung des Einflusses unterschiedlicher dynamischer Simulationsparameter auf die statistischen Parameter	45
6.4. Vergleich der Datenparameter der Pilotsimulationen (Ecken des Parameter-raumes) mit denen der IRIS-Patienten	47
6.5. Darstellung der Parameterlandschaft für den statistischen Parameter A	50
6.6. Ausschnitt des linken unteren Feldes der Abbildung A.1 zur besseren Veranschaulichung	51

6.7. Vergleich der Datenparameter nach biexponentieller Auswertung aller Simulationen des Parameterraumes mit den Datenparametern der IRIS-Patienten . . .	52
6.8. Beispielhafte Rangliste zur Ermittlung der Simulation mit dem geringsten Abstand zum Therapieverlauf eines Patienten	56
6.9. Vergleich der unterschiedlichen Anpassungsverfahren der IRIS-Patienten . . .	57
6.10. Vergleich der unterschiedlichen Anpassungsverfahren für beide Therapieabschnitte	58
6.11. Beispielpatient mit dazu passenden <i>in silico</i> Therapiesimulationen	62
6.12. Verläufe der Tumorlast im Stammzellkompartiment für die 1000 <i>in silico</i> Therapiesimulationen, deren Verlauf der Tumorlast im peripheren Blut am nächsten am Therapieverlauf des Beispielpatienten aus Abbildung 6.11 liegen	63
6.13. Verteilung des simulierten Therapieendes (entspricht 200 leukämischen Stammzellen) für die 1000 <i>in silico</i> Therapieverläufe, deren Verlauf der Tumorlast im peripheren Blut am nächsten am Therapieverlauf des Beispielpatienten aus Abbildung 6.11 liegen	64
6.14. Die 100 am nächsten am Patienten aus Abbildung 6.11 liegenden Tumorlastverläufe auf Stammzellebene nach den konkreten Werten der unterschiedlichen Modellparameter eingefärbt	65
6.15. Kurve, die den Anteil aller IRIS-Patienten des deutschen Armes ($n = 69$) zeigt, deren Simulation zu einem bestimmten Zeitpunkt die Zahl von 200 leukämischen Stammzellen erreicht haben	67
A.1. Darstellung der Parameterlandschaft für den statistischen Parameter α	79
A.2. Darstellung der Parameterlandschaft für den statistischen Parameter B	80
A.3. Darstellung der Parameterlandschaft für den statistischen Parameter β	81
A.4. Verfahrensvergleich für alle verwendeten Patienten der IRIS-Studie.	82
A.5. Fortsetzung der Abbildung zum Verfahrensvergleich für alle verwendeten Patienten der IRIS-Studie.	83
A.6. Fortsetzung der Abbildung zum Verfahrensvergleich für alle verwendeten Patienten der IRIS-Studie.	84

Tabellenverzeichnis

2.1. Definitionen des Ansprechens auf eine Therapie mit TKI.	10
6.1. Grenzen des Parameterraumes	46
6.2. Mittelwerte der mittleren Abstände zwischen Patientendaten und den Simulationen, die mittels unterschiedlicher Verfahren ermittelt wurden.	59
A.1. Genutzte Parameterkonfiguration des Stammzellmodells	78

Danksagungen

Schreiben muss man eine solche Arbeit zwar allein, es gibt aber viele Personen, die maßgeblich an ihrem Gelingen beteiligt waren. Diesen Menschen möchte ich an dieser Stelle meinen Dank aussprechen.

An erster Stelle sei mein Doktorvater Prof. Dr. rer. med. Ingo Röder genannt, der mir das Thema angeboten hat, nachdem das zuerst in Auge gefasste Themenfeld nicht mehr zur Verfügung stand. Er hatte immer ein offenes Ohr für meine Fragen und hat mich mit vielen unschätzbaren Hinweisen unterstützt.

Ein großer Dank gilt auch meinem Mitbetreuer Dr. rer. med. Ingmar Glauche, der immer kurzfristig für mich zu erreichen war. Er hat mich stets mit konstruktiver Kritik unterstützt und mir geholfen unbewältigbar scheinende Probleme zu überwinden. Ich danke ihm, dass er mich auf eine so herausragende Weise betreut hat.

Weiterhin danken möchte ich Christoph Baldow, ohne den die programmiertechnische Umsetzung der Arbeit so nicht möglich gewesen wäre. Er hat in vielen Stunden Arbeit die Simulationen am Zentrum für Informationsdienst und Hochleistungsrechnen der TU Dresden (ZIH) begleitet, mich bei der Auswertung unterstützt und den Anstoß für die neu entwickelten Verfahren der Therapieidentifikation gegeben. Ich danke ihm für die vielen schnell beantworteten Fragen, die oft gelesenen Manuskripte und die vielen Anregungen, welche die Arbeit maßgeblich beeinflusst haben.

