

Auswertung und Planung von Dosis-Findungsstudien mit aktiver Kontrolle

Inauguraldissertation
zur Erlangung des akademischen Grades
eines Doktors der Naturwissenschaften
der Universität Mannheim

vorgelegt von
Hans-Joachim Helms
aus Celle

Mannheim, 2014

Dekan: Prof. Dr. Heinz Jürgen Müller, Universität Mannheim
Referent: Prof. Dr. Martin Schlather, Universität Mannheim
Korreferent: Prof. Dr. Tim Friede, Universität Göttingen

Tag der mündlichen Prüfung: 04.06.2014

Zusammenfassung

In der klinischen Entwicklung von Arzneimitteln spielt die Dosis-Findung eine wichtige Rolle. Die vorliegende Arbeit fokussiert sich auf die Auswertung und Planung von Dosis-Findungsstudien mit aktiver Kontrolle. In einer Dosis-Findungsstudie mit aktiver Kontrolle wird zusätzlich zu den verschiedenen Dosisstufen der neuen Substanz ein bereits zugelassenes Medikament, die sogenannte aktive Kontrolle, untersucht. Der Fokus liegt dabei auf der Bestimmung des Dosis-Bereiches der neuen Substanz, der eine überlegene Wirkung gegenüber der aktiven Kontrolle sicherstellt. In diesem Zusammenhang ist es von primärem Interesse die kleinste Dosis zu bestimmen, die zum gleichen erwarteten Wirkungsgrad wie die aktive Kontrolle führt. Zusätzlich zu der Schätzung der Zieldosis werden auch Konfidenzintervalle der Zieldosis benötigt, um die Güte der Schätzung beschreiben zu können. Neben der Auswertung ist die Planung derartiger Studien von großem Interesse. Besonders die Berechnung der benötigten Fallzahl und die Bestimmung der zu untersuchenden Dosisstufen stellen unter praktischen Gesichtspunkten eine große Herausforderung dar. Die vorliegende Arbeit behandelt die verschiedenen Problemstellungen in drei Abschnitten.

Zuerst wird das statistische Modell eingeführt und die notwendigen Parameterschätzer sowie deren asymptotische Eigenschaften basierend auf der Maximum-Likelihood Theorie für allgemeine Dosis-Wirkungskurven beschrieben. Der erste Abschnitt befasst sich mit der Schätzung der Zieldosis und der Bestimmung von zugehörigen Konfidenzintervallen im Falle von linearen Wirkungskurven. Durch die Annahme einer linearen Wirkungsfunktion ist es möglich, Resultate aus der Theorie der linearen Modelle auf das Problem der Dosis-Findung zu übertragen. Dies ermöglicht eine übersichtliche Darstellung des statistischen Modells sowie eine explizite Berechnung der Parameterschätzer und ihrer Verteilungen. Basierend auf diesen Ergebnissen werden vier verschiedene Verfahren zur Berechnung eines Konfidenzintervalls der Zieldosis hergeleitet und durch Simulationsstudien miteinander verglichen. Dabei liegt der Schwerpunkt auf der Untersuchung der Überdeckungswahrscheinlichkeit und der medianen Länge des Konfidenzintervalls. Es kann gezeigt werden, dass alle Verfahren die Überdeckungswahrscheinlichkeit gut einhalten, wobei das Konfidenzintervall basierend auf der Normalapproximation der Δ -Regel zu konservativeren Ergebnissen führt.

Mit dem Fall einer unbekanntem nichtlinearen Wirkungsfunktion beschäftigt sich der zweite Abschnitt. Um nicht von der Wahl der zugrundeliegenden Wirkungsfunktion abhängig zu sein, wird die Schätzung der Zieldosis für unbekanntem Dosis-Wirkungsfunktionen mit Hilfe linearer und kubischer Spline-Interpolationen hergeleitet und untersucht. Zur Konstruktion eines Konfidenz-

intervalls der Zieldosis wird ein parametrischer Bootstrap in Kombination mit den Interpolationsverfahren vorgestellt. Die Eigenschaften der Spline-Verfahren werden in Simulationsstudien mit parametrischen Regressionen verglichen. Die zum Teil substantielle Verzerrung der Spline-basierten Zieldosissschätzer, die nicht asymptotisch gegen null geht, kann zu Problemen führen und die Überdeckungswahrscheinlichkeit der Interpolationsverfahren negativ beeinflussen.

Um dieses Problem lösen zu können, werden im dritten Abschnitt die Aspekte der Designplanung von Dosis-Findungsstudien mit dem Fokus auf Dosisallokation untersucht. Es werden verschiedene Bedingungen zur Berechnung einer günstigen Dosisallokation vorgestellt und in einem neuen Kriterium kombiniert. Dieses Kriterium bestimmt für die gewünschte Spline-Interpolation die Dosisallokation, die die maximale absolute Verzerrung des Zieldosissschätzers minimiert. Zusätzlich wird in diesem Abschnitt die Berechnung einer Fallzahlplanung für Dosis-Findungsstudien mit aktiver Kontrolle behandelt. Um dabei die Präzision des Zieldosissschätzers in der Planungsphase kontrollieren zu können, werden zwei verschiedene Kriterien zur Fallzahlplanung untersucht, die die Breite des Konfidenzintervalls der Zieldosis beschränken. Im Falle der Fallzahl N_E soll die Intervallbreite in Erwartung nicht breiter als die vordefinierte Konstante $2c$ sein. Für die Fallzahl N_γ soll die Intervallbreite mit einer Wahrscheinlichkeit von γ nicht breiter als die Konstante $2c$ sein. Basierend auf der Regression mittels Maximum-Likelihood Methode und dem Konfidenzintervall der Δ -Regel werden die Fallzahlen N_E und N_γ für beliebige Dosis-Wirkungskurven hergeleitet und für lineare Dosis-Wirkungsfunktionen explizit berechnet. In Simulationsstudien kann gezeigt werden, dass die Fallzahl N_E die mittlere Breite des Konfidenzintervalls sehr gut kontrolliert. Ebenfalls führt die Fallzahl N_γ in $100 \cdot \gamma\%$ der Simulationen zu Intervallbreiten $\leq 2c$. Beide Fallzahlformeln erfüllen ihre jeweiligen Bedingungen und sind somit geeignet, die Güte der Zieldosissschätzung in der Planungsphase zu kontrollieren.

Danksagung

Die vorliegende Arbeit entstand in Kooperation mit dem Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM) im Rahmen des Projektes „Auswertung und Bewertung von Dosis-Findungsstudien mit dem Ziel der modellbasierten Dosiskalibrierung“. Für die Förderung und gute Zusammenarbeit möchte ich mich herzlich bedanken.

Mein Dank gilt Prof. Dr. Martin Schlather für die Betreuung meiner Arbeit. Für die Übernahme des Korreferats bedanke ich mich herzlich bei Prof. Dr. Tim Friede. Ich bedanke mich für die stetige Unterstützung und Förderung in jeglicher Hinsicht.

Für die wissenschaftliche Unterstützung und die anregenden Diskussionen bedanke ich mich ganz herzlich bei Prof. Dr. Thomas Kneib, Prof. Dr. Robert Schaback und Dr. Norbert Benda.

Weiterer Dank gilt Dr. Christian Röver und PD Dr. Frank Konietschke für ihre ständige Hilfsbereitschaft und die kritische Durchsicht meiner Dissertation.

Ein besonderer Dank gilt Dr. Simon Schneider für die vielen produktiven Diskussionen, hilfreichen Kommentare und Hinweise sowie für das sprachliche Korrekturlesen meiner Arbeit.

Für eine einmalige Arbeitsatmosphäre danke ich meinen Freunden, die mich auf meinem akademischen Weg begleitet haben - allen voran Simon, David, Markus und Christian sowie Team 17 und dem restlichen Institut für Medizinische Statistik.

Weiterhin möchte ich mich bei meinen Freunden Max, Anett, Andre, Stefan, Christopher und Fabian und allen Mitgliedern der Schweden-Wanderung für die unvergessliche Studienzeit in Göttingen bedanken.

Mein spezieller Dank gilt Mandy für ihre anhaltende Unterstützung und Motivation während der gesamten Promotionszeit.

Ohne euch hätte ich es nicht geschafft.

-Finally-
Goh StSCI

Inhaltsverzeichnis

1. Einleitung	1
1.1. Motivation	1
1.2. Beispiele für Dosis-Findungsstudien mit aktiver Kontrolle	4
1.2.1. Dosis-Findungsstudie bei hyperaktiven Blasenproblemen	4
1.2.2. Dosis-Findungsstudie in Typ 2 Diabetes	4
1.3. Aufbau der Arbeit	6
2. Notationen, statistisches Modell, Schätzer und deren Eigenschaften	7
2.1. Notationen	7
2.2. Statistisches Modell	8
2.3. Schätzer und deren asymptotische Eigenschaften	9
2.3.1. Asymptotische Eigenschaften	11
2.4. Definition und Schätzung der Zieldosis	13
2.4.1. Definition der Zieldosis	13
2.4.2. Schätzung der Zieldosis	13
2.4.3. Verteilung des Zieldosis-Schätzers	14
3. Auswertung von Dosis-Findungsstudien mit aktiver Kontrolle	17
3.1. Statistisches Modell und Schätzer für (log-)lineare Dosis-Wirkungskurven	17
3.1.1. Statistisches Modell	17
3.1.2. Die Zieldosis d^*	17
3.1.3. Schätzer der Parameter und der Zieldosis und deren Eigenschaften im linearen Modell	18
3.1.4. Eigenschaften des Schätzers der Zieldosis	19
3.2. Ansätze zur Konstruktion von Konfidenzintervallen für die Zieldosis	20
3.2.1. Normalapproximation mittels Δ -Regel	21
3.2.2. Profile-Likelihood Methode	22
3.2.3. Numerische Berechnung der Intervallgrenzen	24
3.2.4. Parametrischer Bootstrap	26
3.2.5. Erweiterung der Methode von Fisch und Strehlau (1993)	27
3.3. Simulationsstudie für das lineare Modell	28
3.3.1. Verzerrung des Zieldosis-Schätzers	29

3.3.2. Eigenschaften der Konfidenzintervalle	30
3.4. Auswertung: Hyperaktive Blasenprobleme	34
3.5. Zusammenfassung	36
4. Spline-basierte Auswertung von nichtlinearen Dosis-Findungsstudien mit aktiver Kontrolle	39
4.1. Statistisches Modell	39
4.2. Schätzung der Zieldosis durch Spline-Interpolation	39
4.2.1. Kubische Spline-Interpolation	40
4.2.2. Lineare Spline-Interpolation	43
4.3. Eigenschaften der Schätzer der Zieldosis	44
4.4. Spline-basiertes Bootstrap-Konfidenzintervall	46
4.5. Simulationen der Spline-Interpolationen für verschiedene Dosis-Wirkungskurven .	47
4.5.1. Eigenschaften der Schätzer der Zieldosis	50
4.5.2. Eigenschaften der Konfidenzintervalle	51
4.5.3. Penalisierter kubischer Glättungspline	53
4.6. Zusammenfassung	55
5. Design- und Fallzahlplanung für Dosis-Findungsstudien mit aktiver Kontrolle	59
5.1. Verzerrungs-reduzierte Designs für die Verwendung Spline-basierter Auswertungen	59
5.1.1. Minimum des maximalen globalen Interpolationsfehlers	60
5.1.2. Dosisstufen mit abstandsgleichen Wirkungsgraden	60
5.1.3. Gleichmäßige Verteilung der Obergrenze des Interpolationsfehlers	61
5.1.4. Das Minimum der maximalen globalen absoluten Verzerrung des Zieldosis-	
sisschätzers	62
5.1.5. Verzerrungs-minimales Design: Simulationsstudie	63
5.1.6. Zusammenfassung	64
5.2. Fallzahlplanung	64
5.2.1. Normalapproximation mittels Δ -Regel	65
5.2.2. Fallzahl zur Kontrolle der erwarteten Länge des Konfidenzintervalls	66
5.2.3. Fallzahlformel zur Kontrolle der Länge des Konfidenzintervalls mit Wahr-	
scheinlichkeit γ	67
5.2.4. Simulationen	72
5.2.5. Empirische Approximationsansätze für N_γ im Falle nichtlinearer Wirkungs-	
kurven	74
5.2.6. Alternative Verfahren zur Fallzahlplanung	78
5.3. Zusammenfassung	79
6. Zusammenfassung, Diskussion und Ausblick	85

A. Definitionen, Lemmata und Sätze	89
A.1. ML-Schätzer und deren asymptotische Eigenschaften	89
A.2. Lineare Modelle Theorie	90
B. Beweise und Berechnungen	95
B.1. Eigenschaften des Schätzers der Zieldosis	95
B.2. Verzerrungs-reduzierte Designs für die Verwendung Spline-basierter Auswertungen	96
B.2.1. Dosisstufen mit abstandsgleichen Wirkungsgraden	96
B.2.2. Die lineare Wirkungskurve	96
B.2.3. Die Emax-Wirkungskurve	96
B.2.4. Die sigmoidale Emax-Wirkungskurve	96
B.2.5. Gleichmäßige Verteilung der Obergrenze des Interpolationsfehlers	97
B.2.6. Die lineare Wirkungskurve	97
B.2.7. Die Emax-Wirkungskurve	97
B.2.8. Die sigmoidale Emax-Wirkungskurve	98
Literaturverzeichnis	99

Abbildungsverzeichnis

1.1. Dosis-Findungsstudie mit aktiver Kontrolle	3
1.2. Die mittlere Reduktion der Blasenentleerungen in 24 Stunden	5
1.3. FPG-Reduktion der Dosisstufen des experimentellen Medikamentes und der aktiven Kontrolle	6
3.1. Verzerrung des Schätzers der Zieldosis für lineare Wirkungsfunktionen	30
3.2. Überdeckungswahrscheinlichkeit der Konfidenzintervalle für lineare Wirkungsfunktionen	32
3.3. Mediane Konfidenzintervallbreite für lineare Wirkungsfunktionen	33
3.4. Empirische Länge der Konfidenzintervallhälften	34
3.5. Überdeckungswahrscheinlichkeit der Konfidenzintervalle für Studien mit $k = 3$ und $k = 4$ Dosisstufen	35
3.6. Robustheit gegenüber nicht-normalverteilten Fehlertermen: Überdeckungswahrscheinlichkeit und mediane Konfidenzintervalllänge	38
4.1. Simulationsszenarien: Dosis-Wirkungsfunktionen	49
4.2. Spline-Interpolation der Emax-Dosis-Wirkungsfunktion und Verzerrung des Schätzers der Zieldosis	50
4.3. Überdeckungswahrscheinlichkeit der Spline-Interpolationen für verschiedene Dosis-Wirkungskurven	52
4.4. Überdeckungswahrscheinlichkeit der Spline-Interpolationen für verschiedene Anzahlen an Dosisstufen	57
5.1. Optimales Design: Spline-Interpolation der Emax-Dosis-Wirkungsfunktion und Verzerrung des Schätzers der Zieldosis	63
5.2. Optimales Design: Überdeckungswahrscheinlichkeit	68
5.3. Lineare Regression: Fallzahl N_E gegen exakte benötigte Fallzahl $N_{\gamma,ex}$	75
5.4. Lineare Regression: Fallzahl N_E gegen Fallzahl $N_{\gamma,ex}$	77
5.5. Boxplot der Quotienten verschiedener Fallzahl-Approximationen	78

Tabellenverzeichnis

3.1. Simulationsszenarien im linearen Modell	29
3.2. Auswertung des Beispiels: Zieldosischätzer mit zugehörigen Konfidenzintervallen	36
4.1. Simulationsszenarien der Spline-Interpolation	48
4.2. Eigenschaften des Glättungspline gegenüber der kubischen Spline-Interpolation .	58
5.1. Simulationsszenarien für die Fallzahlplanung	73
5.2. Eigenschaften der Fallzahlen für lineare Wirkungsfunktionen	81
5.3. Eigenschaften der Fallzahlen für Emax-Wirkungskurven	82
5.4. Eigenschaften der Fallzahlen für die Methode von Fisch und Strehlau	83

1. Einleitung

1.1. Motivation

In der Entwicklung und Untersuchung neuer Arzneimittel und Therapien sind klinische Studien ein wichtiger Bestandteil der Forschung. Die auf dem Weg zur Zulassung und darüber hinaus durchgeführten Studien werden üblicherweise in vier Phasen unterteilt (Sheiner, 1997). Dabei beschreibt jede Phase ein anderes Stadium des Forschungsprozesses mit spezifischen Problemstellungen und Zielsetzungen. Im Folgenden werden die vier verschiedenen Phasen von klinischen Studien kurz vorgestellt.

Bei der Erforschung einer neuen Substanz wird in sogenannten *Phase I* Studien die erstmalige Testung des neuen Präparates am, in der Regel, gesunden Probanden untersucht. Im Fokus stehen dabei die Bestimmung von pharmakokinetischen Merkmalen sowie die Gewinnung von Informationen zur Sicherheit und Toxizität. Speziell wird häufig eine maximal verträgliche Dosis bestimmt, die im weiteren Verlauf als Obergrenze der Applikation des Wirkstoffes dient.

In *Phase II* Studien wird darauf aufbauend die Dosis-Wirkungsbeziehung untersucht und die neue Substanz zum ersten Mal dem möglichen Ziel-Kollektiv (kranken Patienten) verabreicht. Dabei werden die Patienten auf verschiedene Konzentrationen der Medikation randomisiert verteilt. In Pharmaunternehmen ist oftmals eine Unterscheidung in *Phase IIa* (Proof of Concept) und *Phase IIb* (Dosis-Findung) Studien üblich (Chow und Liu, 2004, Kap. 1). Das vorrangige Ziel von Phase II Studien ist die Identifikation des Dosis-Wirkungsprofils und das Finden einer geeigneten Zieldosis. Häufig wird dafür die minimale effektive Dosis verwendet, die eine festgelegte Wirkungssteigerung gegenüber Placebo sichert.

Phase III Studien stellen den konfirmatorischen Teil der klinischen Studien dar und benötigen eine deutlich höhere Anzahl an Patienten im Vergleich zu Phase I oder Phase II Studien. Der Schwerpunkt dieser Studien liegt auf dem Nachweis der Wirkung und Sicherheit der gewählten Dosis gegenüber Placebo oder einer Standardtherapie. Die Studien bilden den Grundstein für die Zulassung der Substanz als pharmazeutisches Produkt und sind daher stark durch Gesetze reglementiert und kontrolliert. Auf nationaler Ebene ist die Einhaltung der Richtlinien auf dem Weg zur Zulassung durch das Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM) geregelt.

Nachdem das Medikament zugelassen ist, werden in *Phase IV* Studien die Wirkung, die auftretenden Komplikationen und Nebenwirkungen sowie weitere Fragestellungen im Langzeitverlauf untersucht. Der Fokus dieser Arbeit liegt in der Auswertung und Planung einer speziellen Art

von Phase II Studien, nämlich Dosis-Findungsstudien mit aktiver Kontrolle.

In einer allgemeinen Phase II Dosis-Findungsstudie sind die primären Ziele die Beschreibung der Dosis-Wirkungsbeziehung und die Bestimmung einer Zieldosis für den Übergang zu Phase III. Die Zieldosis wird dabei oftmals durch einen vorgegebenen klinisch relevanten Effekt gegenüber Placebo definiert. Ein Überblick über die verschiedenen Methoden zur Planung und Auswertung derartiger Studien wird in Ruberg (1995a,b) und Bornkamp et al. (2007) gegeben. Der traditionelle ANOVA-Ansatz (Analysis of Variance) verwendet Kontrast-Tests und Niveau-Adjustierung (z. B. Dunnett, 1955), um die Unterschiede in der Wirkung der Dosisstufen gegenüber Placebo zu untersuchen. Aus der Menge der untersuchten Dosisstufen mit signifikantem Unterschied zu Placebo wird die kleinste Dosis ausgewählt, die einen klinisch relevanten Unterschied gegenüber Placebo sichert. Eine weitere Methode wird in Bretz et al. (2005) und Pinheiro et al. (2006a) vorgestellt, die das formale Testen der Dosisstufen mit der Modellierung der Dosis-Wirkungskurve zur Bestimmung der Zieldosis verbindet. Dieses Verfahren wird durch die Kombination von multiplen Vergleichen (multiple comparison procedures (MCP)) mit der Modellierung auch als MCP-Mod bezeichnet (Bretz et al., 2008; Pinheiro et al., 2006b; EMA, 2014). Wird nur die Modellierung der Wirkungskurve zum Finden der Zieldosis verwendet, so liegt ein Kalibrierungsproblem vor. Zur Lösung von Kalibrierungsproblemen muss zwischen der klassischen und der inversen Regression unterschieden werden, da diese im Allgemeinen zu unterschiedlichen Schätzern führen.

Als einer der ersten Statistiker befasste sich Eisenhart (1939) mit der Problematik, sich zwischen den beiden Ansätzen entscheiden zu müssen. Dabei untersuchte er lineare Kalibrierungsprobleme und kam zu der Erkenntnis, dass eine inverse Regression nur verwendet werden sollte, falls der Regressor in der klassischen Regression zufällig ist. Werden hingegen, wie in Dosis-Findungsstudien die Regel, die Dosisstufen im Vorfeld durch das Studien-Team bestimmt und somit nicht zufällig gewählt, sollte auf die inverse Regression verzichtet und stattdessen die klassische Regression verwendet werden. Diese Ergebnisse wurden von Krutchkoff (1967) aufgegriffen, diskutiert und weiterentwickelt (siehe z. B. Krutchkoff, 1967; Williams, 1969; Berkson, 1969; Kannan et al., 2007). Ein systematischer Überblick der historischen Entwicklung im Themengebiet der Kalibrierung kann zum Beispiel in Thonnard (2006, Kap. 3) gefunden werden. Um das Kalibrierungsproblem zu lösen, kann die Wirkungsfunktion durch die klassische Regression mit einer parametrischen Funktion angenähert und die Zieldosis, die zu einem klinisch relevanten Effekt gegenüber Placebo führt, bestimmt werden. Zu dieser Problematik existiert eine Vielzahl an Literatur in den Bereichen der Qualitätskontrolle und der Dosis-Findung (Hsu und Berger, 1999; Tamhane und Logan, 2002; Morales et al., 2006; Budtz-Jørgensen, 2007). Diese Ansätze haben ihren Schwerpunkt in der Bestimmung der Zieldosis, basierend auf einer festen Wirkungsdifferenz gegenüber Placebo. Allerdings ist der Gebrauch von Placebo in einigen Fällen unethisch, wenn zum Beispiel bereits eine Standardtherapie existiert. Daher besteht ein gesteigertes Interesse an der Betrachtung von Dosis-Wirkungsstudien mit aktiver Kontrolle und einer Bestimmung der Zieldosis in Abhängigkeit dieses aktiven Komparators.

In einer Dosis-Findungsstudie mit aktiver Kontrolle wird zusätzlich zu den verschiedenen Do-

sisstufen der neuen Substanz (mit oder ohne Placebo) ein bereits zugelassenes Medikament untersucht. Dies führt häufig zu einer Änderung in der Zielsetzung der Studie. Der Fokus liegt dann nicht mehr auf dem Finden einer klinisch relevant überlegenen Dosis gegenüber Placebo, sondern beispielsweise auf der Bestimmung des Dosis-Bereiches der neuen Substanz, der eine überlegene Wirkung gegenüber der aktiven Kontrolle sicherstellt. In diesem Zusammenhang ist es von primärem Interesse die kleinste Dosis so genau wie möglich zu bestimmen, die zum gleichen erwarteten Wirkungsgrad wie der aktive Komparator führt. Zusätzlich zur Schätzung dieser Zieldosis werden auch entsprechende Konfidenzintervalle der Zieldosis benötigt, um die Güte der Schätzung beschreiben zu können. Diese Problemstellung wurde in der Literatur bisher noch nicht detailliert behandelt; Ausnahmen bilden die Arbeiten von Källén und Larsson (1999), Dilleen et al. (2003) und Dette et al. (2014). In Abbildung 1.1 ist die Bestimmung der Zieldosis mit zugehörigem Konfidenzintervall in einer Dosis-Findungsstudie mit aktiver Kontrolle exemplarisch dargestellt.