Auch den anderen Kollegen des Institutes für medizinische Informatik und Biometrie möchte ich für Ihre Unterstützung danken. Namentlich genannt seien Lars Thielecke und Matthias Kuhn, die mir in der Programmierung des Modells und der mathematischen Auswertung der Daten halfen.

Des weiteren danken möchte ich Matthias Horn, der mit seiner Arbeit den Grundstein für die vorliegenden Betrachtungen geschaffen hat. Er überließ mir seine Skripte und Daten und hat mir damit die Einarbeitung vereinfacht.

Außerdem gilt mein Dank dem ZIH für die unkomplizierte Bereitstellung der Hochleistungsrechenkapazitäten, die überhaupt erst die Erstellung der Datengrundlage für die vorliegende Arbeit ermöglichten.

Ein weiterer Dank gilt Volker Fickert, meinem früheren Lehrer für Informatik und Physik, der den Grundstein für meine Arbeit an und mit Computern gelegt hat. Ihm verdanke ich die Nutzung von \LaTeX , die zwar viel Einarbeitung brauchte, mich jedoch vor vielen Problemen bewahrt hat.

Ich danke meiner Familie und meinen Schwiegereltern für den Rückhalt, den ich in all den Jahren des Studiums von ihnen erhalten habe. Sie standen stets mit gutem Rat an meiner Seite.

Mein größter Dank gilt meiner Verlobten Christiane Strobel für ihren Beistand beim Verfassen dieser Arbeit. Sie hat mich angespornt, zum Durchhalten animiert und ist immer für mich da. Ich freue mich auf unsere gemeinsame Zukunft.

Erklärungen zur Eröffnung des Promotionsverfahrens

1. Hiermit versichere ich, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe; die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht.

2. Bei der Auswahl und Auswertung des Materials sowie bei der Herstellung des Manuskripts habe ich Unterstützungsleistungen von folgenden Personen erhalten:

Christoph Baldow,
Ingmar Glauche,
Matthias Horn,
Matthias Kuhn,
Rick Proschmann,
Ingo Röder,
Lars Thielecke

3. Weitere Personen waren an der geistigen Herstellung der vorliegenden Arbeit nicht beteiligt. Insbesondere habe ich nicht die Hilfe eines kommerziellen Promotionsberaters in Anspruch genommen. Dritte haben von mir weder unmittelbar noch mittelbar geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen.

4. Die Arbeit wurde bisher weder im Inland noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

5. Die Inhalte der Dissertation flossen in folgende Veröffentlichungen ein:

- Wilsede Meeting 2014, Poster S-I-04: On the trace of hidden cells: Application of mathematical modelling to predict the outcome of individual CML patients under TKI treatment

- Jahrestagung der Deutschen, Österreichischen und Schweizerischen Gesellschaften für Hämatologie und Medizinische Onkologie 2014; freier Vortrag V667: On the trace of hidden cells: Application of mathematical modelling to predict the outcome of individual CML patients under TKI treatment (Abstract siehe Bokemeyer und Schmitz (2014))
- ASH 2014, Poster: Model-Based Characterization of the Molecular Response Dynamics of Tyrosine Kinase Inhibitor (TKI)-Treated CML Patients – a Comparison of Imatinib and Dasatinib First-Line Therapy (Abstract siehe Glauche et al. (2014))
- Haematologica: Response dynamics of paediatric patients with chronic myeloic leukaemia on imatinib therapy: statistical analysis and model-based predictions (Proschmann et al. 2016)

6. Ich bestätige, dass es keine zurückliegenden erfolglosen Promotionsverfahren gab.

7. Ich bestätige, dass ich die Promotionsordnung der Medizinischen Fakultät der Technischen Universität Dresden anerkenne.

8. Ich habe die Zitierrichtlinien für Dissertationen an der Medizinischen Fakultät der Technischen Universität Dresden zur Kenntnis genommen und befolgt.

9. Ich bestätige die Einhaltung der folgenden aktuellen gesetzlichen Vorgaben im Rahmen meiner Dissertation:

- Die Einhaltung von Datenschutzbestimmungen der Medizinischen Fakultät und des Universitätsklinikums Carl Gustav Carus.
- Ein Votum der Ethikkommission sowie die Einhaltung von Bestimmungen des Tierschutzgesetzes und des Gentechnikgesetzes waren im Rahmen meiner Dissertation nicht notwendig.

Tino Rothe

Dresden, den 20.12.2016