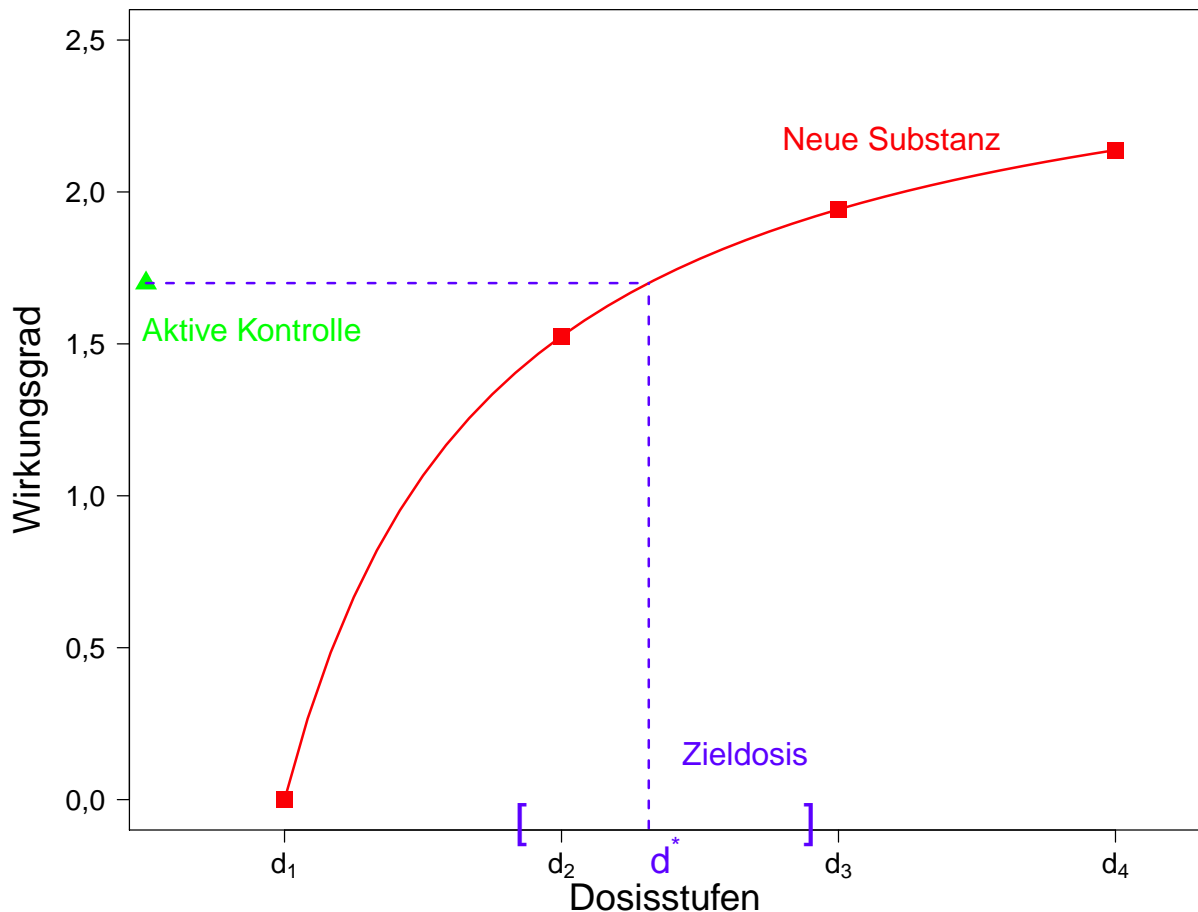


Abbildung 1.1.: Exemplarische Darstellung der Bestimmung der Zieldosis mit entsprechendem $(1 - \alpha)$ -Konfidenzintervall in einer Dosis-Findungsstudie mit aktiver Kontrolle.

Neben der Auswertung ist die Planung derartiger Studien für den Anwender von großem In-

teresse. Besonders die Berechnung der benötigten Fallzahl für ein gewähltes Güte-Kriterium und die Allokation der zu untersuchenden Dosisstufen stellen unter praktischen Gesichtspunkten eine große Herausforderung dar. Die vorliegende Arbeit beschäftigt sich mit der Schätzung der Zieldosis, der Konstruktion von zugehörigen Konfidenzintervallen und den Planungsaspekten von Dosis-Findungsstudien mit aktiver Kontrolle. Dabei werden existierende Verfahren auf das vorgestellte Problem übertragen und neue Ansätze entwickelt, um diesen wichtigen Forschungsbereich im Gebiet der klinischen Studien weiterzuentwickeln.

1.2. Beispiele für Dosis-Findungsstudien mit aktiver Kontrolle

Das Design von Dosis-Findungsstudien mit aktiver Kontrolle ist sehr vielfältig, da es in fast allen pharmakologischen Bereichen Anwendung findet. Je nach Anwendungsgebiet variieren die Anzahl an untersuchten Dosisstufen sowie die Anzahl an rekrutierten Probanden. Üblicherweise beinhaltet eine Dosis-Findungsstudie mit aktiver Kontrolle drei bis fünf Dosen der experimentellen Medikation, einen aktiven Komparator und gegebenenfalls eine Placebo-Kontrolle. Im Folgenden werden zwei motivierende Beispiele von Dosis-Findungsstudien mit aktiver Kontrolle vorgestellt.

1.2.1. Dosis-Findungsstudie bei hyperaktiven Blasenproblemen

Das erste Beispiel ist eine doppelblinde, sechsarmige Phase II Dosis-Findungsstudie mit aktiver Kontrolle, die den Effekt von Solifenacin an Patienten mit idiopathischer Detrusorhyperaktivität (hyperaktiver Blase) nach sechs Wochen von Baseline untersucht (Chapple et al., 2004). Es wurden insgesamt $N = 255$ Patienten im gleichen Verhältnis zu Placebo (0 mg/Tag), einer von vier Dosisstufen Solifenacin (2,5; 5; 10 oder 20 mg/Tag) oder zur aktiven Kontrolle Tolterodine (2 mg zweimal täglich) randomisiert verteilt. Der primäre Endpunkt, der im motivierenden Beispiel verwendet wird, ist die Reduktion der mittleren Anzahl an Blasenentleerungen in 24 Stunden (*Voids/24h*). Da in Chapple et al. (2004) für die verschiedenen Dosen die Mittelwerte der Reduktion der *Voids/24h* ohne Standardabweichung (*SD*) angegeben wurden, musste diese aus den berichteten p-Werten der Teststatistiken rekonstruiert werden. Nach Abschätzung aus den berichteten p-Werten liegt die *SD* im Bereich [1,9; 2,5]. Für Simulationen und Auswertungen wird im Folgenden von einer Standardabweichung $\sigma = 2$ ausgegangen. Die Ergebnisse der Studie sind in Abbildung 1.2 auf der $\log(1+\text{Dosis})$ -Skala als Mittelwert \pm Standardfehler (*SE*) dargestellt.

1.2.2. Dosis-Findungsstudie in Typ 2 Diabetes

Als zweites Beispiel dient eine fünfarmige Dosis-Findungsstudie zur Behandlung von Typ 2 Diabetes von Nauck et al. (2009). Die Studie war randomisiert, doppelblind, und lief über einen Zeitraum von 26 Wochen. Weiterhin enthielt die Studie eine Placebo- und eine aktive Kontrolle. Die Gesamtanzahl von $N = 1.091$ Patienten wurde zufällig im Verhältnis (2 : 2 : 2 : 1 : 2) auf Liraglutide (0,6; 1,2 oder 1,8 mg/Tag), auf Placebo (0 mg/Tag) oder die aktive

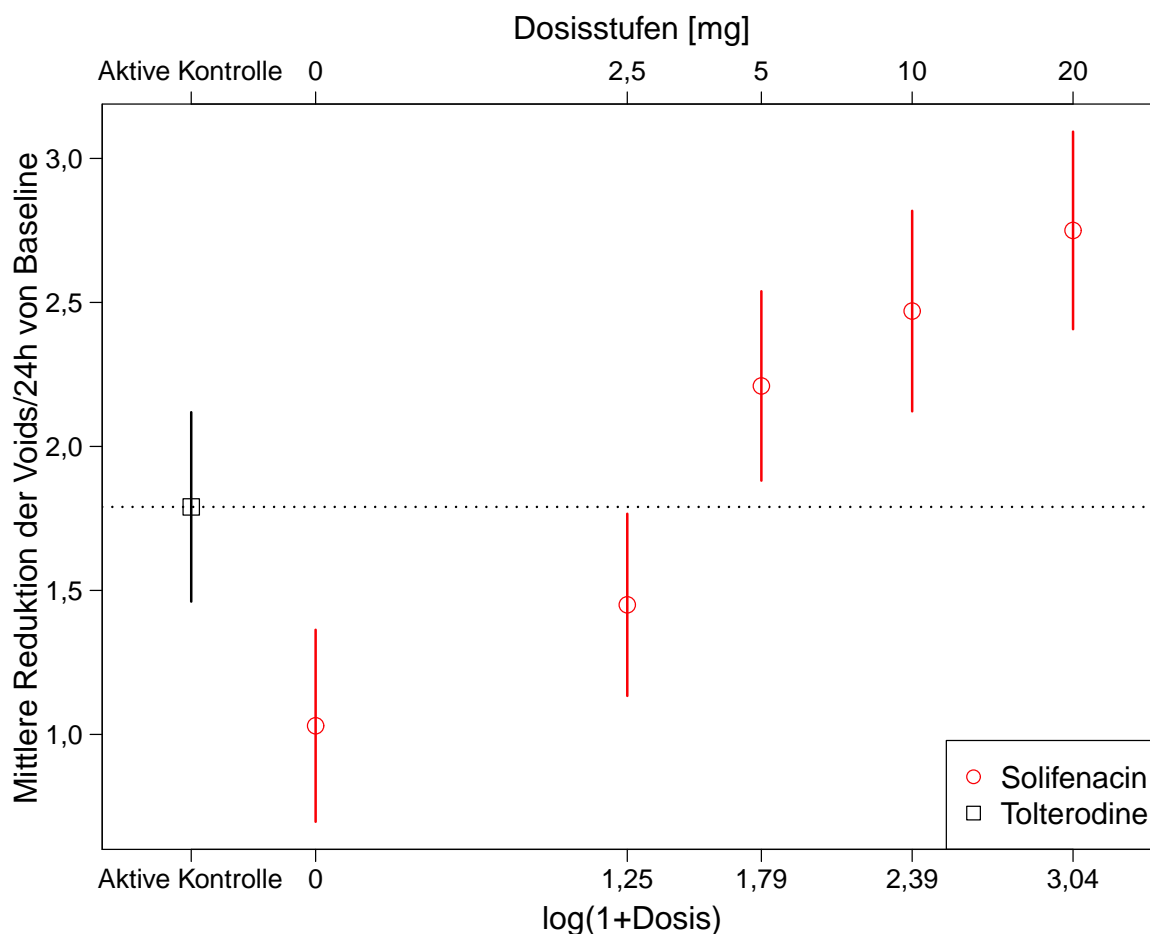


Abbildung 1.2.: Die mittlere Reduktion der Blasenentleerungen in 24 Stunden (*Voids/24h*) von Baseline bis zum Ende der Studie für die Dosisstufen des experimentellen Medikamentes und der aktiven Kontrolle dargestellt als Mittelwert \pm Standardfehler (*SE*) (Chapple et al., 2004).

Kontrolle Glimepiride verteilt. Der primäre Endpunkt der Studie war die prozentuale Veränderung des Glykohämoglobin-Spiegels HbA_{1c} am Ende der Studie im Vergleich zu Baseline. Als einer der sekundären Endpunkte wurde die Veränderung des Nüchtern-Plasmaglukose-Levels (*fasting plasma glucose* (FPG)) in mmol/l am Ende der Studie untersucht, welche im Folgenden als Zielgröße des Beispiels verwendet wird. In der Studie wurden nur die Standardabweichungen (*SD*) der FPG-Werte zu Beginn und am Ende der Behandlungsphase angegeben. Da die *SD* der Reduktion der FPG-Werte nicht in der Publikation enthalten war, wird in Anlehnung an die berichteten Standardabweichungen in den Simulationen von einer *SD* $\sigma = 1,8 \text{ mmol/l}$ ausgegangen. Die Ergebnisse der Studie sind in Form von Differenzmittelwerten und angenommenen Standardfehlern (*SE*) in Abbildung 1.3 dargestellt.

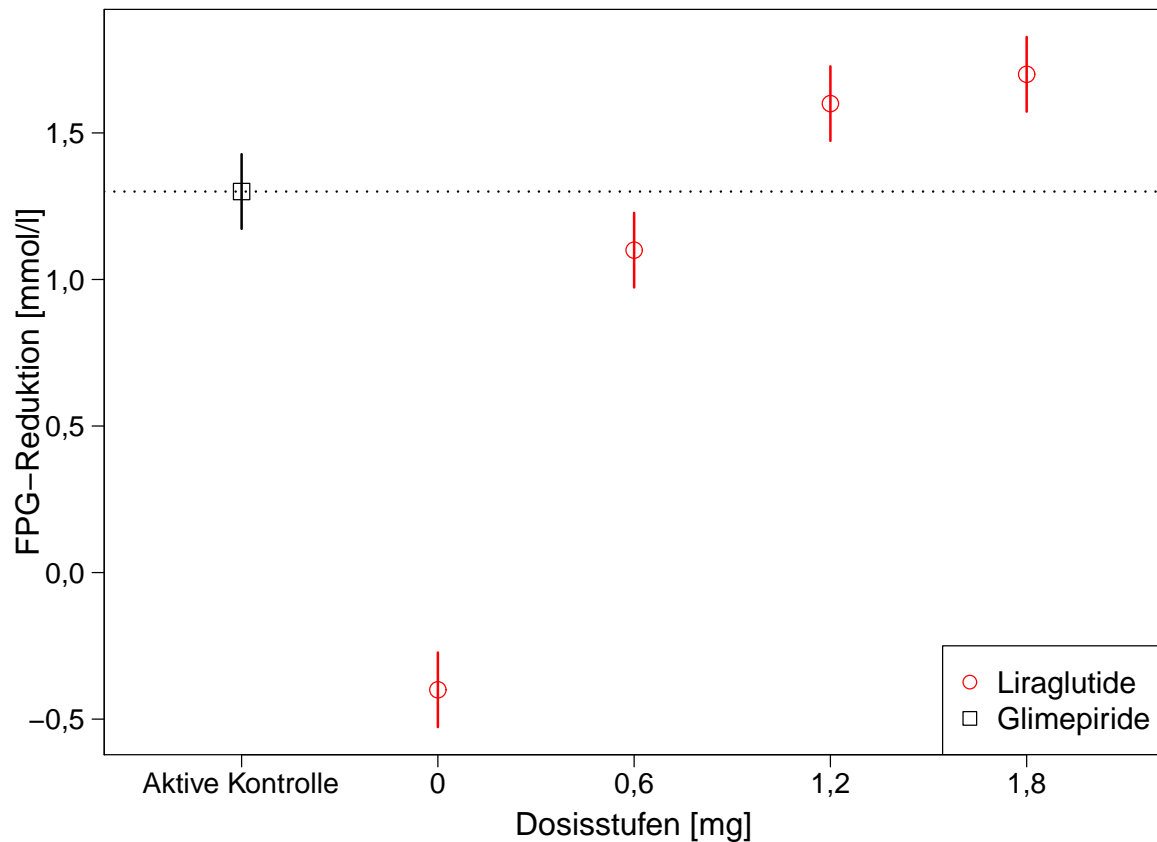


Abbildung 1.3.: FPG-Reduktion der Dosisstufen des experimentellen Medikamentes und der aktiven Kontrolle dargestellt als Mittelwert \pm angenommen Standardfehler (SE) am Ende der Studie (Nauck et al., 2009).

1.3. Aufbau der Arbeit

In Kapitel 2 werden die verwendeten Notationen und Modellannahmen eingeführt sowie die notwendigen Parameterschätzer und deren asymptotische Eigenschaften beschrieben. Kapitel 3 befasst sich mit der Schätzung der Zieldosis und der Bestimmung von zugehörigen Konfidenzintervallen im Falle von linearen Wirkungskurven. Die Herleitung von Spline-basierten Schätzern und Konfidenzintervallen, die ohne Annahmen an die Gestalt der Wirkungsfunktion auskommen, werden in Kapitel 4 dargestellt. Die Planung von Dosis-Findungsstudien mit aktiver Kontrolle wird in Kapitel 5 thematisiert. Es werden zum einen Verfahren zur optimalen Dosisallokation von Spline-basierten Auswertungen entwickelt und zum anderen Methoden zur Fallzahlkalkulation hergeleitet, die die Länge des Konfidenzintervalls der Zieldosis beschränken. Eine Zusammenfassung und ein Ausblick auf zukünftige Forschungsprojekte bilden den Abschluss dieser Arbeit. Aus Gründen der Leserlichkeit werden technische Beweise und Definitionen in den Anhang ausgelagert.

2. Notationen, statistisches Modell, Schätzer und deren Eigenschaften

In diesem Kapitel wird die verwendete Notation eingeführt und das statistische Modell für Dosis-Findungsstudien mit aktiver Kontrolle vorgestellt. Weiterhin werden die notwendigen Parameterschätzer und deren asymptotische Eigenschaften beschrieben.

2.1. Notationen

Zunächst werden die elementaren Notationen, die in den folgenden Kapiteln verwendet werden, eingeführt. KONSTANTE SKALARE werden mit lateinischen oder griechischen Kleinbuchstaben bezeichnet, z. B. a , α . VEKTOREN sollen im Allgemeinen mit kleinen, fettgedruckten, lateinischen oder griechischen Buchstaben bezeichnet werden, z. B. \mathbf{a} , $\boldsymbol{\alpha}$. Des Weiteren sei $\mathbf{1}_n$ der n -dimensionale Einser-Vektor und $\mathbf{0}_n$ der n -dimensionale Null-Vektor. MATRIZEN werden mit großen, fettgedruckten, lateinischen oder griechischen Buchstaben bezeichnet, z. B. \mathbf{A} , $\boldsymbol{\Sigma}$. Insbesondere bezeichne \mathbf{I}_n die $n \times n$ Einheitsmatrix. Weiterhin sei \mathbf{A}' die TRANSPONIERTE und \mathbf{A}^{-1} die INVERSE der Matrix \mathbf{A} und bezeichne $\text{Sp}(\mathbf{A})$ die SPUR, $r(\mathbf{A})$ den RANG und $\det(\mathbf{A})$ die DETERMINANTE der Matrix \mathbf{A} . ZUFALLSVARIABLEN werden, wenn nicht anders angegeben, mit lateinischen Großbuchstaben bezeichnet und ZUFALLSVEKTOREN werden zusätzlich in Fettdruck dargestellt (Y_{kl} , \mathbf{Y}_k). Alle INDIZES von Vektoren, Matrizen und Zufallsvariablen werden mit kleinen lateinischen Buchstaben wie zum Beispiel k , l , r , s bezeichnet. Neben den Standard-Matrizenoperationen Addition (+) und Multiplikation (\cdot) wird die DIREKTE SUMME (\oplus) zweier Matrizen \mathbf{A} und \mathbf{B} benötigt und wie folgt definiert:

$$\mathbf{A} \oplus \mathbf{B} = \left(\begin{array}{c|c} \mathbf{A} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{B} \end{array} \right).$$

Der Ausdruck $A \stackrel{!}{=} B$ beschreibt die Bedingung, dass A gleich B sein soll. Im weiteren Verlauf werden die Konvergenzen wie folgt dargestellt. Für eine Folge von Zufallsvariablen (X_n) bezeichne $X_n \xrightarrow{P} X$ die stochastische Konvergenz, $X_n \xrightarrow{\mathcal{L}} X$ die Verteilungskonvergenz und $X_n \xrightarrow{f.s.} X$ die fast sichere Konvergenz gegen X .

2.2. Statistisches Modell

In einer Dosis-Findungsstudie sollen k Dosisstufen $0 \leq d_1 < d_2 < \dots < d_k$ und eine aktive Kontrolle (ak) untersucht werden. Die Zufallsvariable des j -ten Individuums in der i -ten Dosisstufe wird wie folgt dargestellt:

$$Y_{ij} = f_{\boldsymbol{\theta}}(d_i) + \epsilon_{ij},$$

wobei $f_{\boldsymbol{\theta}}(d)$ die erwartete Dosis-Wirkungskurve mit dem zugehörigen Parametervektor $\boldsymbol{\theta} = (\theta_0, \theta_1, \dots, \theta_b)' \in \boldsymbol{\Theta}$ darstellt und $\epsilon_{ij} \sim \mathcal{N}(0, \sigma_i^2)$, $i = 1, \dots, k$, $j = 1, \dots, n_i$, den unabhängigen Fehlerterm beschreibt.

Die Zufallsvariable des j' -ten Individuums in der aktiven Kontrolle lässt sich durch

$$Y_{ak,j'} = \mu + \epsilon_{ak,j'}$$

mit Erwartungswert μ und Fehlerterm $\epsilon_{ak,j'} \sim \mathcal{N}(0, \sigma_{ak}^2)$, $j' = 1, \dots, n_{ak}$, beschreiben. Weiterhin seien $\sigma_i > 0$, $\sigma_{ak} > 0$ und $n_d = \sum_{i=1}^k n_i \in \mathbb{N}$ die Fallzahl der k Dosisstufen sowie $N = n_d + n_{ak} \in \mathbb{N}$ die Gesamtfallzahl der Studie. Entsprechend den verschiedenen Fallzahlen werden die Gewichte $w_i = n_i/N$, $i = 1, \dots, k$, $w_d = n_d/N$ und $w_{ak} = n_{ak}/N$ definiert, so dass $w_d + w_{ak} = 1$ gilt (Dette et al., 2008). Um eine übersichtliche Struktur zu erhalten, werden die Zufallsvariablen in der folgenden Vektorschreibweise zusammengefasst. $\mathbf{Y}_i = (Y_{i1}, \dots, Y_{i,n_i})'$, $i = 1, \dots, k$, beschreibt den Zufallsvektor der i -ten Dosisstufe, $\mathbf{Y}_d = (\mathbf{Y}'_1, \dots, \mathbf{Y}'_k)'$ den Zufallsvektor der k Dosisstufen, $\mathbf{Y}_{ak} = (Y_{ak,1}, \dots, Y_{ak,n_{ak}})'$ den Vektor der aktiven Kontrolle und $\mathbf{Y} = (\mathbf{Y}'_d, \mathbf{Y}'_{ak})'$ den Vektor aller Zufallsvariablen. Entsprechend ergeben sich die Vektoren der Fehlerterme mit den zugehörigen multivariaten Normalverteilungen $\boldsymbol{\epsilon}_i \sim \mathcal{N}(\mathbf{0}, \sigma_i^2 \mathbf{I}_{n_i})$, $\boldsymbol{\epsilon}_{ak} \sim \mathcal{N}(\mathbf{0}, \sigma_{ak}^2 \mathbf{I}_{n_{ak}})$ und $\boldsymbol{\epsilon}_d \sim \mathcal{N}(\mathbf{0}, \bigoplus_{i=1}^k \sigma_i^2 \mathbf{I}_{n_i})$ sowie $\boldsymbol{\epsilon} = (\boldsymbol{\epsilon}'_d, \boldsymbol{\epsilon}'_{ak})' \sim \mathcal{N}(\mathbf{0}, \bigoplus_{i=1}^k \sigma_i^2 \mathbf{I}_{n_i} \oplus \sigma_{ak}^2 \mathbf{I}_{n_{ak}})$. Dann lässt sich das Dosis-Wirkungsmodell durch $\mathbf{Y}_d = f_{\boldsymbol{\theta}}(\mathbf{d}) + \boldsymbol{\epsilon}_d$ mit $f_{\boldsymbol{\theta}}(\mathbf{d}) = (f_{\boldsymbol{\theta}}(d_1)\mathbf{1}'_{n_1}, \dots, f_{\boldsymbol{\theta}}(d_k)\mathbf{1}'_{n_k})'$ beschreiben.

Für die Betrachtungen der Dosis-Wirkungskurve wird sich auf stetige Funktionen $f_{\boldsymbol{\theta}}(d) : D \rightarrow \mathbb{R}$ mit $\boldsymbol{\Theta} \subset \mathbb{R}^{b+1}$ kompakt und $D = [d_1, d_k] \subset \mathbb{R}^+$ beschränkt, so dass $\sup_{d \in D} |f_{\boldsymbol{\theta}}(d)| \leq \kappa < \infty$. Weiterhin wird vorausgesetzt, dass neben der Dosis-Wirkungsfunktion $f_{\boldsymbol{\theta}}(d)$ auch die Ableitungen bis zum vierten Grad stetig sind und existieren. Im Folgenden wird der Spezialfall gleicher Varianzen der Fehlerterme $\sigma_i = \sigma_{ak} = \sigma > 0$ betrachtet. Der allgemeine Fall ungleicher Varianzen wird in Kapitel 4 behandelt.

Selbst wenn für $f_{\boldsymbol{\theta}}(d)$ eine Klasse an Wirkungsfunktionen (z. B. eine lineare Funktion) bekannt ist, sind im Allgemeinen der Parametervektor der erwarteten Wirkungsfunktion $\boldsymbol{\theta}$ und der Erwartungswert der Wirkung der aktiven Kontrolle μ unbekannt und müssen aus den Realisationen des Zufallsvektors \mathbf{Y} geschätzt werden. Der folgende Abschnitt definiert die benötigten Schätzer mit Hilfe der Maximum-Likelihood Methode und beschreibt ihre asymptotischen Eigenschaften.

2.3. Schätzer und deren asymptotische Eigenschaften

In diesem Abschnitt werden Schätzer für die unbekannt Parameter der Dosis-Wirkungsfunktion $f_{\boldsymbol{\theta}}(d)$ und der Wirkung der aktiven Kontrolle μ auf Grundlage der Maximum-Likelihood (ML) Theorie bestimmt und anschließend deren Eigenschaften vorgestellt. Im Folgenden wird davon ausgegangen, dass die Klasse der Wirkungsfunktionen vorgegeben oder bekannt ist und lediglich der Parametervektor $\boldsymbol{\theta}$ geschätzt werden muss. Durch die in Abschnitt 2.2 vorgestellte Annahme normalverteilter unabhängiger Fehlerterme mit gleicher Varianz $\sigma^2 > 0$ sind die präsentierten Resultate des ML-Ansatzes identisch mit dem Ansatz der kleinsten Quadrate (z. B. Ferguson, 1996). Zunächst wird die Likelihood Funktion des gesamten Zufallsvektors $\mathbf{Y} = \mathbf{y}$ aufgestellt:

$$\begin{aligned} L(\boldsymbol{\theta}, \mu, \sigma^2 | \mathbf{Y} = \mathbf{y}) &= L(\boldsymbol{\theta}, \mu, \sigma^2 | \mathbf{Y}) \\ &= \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right)^N \exp \left(-\frac{1}{2\sigma^2} \left[\sum_{i=1}^k \sum_{j=1}^{n_i} (Y_{ij} - f_{\boldsymbol{\theta}}(d_i))^2 + \sum_{j'=1}^{n_{ak}} (Y_{ak,j'} - \mu)^2 \right] \right) \end{aligned} \quad (2.1)$$

mit entsprechender Log-Likelihood Funktion

$$\begin{aligned} l &= \ln(L(\boldsymbol{\theta}, \mu, \sigma^2 | \mathbf{Y})) \\ &= -\frac{N}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \left[\sum_{i=1}^k \sum_{j=1}^{n_i} (Y_{ij} - f_{\boldsymbol{\theta}}(d_i))^2 + \sum_{j'=1}^{n_{ak}} (Y_{ak,j'} - \mu)^2 \right]. \end{aligned} \quad (2.2)$$

Zur Berechnung der benötigten Schätzer muss das Maximum der Likelihood Funktion bezüglich des entsprechenden Parameters bestimmt werden. Dafür wird die erste Ableitung der Log-Likelihood Funktion gleich null gesetzt und gelöst. Um ein globales Maximum der Likelihood Funktion des Schätzers sicherzustellen, muss zusätzlich die zweite Ableitung der Likelihood Funktion für den Schätzer negativ sein.

Definition 2.1 *Der ML-Schätzer des Parametervektors $\hat{\boldsymbol{\theta}}$ ist mit den Annahmen aus Abschnitt 2.2 die Lösung des Gleichungssystems:*

$$\begin{aligned} \frac{\partial l}{\partial \boldsymbol{\theta}} &= \left(\frac{\partial l}{\partial \theta_0}, \dots, \frac{\partial l}{\partial \theta_b} \right)' = \mathbf{0} \text{ mit} \\ \frac{\partial l}{\partial \theta_s} &= \sum_{i=1}^k \sum_{j=1}^{n_i} (Y_{ij} - f_{\boldsymbol{\theta}}(d_i)) \cdot \partial f_{\boldsymbol{\theta}} / \partial \theta_s |_{d_i} = 0, s = 0, \dots, b. \end{aligned}$$

Ob dieses Gleichungssystem analytisch gelöst werden kann, hängt von der Form der Wirkungsfunktion $f_{\boldsymbol{\theta}}(d)$ ab. Falls dies nicht möglich ist, muss auf numerische Lösungsverfahren zur Null-

stellensuche zurückgegriffen werden. In der Statistiksoftware SAS 9.3 kann dafür zum Beispiel die Prozedur NLMIXED verwendet werden, die auf die Levenberg-Marquardt Kleinste-Quadrate-Methode zurückgreift (Helms et al., 2012). Neben dem Parametervektor $\boldsymbol{\theta}$ muss auch die Wirkung der aktiven Kontrolle geschätzt werden. Das folgende Korollar liefert einen Schätzer für μ .

Korollar 2.1 *Ein erwartungstreuer und stark konsistenter Schätzer der Wirkung der aktiven Kontrolle μ ist durch*

$$\hat{\mu} = \bar{Y}_{ak\cdot} = \frac{1}{n_{ak}} \sum_{j'=1}^{n_{ak}} Y_{ak,j'} \quad (2.3)$$

gegeben. Der Schätzer ist mit Erwartungswert μ und Varianz σ^2/n_{ak} normalverteilt; $\hat{\mu} \sim \mathcal{N}(\mu, \sigma^2/n_{ak})$.

BEWEIS.

Der Schätzer $\hat{\mu}$ ergibt sich sofort aus der Lösung der Gleichung $\partial l / \partial \mu = \sum_{j'=1}^{n_{ak}} (Y_{ak,j'} - \mu) = 0$. Die Verteilung von $\hat{\mu}$ folgt direkt aus den Modellannahmen (Abschnitt 2.2). Die Erwartungstreue und starke Konsistenz (Def. A.2) kann durch $E(\hat{\mu}) = 1/n_{ak} \cdot \sum_{j'=1}^{n_{ak}} \mu = \mu$ und mit Hilfe des starken Gesetzes der großen Zahlen gezeigt werden (siehe z. B. Dehling und Haupt, 2004, S. 292-293). \square

Neben den vorgestellten Parameterschätzern für die Dosis-Wirkungskurve und Wirkung der aktiven Kontrolle wird zusätzlich ein Schätzer für die unbekannte Varianz der Fehlerterme benötigt. Dieser ergibt sich aus folgender Gleichung:

$$\frac{\partial l}{\partial \sigma^2} = -\frac{N}{2} \frac{1}{\sigma^2} + \frac{1}{2\sigma^4} \left[\sum_{i=1}^k \sum_{j=1}^{n_i} (Y_{ij} - f_{\boldsymbol{\theta}}(d_i))^2 + \sum_{j'=1}^{n_{ak}} (Y_{ak,j'} - \mu)^2 \right] = 0.$$

Als Lösung ergibt sich

$$\tilde{\sigma}^2 = \frac{1}{N} \left[\sum_{i=1}^k \sum_{j=1}^{n_i} (Y_{ij} - f_{\boldsymbol{\theta}}(d_i))^2 + \sum_{j'=1}^{n_{ak}} (Y_{ak,j'} - \mu)^2 \right]. \quad (2.4)$$

Wie zuvor beschrieben, sind der Parametervektor $\boldsymbol{\theta}$ und die erwartete Wirkung der aktiven Kontrolle μ unbekannt, so dass sich der Varianzschätzer nicht direkt bestimmen lässt. Werden die unbekannt Parameter durch ihre ML-Schätzer ersetzt, die unabhängig von σ^2 geschätzt werden, so lässt sich der Varianzschätzer wie folgt definieren.

Definition 2.2 *Seien $\hat{\boldsymbol{\theta}}$ und $\hat{\mu}$ ML-Schätzer für $\boldsymbol{\theta}$ und μ , dann ist der Maximum-Likelihood*

Schätzer $\hat{\sigma}^2$ der Varianz des Fehlerterms mittels Gl. (2.4) wie folgt gegeben:

$$\hat{\sigma}^2 = \frac{1}{N} \left[\sum_{i=1}^k \sum_{j=1}^{n_i} (Y_{ij} - f_{\hat{\theta}}(d_i))^2 + \sum_{j'=1}^{n_{ak}} (Y_{ak,j'} - \hat{\mu})^2 \right].$$

Dieser Varianzschätzer ist nicht erwartungstreu. Falls die Verzerrung in Abhängigkeit der Wirkungsfunktion bestimmt werden kann, besteht die Möglichkeit einen erwartungstreuen Schätzer durch Bias-Korrektur herzuleiten (siehe z. B. Kapitel 3).

2.3.1. Asymptotische Eigenschaften

Im folgenden Abschnitt werden die asymptotischen Eigenschaften des ML-Schätzers des Parametervektors θ präsentiert. Dafür werden einige Eigenschaften hinsichtlich des asymptotischen Verhaltens der Fallzahlen vorausgesetzt.

Voraussetzung 2.1

Geht die Gesamtfallzahl N in den asymptotischen Betrachtungen gegen unendlich, dann konvergieren die Fallzahlen pro Dosisstufe reziprok; d.h. für $N \rightarrow \infty$ folgt $n_i/N = w_i \rightarrow \kappa_i$ mit $\kappa_i \in (0, 1)$, $i = 1, \dots, k$ und $n_{ak}/N = w_{ak} \rightarrow \kappa_{ak} \in (0, 1)$.

Zuerst wird die starke Konsistenz des ML-Schätzers nachgewiesen.

Korollar 2.2 (Starke Konsistenz der Maximum-Likelihood Schätzer)

Sei $\mathbf{Y}_d = f_{\theta}(\mathbf{d}) + \epsilon$ ein Zufallsvektor mit $\theta \in \Theta$ und normalverteiltem Fehlerterm $\epsilon \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I}_{n_d})$, der die Annahmen aus Abschnitt 2.2 erfüllt. Weiterhin seien die Regularitätsbedingungen aus Satz A.1 (im Anhang) erfüllt, dann ist $\hat{\theta}$ ein konsistenter Schätzer für den Parametervektor θ im starken Sinne und es gilt:

$$\hat{\theta} \xrightarrow{f.s.} \theta.$$

BEWEIS. Unter den in Abschnitt 2.2 gemachten Modellannahmen sind die Regularitätsbedingungen von Satz A.1 (im Anhang) erfüllt, aus welchem die starke Konsistenz direkt folgt. \square

Mit Hilfe der starken Konsistenz kann in einem nächsten Schritt die asymptotische Normalität von $\hat{\theta}$ gezeigt werden. Dafür wird zuerst die erwartete Fisher-Informationsmatrix definiert.

Definition 2.3 (Fisher-Informationsmatrix) Betrachtet wird das statistische Modell aus Abschnitt 2.2 mit dem Zufallsvektor $\mathbf{Y}_d = f_{\theta}(\mathbf{d}) + \epsilon$ und sei $l = \ln[L(\theta, \mu, \sigma^2 | \mathbf{Y})]$ die zugehörige Log-Likelihood Funktion aus Gl. (2.2), dann kann die erwartete Fisher-Informationsmatrix des Parametervektors θ wie folgt definiert werden (Seber und Wild, 1989, S. 33):

$$\mathcal{F}(\theta) = -E \left(\frac{\partial^2 l}{\partial \theta \partial \theta'} \right).$$

Eine alternative Darstellung von $\mathcal{F}(\boldsymbol{\theta})$ kann über den Gradienten der Wirkungsfunktion $f_{\boldsymbol{\theta}}(d)$ bezüglich des Vektors $\boldsymbol{\theta}$ definiert werden (Seber und Wild, 1989, S. 191-193):

$$\Delta f_{\boldsymbol{\theta}}(d) = \left(\frac{\partial f_{\boldsymbol{\theta}}}{\partial \theta_0} \quad \frac{\partial f_{\boldsymbol{\theta}}}{\partial \theta_1} \quad \dots \quad \frac{\partial f_{\boldsymbol{\theta}}}{\partial \theta_b} \right) \Big|_d \quad \forall d \in [d_1, d_k].$$

Sei weiterhin

$$\begin{aligned} F_{\boldsymbol{\theta}} &= \frac{1}{N} \sum_{i=1}^k n_i \Delta f_{\boldsymbol{\theta}}(d_i)' \cdot \Delta f_{\boldsymbol{\theta}}(d_i) \\ &= \sum_{i=1}^k w_i \cdot \Delta f_{\boldsymbol{\theta}}(d_i)' \cdot \Delta f_{\boldsymbol{\theta}}(d_i), \end{aligned}$$

dann lässt sich die erwartete Fisher-Informationsmatrix wie folgt darstellen (siehe z. B. Dette et al., 2008):

$$\mathcal{F}(\boldsymbol{\theta}) = \frac{\sigma^2}{N} \cdot F_{\boldsymbol{\theta}}^{-1}. \quad (2.5)$$

Die asymptotische Normalität des ML-Schätzers $\widehat{\boldsymbol{\theta}}$ kann darauf aufbauend mit Hilfe des folgenden Korollars gezeigt werden.

Korollar 2.3 (Asymptotische Normalität der ML-Schätzer) Sei $\mathbf{Y}_d = f_{\boldsymbol{\theta}}(d) + \epsilon$ ein Zufallsvektor, welcher dem betrachteten Modell aus Abschnitt 2.2 genügt und sei $\widehat{\boldsymbol{\theta}}$ ein konsistenter Schätzer für den Parametervektor $\boldsymbol{\theta} \in \widetilde{\Theta}$ im starken Sinne; speziell seien die Bedingungen aus Satz A.1 erfüllt. Sei weiter $\widetilde{\Theta} \subset \Theta$ offen und $\mathcal{F}(\boldsymbol{\theta})$ die positiv definite (p. d.) Fisher-Informationsmatrix des Parametervektors $\boldsymbol{\theta}$ aus Gl. (2.5). Dann gilt für den Vektor der Parameterschätzer $\widehat{\boldsymbol{\theta}}$ mit Voraussetzung 2.1:

$$\sqrt{N} (\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{L} \mathcal{N}(\mathbf{0}, \mathcal{F}(\boldsymbol{\theta})^{-1}).$$

BEWEIS. Unter den Modellannahmen und der Bedingung, dass die Fisher-Informationsmatrix des Parametervektors $\mathcal{F}(\boldsymbol{\theta})$ existiert und positiv definit ist, folgt die asymptotische Normalität aus dem Satz von Cramér (Satz A.2 im Anhang). \square

Im weiteren Verlauf wird eine Definition für die Zieldosis sowie ein entsprechender ML-Schätzer gegeben und dessen asymptotische Verteilung hergeleitet.

2.4. Definition und Schätzung der Zieldosis

2.4.1. Definition der Zieldosis

Wie in Kapitel 1 beschrieben, ist neben der Identifikation des Dosis-Wirkungsprofils das Finden der Zieldosis, die zum gleichen Wirkungsgrad wie der aktive Komparator führt, von primärem Interesse. Eine formale Definition der Zieldosis wird in nachstehender Definition gegeben.

Definition 2.4 (Die Zieldosis d^*) Sei $f_{\theta}(d)$ die zugrundeliegende Dosis-Wirkungskurve auf $D = [d_1, d_k]$ und μ der Erwartungswert der Wirkung der aktiven Kontrolle. Dann lässt sich die Zieldosis wie folgt definieren:

$$\begin{aligned} d^* &= \min_{d \in D} \{d \mid f_{\theta}(d) = \mu\} \text{ oder, falls } f_{\theta}^{-1} \text{ existiert,} \\ d^* &= f_{\theta}^{-1}(\mu). \end{aligned} \quad (2.6)$$

Sollte $f_{\theta}(d)$ keine analytische Umkehrfunktion besitzen, so muss die Gleichung numerisch gelöst werden. Falls mehrere Lösungen existieren, wird durch die Bedingung der minimalen Dosis (Gl. 2.6) eine Eindeutigkeit der Zieldosis garantiert. In den folgenden Betrachtungen wird davon ausgegangen, dass die Umkehrfunktion f_{θ}^{-1} existiert.

2.4.2. Schätzung der Zieldosis

Um die Gleichung der Zieldosis aus Gl. (2.6) lösen zu können, werden die Schätzer des unbekannten Parametervektors θ und der erwarteten Wirkung der aktiven Kontrolle μ benötigt (siehe Abs. 2.3). Die vorgestellten konsistenten ML-Schätzer der unbekannten Parameter $\hat{\theta}$ (Def. 2.1) und $\hat{\mu}$ (Kor. 2.1) werden in die obige Gleichung eingesetzt. Der Schätzer der Zieldosis d^* ist durch

$$\hat{d}^* = \min_{d \in D} \{d \mid f_{\hat{\theta}}(d) = \hat{\mu}\} = f_{\hat{\theta}}^{-1}(\hat{\mu}) \quad (2.7)$$

definiert. Die Konsistenz des Zieldosis-Schätzers wird im folgenden Korollar gezeigt.

Korollar 2.4 Sei $f_{\theta}(d)$ die zugrundeliegende Dosis-Wirkungskurve auf $D = [d_1, d_k]$ mit stetiger Umkehrfunktion f_{θ}^{-1} und sei μ der Erwartungswert der Wirkung der aktiven Kontrolle. Weiter seien $\hat{\theta}$ und $\hat{\mu}$ die konsistenten ML-Schätzer der unbekannten Parameter θ und μ (Abs. 2.3). Dann ist der Schätzer \hat{d}^* aus Gleichung (2.7) ein konsistenter Schätzer der Zieldosis d^* .

BEWEIS. Die Konsistenz des Schätzers der aktiven Kontrolle $\hat{\mu}$ und des Schätzers des Parametervektors $\hat{\theta}$ wurde bereits in Korollar 2.1 und Korollar 2.2 gezeigt. Unter der Annahme, dass die Umkehrfunktion der Dosis-Wirkungsfunktion f_{θ}^{-1} existiert und stetig ist, folgt die Konsistenz des Zieldosis-Schätzers aus dem Satz von Slutsky (siehe z. B. Ferguson, 1996, S. 39-43). \square

2.4.3. Verteilung des Zieldosis-Schätzers

Für die meisten Wirkungskurven kann die Verteilung des Schätzers \widehat{d}^* nicht direkt berechnet werden, so dass im Folgenden die asymptotische Verteilung des Zieldosis-Schätzers hergeleitet wird. Dazu werden zuerst die ML-Schätzer $\widehat{\boldsymbol{\theta}}$ und $\widehat{\mu}$, die zur Bestimmung des Zieldosis-Schätzers benötigt werden in einem Vektor $\widehat{\boldsymbol{\theta}}_{d^*} = (\widehat{\boldsymbol{\theta}}', \widehat{\mu})'$ zusammengefasst und dessen gemeinsame approximative Verteilung angegeben:

$$\widehat{\boldsymbol{\theta}}_{d^*} = \begin{pmatrix} \widehat{\boldsymbol{\theta}} \\ \widehat{\mu} \end{pmatrix} \dot{\sim} \mathcal{N}(\boldsymbol{\theta}_{d^*}, \boldsymbol{\Sigma}), \quad (2.8)$$

wobei $\boldsymbol{\theta}_{d^*} = (\boldsymbol{\theta}', \mu)'$ den Erwartungswertvektor und $\boldsymbol{\Sigma} = (\boldsymbol{\Sigma}_{\boldsymbol{\theta}} \oplus \sigma^2/n_{ak}) = \sigma^2/N(\widetilde{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}} \oplus 1/w_{ak})$ die Kovarianzmatrix darstellen. Falls die Kovarianzmatrix der Parameterschätzer der Wirkungsfunktion $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$ nicht analytisch bestimmt werden kann, wird sie durch die erwartete Fisher-Informationsmatrix $\mathcal{F}(\boldsymbol{\theta})$ aus Gl. (2.5) ersetzt. Es ergibt sich die folgende Darstellung der Kovarianzmatrix $\boldsymbol{\Sigma}$:

$$\boldsymbol{\Sigma} = \frac{\sigma^2}{N} (F_{\boldsymbol{\theta}}^{-1} \oplus 1/w_{ak}).$$

Weiterhin können die Zieldosis und der Zieldosis-Schätzer bei Existenz der Umkehrfunktion der Wirkungskurve als eine Funktion von $\boldsymbol{\theta}_{d^*}$ beziehungsweise $\widehat{\boldsymbol{\theta}}_{d^*}$ geschrieben werden (z. B. $d^* = f_{\boldsymbol{\theta}}^{-1}(\mu) = g(\boldsymbol{\theta}_{d^*})$). Unter der Annahme der stetigen Differenzierbarkeit von $g(\boldsymbol{\theta}_{d^*})$ ist der Gradient von $g(\boldsymbol{\theta}_{d^*})$ durch

$$\Delta g(\boldsymbol{\theta}_{d^*}) = \left(\frac{\partial g}{\partial \theta_0}, \dots, \frac{\partial g}{\partial \theta_b}, \frac{\partial g}{\partial \mu} \right) \quad (2.9)$$

gegeben. Dann kann die asymptotische Verteilung des Zieldosis-Schätzers mit Hilfe des folgenden Korollars bestimmt werden.

Korollar 2.5 Sei $\widehat{\boldsymbol{\theta}}_{d^*}$ der Parameterschätzer aus Gl. (2.8) für $\boldsymbol{\theta}_{d^*}$ und $\widehat{d}^* = g(\widehat{\boldsymbol{\theta}}_{d^*})$ der konsistente Schätzer (Kor. 2.4) der Zieldosis $d^* = g(\boldsymbol{\theta}_{d^*})$, dann gilt unter Voraussetzung 2.1:

$$\sqrt{N} (\widehat{d}^* - d^*) \xrightarrow{L} \mathcal{N}(0, \tau^2)$$

mit

$$\tau^2 = N \cdot \text{Var}(\widehat{d}^*) = \sigma^2 \Delta g(\boldsymbol{\theta}_{d^*}) \left(F_{\boldsymbol{\theta}}^{-1} \oplus \frac{1}{w_{ak}} \right) \Delta g(\boldsymbol{\theta}_{d^*})'. \quad (2.10)$$

BEWEIS. Der Beweis folgt aus dem Satz von Cramér (Ferguson, 1996, S. 45). \square

Mit diesem Resultat lässt sich ein $(1 - \alpha)$ -Konfidenzintervall für die Zieldosis konstruieren.

Korollar 2.6 Sei $\hat{\boldsymbol{\theta}}_{d^*}$ der Vektor der Parameterschätzer aus Gl. (2.8) und $\hat{\sigma}^2$ der in Abschnitt 2.3 eingeführte ML-Schätzer der Varianz der Fehlerterme, dann ergibt sich der konsistente Schätzer $\hat{\tau}^2$ der unbekanntes Varianz τ^2 durch das Einsetzen der ML-Schätzer in Gl. (2.10).

BEWEIS. Die Konsistenz des Plug-in Schätzers $\hat{\tau}^2$ folgt analog zum Beweis der Konsistenz des Zieldosissschätzers \hat{d}^* (Kor. 2.4) mit Hilfe des Satzes von Slutsky (siehe Ferguson, 1996, S. 39-43). \square

Nachdem die asymptotische Verteilung des Zieldosissschätzers \hat{d}^* bestimmt werden konnte, ist es im Folgenden möglich ein zweiseitiges $(1 - \alpha)$ -Konfidenzintervall der Zieldosis zu konstruieren. Dabei wird die Verteilung des Zieldosissschätzers durch die asymptotische Normalverteilung (Kor. 2.5) approximiert und die unbekanntes Varianz des Zieldosissschätzers τ^2 durch den Plug-in Schätzer $\hat{\tau}^2$ aus Korollar 2.6 ersetzt. Dann kann ein $(1 - \alpha)$ -Konfidenzintervall der Zieldosis wie folgt konstruiert werden:

$$KI_{\Delta} = \left[\hat{d}^* - \frac{u_{1-\frac{\alpha}{2}}}{\sqrt{N}} \sqrt{\hat{\tau}^2}, \hat{d}^* + \frac{u_{1-\frac{\alpha}{2}}}{\sqrt{N}} \sqrt{\hat{\tau}^2} \right], \quad (2.11)$$

wobei $u_{1-\frac{\alpha}{2}}$ das $(1 - \alpha/2)$ -Quantil der Standardnormalverteilung ist. Alternativ wäre es möglich anstelle der Normalapproximation zum Beispiel eine t-Approximation zu verwenden, da die unbekanntes Varianz geschätzt werden muss. Allerdings sind in Dosis-Findungsstudien mit aktiver Kontrolle in der Regel Gesamtfallzahlen $N > 50$ zu erwarten (siehe Abs. 1.2), so dass die Unterschiede zwischen den Quantilen der beiden Verteilungsapproximationen sehr gering sind. Im folgenden Kapitel werden Dosis-Findungsstudien mit einer linearen oder log-linearen Dosis-Wirkungsfunktion betrachtet und verschiedene Ansätze zur Konstruktion eines Konfidenzintervalls für die Zieldosis vorgestellt.

3. Auswertung von Dosis-Findungsstudien mit aktiver Kontrolle und linearen Dosis-Wirkungskurven

In diesem Kapitel werden lineare oder linearisierte Funktionen f_{θ} z. B. durch Logarithmierung der Dosisstufen als Dosis-Wirkungskurve angenommen, wodurch es möglich ist, Resultate aus der Theorie der linearen Modelle auf das Problem der Dosis-Findung zu übertragen und eine übersichtliche Darstellung zu ermöglichen. Unter den Modellannahmen aus Abschnitt 2.2 sind die Schätzer der Maximum-Likelihood Theorie identisch mit den Kleinste-Quadrate-Schätzern.

3.1. Statistisches Modell und Schätzer für (log-)lineare Dosis-Wirkungskurven

3.1.1. Statistisches Modell

Die Dosis-Wirkungskurve sei im Folgenden durch $f_{\theta}(d) = \theta_0 + \theta_1 d$ gegeben. Das statistische Modell lässt sich mit Hilfe der Vektorschreibweise wie folgt darstellen:

$$\mathbf{Y}_d = \mathbf{X}\boldsymbol{\theta} + \boldsymbol{\epsilon}_d \tag{3.1}$$

$$\mathbf{Y}_{ak} = \mu \mathbf{1}_{n_{ak}} + \boldsymbol{\epsilon}_{ak}$$

mit Parametervektor $\boldsymbol{\theta} = (\theta_0, \theta_1)'$, Designmatrix $\mathbf{X} = (\mathbf{X}'_1 | \dots | \mathbf{X}'_k)'$ und Designmatrix der i -ten Dosisstufe $\mathbf{X}_i = \mathbf{1}_{n_i} \otimes (1 : d_i)$, $i = 1, \dots, k$, sowie den unabhängigen Fehlertermen $\boldsymbol{\epsilon}_d \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I}_{n_d})$ und $\boldsymbol{\epsilon}_{ak} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I}_{n_{ak}})$ (analog zu Abs. 2.2).

3.1.2. Die Zieldosis d^*

Durch die Annahme einer linearen Wirkungsfunktion der Form $f_{\theta}(d) = \theta_0 + \theta_1 d$ vereinfacht sich die Definition der Zieldosis d^* aus Gl. (2.6) mit $\theta_1 \neq 0$ zu:

$$d^* = \frac{\mu - \theta_0}{\theta_1}. \tag{3.2}$$

Die Schätzer der unbekannt Parameter $\boldsymbol{\theta}$ und μ werden im folgenden Abschnitt mittels der Theorie linearer Modelle vorgestellt. Für die weiteren Betrachtungen wird $\theta_1 \neq 0$ angenommen, um die Existenz der Zieldosis zu gewährleisten.

3.1.3. Schätzer der Parameter und der Zieldosis und deren Eigenschaften im linearen Modell

Im Falle einer linearen Wirkungskurve kann die Verteilung der Parameterschätzer direkt bestimmt werden, ohne auf asymptotische Resultate zurückgreifen zu müssen. Weiterhin sind die Schätzer unverzerrt und besitzen minimale Varianz unter allen erwartungstreuen Schätzern (Ravishanker und Dey, 2002, Kap. 4).

Korollar 3.1 Sei $\mathbf{Y}_d = \mathbf{X}\boldsymbol{\theta} + \boldsymbol{\epsilon}$ der Zufallsvektor aus Gl. (3.1) mit $\boldsymbol{\epsilon}_d \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I}_{n_d})$ und $\det(\mathbf{X}'\mathbf{X}) \neq 0$, dann ist der Schätzer des Parametervektors $\boldsymbol{\theta}$ gegeben durch

$$\hat{\boldsymbol{\theta}} = (\hat{\theta}_0, \hat{\theta}_1)' = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{Y}_d. \quad (3.3)$$

Der Schätzer ist multivariat normalverteilt mit Erwartungswertvektor $\boldsymbol{\theta}$ und Kovarianzmatrix $\sigma^2(\mathbf{X}'\mathbf{X})^{-1}$:

$$\hat{\boldsymbol{\theta}} \sim \mathcal{N}(\boldsymbol{\theta}, \sigma^2 (\mathbf{X}'\mathbf{X})^{-1}). \quad (3.4)$$

Weiterhin ist dieser Schätzer erwartungstreu und besitzt minimale Varianz unter allen erwartungstreuen Schätzern.

BEWEIS. Aus der Definition der Designmatrix \mathbf{X} folgt $\det(\mathbf{X}'\mathbf{X}) \neq 0$ und die Existenz von $(\mathbf{X}'\mathbf{X})^{-1}$, so dass die Bedingung $\mathbf{X}'(\sigma^2 \mathbf{I}_{n_d}) [\mathbf{I}_{n_d} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'] = 0$ erfüllt ist. Die Erwartungstreue und die Eigenschaft der minimalen Varianz der Schätzers folgen dann direkt aus dem Satz von Gauss-Markov-Aitken (Satz A.6 im Anhang). Somit ist der Schätzer $\hat{\boldsymbol{\theta}}$ BLUE (best linear unbiased estimator) für den Parametervektor $\boldsymbol{\theta}$. Die Normalverteilung des Schätzers folgt direkt mit Hilfe von Definition A.3 aus Satz A.4. \square

Der Schätzer des Effekts der aktiven Kontrolle wurde bereits mit den entsprechenden Eigenschaften in Korollar 2.1 definiert, kann aber auch analog mit der Designmatrix $\mathbf{1}_{n_{ak}}$ hergeleitet werden

$$\hat{\mu} = (\mathbf{1}'_{n_{ak}} \mathbf{1}_{n_{ak}})^{-1} \mathbf{1}_{n_{ak}} \mathbf{Y}_{ak} = \frac{1}{n_{ak}} \sum_{j=1}^{n_{ak}} Y_{ak,j} = \bar{Y}_{ak} \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n_{ak}}\right) \quad (3.5)$$

und ist nach der Argumentation von Korollar 3.1 ebenfalls ein erwartungstreu Schätzer für μ und BLUE (Satz A.6). Zusätzlich kann die Varianz der Fehlerterme mit Hilfe der Theorie linearer

Modelle erwartungstreu geschätzt werden. Dazu werden die Informationen aller Zufallsvariablen verwendet werden.

Korollar 3.2 Sei $\mathbf{Y} = (\mathbf{Y}'_d, \mathbf{Y}'_{ak})' = \mathbf{X}_t \cdot (\boldsymbol{\theta}', \mu) + \boldsymbol{\epsilon}$ der Vektor aller Zufallsvariablen mit entsprechender Designmatrix $\mathbf{X}_t = (\mathbf{X} \oplus \mathbf{1}_{n_{ak}})$ mit $\det(\mathbf{X}'_t \mathbf{X}_t) \neq 0$ (siehe Abs. 3.1.1) und zugehörigem Fehlerterm $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I}_N)$, dann kann der erwartungstreue Schätzer der Varianz wie folgt definiert werden:

$$\hat{\sigma}^2 = \frac{1}{N-3} \mathbf{Y}' \left(\mathbf{I}_N - \mathbf{X}_t (\mathbf{X}'_t \mathbf{X}_t)^{-1} \mathbf{X}'_t \right) \mathbf{Y}. \quad (3.6)$$

Für die Verteilung von $\hat{\sigma}^2$ gilt:

$$\hat{\sigma}^2 \sim \frac{\sigma^2}{N-3} \cdot \chi_{N-3}^2. \quad (3.7)$$

BEWEIS. Mit $r(\mathbf{X}_t) = 3$, dem Rang der Matrix \mathbf{X}_t , folgt aus Satz A.7 (im Anhang) die Erwartungstreue sowie die Verteilung des Varianzschätzers. \square

Dieser Varianzschätzer ist nicht identisch mit dem verzerrten Schätzer der ML-Theorie. Nachdem alle unbekannt Parameter geschätzt sind, kann im Folgenden ein konsistenter Zieldosischätzer für das lineare Modell durch Einsetzen der Parameterschätzer $\hat{\boldsymbol{\theta}}$ und $\hat{\mu}$ aus Gl. (3.1) und Gl. (3.5) in Gl. (3.2) angegeben werden:

$$\hat{d}^* = \frac{(\hat{\mu} - \hat{\theta}_0)}{\hat{\theta}_1}. \quad (3.8)$$

Die Konsistenz ist dann durch Korollar 2.4 gewährleistet. Sollte der Schätzer der Steigung $\hat{\theta}_1 = 0$ sein, kann der Zieldosischätzer \hat{d}^* nicht berechnet werden. Da die Zufallsvariablen $\hat{\theta}_0$ und $\hat{\theta}_1$ korreliert sind, kann die Verteilung von \hat{d}^* selbst im linearen Fall nicht analytisch berechnet werden. Die Eigenschaften dieses Punktschätzers werden im nächsten Abschnitt beschrieben.

3.1.4. Eigenschaften des Schätzers der Zieldosis

Im Folgenden werden der Erwartungswert und eine mögliche Verzerrung des Zieldosischätzers untersucht. Da sich der Schätzer $\hat{d}^* = (\hat{\mu} - \hat{\theta}_0)/\hat{\theta}_1$ als Quotient von abhängigen Zufallsvariablen darstellt, ist es nicht möglich den Erwartungswert direkt zu berechnen. Stattdessen wird eine Taylor-Approximation zweiter Ordnung verwendet, um den Erwartungswert anzunähern. Wie bereits in Abschnitt 2.4.3 definiert, lässt sich die Zieldosis d^* als Funktion des Parametervektors $\boldsymbol{\theta}_{d^*}$ mit $d^* = g(\boldsymbol{\theta}_{d^*})$ schreiben. Für die Berechnung der Taylor-Approximation werden der Gra-

dient $\Delta g(\boldsymbol{\theta}_{d^*})$ aus Gl. (2.9) und die Hessematrix $\mathbf{H}(\boldsymbol{\theta}_{d^*}) = \partial^2 g / \partial \boldsymbol{\theta}_{d^*} \partial \boldsymbol{\theta}_{d^*}'$ von g an der Stelle $\boldsymbol{\theta}_{d^*}$ benötigt, die im Falle einer linearen Wirkungsfunktion wie folgt dargestellt werden können:

$$\Delta g(\boldsymbol{\theta}_{d^*}) = \left(-\frac{1}{\theta_1}, \quad -\frac{(\mu - \theta_0)}{\theta_1^2}, \quad \frac{1}{\theta_1} \right), \quad \mathbf{H}(\boldsymbol{\theta}_{d^*}) = \begin{pmatrix} 0 & \frac{1}{\theta_1^2} & 0 \\ \frac{1}{\theta_1^2} & \frac{2 \cdot (\mu - \theta_0)}{\theta_1^3} & -\frac{1}{\theta_1^2} \\ 0 & -\frac{1}{\theta_1^2} & 0 \end{pmatrix} \quad (3.9)$$

mit $\theta_1 \neq 0$. Dann folgt die Approximation des Erwartungswertes der Zieldosis mit Satz B.1 (siehe Anhang):

$$E(\hat{d}^*) \approx d^* + \frac{1}{2} \text{Sp}[\mathbf{H}(\boldsymbol{\theta}_{d^*})\boldsymbol{\Sigma}].$$

Die Verzerrung des Zieldosischätzers (Bias) ist dann durch $\text{Bias}(\hat{d}^*) = E(\hat{d}^*) - d^* = E(\hat{d}^*) - d^*$ gegeben. Somit ergibt sich durch die Verwendung der obigen Resultate die folgende Approximation:

$$\begin{aligned} \text{Bias}(\hat{d}^*) &\approx \frac{1}{2} \text{Sp}[\mathbf{H}(\boldsymbol{\theta}_{d^*})\boldsymbol{\Sigma}] = \frac{1}{2} \text{Sp} \left[\sigma^2 \mathbf{H}(\boldsymbol{\theta}_{d^*}) \left((\mathbf{X}'\mathbf{X})^{-1} \oplus \frac{1}{n_{ak}} \right) \right] \\ &= \frac{\sigma^2}{N\theta_1^2} \frac{\sum_{i=1}^k w_i [d^* - d_i]}{\left(\sum_{i=1}^k w_i \right) \left(\sum_{i=1}^k w_i d_i^2 \right) - \left(\sum_{i=1}^k w_i d_i \right)^2}, \end{aligned}$$

wobei w_i das in Abschnitt 2.2 definierte Gewicht der einzelnen Dosisstufen und \mathbf{X} die Designmatrix aus Gl. (3.1) bezeichnet. Da die Bias-Approximation in geschlossener Form dargestellt werden kann, ist es möglich die folgenden Eigenschaften abzuleiten. Für $d^* = (\sum_{i=1}^k w_i d_i) / (1 - w_{ak})$ ist die approximative Verzerrung gleich null. Weiterhin ist ersichtlich, dass die Verzerrung von der Varianz des Fehlerterms σ^2 , der Steigung der Wirkungsfunktion θ_1 und der Fallzahl N abhängt. Insbesondere geht die approximative Verzerrung für $N \rightarrow \infty$ gegen 0. Die Güte dieser Approximation wird für endliche Fallzahlen in Abschnitt 3.3 mittels Simulationsstudien untersucht.

3.2. Ansätze zur Konstruktion von Konfidenzintervallen für die Zieldosis

In diesem Abschnitt werden verschiedene Ansätze zur Berechnung eines Konfidenzintervalls (KI) für die Zieldosis vorgestellt. Alle Verfahren verwenden den in Gl. (3.8) definierten Schätzer der Zieldosis \hat{d}^* zur Konstruktion des jeweiligen Konfidenzintervalls.

3.2.1. Normalapproximation mittels Δ -Regel

Die Normalapproximation mittels Δ -Regel wird im Folgenden auf den linearen Fall angewendet. Die Parameterschätzer $\hat{\boldsymbol{\theta}}$ und $\hat{\mu}$ aus Gl. (3.3) und Gl. (3.5), die zur Bestimmung des Zieldosis-Schätzers \hat{d}^* benötigt werden, besitzen nach Definition A.3 die folgende gemeinsame Verteilung

$$\hat{\boldsymbol{\theta}}_{d^*} = \begin{pmatrix} \hat{\theta}_0 \\ \hat{\theta}_1 \\ \hat{\mu} \end{pmatrix} \sim \mathcal{N}(\boldsymbol{\theta}_{d^*}, \boldsymbol{\Sigma})$$

mit Erwartungswertvektor $\boldsymbol{\theta}_{d^*} = (\theta_0, \theta_1, \mu)'$ und Kovarianzmatrix $\boldsymbol{\Sigma} = \sigma^2 \cdot ((\mathbf{X}'\mathbf{X})^{-1} \oplus 1/n_{ak})$. Weiter ist die Zieldosis im Falle einer linearen Wirkungsfunktion eindeutig bestimmt und kann in der Form $d^* = g(\boldsymbol{\theta}_{d^*}) = (\mu - \theta_0)/\theta_1$ geschrieben werden (siehe Abs. 3.1.4). Der zugehörige Gradient $\Delta g(\boldsymbol{\theta}_{d^*})$ der Funktion $g(\boldsymbol{\theta}_{d^*})$ ist in Gl. (3.9) gegeben. Im Falle einer linearen Dosis-Wirkungsfunktion folgt die asymptotische Verteilung des Zieldosis-Schätzers aus Korollar 2.5, mit Voraussetzung 2.1:

$$\sqrt{N} (\hat{d}^* - d^*) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \tau^2) \quad (3.10)$$

mit

$$\begin{aligned} \tau^2 &= \sigma^2 \Delta g(\boldsymbol{\theta}_{d^*}) \left(\tilde{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}} \oplus 1/w_{ak} \right) \Delta g(\boldsymbol{\theta}_{d^*})' \\ &= \frac{\sigma^2}{\theta_1^2} \left[\frac{\left(\sum w_i [d_i - d^*]^2 \right)}{\left((\sum w_i) (\sum w_i d_i^2) - (\sum w_i d_i)^2 \right)} + \frac{1}{w_{ak}} \right]. \end{aligned} \quad (3.11)$$

Durch Einsetzen der Schätzer \hat{d}^* aus Gl. (3.2), $\hat{\boldsymbol{\theta}}$ aus Gl. (3.3) und $\hat{\sigma}^2$ aus Gl. (3.6) in Gl. (3.11) wird der Plug-in Schätzer $\hat{\tau}^2$ für die unbekannte Varianz τ^2 generiert:

$$\hat{\tau}^2 = \frac{\hat{\sigma}^2}{\hat{\theta}_1^2} \left[\frac{\left(\sum w_i [d_i - \hat{d}^*]^2 \right)}{\left((\sum w_i) (\sum w_i d_i^2) - (\sum w_i d_i)^2 \right)} + \frac{1}{w_{ak}} \right]. \quad (3.12)$$

Nach Korollar 2.6 und dem Satz von Slutsky (Ferguson, 1996, S. 39-43) ist $\hat{\tau}^2$ ein konsistenter Schätzer für die Varianz τ^2 . Ein $(1 - \alpha)$ -Konfidenzintervall der Zieldosis kann analog zu Gl. (2.11) konstruiert werden:

$$KI_{\Delta} = \left[\hat{d}^* - u_{1-\frac{\alpha}{2}} \sqrt{\hat{\tau}^2}, \hat{d}^* + u_{1-\frac{\alpha}{2}} \sqrt{\hat{\tau}^2} \right]$$

mit dem $(1 - \alpha/2)$ -Quantil der Standardnormalverteilung $u_{1-\frac{\alpha}{2}}$.

3.2.2. Profile-Likelihood Methode

Der Ansatz der Profile-Likelihood (PL) Methode verwendet nicht die Normalapproximation mit Hilfe der Δ -Regel, sondern die Likelihood Funktion selbst, um ein Konfidenzintervall für die Zieldosis zu konstruieren. Dafür wird in der Likelihood Funktion die Zieldosis d^* durch eine feste, bekannte Konstante ersetzt und anschließend wird das Maximum der Likelihood Funktion in Abhängigkeit dieser festen Dosis berechnet (Pawitan, 2001, S. 61-64). Dieses Vorgehen wird für eine Vielzahl von $d \in D = [d_1, d_k]$ wiederholt. Die Ergebnisse der einzelnen Likelihood Funktionen ergeben die Profile-Likelihood. Die entsprechende Likelihood Funktion folgt aus Gl. (2.1) mit $\mu = \theta_0 + \theta_1 d^*$:

$$\begin{aligned} L(d^*, \boldsymbol{\theta}, \sigma^2) &= L(d^*, \boldsymbol{\theta}, \sigma^2 | \mathbf{Y}) \\ &= \gamma^N \exp \left[-\frac{\sum_{i=1}^k \sum_{j=1}^{n_i} (Y_{ij} - \mathbf{X}_i \boldsymbol{\theta})^2 + \sum_{j'=1}^{n_{ak}} (Y_{ak,j'} - \theta_0 - \theta_1 d^*)^2}{2\sigma^2} \right], \end{aligned} \quad (3.13)$$

wobei \mathbf{X}_i die Designmatrix der i -ten Dosisstufe aus Gl. (3.1) und $\gamma = 1/\sqrt{2\pi\sigma^2}$ darstellen. Wird nun die Zieldosis d^* durch eine bekannte Konstante \tilde{d} ersetzt, können die Schätzer der verbleibenden Parameter in Abhängigkeit von \tilde{d} berechnet werden. Die Profile-Likelihood ist für festes \tilde{d} durch

$$L(\tilde{d}) = \max_{\boldsymbol{\theta}, \sigma^2} L(\tilde{d}, \boldsymbol{\theta}, \sigma^2)$$

definiert. Das lineare Modell aus Abschnitt 3.1.1 kann unter der Annahme einer festen Dosis \tilde{d} wie folgt beschrieben werden:

$$\begin{aligned} \mathbf{Y} &= \begin{pmatrix} \mathbf{Y}_d \\ \mathbf{Y}_{ak} \end{pmatrix} = \begin{pmatrix} \mathbf{X} \\ \mathbf{X}_{ak}(\tilde{d}) \end{pmatrix} \begin{pmatrix} \theta_0 \\ \theta_1 \end{pmatrix} + \begin{pmatrix} \boldsymbol{\epsilon}_d \\ \boldsymbol{\epsilon}_{ak} \end{pmatrix} \\ &= \mathbf{X}_p(\tilde{d})\boldsymbol{\theta} + \boldsymbol{\epsilon}, \end{aligned}$$

mit der Designmatrix $\mathbf{X}_{ak}(\tilde{d}) = (\mathbf{1}_{n_{ak}}; \tilde{d}\mathbf{1}_{n_{ak}})$ der Dosis \tilde{d} , so dass $\det(\mathbf{X}_p(\tilde{d})'\mathbf{X}_p(\tilde{d})) \neq 0$ gilt. Mit Hilfe dieser Notation und den in Abschnitt 3.1.3 präsentierten Parameterschätzern ist es möglich die Schätzer für $\boldsymbol{\theta}$ und σ^2 in Abhängigkeit von \tilde{d} zu berechnen:

$$\hat{\boldsymbol{\theta}}_p(\tilde{d}) = \left(\mathbf{X}_p(\tilde{d})'\mathbf{X}_p(\tilde{d}) \right)^{-1} \mathbf{X}_p(\tilde{d})'\mathbf{Y}$$

und

$$\begin{aligned}\hat{\sigma}_p^2(\tilde{d}) &= \frac{1}{N-2} \cdot \mathbf{Y}' \left(\mathbf{I}_N - \mathbf{X}_p(\tilde{d}) \left[\mathbf{X}_p(\tilde{d})' \mathbf{X}_p(\tilde{d}) \right]^{-1} \mathbf{X}_p(\tilde{d})' \right) \mathbf{Y} \\ &= \frac{1}{N-2} \sum_{i=1}^N \left(Y_i - \mathbf{X}_{p,i}(\tilde{d}) \cdot \hat{\boldsymbol{\theta}}_p(\tilde{d}) \right)^2,\end{aligned}$$

wobei $\mathbf{X}_{p,i}(\tilde{d})$ die i -te Zeile der Matrix $\mathbf{X}_p(\tilde{d})$ beschreibt. Dann ergibt sich die Profile-Likelihood-Gleichung für die Dosis \tilde{d} durch Einsetzen der berechneten Schätzer:

$$\begin{aligned}L(\tilde{d}) &= \frac{1}{\left(\sqrt{2\pi\hat{\sigma}_p^2(\tilde{d})} \right)^N} \cdot \exp \left[-\frac{\sum_{i=1}^N \left(Y_i - \mathbf{X}_{p,i}(\tilde{d}) \hat{\boldsymbol{\theta}}_p(\tilde{d}) \right)^2}{2\hat{\sigma}_p^2(\tilde{d})} \right] \\ &= \frac{1}{\left(\sqrt{2\pi\hat{\sigma}_p^2(\tilde{d})} \right)^N} \cdot \exp \left[-\frac{(N-2)}{2} \right].\end{aligned}$$

Das Maximum der Profile-Likelihood $L_{\max} = \max_{\tilde{d}} L(\tilde{d})$ liegt beim ML-Schätzer $\hat{d}^* = (\hat{\mu} - \hat{\theta}_0) / \hat{\theta}_1$ und es ergibt sich $L_{\max} = L(\hat{d}^*)$. Um das Konfidenzintervall der Zieldosis zu bestimmen, wird die Profile-Likelihood mittels L_{\max} normiert, so dass für den Quotienten $L(\tilde{d}) / L(\hat{d}^*) \in [0, 1]$ gilt. Dann hat die Wilks Likelihood Ratio Statistik W (siehe z. B. Pawitan, 2001, S. 36-37) unter $H_0 : \tilde{d} = d^*$ die folgende Verteilung:

$$W = 2 \log \frac{L(\hat{d}^*)}{L(\tilde{d})} \xrightarrow{\mathcal{L}} \chi_1^2.$$

Daraus folgt $P(L(\tilde{d}) / L(\hat{d}^*) > q_{1-\alpha}) = 1 - \alpha$ mit $q_{1-\alpha} = e^{-\frac{1}{2}\chi_{1,1-\alpha}^2}$ und das $(1 - \alpha)$ -Konfidenzintervall für die Zieldosis

$$KI_{PL} = \left\{ \tilde{d} \mid \frac{L(\tilde{d})}{L(\hat{d}^*)} > q_{1-\alpha} \right\}. \quad (3.14)$$

Im Falle eines 95%-Konfidenzintervalls sind alle \tilde{d} enthalten, die die Bedingung $L(\tilde{d}) / L(\hat{d}^*) > 0,1465$ erfüllen. Im folgenden Abschnitt werden numerische Verfahren zur Berechnung der Intervallgrenzen vorgestellt.

3.2.3. Numerische Berechnung der Intervallgrenzen

Um die Intervallgrenzen numerisch bestimmen zu können, muss zuerst der Dosisbereich so eingeschränkt werden, dass jeweils nur die gewünschte Intervallgrenze enthalten ist. Damit ergeben sich die folgenden Gleichungen zur Berechnung der Intervallgrenzen aus Gl. (3.14) der Profile-Likelihood Methode. Für die untere Konfidenzintervallgrenze gilt:

$$KI_u(\tilde{d}) = \frac{L(\tilde{d})}{L(\hat{d}^*)} - q_{1-\alpha} \stackrel{!}{=} 0 \text{ mit } \tilde{d} \in [d_1 - v, \hat{d}^*]$$

und für die obere Konfidenzintervallgrenze:

$$KI_o(\tilde{d}) = \frac{L(\tilde{d})}{L(\hat{d}^*)} - q_{1-\alpha} \stackrel{!}{=} 0 \text{ mit } \tilde{d} \in [\hat{d}^*, d_k + v], \quad (3.15)$$

wobei $v \geq 0$ eine Suche über den betrachteten Dosisbereich D hinaus ermöglicht, falls eine Extrapolation zugelassen wird. In allen betrachteten Szenarien wurde $v = 0$ gewählt. Das Lösen dieser beiden Gleichungen erfolgt mit Hilfe des Intervallhalbierungsverfahrens, das im Folgenden für die untere (bzw. obere) Intervallgrenze allgemein vorgestellt wird.

Algorithmus zur Intervallhalbierung

Input: Zielfunktion $KI_u(\cdot)$ (bzw. $KI_o(\cdot)$) und
die Grenzen des zu untersuchenden Intervalls: $[u, o]$ mit $0 \leq u < o$.

Sei $t = 0$

1. Definition des Startwertes: $x = (o - u)/2$.
2. Berechnung von $KI_u(x)$ (bzw. $KI_o(x)$) und Überprüfung, ob
 $|KI_u(x)| < 0,001$ und $KI_u(x) > 0$ (bzw. $|KI_o(x)| < 0,001$ und $KI_o(x) > 0$) gilt.
3. Falls 2. erfüllt ist \rightarrow Abbruch des Algorithmus und wähle $d_u = x$ (bzw. $d_o = x$) als
untere (bzw. obere) Grenze.
4. Falls 2. nicht erfüllt ist \rightarrow Anpassen des Suchbereichs und Neubestimmung von x .

Falls $KI_u(x) < 0$ (bzw. $KI_o(x) > 0$): wähle $u = x$ als untere Grenze.

Falls $KI_u(x) \geq 0$ (bzw. $KI_o(x) \leq 0$): wähle $o = x$ als obere Grenze.

5. Setze $t = t + 1$. Der neue Wert für x ist: $x = (o - u)/2$.
6. Wiederhole die Schritte 2 bis 5 bis 3. erfüllt ist.
7. Falls nach $t = 30$ Iterationen kein Ergebnis gefunden wurde, bricht der Algorithmus
ab und verwendet die triviale Intervallgrenze $d_u = 0$ (bzw. $d_o = \infty$, was zu
einseitigen Intervallen führt).

Output: t und d_u (bzw. d_o).

Das Abbruchkriterium wurde nach $t = 30$ Iterationen gewählt, um zu gewährleisten, dass die Intervalllänge des Suchbereiches kleiner als die geforderte Genauigkeit aus Schritt 2 ist: $o - u < 0,001$. Nach $t = 30$ Iterationen ist die ursprüngliche Intervalllänge aus Schritt 1 um den Faktor $(1/2)^{30}$ verringert, so dass die Bedingung für Start-Intervalllängen kleiner 10^6 erfüllt ist. Sollten größere Dosisbereiche betrachtet werden, muss t entsprechend angepasst werden.

Neben dem Intervallhalbierungsverfahren existiert eine Vielzahl an verschiedenen Methoden zum Lösen der obigen Gleichungen, die in den meisten statistischen Software-Paketen (SAS/IML, R) implementiert sind. Zum Beispiel können die Gleichungen mit Hilfe des Newton-Raphson Verfahrens gelöst werden (Schöbel, 2007, S. 115-118), das in SAS/IML durch die Routine NLPNRA aufgerufen werden kann (SAS-Institute, 2008a, Kap. 11). Alternativ können die Intervallgrenzen mit Hilfe von Gl. (3.15) unter Verwendung der Levenberg-Marquardt Kleinste-Quadrate-Methode bestimmt werden, die ebenfalls in SAS/IML in der Prozedur NLPLM nach Moré (1978) implementiert ist (siehe SAS-Institute, 2008a, Kap. 11). Allerdings benötigen beide Verfahren die Exis-

tenz der ersten und zweiten Ableitung der Zielfunktion. Falls diese Verfahren zu keiner gültigen numerischen Lösung kommen, werden die Intervallgrenzen ebenfalls auf $d_u = 0$ (bzw. $d_o = \infty$) gesetzt.

3.2.4. Parametrischer Bootstrap

Wie bereits in Abschnitt 3.1.3 gezeigt, ist die Verteilung aller Parameterschätzer ($\hat{\mu}$ und $\hat{\theta}$) bekannt, die zur Berechnung des Schätzers der Zieldosis d^* benötigt werden. Daher ist es möglich einen parametrischen Bootstrap (Efron, 1979) basierend auf Monte-Carlo-Simulationen (MC) für die Konstruktion eines Konfidenzintervalls der Zieldosis d^* zu verwenden. Alternativ stehen dem Anwender eine Vielzahl anderer Bootstrap-Verfahren wie zum Beispiel Efron's nichtparametrischer Resampling-Bootstrap, der gewichtete oder der sogenannte Wild-Bootstrap zur Verfügung. Eine Übersicht wird zum Beispiel in Janssen und Pauls (2003) gegeben. Die Idee beim parametrischen Bootstrap mit MC ist die Annäherung der unbekanntenen Verteilung des Zieldosis-Schätzers \hat{d}^* durch Simulation. Dafür werden basierend auf den Verteilungen der Schätzer $\hat{\mu}$ und $\hat{\theta}$ neue Sets von Parameterschätzern generiert. Die entsprechenden Bootstrap-Parameterschätzer lassen sich wie folgt definieren:

$$\hat{\theta}^b = \hat{\theta} + \mathbf{Z}_\theta^b \text{ mit } \mathbf{Z}_\theta^b \sim \mathcal{N}\left(\mathbf{0}, \hat{\sigma}^2 (\mathbf{X}'\mathbf{X})^{-1}\right)$$

und

$$\hat{\mu}^b = \hat{\mu} + Z_\mu^b \text{ mit } Z_\mu^b \sim \mathcal{N}\left(0, \frac{\hat{\sigma}^2}{n_{ak}}\right), \quad b = 1, \dots, n_{boot},$$

wobei $\mathbf{Z}_\theta^b, Z_\mu^b$ die gebootstrapteten Fehlerterme und $\hat{\sigma}^2$ den Varianzschätzer aus Gl. (3.6) beschreiben. Dann können in einem weiteren Schritt die Bootstrap-Schätzer der Zieldosis durch:

$$\hat{d}^{*b} = \frac{(\hat{\mu}^b - \hat{\theta}_0^b)}{\hat{\theta}_1^b}, \quad b = 1, \dots, n_{boot}$$

berechnet werden. Sei $\hat{\mathbf{d}}^{*b}$ der geordnete Vektor aller Bootstrap-Schätzer der Zieldosis, so dass ohne Einschränkung der Allgemeinheit das erste Vektorelement von $\hat{\mathbf{d}}^{*b}$ das Minimum und das letzte Vektorelement das Maximum über alle Bootstrap-Schätzer beschreibt. In einem anschließenden Schritt kann das $\alpha/2$ und $(1 - \alpha/2)$ -Quantil von $\hat{\mathbf{d}}^{*b}$ zum Beispiel mit Hilfe der Quantilmethode abgelesen werden, um das zugehörige Konfidenzintervall zu bestimmen

$$KI_{boot} = \left[\hat{\mathbf{d}}^{*b} \left[n_{boot} \cdot \frac{\alpha}{2} \right], \hat{\mathbf{d}}^{*b} \left[n_{boot} \cdot \left(1 - \frac{\alpha}{2}\right) \right] \right].$$

Zusätzlich existiert eine große Auswahl an Verfahren zur Bestimmung des Bootstrap-Konfidenzintervalls in der Literatur, wie zum Beispiel das Bias-Corrected- oder der Accelerated-Bootstrap-KI für Bootstrap-Verfahren mit und ohne MC (für weitere Erläuterungen siehe Efron, 1987).

3.2.5. Erweiterung der Methode von Fisch und Strehlau (1993)

Eine weitere Möglichkeit, ein Konfidenzintervall für die Zieldosis d^* zu bestimmen, ist eine Anwendung der folgenden Erweiterung der Methode von Fisch und Strehlau (1993), die für Kalibrierungsprobleme vorgeschlagen wurde. Dafür wird der Erwartungswert der Wirkung der aktiven Kontrolle in Gl. (3.1) durch $\mu = \theta_0 + \theta_1 d^*$ ersetzt, was das lineare Modell in ein nichtlineares transformiert. Anstatt eine klassische nichtlineare Regression durchzuführen, kann die ML-Gleichung des Schätzers der Zieldosis \hat{d}^* durch Ableiten der Log-Likelihood $\ln[L(d^*, \boldsymbol{\theta}, \sigma^2)]$ aus Gl. (3.13) nach d^* bestimmt und in die folgende Form umgeschrieben werden:

$$\bar{Y}_{ak\cdot} - \hat{\theta}_0 - \hat{\theta}_1 d = 0.$$

Hierbei wird die Dosis d als fest aber unbekannt angesehen. Alle notwendigen Parameterschätzer wurden bereits in Abschnitt 3.1.3 beschrieben und sind identisch mit den ML-basierten Schätzern. Die linke Seite der Gleichung kann als eine neue Zufallsvariable $V(d)$ mit folgender Verteilung

$$\begin{aligned} V(d) &= \bar{Y}_{ak\cdot} - \hat{\theta}_0 - \hat{\theta}_1 d \\ &= (\theta_0 - \hat{\theta}_0) + (\theta_1 - \hat{\theta}_1) \cdot d + \bar{\epsilon}_{ak\cdot} \sim \mathcal{N}(0, \tau(d)^2) \end{aligned}$$

gesehen werden, da $\bar{Y}_{ak\cdot} = \mu + \bar{\epsilon}_{ak\cdot}$ und $\mu = \theta_0 + \theta_1 d$ gilt. Die Varianz $\tau(d)^2$ von $V(d)$ kann durch

$$\begin{aligned} \tau(d)^2 &= \text{Var}(\hat{\theta}_0 + \hat{\theta}_1 \cdot d) + \text{Var}(\bar{\epsilon}_{ak\cdot}) \\ &= \sigma^2 \cdot \left[\frac{1}{n_{ak}} + \frac{1}{n_d} + (d - \bar{d})^2 / S_{dd} \right] \end{aligned}$$

mit $\bar{d} = \frac{1}{n_d} \sum_{i=1}^k n_i d_i$ und $S_{dd} = \sum_{i=1}^k n_i (d_i - \bar{d})^2$ bestimmt werden (Fisch und Strehlau, 1993). Der einzige unbekannt Parameter in $\tau(d)^2$ ist die Varianz des Fehlerterms σ^2 , die durch $\hat{\sigma}^2$ Gl. (3.6) geschätzt werden kann. Abschließend wird die neue Zufallsvariable $W(d)$ durch

$$\begin{aligned} W(d) &= (V(d) / \tau(d)) (\sigma^2 / \hat{\sigma}^2)^{-1/2} \\ &= \frac{V(d)}{\left(\hat{\sigma} \cdot \left[\frac{1}{n_{ak}} + \frac{1}{n_d} + (d - \bar{d})^2 / S_{dd} \right] \right)} \end{aligned}$$

mit $\tau(d) = \sqrt{\tau(d)^2}$ definiert. Für die Verteilung von $W(d)$ gilt

$$W(d) \sim t_{N-3},$$

wobei t_{N-3} die zentrale t-Verteilung mit $N - 3$ Freiheitsgraden beschreibt. Das Quantil der t -Verteilung der Variablen $W(d)$ hängt lediglich von der Gesamtfallzahl N ab. Aus diesem Grund ist es möglich ein Konfidenzintervall für die Zieldosis d^* zu berechnen, indem die Grenzen des $(1 - \alpha)$ -Intervalls für $W(d)$ in Abhängigkeit von d numerisch berechnet werden. Daraus ergibt sich das folgende $(1 - \alpha)$ -Konfidenzintervall für d^* mit $w(d)$ als Realisation von $W(d)$:

$$KI_{F\&S} = \left\{ d \mid -t_{N-3, 1-\frac{\alpha}{2}} \leq w(d) \leq t_{N-3, 1-\frac{\alpha}{2}} \right\}. \quad (3.16)$$

Diese Art der Konstruktion führt nicht immer zu endlichen Konfidenzintervallen. Die Bedingung für das Auftreten endlicher Konfidenzintervalle ist in Fisch und Strehlau (1993) mit Hilfe des linearen Bestimmtheitsmaßes $R^2 = \hat{\theta}_1^2 / S_{dd} \cdot S_{YY}$ mit $S_{YY} = \sum_{i=1}^k \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_d)^2$ und $\bar{Y}_d = 1/n_d \sum_{i=1}^k \sum_{j=1}^{n_i} Y_{ij}$ dargestellt. Falls die Bedingung $R^2 > t_{N-3, 1-\alpha/2} / (N-3 + t_{N-3, 1-\alpha/2})$ erfüllt ist, so ist das $(1 - \alpha)$ -Konfidenzintervall $KI_{F\&S}$ aus Gl. (3.16) endlich.

Zur numerischen Berechnung der Intervallgrenzen wird eine analoge Vorgehensweise wie in Abschnitt 3.2.3 gewählt. Zuerst muss der Dosisbereich wieder so eingeschränkt werden, dass jeweils nur die gewünschte Intervallgrenze enthalten ist. Dabei ist zu beachten, dass per Konstruktion $w(\hat{d}^*) = 0$ gilt und dass aus $d > \hat{d}^* \rightarrow w(d) < 0$ folgt und umgekehrt. Damit ergeben sich die folgenden Gleichungen zur Berechnung der Intervallgrenzen. Für die untere Konfidenzintervallgrenze:

$$KI_u(d) = w(d) - t_{N-3, 1-\frac{\alpha}{2}} \stackrel{!}{=} 0 \text{ mit } d \in [d_1 - v, \hat{d}^*]$$

und für die obere Konfidenzintervallgrenze:

$$KI_o(d) = w(d) + t_{N-3, 1-\frac{\alpha}{2}} \stackrel{!}{=} 0 \text{ mit } d \in [\hat{d}^*, d_k + v]$$

mit $v \geq 0$. Zum Lösen der beiden Gleichungen wird der Algorithmus zur Intervallhalbierung aus Abschnitt 3.2.3 verwendet.

3.3. Simulationsstudie für das lineare Modell

In diesem Abschnitt wird der Schätzer der Zieldosis \hat{d}^* bezüglich seiner Verzerrung untersucht. Zusätzlich werden die verschiedenen Methoden zur Konfidenzintervallberechnung aus Abschnitt 3.2 durch umfangreiche Simulationsstudien in Bezug auf Überdeckungswahrscheinlichkeit und medianer Konfidenzintervalllänge verglichen. Motiviert durch die Beispielstudie in Ab-

schnitt 1.2.1 wird im Folgenden ein sechsamiges Studiendesign angenommen, wobei Placebo als experimentelle Dosis mit Dosislevel $d_1 = 0$ angesehen wird. Des Weiteren seien die $k = 5$ untersuchten Dosisstufen äquidistant auf dem Intervall $[0, 1]$ verteilt, so dass $(d_1, \dots, d_5) = (0; 0,25; 0,5; 0,75; 1)$. Weiterhin gelte für die Fallzahlen pro Dosisstufe $n_i = n$ für $i = 1, \dots, 5$, sowie $n_{ak} = 2n$ für die Fallzahl der aktiven Kontrolle. Die Wirkung der aktiven Kontrolle wird auf $\mu = 1$ und der Achsenabschnitt der linearen Wirkungsfunktion auf $\theta_0 = 0$ gesetzt. In allen Simulationen ist die Anzahl der Wiederholungen und die Anzahl der Bootstrapsimulationen pro Simulationsdurchlauf durch $n_{sim} = n_{boot} = 10.000$ gegeben. Die Zufallszahlen für die verschiedenen Simulationsszenarien (Tab. 3.1) werden mit Hilfe von SAS 9.3 erzeugt. Dafür können die Zufallszahlen entweder direkt aus den Verteilungen der Parameterschätzern (Abs. 3.1.3) oder aus der Verteilung der N Patientendaten aus Gl. (3.1) gezogen werden (siehe Helms et al., 2012, zur Implementation in SAS 9.3).

Tabelle 3.1.: Simulationsszenarien für verschiedene Varianzen σ^2 , unterschiedliche Steigungen θ_1 , mit zugehörigen Zieldosen d^* und diversen Fallzahlen n pro Gruppe.

σ^2	θ_1	d^*	n
1	1,25	0,8	10, 12, ..., 30, 40, ..., 100
	2	0,5	10, 12, ..., 30, 40, ..., 100
	5	0,2	10, 12, ..., 30, 40, ..., 100
2	1,25	0,8	10, 12, ..., 30, 40, ..., 100
	2	0,5	10, 12, ..., 30, 40, ..., 100
	5	0,2	10, 12, ..., 30, 40, ..., 100

3.3.1. Verzerrung des Zieldosischätzers

In diesem Abschnitt werden die Eigenschaften des Schätzers der Zieldosis \hat{d}^* untersucht. Dazu wird die analytische Approximation der Verzerrung (Bias) aus Abschnitt 3.1.4 der empirischen Verzerrung des Schätzers gegenübergestellt. Abbildung 3.1 zeigt die approximierten und empirischen Verzerrungen in den verschiedenen Simulationsszenarien aus Tabelle 3.1. Es ist ersichtlich, dass die approximative Verzerrung die empirische Verzerrung sehr gut annähert und dass für wachsende Fallzahlen beide Verzerrungen gegen null gehen. Es kann also empirisch bestätigt werden, dass der Zieldosischätzer konsistent und somit asymptotisch erwartungstreu ist.

In den Szenarien mit Steigung $\theta_1 = 2$ ist die approximative Verzerrung gleich null, da die Bedingung $d^* = (\sum_{i=1}^k w_i d_i) / (1 - w_{ak}) = 1/k \sum_{i=1}^k d_i = 0,5$ aus Abschnitt 3.1.4 erfüllt ist. Die empirische Verzerrung ist in diesen Szenarien ebenfalls nahezu null. Allerdings ist es nicht empfehlenswert die Bias-Approximation unter Verwendung der Plug-in Schätzer aus Abschnitt 3.1.3 zu bestimmen, um damit die Verzerrung des Zieldosischätzers zu korrigieren. Diese Schätzung der approximativen Verzerrung führt wie erwartet in vielen Situationen, besonders für kleine Fallzahlen, zu massiven Schwankungen und zu einer erhöhten empirischen Verzerrung. Falls die

approximative Verzerrung bekannt ist, kann die Verzerrung des Zieldosissschätzers damit reduziert werden. In den betrachteten Szenarien wurde auf eine Korrektur der Verzerrung verzichtet.

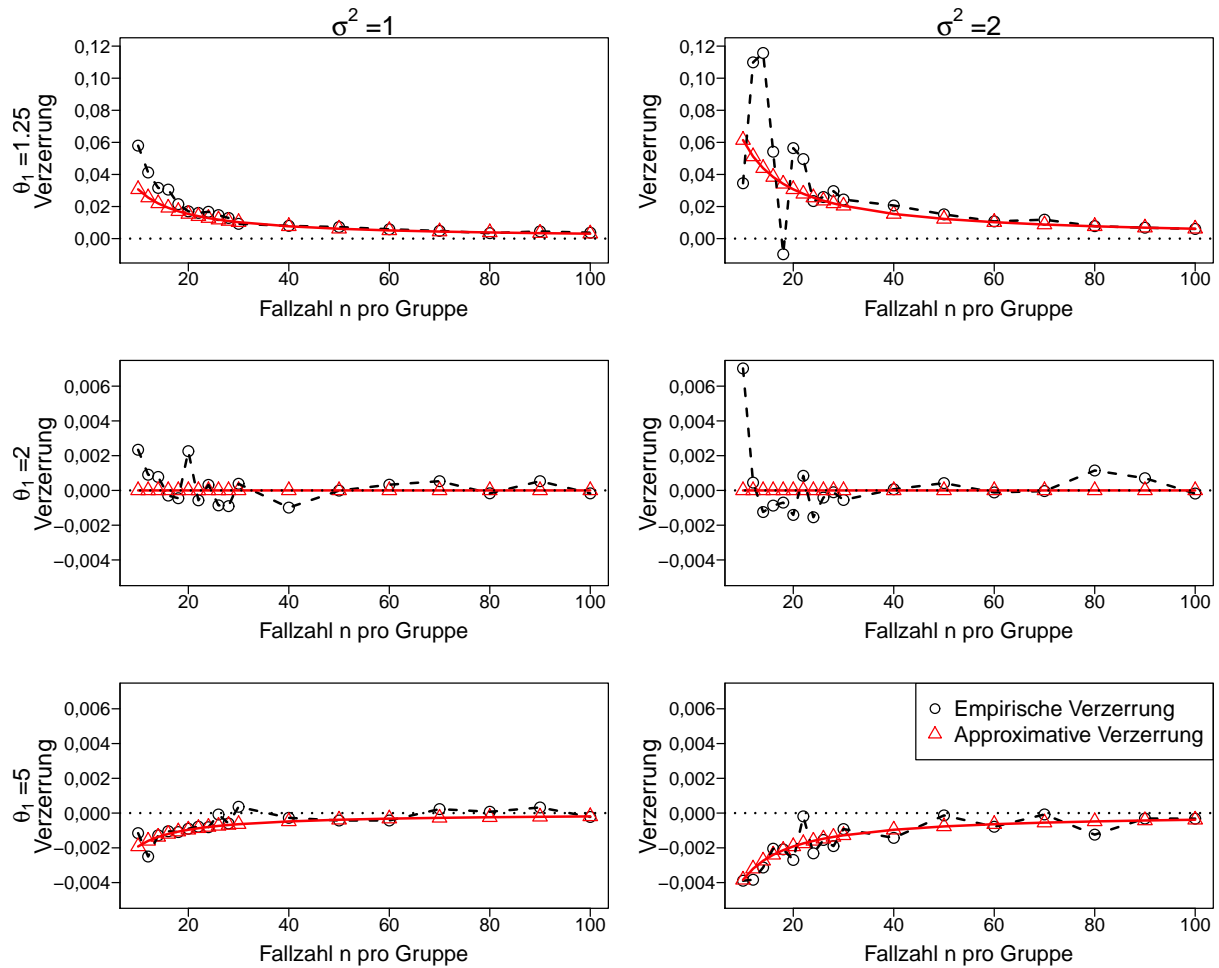


Abbildung 3.1.: Approximative und empirische Verzerrung des Zieldosissschätzers für $k = 5$ Dosisstufen aufgetragen gegen die Fallzahlen n pro Gruppe für verschiedene Varianzen σ^2 und Steigungen θ_1 der Dosis-Wirkungsgeraden.

3.3.2. Eigenschaften der Konfidenzintervalle

In diesem Abschnitt werden die Eigenschaften der verschiedenen Methoden zur Konstruktion eines Konfidenzintervalls mit Hilfe der Simulationsszenarien untersucht, die in Tabelle 3.1 zusammengefasst sind. Zunächst wird die Überdeckungswahrscheinlichkeit mit einem nominalen Niveau von $(1 - \alpha) = 0,95$ untersucht, indem die Überdeckungswahrscheinlichkeit empirisch aus den n_{sim} Simulationen berechnet wird. Um entscheiden zu können, ob die beobachteten Überdeckungswahrscheinlichkeiten systematisch vom nominalen Level abweichen oder nur Simulationsfehler sind, wird im Folgenden der $(1 - \alpha_z)$ -Zufallsstreifen der empirischen Überdeckungswahrscheinlichkeit \hat{p} eingeführt (siehe Helms, 2011, S. 53-54). Da sich \hat{p} als relative Häufigkeit von n_{sim} Bernoulli-verteilen Zufallsvariablen mit Erfolgswahrscheinlichkeit $(1 - \alpha)$

beschreiben lässt, folgt die allgemeine Darstellung des Zufallsstreifens mit Hilfe des zentralen Grenzwertsatzes (siehe z. B. Dehling und Haupt, 2004, S. 213-226)

$$ZS_{\hat{p}} = \left[(1 - \alpha) - \frac{1}{\sqrt{n_{sim}}} \cdot \sqrt{\alpha(1 - \alpha)} \cdot u_{1 - \frac{\alpha_z}{2}}; (1 - \alpha) + \frac{1}{\sqrt{n_{sim}}} \cdot \sqrt{\alpha(1 - \alpha)} \cdot u_{1 - \frac{\alpha_z}{2}} \right]$$

wobei $u_{1 - \frac{\alpha_z}{2}}$ das $1 - \frac{\alpha_z}{2}$ -Quantil der Standardnormalverteilung ist. Damit ergibt sich für die durchgeführten Simulationen der 99%-Zufallsstreifen des Simulationsfehlers:

$$ZS_{\hat{p}} = [0,94438; 0,95561], \quad (3.17)$$

zum Zufallslevel $\alpha_z = 0,01$.

Im Folgenden werden Überdeckungswahrscheinlichkeiten außerhalb des 99%-Intervalls des Simulationsfehlers, was einer Abweichung von rund $\pm 0,5\%$ entspricht, als systematisch verschieden vom nominalen 95% Niveau angesehen (siehe auch Friede et al., 2007). In Abbildung 3.2 sind die empirischen Überdeckungswahrscheinlichkeiten der untersuchten Methoden dargestellt, wobei die gepunkteten Linien das 99%-Intervall des Simulationsfehlers beschreiben. Es ist ersichtlich, dass die Profile-Likelihood Methode (*PL*) (Abs. 3.2.2), der parametrische Bootstrap (Abs. 3.2.4) sowie die Methode von Fisch und Strehlau (*F&S*) (Abs. 3.2.5) das nominale 95% Niveau in fast allen Situationen sehr gut einhalten. Weiterhin kann beobachtet werden, dass die Δ -Regel (Abs. 3.2.1), speziell für kleine Fallzahlen, zu konservativen Überdeckungswahrscheinlichkeiten führt (Abb. 3.2).

Als Zweites wird neben der Überdeckungswahrscheinlichkeit die Länge der Konfidenzintervalle untersucht. Da eine geschlossene Darstellung des Konfidenzintervalls nur für die Δ -Regel existiert, wird die empirische Länge mit Hilfe einer Simulationsstudie untersucht (Tab. 3.1). Um den Einfluss von numerischen, beziehungsweise simulationsbedingten Ausreißern in den Konfidenzintervalllängen zu reduzieren, wird anstelle der mittleren Intervalllänge die mediane Konfidenzintervalllänge betrachtet. Aus Abbildung 3.3 ist ersichtlich, dass die medianen Intervalllängen für alle vier Methoden sehr ähnlich sind und mit wachsender Fallzahl kleiner werden. Allerdings ist in allen Situationen die mediane Intervallbreite der Δ -Regel kleiner als die Breite der Intervalle der anderen Methoden, obwohl die Δ -Regel eher konservative Überdeckungswahrscheinlichkeiten aufweist (Abb. 3.2). Eine detaillierte Erläuterung folgt anhand eines ausgewählten Szenarios.

Um das genannte Verhalten der medianen Intervalllänge der Konfidenzintervalle zu erklären, wird das Szenario mit $\theta_1 = 1,25$, $\sigma^2 = 1$ und $n = 20$ pro Gruppe gewählt (Tab. 3.1). Dazu werden die empirischen Längen in untere und obere Konfidenzintervallhälfte geteilt. Die Ergebnisse der Simulation sind in Abbildung 3.4 zu finden. Es ist ersichtlich, dass die mediane Länge der unteren Konfidenzintervallhälfte der Δ -Regel größer ist als die medianen Längen der anderen drei Methoden. Ein umgekehrtes Verhalten kann für die obere mediane Intervallhälfte beobachtet werden. Ein Grund dafür ist, dass die Δ -Regel als einzige Methode symmetrische Konfidenzintervalle per Definition erzeugt. Dies führt in Relation zu den anderen Verfahren zu einer leicht längeren

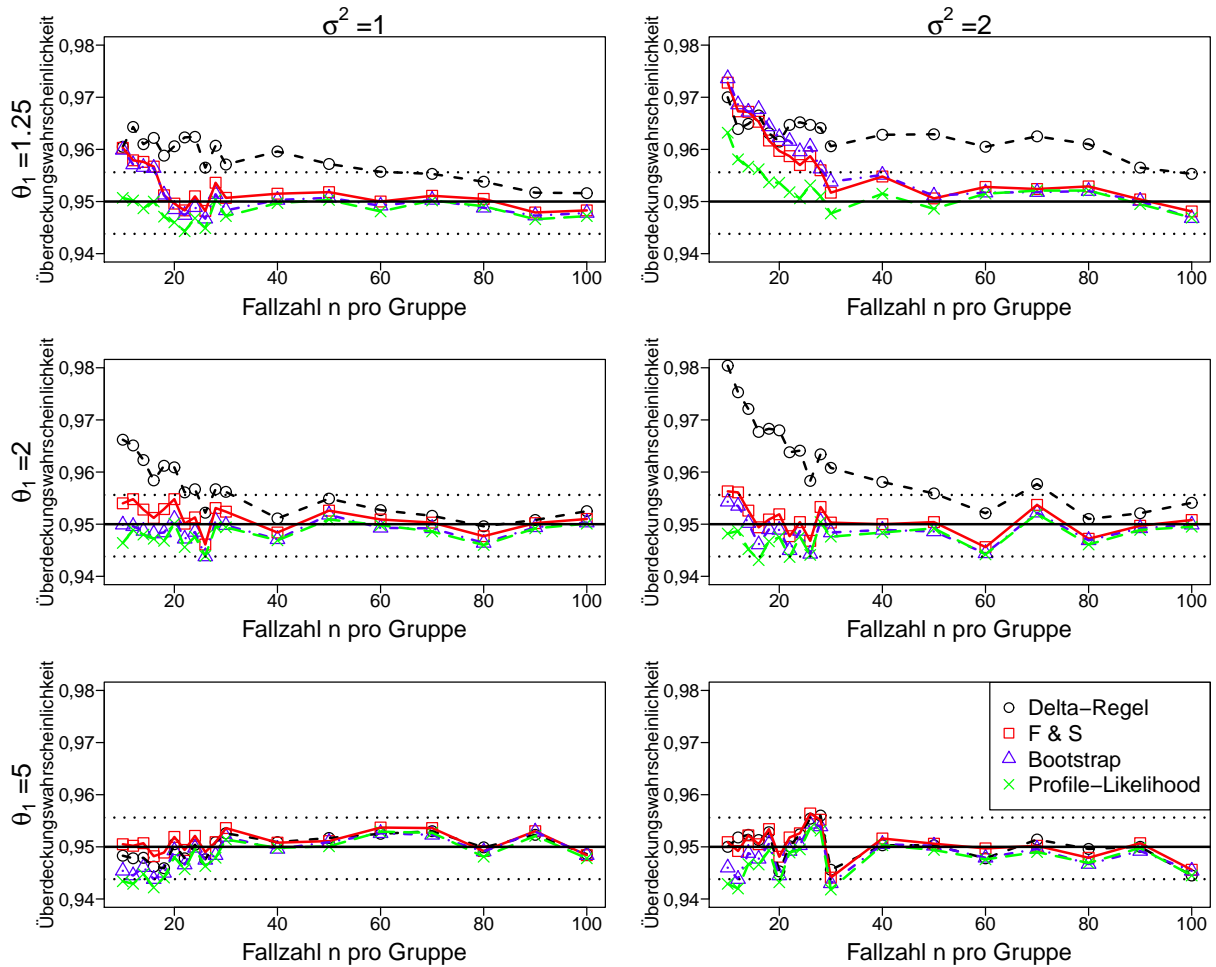


Abbildung 3.2.: Überdeckungswahrscheinlichkeit der Δ -Regel, der Methode von Fisch und Strehlau (*F&S*), des parametrischen Bootstraps und der Profile-Likelihood Methode für $k = 5$ Dosisstufen aufgetragen gegen die Fallzahlen n pro Gruppe für verschiedene Varianzen σ^2 und Steigungen θ_1 der Dosis-Wirkungsgeraden. Die gepunkteten Linien bezeichnen das 99%-Intervall des Simulationsfehlers.

unteren Intervallhälfte und einer kürzen oberen Intervallhälfte.

Zusätzlich zu den präsentierten Simulationsergebnissen wurden Dosis-Findungsstudien mit $k = 4$ und $k = 3$ äquidistanten Dosisstufen auf dem Dosisbereich $[0, 1]$ für die Simulationsszenarien aus Tabelle 3.1 untersucht (siehe ergänzendes Material Helms et al., 2014a). Exemplarisch sind die Ergebnisse der Überdeckungswahrscheinlichkeit in Abbildung 3.5 dargestellt. Es ist ersichtlich, dass die Überdeckungswahrscheinlichkeiten der vorgestellten Methoden für $k = 4$ und $k = 3$ nahezu identisch zu den Überdeckungswahrscheinlichkeiten für $k = 5$ Dosisstufen (Abb. 3.2) sind. Die vorgestellten Verfahren halten in den untersuchten Szenarien die nominale Überdeckungswahrscheinlichkeit auch für geringere Anzahlen an Dosisstufen gut ein. Für die Verzerrung des Zieldosissschätzers und die mediane Konfidenzintervalllänge in den Simulationsszenarien mit $k = 4$ und $k = 3$ Dosisstufen ist ein ähnliches Verhalten zu beobachten, weshalb an dieser Stelle für weitere Ergebnisse auf Helms et al. (2014a) verwiesen wird.

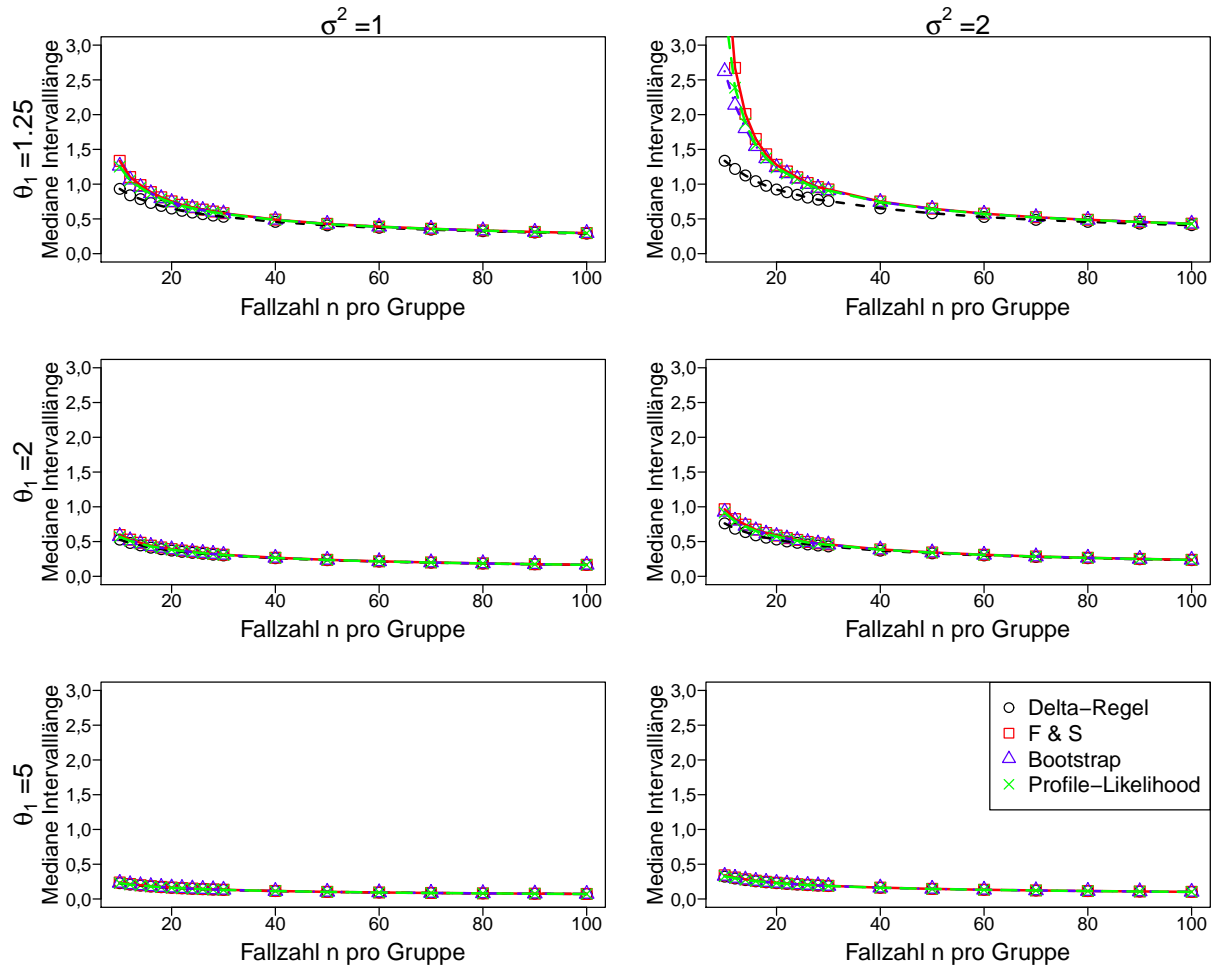


Abbildung 3.3.: Mediane Intervallbreite der Δ -Regel, der Methode von Fisch und Strehlau ($F\&S$), des parametrischen Bootstraps und der Profile-Likelihood Methode für $k = 5$ Dosisstufen aufgetragen gegen die Fallzahlen n pro Gruppe für verschiedene Varianzen σ^2 und Steigungen θ_1 der Dosis-Wirkungsgeraden.

Robustheit bei nicht-normalverteilten Fehlertermen

In diesem Abschnitt werden die verschiedenen Methoden zur Konfidenzintervallberechnung auf ihre Robustheit bezüglich der Annahme des normalverteilten Fehlerterms untersucht. Dafür wird eine Auswahl der Szenarien aus Tabelle 3.1 mit verschiedenen nicht-normalverteilten Fehlertermen simuliert, wobei die in Abschnitt 3.1.3 und 3.2 beschriebenen Verfahren unter der Annahme von normalverteilten Fehlertermen auf die erzeugten Daten angewendet werden. Als nicht-normale Verteilungen werden eine t -Verteilung mit 4 Freiheitsgraden sowie eine Log-Normalverteilung verwendet. Alle Fehlerterme werden so skaliert, dass der Erwartungswert gleich null und die Varianz gleich σ^2 aus Tabelle 3.1 ist. Die Ergebnisse dieser Simulationsstudie sind in Abbildung 3.6 dargestellt. Es zeigt sich, dass die untersuchten Methoden auch mit t - oder Log-Normalverteilten Fehlertermen eine gute Überdeckungswahrscheinlichkeit aufweisen. Allerdings ist eine leichte Verringerung der medianen Konfidenzintervalllänge für alle Methoden, mit

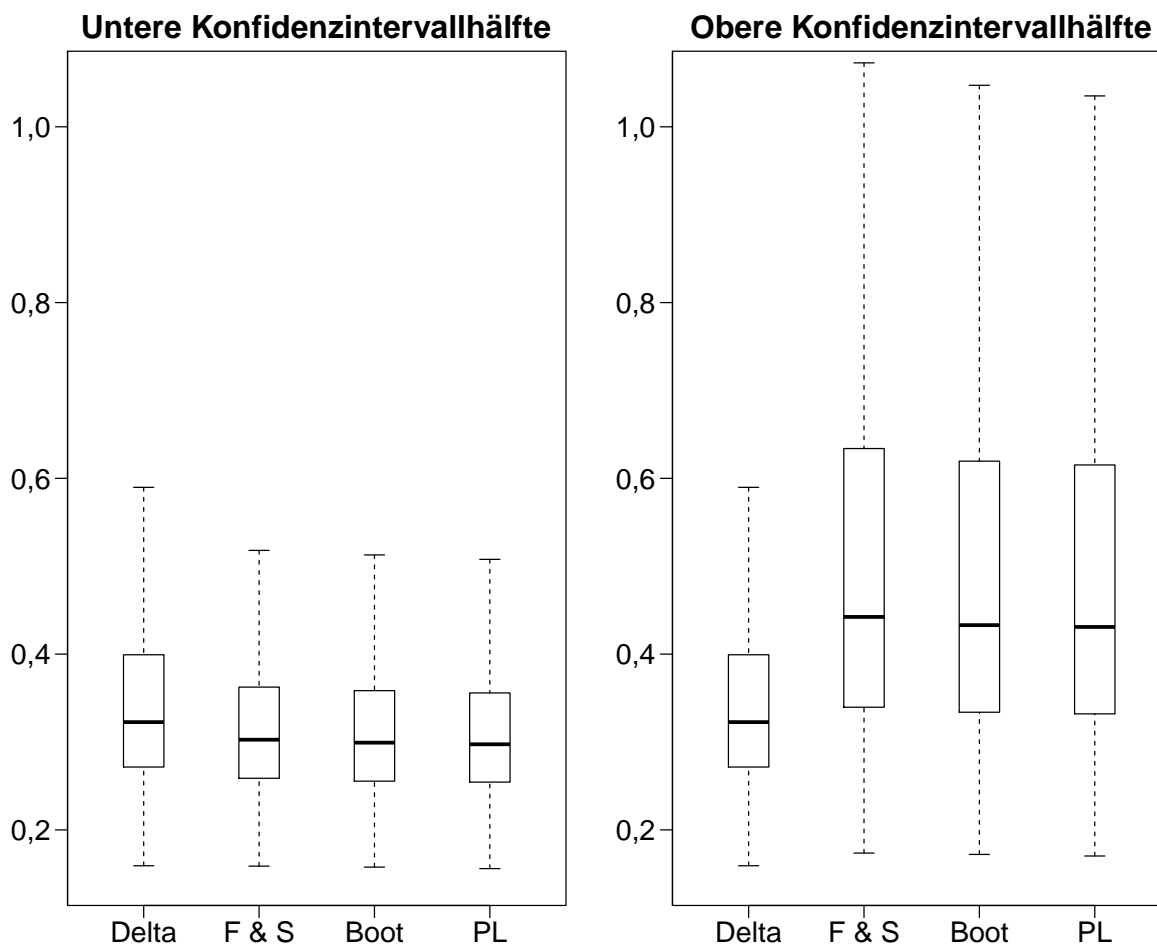


Abbildung 3.4.: Empirische Länge der unteren und oberen Intervallhälfte der Δ -Regel (*Delta*), der Methode von Fisch und Strehlau (*F&S*), des parametrischen Bootstraps (*Boot*) und der Profile-Likelihood Methode (*PL*) für $k = 5$ Dosisstufen, einer Varianz $\sigma^2 = 1$, einer Steigung $\theta_1 = 1.25$ der Dosis-Wirkungsgeraden und Fallzahl $n = 20$ pro Gruppe.

Ausnahme der Δ -Regel, zu beobachten. Basierend auf den präsentierten Ergebnissen erscheinen die Methoden robust gegenüber Abweichungen der Annahme eines normalverteilten Fehlerterms. Dies kann durch die asymptotischen Eigenschaften der Punktschätzer (siehe Abs. 2.3) für $N \rightarrow \infty$, in Kombination mit der Tatsache, dass die minimal untersuchte Gesamtfallzahl $N = 70$ ist, begründet werden.

3.4. Auswertung: Hyperaktive Blasenprobleme

Im Folgenden werden die präsentierten Verfahren verwendet, um das Beispiel aus Abschnitt 1.2.1 auszuwerten. Die eingeschlossenen Patientenzahlen für Placebo und die vier Dosisstufen Solifenacin (0; 2,5; 5; 10; 20 mg) sind $(n_1, \dots, n_5)' = (36, 40, 37, 33, 34)'$ sowie für die aktive Kontrolle (2 mg Tolterodine) $n_{ak} = 37$. Der primäre Endpunkt ist die Reduktion der mittleren

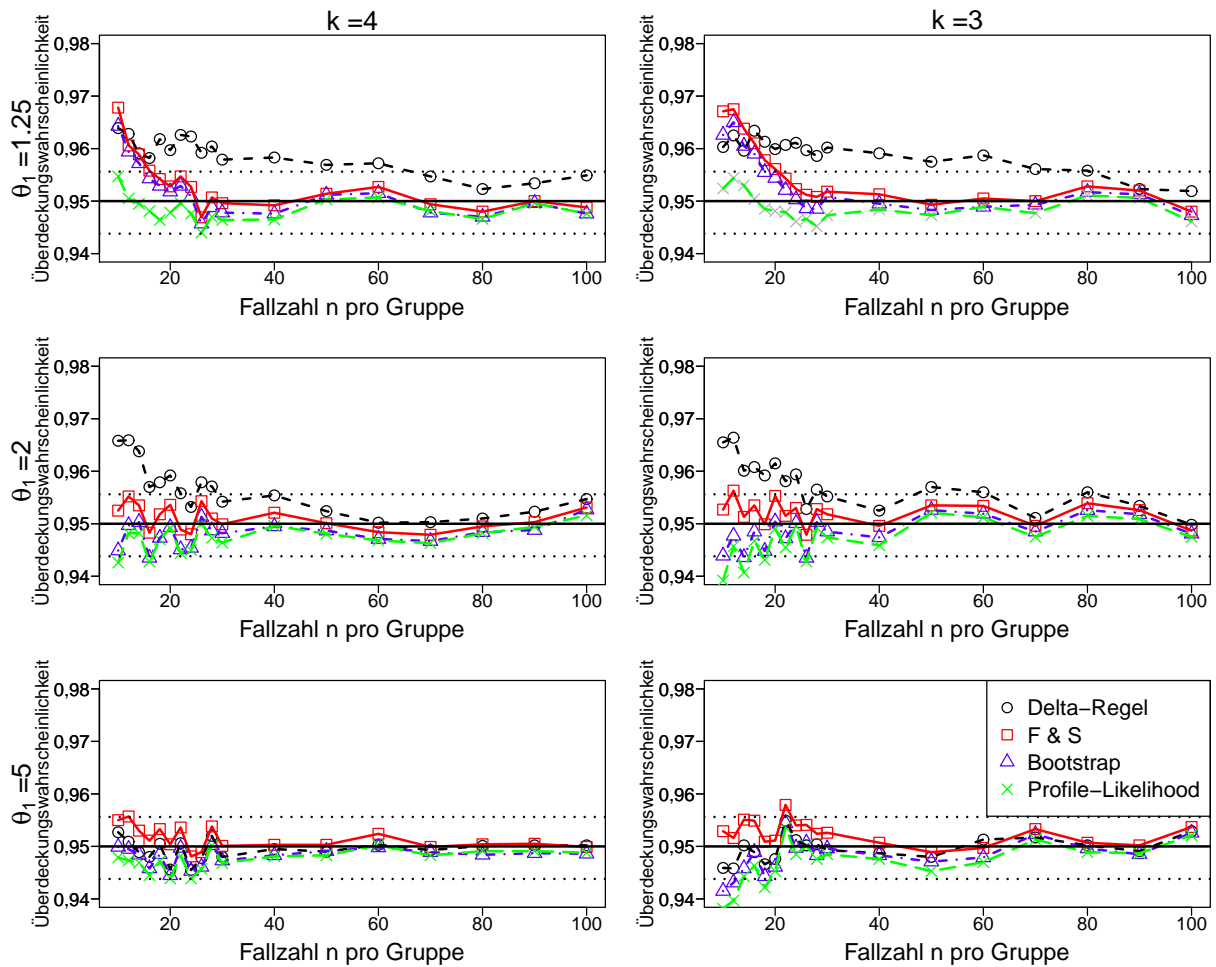


Abbildung 3.5.: Überdeckungswahrscheinlichkeit der Δ -Regel, der Methode von Fisch und Strehlau ($F\&S$), des parametrischen Bootstraps und der Profile-Likelihood Methode für $k = 3$ und $k = 4$ Dosisstufen aufgetragen gegen die Fallzahlen n pro Gruppe für die Varianz $\sigma^2 = 1$ sowie für verschiedene Steigungen θ_1 der Dosis-Wirkungsgeraden. Die gepunkteten Linien bezeichnen das 99%-Intervall des Simulationsfehlers.

Anzahl an Blasenentleerungen in 24 Stunden ($Voids/24h$) nach sechs Wochen von Baseline. Die Ergebnisse sind in Abbildung 1.2 auf der $\log(1 + d)$ -Dosis dargestellt, auf welcher sie auch ausgewertet werden. Selbst wenn nur die Mittelwerte pro Dosisstufe mit entsprechenden Fallzahlen und die zugehörige Varianz des Fehlerterms $\sigma^2 = 2$ zur Verfügung stehen, kann die Zieldosis basierend auf den Mittelwerten geschätzt werden. Weiterhin benötigen die Δ -Regel, die Methode von Fisch und Strehlau sowie der parametrische Bootstrap lediglich die Parameterschätzer, die Fallzahlen und σ^2 , um ein Konfidenzintervall für die Zieldosis zu bestimmen. Die Ergebnisse der verschiedenen Konfidenzintervallmethoden sind in Tabelle 3.2 mit einem Zieldosissschätzer von $\hat{d}^* = 3,011 \text{ mg}$ dargestellt. Für die Methode von Fisch und Strehlau, den parametrischen Bootstrap und den Profile-Likelihood Ansatz wird die untere Intervallgrenze auf null gesetzt, um positive Dosisbereiche zu garantieren. Da es nicht möglich ist, das Profile-Likelihood-basierte

Intervall allein auf Basis der Mittelwerte zu bestimmen, wird dafür ein Datensatz aus einzelnen Patientendaten mit entsprechenden Fallzahlen erzeugt, der die gleichen Mittelwerte und eine entsprechende Varianz σ^2 aufweist. Es ist ersichtlich, dass die verschiedenen Methoden zu ähnlichen Konfidenzintervallen führen. Mit einer totalen Fallzahl von $N = 217$ konnte nur die maximale Dosis als mögliche Zieldosis ausgeschlossen werden. Zur Berechnung der Zieldosis und zur Anwendung der verschiedenen Methoden wurden zwei SAS-Makros entwickelt, wobei das Makro „*DF_AC_LIN_MEAN*“ im Falle von Mittelwerten und das Makro „*DF_AC_LIN*“ im Falle von vorliegenden Rohdaten verwendet werden kann. Beide Makros wurden im ergänzenden Material von Helms et al. (2014a) veröffentlicht.

Tabelle 3.2.: Die 95%-Konfidenzintervalle (95%-KI) der Δ -Regel, der Methode nach Fisch und Strehlau (*F&S*), des parametrischen Bootstraps (*Bootstrap*) und dem Profile-Likelihood Ansatz (*PL*) für den Zieldosissschätzer $\hat{d}^* = 3,011 \text{ mg}$, basierend auf den in Abbildung 1.2 präsentierten Daten.

KI-Methode	95%-KI	KI-Länge
Δ -Regel	[0,236; 12,001]	11,765
<i>F&S</i>	[0,000; 13,237]	13,237
Bootstrap	[0,000; 12,671]	12,671
PL	[0,000; 13,018]	13,018

3.5. Zusammenfassung

In diesem Kapitel wurden der Schätzer der Zieldosis sowie verschiedene Methoden zur Berechnung des Konfidenzintervalls unter der Annahme einer linearen Wirkungsfunktion präsentiert. Die Parameterschätzung erfolgte auf Grundlage der Theorie linearer Modelle, die unter den Annahmen aus Abschnitt 2.2 und Gl. (3.1) identisch mit der Berechnung der ML-Schätzer aus Abschnitt 2.3 sind. Es konnte gezeigt werden, dass der Zieldosissschätzer für endliche Fallzahlen leicht verzerrt ist und dass die Verteilung des Schätzers selbst im Falle einer linearen Wirkungsfunktion nicht explizit bestimmt werden kann. Die Verzerrung kann nur approximiert und empirisch bestimmt werden, wobei ersichtlich wird, dass neben der Fallzahl N auch die Steigung der Wirkungsfunktion θ_1 sowie die Variation des Fehlerterms σ^2 einen Einfluss auf den Grad der Verzerrung haben. Falls $d^* = (\sum_{i=1}^k w_i d_i) / (1 - w_{ak})$ als Zieldosis gilt, ist die approximative Verzerrung gleich null und somit unabhängig von der Fallzahl, der Steigung und der Varianz des Fehlerterms. Speziell geht die empirische und die approximative Verzerrung für wachsende Fallzahlen N gegen null. Dies zeigt empirisch die Konsistenz von \hat{d}^* .

In den Simulationsstudien konnte gezeigt werden, dass das Konfidenzintervall basierend auf Normalapproximation mittels der Δ -Regel zwar direkt konstruiert werden kann, speziell für kleinere Fallzahlen aber zu einer relativ konservativen Überdeckungswahrscheinlichkeit führt. Die Metho-

de von Fisch und Strehlau, der Profile-Likelihood Ansatz und der parametrische Bootstrap liegen in fast allen Simulationsszenarien innerhalb des Zufallsstreifens der nominalen Überdeckungswahrscheinlichkeit von 95%. Allerdings benötigen die Methode von Fisch und Strehlau und der PL-Ansatz numerische Optimierung sowie der parametrische Bootstrap n_{boot} zusätzliche Simulationen zur Berechnung der Konfidenzintervallgrenzen. Der PL-Ansatz und der parametrische Bootstrap unterscheiden sich bezüglich der Überdeckungswahrscheinlichkeit in den untersuchten Szenarien kaum von den Ergebnissen der Methode von Fisch und Strehlau, tendieren allerdings zu einer leicht liberaleren Überdeckungswahrscheinlichkeit. Die mediane Konfidenzintervalllänge ist für alle vier Methoden ähnlich, wobei die Δ -Regel in allen betrachteten Situationen eine leicht geringere mediane Intervalllänge als die anderen Methoden aufweist. Ein Grund für die geringere Intervalllänge bei gleichzeitiger konservativer Überdeckungswahrscheinlichkeit ist, dass die Δ -Regel als einzige Methode per Definition symmetrische KI erzeugt. Daher ist die Bewertung der Methoden basierend auf der Intervalllänge allein nicht zu empfehlen. Zuerst muss sichergestellt werden, dass die Methoden die Überdeckungswahrscheinlichkeit gut einhalten und danach wird von den verbleibenden Verfahren die Methode mit dem kürzesten Intervall gewählt. „Short intervals are desirable when they cover the true parameter value but not necessarily otherwise“ (Lehmann, 1959). Weiterhin konnte gezeigt werden, dass die hier dargestellten Ergebnisse bezüglich Verzerrung des Zieltosschätzers, Überdeckungswahrscheinlichkeit und medianer Konfidenzintervalllänge auch für Versuchsdesigns mit drei oder vier Dosisstufen und aktiver Kontrolle gelten (siehe Helms et al., 2014a). Diese Ergebnisse wurden exemplarisch für die Überdeckungswahrscheinlichkeit präsentiert. Zusätzlich wurde der Einfluss von nicht-normalverteilten Fehlertermen auf die Güte der Intervallberechnung mit Hilfe einer Simulationsstudie untersucht. In den betrachteten Situationen, in denen t - beziehungsweise Log-normalverteilte Fehler vorlagen, konnte kein Einfluss auf die Überdeckungswahrscheinlichkeit und die mediane Intervalllänge festgestellt werden.

Sollte die zugrundeliegende Wirkungsfunktion von einem linearen Modell abweichen und sich auch nicht entsprechend transformieren lassen, ist es notwendig, entsprechende nichtlineare Modelle anzupassen. Mit dem Fall einer unbekanntem nichtlinearen Wirkungsfunktion beschäftigt sich das folgende Kapitel.

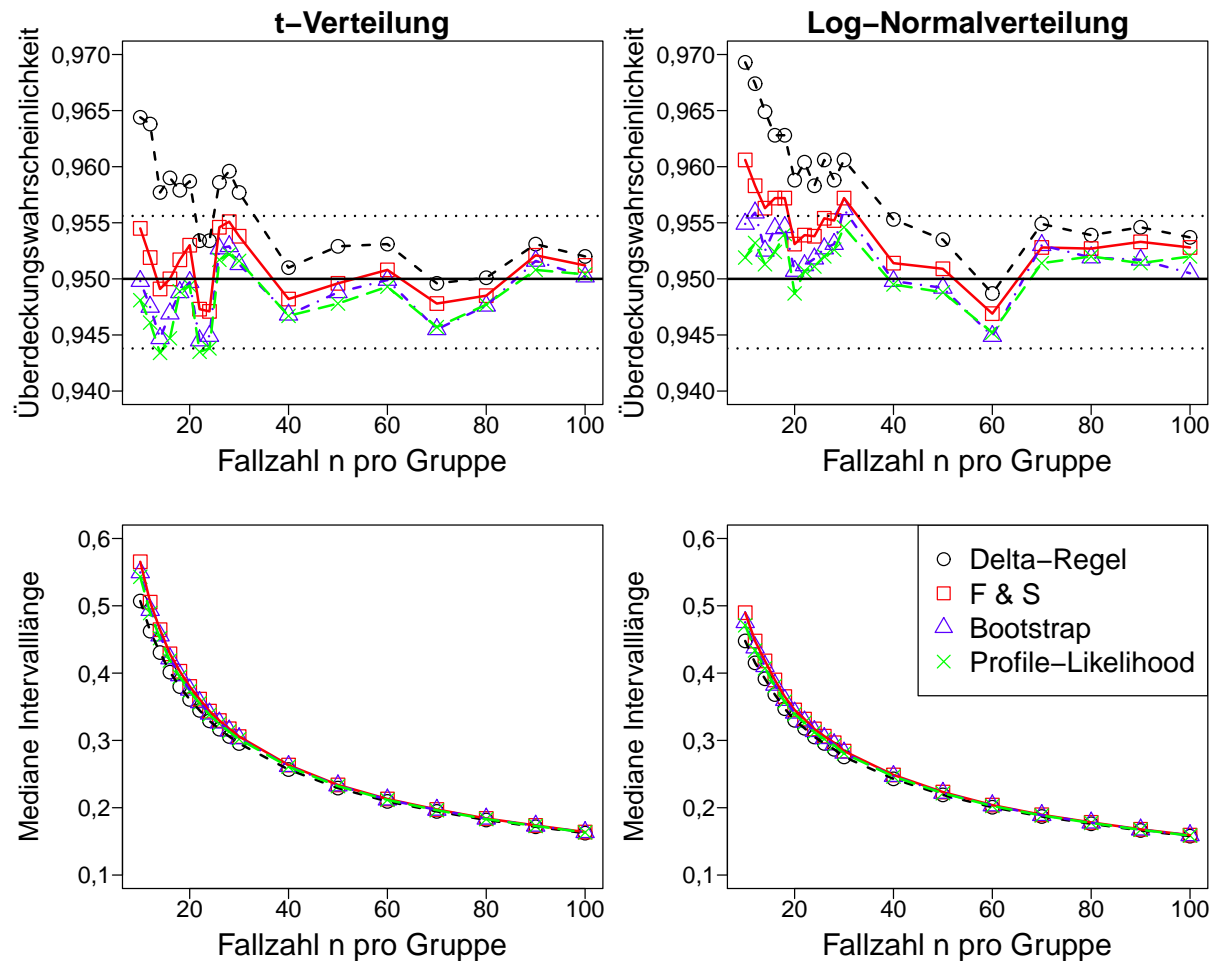


Abbildung 3.6.: Empirische Überdeckungswahrscheinlichkeit und mediane Konfidenzintervalllänge der Δ -Regel, der Methode von Fisch und Strehlau (*F&S*), des parametrischen Bootstraps und der Profile-Likelihood Methode aufgetragen gegen die Fallzahlen n pro Gruppe für *t*-verteilte und Log-Normalverteilte Fehlerterme für $k = 5$ Dosisstufen, mit Varianz $\sigma^2 = 1$ und der Steigung $\theta_1 = 2$ der Dosis-Wirkungsgeraden. Die gepunkteten Linien bezeichnen wieder das 99%-Intervall des Simulationsfehlers.

4. Spline-basierte Auswertung von nichtlinearen Dosis-Findungsstudien mit aktiver Kontrolle

In diesem Kapitel werden Ansätze zur Spline-basierten Auswertung von nichtlinearen Dosis-Findungsstudien mit aktiver Kontrolle entwickelt und ihre Eigenschaften untersucht. Da die zugrundeliegende Klasse der Dosis-Wirkungsfunktion $f_{\boldsymbol{\theta}}(d)$ (z. B. linear) im Allgemeinen unbekannt ist und somit eine Unsicherheit über die Wahl einer geeigneten Wirkungsfunktion besteht, werden im Folgenden flexible Spline-Verfahren zur Modellierung der Dosis-Wirkungsbeziehung vorgestellt und deren Eigenschaften untersucht. Die Verwendung dieses Ansatzes eröffnet die Möglichkeit, die Zieldosis d^* ohne Annahmen an die Klasse der Dosis-Wirkungsfunktion zu schätzen. Zusätzlich wird unter diesen Bedingungen ein Konfidenzintervall für die Zieldosis entwickelt.

4.1. Statistisches Modell

Es wird das statistische Modell aus Abschnitt 2.2 ohne die Annahme gleicher Varianzen der Fehlerterme verwendet. Das Modell lässt sich wie folgt darstellen:

$$\begin{aligned} Y_{ij} &= f_{\boldsymbol{\theta}}(d_i) + \epsilon_{ij} \\ Y_{ak,j'} &= \mu + \epsilon_{ak,j'}, \end{aligned}$$

wobei $f_{\boldsymbol{\theta}}(d)$ die erwartete Dosis-Wirkungskurve und μ die erwartete Wirkung der aktiven Kontrolle beschreibt. Die Fehlerterme seien unabhängig und identisch normalverteilt $\epsilon_{ij} \sim \mathcal{N}(0, \sigma_i^2)$, $i = 1, \dots, k$, $j = 1, \dots, n_i$ und $\epsilon_{ak,j'} \sim \mathcal{N}(0, \sigma_{ak}^2)$, $j' = 1, \dots, n_{ak}$, mit $\sigma_i > 0$ und $\sigma_{ak} > 0$.

4.2. Schätzung der Zieldosis durch Spline-Interpolation

Im ersten Schritt wird ein Schätzer für die Zieldosis d^* entwickelt, der ohne weitere Annahmen an die Form der Dosis-Wirkungsfunktion die Gleichung $d^* = \min_{d \in D} \{d \mid f_{\boldsymbol{\theta}}(d) = \mu\}$ löst. Der Schätzer der Wirkung der aktiven Kontrolle $\hat{\mu}$ kann mit Hilfe von Gl. (2.3), losgelöst von $f_{\boldsymbol{\theta}}(d)$, durch die Maximum-Likelihood Methode bestimmt werden. Für die Schätzung der Wirkungsfunktion wird eine stetige und nach Möglichkeit glatte Interpolations- oder Glättungsfunktion benötigt, wobei lediglich an den k Dosisstufen $0 \leq d_1 < d_2 < \dots < d_k$ Informationen vorliegen. Zur

Konstruktion der Interpolation werden die Mittelwerte der Zufallsvariablen bezüglich der k Dosisstufen $\bar{Y}_i = 1/n_i \sum_{j=1}^{n_i} Y_{ij} = \widehat{f_{\theta}}(d_i)$, $i = 1, \dots, k$, als Schätzer der Wirkung verwendet. Im Folgenden werden ein kubisches und ein lineares Spline-Interpolationsverfahren vorgestellt. Die entsprechenden Spline-Funktionen werden durch $s_q(\cdot)$, $q = 1, 3$, beschrieben, wobei q den Grad des Polynomzuges angibt. Zunächst werden die notwendigen Notationen und Definitionen der Spline-Theorie im Kontext der Dosis-Findung eingeführt.

Definition 4.1 (Knotenvektor) Ein Vektor $\mathbf{d}_v = (d_{v1}, \dots, d_{vm})'$, für dessen Elemente $d_{v1} \leq d_{v2} \leq \dots \leq d_{vm}$, $d_{vi} \in \mathbb{R}$ gilt, wird als Knotenvektor bezeichnet und die Elemente d_{vi} entsprechend als Knoten.

Mit Hilfe der Definition 4.1 wird der Begriff der Basisfunktion eingeführt.

Definition 4.2 (B-Spline Basisfunktion) Sei $\mathbf{d}_v = (d_{v1}, \dots, d_{vm})'$ ein Knotenvektor (siehe Def. 4.1), dann ist die i -te Basisfunktion q -ten Grades $B_{q,i}(d)$ rekursiv gegeben durch

$$B_{0,i}(d) = \begin{cases} 1, & d \in [d_{v,i}, d_{v,i+1}); \text{ für } q = 0 \text{ und } 0 \leq i \leq m - q - 1 \\ 0, & \text{sonst} \end{cases} \quad (4.1)$$

und für $q \geq 1$ durch

$$B_{q,i}(d) = \frac{d - d_{v,i}}{d_{v,i+q} - d_{v,i}} B_{q-1,i}(d) + \frac{d_{v,i+q+1} - d}{d_{v,i+q+1} - d_{v,i+1}} B_{q-1,i+1}(d). \quad (4.2)$$

Des Weiteren werden die Ableitungen der Basisfunktionen benötigt. Diese sind durch das folgende Korollar gegeben.

Lemma 4.1 Für $i = 1, \dots, m$, und $q \geq 1$ ist die Ableitung der Basisfunktion $B_{q,i}(d)$ durch

$$B_{q,i}(d)' = \frac{q}{d_{v,i+q} - d_{v,i}} B_{q-1,i}(d) + \frac{q}{d_{v,i+q+1} - d_{v,i+1}} B_{q-1,i+1}(d) \quad (4.3)$$

gegeben.

BEWEIS. Der Beweis folgt durch Induktion nach q (Smajic, 2009; Piegl und Tiller, 1997). \square

Die eingeführten Definitionen bilden die Grundlage für die Konstruktion der linearen und kubischen Spline-Funktionen, die im nächsten Abschnitt diskutiert werden.

4.2.1. Kubische Spline-Interpolation

Um die benötigten Basispolynome $B_{3,i}(d)$ des kubischen Splines zu berechnen, wird ein Knotenvektor \mathbf{d}_v^c mit $k+6$ Knoten benötigt, der durch mehrfache Verwendung der k Dosisstufen erzeugt wird (Floater, 2007, S. 110):

$$\mathbf{d}_v^c = (d_{v,1}, \dots, d_{v,k+6}) = (\overbrace{d_1, \dots, d_1}^q, d_1, d_2, \dots, d_{k-1}, d_k, \overbrace{d_k, \dots, d_k}^q)', \quad (4.4)$$

wobei $q = 3$ den Grad der verwendeten Spline-Funktion angibt. Die Art der Darstellung wird häufig verwendet, hat allerdings einige numerische Nachteile, da in den Berechnungen der Basisfunktionen der Fall $d_{v,i} - d_{v,j} = 0$, $i \neq j$ vorkommt. Das bedeutet, dass der Divisor in Definition 4.2 null sein kann. Dieses Problem wird umgangen, indem „jeder durch null geteilte Ausdruck auf null gesetzt wird“ (Floater, 2007, S. 28).

Analytisch entstehen somit keine Probleme, allerdings führt diese Regel in der computergestützten Implementierung der Berechnung zu einer unnötigen Komplexität und einem vermeidbaren numerischen Mehraufwand. Daher wird der Knotenvektor in der vorliegenden Arbeit durch die Verwendung der mittleren Differenz h zwischen den k Dosisstufen berechnet, mit $h = \frac{1}{k-1} \sum_{i=1}^{k-1} (d_{i+1} - d_i)$. Der Knotenvektor ist gegeben durch:

$$\mathbf{d}_v = (d_{-q-1}, \dots, d_0, d_1, \dots, d_k, d_{k+1}, \dots, d_{k+q})', \quad (4.5)$$

wobei $q \geq 1$ den Grad der verwendeten Spline-Funktion $s_q(d)$ angibt mit $d_i = d_1 + (i-1)h$, $i = -q-1, \dots, 0$, sowie $d_{k+j} = d_k + jh$, $j = 1, \dots, q$. Diese Darstellung des Knotenvektors sichert $d_i - d_j \neq 0$, $\forall i \neq j = -q-1, -q-2, \dots, q-1, q$, und führt auf dem Intervall $D = [d_1, d_k]$ zu äquivalenten Basisfunktionen vergleichbar zu dem Knotenvektor aus Gl. (4.4). Mit der gegebenen Notation ist die Definition einer kubischen Spline-Interpolation für Dosis-Findungsstudien im Folgenden gegeben.

Definition 4.3 (Kubische Spline-Funktion) Seien $(c_i)_{i=1}^{k+2} \in \mathbb{R}$ unbekannte Koeffizienten, $q = 3$ der Grad der Spline-Funktion, \mathbf{d}_v^c der in Gl. (4.5) definierte zugehörige Knotenvektor und $d_1 < \dots < d_k$ die untersuchten Dosisstufen mit zugehörigen Mittelwerten \bar{Y}_i , dann lässt sich die kubische Spline-Funktion mit natürlichen Randbedingungen mit Hilfe der kubischen Basisfunktion $B_{3,i}(d)$ (Def. 4.2) wie folgt definieren:

$$s_3(d) = \sum_{i=1}^{k+2} c_i B_{3,i}(d), \quad (4.6)$$

so dass die k Interpolationsbedingungen $s_3(d_i) = \bar{Y}_i$ für $i = 1, \dots, k$, sowie die zwei natürlichen Randbedingungen $s_3''(d_1) = s_3''(d_k) = 0$ erfüllt sind.

Der definierte kubische Spline mit natürlicher Randbedingung (natürlicher kubischer Spline) wird durch eine Summe von lokal definierten Basispolynomen (der B-Spline Basis) generiert (Floater, 2007, S. 116-117), verbindet jedes benachbarte Dosispaar d_i, d_{i+1} , $i = 1, \dots, k-1$, mit einem kubischen Polynom und hat stetige erste und zweite Ableitung auf $D = [d_1, d_k]$, was den

glatten Verlauf der Kurve garantiert. Um die $k + 2$ Unbekannten $(c_i)_{i=1}^{k+2}$ zu bestimmen, müssen die k Interpolationsbedingungen $s_3(d_i) = \bar{Y}_i$, sowie die zwei natürlichen Randbedingungen $s_3''(d_1) = s_3''(d_k) = 0$ erfüllt sein. Die Berechnung der notwendigen zweiten Ableitungen $s_3''(d_1)$ und $s_3''(d_k)$ kann unter Ausnutzung der Eigenschaften der Basisfunktionen vorgenommen werden. Für die zweite Ableitung der Spline-Funktion gilt:

$$s_3''(d) = \sum_{i=1}^{k+2} c_i B_{3,i}''(d),$$

wobei die r -te Ableitung der Basisfunktion q -ten Grades aus Gl. (4.3) als

$$\frac{\partial^r B_{q,i}}{\partial d^r} = B_{q,i}^{(r)}(d) = q \cdot \left(\frac{B_{q-1,i}^{(r-1)}(d)}{d_{v,i+q} - d_{v,i}} - \frac{B_{q-1,i+1}^{(r-1)}(d)}{d_{v,i+q+1} - d_{v,i+1}} \right) \text{ für } r \geq 1$$

berechnet werden kann. Die Randbedingungen werden als 1-te und $k + 2$ -te Bedingung in das resultierende lineare Gleichungssystem eingefügt, so dass $\mathbf{A}\mathbf{c} = \mathbf{F}$ mit $\mathbf{F} = (0, \bar{Y}_1, \dots, \bar{Y}_k, 0)'$ und $\mathbf{c} = (c_1, \dots, c_{k+2})'$ und zugehöriger Designmatrix

$$\mathbf{A} = \begin{pmatrix} B_{3,1}''(d_1) & B_{3,2}''(d_1) & \dots & B_{3,k+2}''(d_1) \\ B_{3,1}(d_1) & B_{3,2}(d_1) & \dots & B_{3,k+2}(d_1) \\ \vdots & \vdots & \vdots & \vdots \\ B_{3,1}(d_k) & B_{3,2}(d_k) & \dots & B_{3,k+2}(d_k) \\ B_{3,1}''(d_k) & B_{3,2}''(d_k) & \dots & B_{3,k+2}''(d_k) \end{pmatrix} \text{ gilt.} \quad (4.7)$$

Die Matrix \mathbf{A} ist invertierbar, so dass das Gleichungssystem eindeutig lösbar ist (siehe Schoenberg-Whitney Theorem, de Boor, 2001, S. 171-175). Die Koeffizienten \mathbf{c} lassen sich daher wie folgt darstellen:

$$\mathbf{c} = \mathbf{A}^{-1}\mathbf{F}. \quad (4.8)$$

Werden diese Koeffizienten \mathbf{c} in die Spline-Funktion s_3 eingesetzt, kann die Spline-Interpolation der Wirkungsfunktion für beliebige Dosislevel $d \in D = [d_1, d_k]$ bestimmt werden. Der Schätzer der Zieldosis \hat{d}^* ergibt sich nach Definition 2.4 durch numerisches Lösen der Gleichung $\hat{d}^* = \min_{d \in D} \{s_3(d) = \hat{\mu}\}$.

Aufgrund der Tatsache, dass weder für die zugrundeliegende Dosis-Wirkungsfunktion f_{θ} noch für die Spline-Interpolation s_3 ein monotoner Verlauf gefordert wird, kann die Gleichung $s_3(d) = \hat{\mu}$ mehrere gültige Lösungen besitzen. Daher wird vor der numerischen Berechnung der Dosisbe-

reich eingeschränkt, so dass nur die gesuchte Lösung (die minimale Dosis) enthalten ist. Dafür wird in einem ersten Schritt auf dem Dosisbereich $D = [d_1, d_k]$ ein Grid $\mathbf{G} = (\tilde{d}_1, \dots, \tilde{d}_a)'$ mit a äquidistanten Elementen erzeugt ($\Delta h = \tilde{d}_{j'+1} - \tilde{d}_{j'}$, $\forall j' = 1, \dots, a - 1$) und der Spline an diesen Stellen ausgewertet. Anschließend muss die niedrigste Dosis $\tilde{d}_{i'}$ bestimmt werden, so dass $s_3(\tilde{d}_{i'}) \leq \hat{\mu} \leq s_3(\tilde{d}_{i'+1})$ (oder $s_3(\tilde{d}_{i'}) \geq \hat{\mu} \geq s_3(\tilde{d}_{i'+1})$) gilt. Somit liegt der gesuchte Zieldosis-schätzer, falls ein solches $i' \in \{1, \dots, a - 1\}$ existiert, im Dosisbereich zwischen dem i' -ten und dem $i' + 1$ -ten Grid-Element. Um diese Eigenschaft zu gewährleisten, muss das Grid fein genug gewählt werden, um einen monotonen Verlauf der Spline-Funktion auf den Teilbereichen zu garantieren (in den Simulationen wurde $\Delta h = 0,1$ gewählt). Zur Berechnung des Zieldosis-schätzers wird der in Abschnitt 3.2.3 vorgestellte Algorithmus zur Intervallhalbierung mit der Zielfunktion $O_{d^*}(d) = s_3(d) - \hat{\mu} \stackrel{!}{=} 0$ mit $d \in [\tilde{d}_{i'}, \tilde{d}_{i'+1}]$ unter Beachtung der Monotonie (fallend oder steigend) angewendet.

Alternativ zur verwendeten B-Spline Darstellung (Def. 4.2) der kubischen Spline-Funktion kann auf andere Darstellungen der Polynombasis zurückgegriffen werden. Zum Beispiel kann die Monom-Basis $\{1, x, x^2, x^3\}$ auf jedem Teilintervall $[d_i, d_{i+1}]$, $i = 1, \dots, k - 1$, verwendet werden, um die die Spline-Funktion pro Teilintervall wie folgt darzustellen:

$$s_{3,i}(d) = a_{0,i} + a_{1,i}(d - d_i) + a_{2,i}(d - d_i)^2 + a_{3,i}(d - d_i)^3 \text{ für } i = 1, \dots, k - 1.$$

Das Finden der unbekanntenen Koeffizienten $a_{0,i}, \dots, a_{3,i}$ lässt sich dann auf das Lösen eines Tri-diagonalsystems zurückführen (siehe z. B. Stoer, 1993, S. 85-91). Nachdem die Koeffizienten berechnet sind, können auf jedem Teilintervall $[d_i, d_{i+1}]$ die cardanischen Formeln (siehe z. B. Jacobson, 2009, S. 210) verwendet werden, um die Nullstellen bzw. Dosisstufen analytisch zu bestimmen, so dass die Gleichungen

$$s_{3,i}(d) = \bar{Y}_{ak}. \text{ für } i = 1, \dots, k - 1$$

erfüllt sind. Anschließend wird aus der Menge der gültigen Dosisstufen das Minimum als Zieldosis-schätzer gewählt. Diese alternative Methode zur Berechnung der Spline-Funktion und zur Bestimmung des Zieldosis-schätzers wurde in der vorliegenden Arbeit nicht weiter betrachtet.

4.2.2. Lineare Spline-Interpolation

Neben dem Ansatz der kubischen Spline-Interpolation wird eine lineare Spline-Interpolation zum Schätzen der Zieldosis zum Vergleich verwendet. Diese Methode wurde bereits von Dilleen et al. (2003) für Dosis-Findungsstudien mit aktiver Kontrolle vorgeschlagen. Die lineare Spline-Interpolation ist eine einfache Methode, um eine stetige Approximation der Wirkungsfunktion zu konstruieren. Dazu werden die Mittelwerte der Wirkungsgrade für benachbarte Dosisstufen durch eine Gerade miteinander verbunden. Der lineare Spline kann im Kontext der Dosis-Findung wie folgt definiert werden.

Definition 4.4 (Lineare Spline-Funktion) Sei $d_1 < \dots < d_k$ ein Set von Dosisstufen mit zugehörigen Mittelwerten des Wirkungsgrades $\bar{Y}_{1\cdot}, \dots, \bar{Y}_{k\cdot}$ und sei \mathbf{d}_v^l ein Knotenvektor nach Gl. 4.5 mit $q = 1$, dann ist die lineare Spline-Funktion definiert durch:

$$s_1(d) = \sum_{i=1}^k \bar{Y}_i \cdot B_{1,i}(d), \quad (4.9)$$

wobei $B_{1,i}(d)$ die lineare Basisfunktion aus Definition 4.2 bezeichnet, so dass die k Interpolationsbedingungen $s_1(d_i) = \bar{Y}_i$ für $i = 1, \dots, k$, erfüllt sind.

Alternativ lässt sich $s_1(d)$ für $d_i \leq d \leq d_{i+1}$, $i \in \{1, \dots, k-1\}$ direkt als Geradengleichung schreiben:

$$s_1(d) = \hat{\theta}_0^{lin} + \hat{\theta}_1^{lin} \cdot d, \quad (4.10)$$

wobei $\hat{\theta}_1^{lin} = (\bar{Y}_{i+1\cdot} - \bar{Y}_{i\cdot}) / (d_{i+1} - d_i)$ die Steigung und $\hat{\theta}_0^{lin} = \bar{Y}_i - \hat{\theta}_1^{lin} d_i$ der Achsenabschnitt ist. Um den Schätzer der Zieldosis mit Hilfe linearer Spline-Interpolation zu berechnen, muss lediglich die niedrigste Dosis $d_{i'}$ bestimmt werden, so dass $\bar{Y}_{i'\cdot} \leq \bar{Y}_{ak\cdot} \leq \bar{Y}_{i'+1\cdot}$ (oder $\bar{Y}_{i'\cdot} \leq \bar{Y}_{ak\cdot} \leq \bar{Y}_{i'+1\cdot}$) gilt. Dann liegt der zugehörige Zieldosissschätzer, falls ein solches $i' \in \{1, \dots, k-1\}$ existiert, zwischen der i' -ten und der $i'+1$ -ten Dosisstufe. Der Zieldosissschätzer kann auf dem Dosisbereich $[d_{i'}, d_{i'+1}]$ mittels der alternativen Schreibweise aus Gl. (4.10) und dem Parameterschätzer des linearen Modells Gl. (3.2) mit $\hat{\theta}_1^{lin} \neq 0$ direkt durch

$$\hat{d}^* = \frac{(\bar{Y}_{ak\cdot} - \hat{\theta}_0^{lin})}{\hat{\theta}_1^{lin}}$$

bestimmt werden. Zur Implementierung der vorgestellten Spline-basierten Schätzer der Zieldosis in SAS 9.3 siehe Helms et al. (2013). Im nachfolgenden Abschnitt werden die Eigenschaften der Zieldosissschätzer bezüglich Erwartungswert und Verzerrung untersucht.

4.3. Eigenschaften der Schätzer der Zieldosis

In diesem Abschnitt werden der Erwartungswert und die Verzerrung der Spline-basierten Zieldosissschätzer untersucht. Die im vorangegangenen Abschnitt hergeleiteten Zieldosissschätzer lassen sich durch numerisches Lösen der Gleichung $s_q(\hat{d}^*) = \bar{Y}_{ak\cdot}$ mit $q = 1$ für den linearen und $q = 3$ für den kubischen Spline bestimmen. Zunächst wird zur Berechnung des Erwartungswertes von \hat{d}^* der Erwartungswert der Spline-Interpolationen benötigt. Dieser ist im folgenden Korollar angegeben.

Satz 4.1 Sei $d_1 < \dots < d_k$ eine Menge von Dosisstufen mit zugehörigen Mittelwerten des Wirkungsgrades $\bar{Y}_{1\cdot}, \dots, \bar{Y}_{k\cdot}$ und es gelte $E(\bar{Y}_{i\cdot}) = f_{\theta}(d_i)$. Zusätzlich bezeichne $s_q(d)$ die Spline-Interpolation vom Polynomgrad $q \in \{1, 3\}$ aus Definition 4.3 und 4.4, dann ergibt sich der Er-

wartungswert der Spline-Funktion $E(s_q(d))$ durch die Interpolation der Dosis-Wirkungsfunktion $f_{\theta}(\cdot)$ an den untersuchten Dosisstufen $d_i, i = 1, \dots, k$.

BEWEIS. Für den Beweis werden die Fälle $q = 1$ und $q = 3$ separat betrachtet.

1.) Linearer Spline ($q = 1$):

Aus Definition 4.4 ergibt sich der Erwartungswert des linearen Splines wie folgt

$$E(s_1(d)) = \sum_{i=1}^k E(\bar{Y}_i) \cdot B_{1,i}(d) = \sum_{i=1}^k f_{\theta}(d_i) \cdot B_{1,i}(d).$$

2.) Kubischer Spline ($q = 3$):

Aus Definition 4.3 ergibt sich der Erwartungswert des kubischen Splines durch

$$E(s_3(d)) = \sum_{i=1}^{k+2} E(c_i) \cdot B_{3,i}(d),$$

wobei $E(c_i)$ das i -te Element des erwarteten Koeffizienten-Vektors $E(\mathbf{c})$ beschreibt. Aus Gl. (4.8) folgt dann

$$E(\mathbf{c}) = \mathbf{A}^{-1}E(\mathbf{F})$$

mit Designmatrix \mathbf{A} aus Gl. (4.7), Bedingungsvektor $\mathbf{F} = (0, \bar{Y}_1, \dots, \bar{Y}_k, 0)'$ und zugehörigem Erwartungsvektor $E(\mathbf{F}) = (0, f_{\theta}(d_1), \dots, f_{\theta}(d_k), 0)'$. Daraus folgt, dass sich die Erwartungswerte der Spline-Interpolationen durch Interpolation der erwarteten Wirkungsfunktion $f_{\theta}(\cdot)$ ergeben. \square

Der Erwartungswert des Zieldosis-Schätzers \hat{d}^* ergibt sich somit für den linearen ($q = 1$) und den kubischen Spline ($q = 3$) durch numerisches Lösen der Gleichung

$$\min_{d \in D} \{E(s_q(d)) = \mu\} \text{ (siehe Satz 4.1).}$$

Dabei kann zur numerischen Berechnung des Erwartungswertes auf die Methoden aus den Abschnitten 4.2.1 und 4.2.2 zurückgegriffen werden. Sei $d_{s_q}^* = E(\hat{d}^*), q \in \{1, 3\}$ die Lösung der vorangestellten Gleichung für die entsprechenden Spline-Interpolationen, dann ist die Verzerrung des Zieldosis-Schätzers für die jeweilige Spline-Interpolation durch

$$\text{Bias} = E(\hat{d}^* - d^*) = d_{s_q}^* - d^*, q \in \{1, 3\}$$

definiert.

Diese Verzerrung hängt von der Güte der Interpolation und somit von der Anzahl der untersuchten Dosisstufen k ab. Je mehr Dosisstufen untersucht werden, desto kleiner wird der Bias. Allerdings ist der Grad der Verzerrung unabhängig von der Fallzahl N , da sich der Erwartungswert der Spline-Funktionen durch Interpolation der erwarteten Wirkungskurve ergibt. Falls eine Verzerrung vorliegt, so geht diese, im Gegensatz zum parametrischen Ansatz, nicht asymptotisch für $N \rightarrow \infty$ gegen null. Die Verzerrung der präsentierten Interpolationen wird in Abschnitt 4.5 exemplarisch für eine Emax-Wirkungskurve berechnet und verglichen. Neben den vorgestellten Schätzern soll auch ein Konfidenzintervall für die Zieldosis konstruiert werden, welches ebenfalls ohne zusätzliche Annahmen an die erwartete Wirkungsfunktion auskommt. Dieses Konfidenzintervall wird im folgenden Abschnitt hergeleitet.

4.4. Spline-basiertes Bootstrap-Konfidenzintervall

Im vorangegangenen Abschnitt wurden Spline-basierte Schätzer zur Berechnung der Zieldosis vorgestellt. In diesem Abschnitt liegt der Fokus auf der Konstruktion eines Konfidenzintervalls für die Zieldosis. Wie zu Beginn des Kapitels beschrieben, sollen keine weiteren Annahmen an die Klasse der Dosis-Wirkungsfunktion gemacht werden. Im weiteren Verlauf ist daher die Verteilungsfunktion von \hat{d}^* unbekannt und es können auch keine asymptotischen Approximationsverfahren (z. B. Abs. 2.4.3) verwendet werden. Um das Problem zu lösen, wird ein ähnlicher Ansatz zu Abschnitt 3.2.4 mittels parametrischem Bootstrap und Monte-Carlo-Simulation (MC) gewählt. Weiterhin sind unter den gemachten Modellannahmen (Abs. 2.2) die Mittelwerte der Wirkung pro Dosislevel erwartungstreue Schätzer für die zugrundeliegende Wirkungsfunktion an der jeweiligen Dosisstufe ($E(\bar{Y}_{i\cdot}) = f_{\theta}(d_i), i = 1, \dots, k$). Die Verteilung der Mittelwerte ist dann bekannt und lässt sich wie folgt beschreiben:

$$\bar{Y}_{i\cdot} = f_{\theta}(d_i) + \bar{\epsilon}_{i\cdot} \sim \mathcal{N}\left(f_{\theta}(d_i), \frac{\sigma_i^2}{n_i}\right) \text{ für } i = 1, \dots, k,$$

und

$$\bar{Y}_{ak\cdot} = \bar{Y}_{k+1\cdot} = \mu + \bar{\epsilon}_{k+1\cdot} \sim \mathcal{N}\left(\mu, \frac{\sigma_{k+1}^2}{n_{k+1}}\right),$$

wobei die aktive Kontrolle (ak) als $k + 1$ -te Dosis der Studie mit $n_{k+1} = n_{ak}$ und $\sigma_{k+1} = \sigma_{ak}$ bezeichnet wird. Allerdings sind die Varianzen der Fehlerterme $\sigma_i^2, i = 1, \dots, k+1$ unbekannt und müssen ebenfalls ohne zusätzliche Annahmen geschätzt werden. Dies ist mit Hilfe der empirischen Varianzschätzer möglich:

$$\hat{\sigma}_i^2 = \frac{1}{(n_i - 1)} \cdot \left(\sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_{i\cdot})^2 \right), i = 1, \dots, k + 1.$$

Wird im statistischen Modell die Annahme gleicher Varianzen der Fehlerterme $\sigma_i = \sigma$ gemacht, so kann alternativ der empirische Varianzschätzer

$$\hat{\sigma}^2 = \frac{1}{(N - (k + 1))} \cdot \left(\sum_{i=1}^{k+1} \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_{i\cdot})^2 \right)$$

verwendet werden. Unter diesen Bedingungen kann ein parametrischer Bootstrap verwendet werden, um das gewünschte Konfidenzintervall zu bestimmen. Dafür werden n_{boot} neue Sets von Wirkungsmittelwerten $\bar{Y}_{i\cdot}^b$, $i = 1, \dots, k + 1$, generiert, indem die Residuen der Mittelwerte unter Verwendung der eben definierten Varianzschätzer $\hat{\sigma}_i^2$ gebootstrapt werden (z. B. Efron, 1979, 1987):

$$\bar{Y}_{i\cdot}^b = \bar{Y}_{i\cdot} + Z_i^b, \quad Z_i^b \sim \mathcal{N}\left(0, \frac{\hat{\sigma}_i^2}{n_i}\right), \quad i = 1, \dots, k + 1, b = 1, \dots, n_{boot}. \quad (4.11)$$

Hierbei stellen Z_i^b die gebootstrappten Residuen und $\hat{\sigma}_i^2/n_i$ die zugehörige residuale Varianz dar. In einem nächsten Schritt wird dann eine der in Abschnitt 4.2.1 vorgestellten Spline-Interpolationen verwendet, um die Zieldosischätzer \hat{d}^{*b} , $b = 1, \dots, n_{boot}$, basierend auf den gebootstrappten Mittelwerten zu berechnen. Alle Schätzer der Zieldosis werden in einem Vektor $\widehat{\mathbf{d}}^{*b}$ der Größe nach aufsteigend sortiert und gespeichert. Dann kann das $\alpha/2$ - und das $(1 - \alpha/2)$ -Quantil des Vektors $\widehat{\mathbf{d}}^{*b}$ mit Hilfe der Quantil-Methode abgelesen werden, um das gewünschte $(1 - \alpha)$ -Konfidenzintervall zu bestimmen (vgl. Abschnitt 3.2.4). Dann lässt sich das Konfidenzintervall wie folgt schreiben:

$$KI_{boot} = \left[\widehat{\mathbf{d}}^{*b} \left[n_{boot} \cdot \frac{\alpha}{2} \right], \widehat{\mathbf{d}}^{*b} \left[n_{boot} \cdot \left(1 - \frac{\alpha}{2}\right) \right] \right].$$

Diese Art der Konfidenzintervallberechnung ist unabhängig von der gewählten Berechnungsmethode des Zieldosischätzers und kann somit für verschiedenste Methoden verwendet werden (siehe z. B. Pinheiro et al., 2014). Speziell kann für jede Art von Spline-basierter Schätzung der Zieldosis ein Konfidenzintervall bestimmt werden (Helms et al., 2014b). Die vorgestellten Methoden zur Zieldosischätzung und zur Konfidenzintervall-Berechnung werden im folgenden Abschnitt mit Hilfe von Simulationsstudien untersucht.

4.5. Simulationen der Spline-Interpolationen für verschiedene Dosis-Wirkungskurven

In diesem Abschnitt werden die Eigenschaften der präsentierten Spline-Interpolationen mit zugehörigen Konfidenzintervallen (KI) mit Hilfe umfangreicher Simulationsstudien untersucht und mit parametrischen Verfahren verglichen. Die klinische Studie von Nauck et al. (2009) (Abs. 1.2.2) motiviert das Design der Dosis-Findungsstudie in den Simulationsszenarien. Daher wird im

Folgendes eine fünfarmige Dosis-Findungsstudie ($k = 4$) mit aktiver Kontrolle (ak) und Placebo ($d_1 = 0$) als minimaler Dosis sowie der Dosisallokation $(d_1, \dots, d_4)' = (0; 0,6; 1,2; 1,8)'$ untersucht. Der betrachtete Endpunkt ist die Reduktion des FPG-Levels [mmol/l] am Ende der Studie im Vergleich zu Studienbeginn. Weiterhin werden die Standardabweichungen auf $\sigma_i = \sigma_{ak} = \sigma = 1,8$ mmol/l und die Fallzahlen pro Dosisstufe auf $n_i = n_{ak} = n$, $i = 1, \dots, 4$, gesetzt. Für die Wirkung der aktiven Kontrolle werden verschiedene Erwartungswerte $\mu = 0,8$ und $\mu = 1,3$, verwendet, um in der Simulationsstudie mehrere Zieldosen zu untersuchen. In allen Simulationsszenarien wird, falls nicht anders angegeben, die Anzahl der Wiederholungen auf $n_{sim} = 10.000$ und die Anzahl der Bootstraps pro Simulationsdurchlauf auf $n_{boot} = 5.000$ gesetzt. Drei verschiedene Klassen von Dosis-Wirkungsfunktionen werden in den Simulationsstudien verwendet: die lineare Dosis-Wirkungsgerade, die dreiparametrische Emax-Wirkungskurve und die sigmoidale Emax-Wirkungskurve, welche auch als logistisches Modell bezeichnet wird (Dragalin et al., 2007). Die lineare Funktion wurde bereits in Kapitel 3 eingeführt und die Emax-Funktion kann wie folgt beschrieben werden:

$$f_{\theta}(d) = \theta_0 + \frac{\theta_1 d}{d + \theta_2}$$

mit Achsenabschnitt θ_0 , Steigung θ_1 und der ED_{50} Dosis θ_2 , welche zu einer Wirkung von $\theta_1/2$ führt. Die sigmoidale Emax-Wirkungsfunktion kann durch:

$$f_{\theta}(d) = \theta_0 + \frac{\theta_1}{1 + (\theta_2/d)^{\theta_3}}$$

mit Achsenabschnitt θ_0 , Steigung θ_1 , der ED_{50} θ_2 und dem Sigmoid-Shape-Parameter θ_3 dargestellt werden. Das Emax-Modell ergibt sich als Spezialfall für $\theta_3 = 1$. Für $d = 0$ gilt $f_{\theta}(0) = \theta_0$. Die Parametrisierung der Wirkungsfunktionen orientiert sich ebenfalls an der Studie von Nauck et al. (2009) und ist in Tabelle 4.1 und Abbildung 4.1 dargestellt. Die Erzeugung der Zufallszahlen erfolgt analog zu Abschnitt 3.3.

Tabelle 4.1.: Simulationsszenarien für lineare, Emax- und sigmoidale Emax-Wirkungsfunktionen mit $\theta_0 = -0,4$, $\sigma = 1,8$ und Fallzahlen $n = 20, 25, \dots, 40, 50, \dots, 100$ pro Gruppe.

Dosis-Wirkungsfunktion	μ	θ_1	θ_2	θ_3	d^*
Linear	0,8	1,25	-	-	0,96
	1,3	1,25	-	-	1,36
Emax	0,8	2,675	0,4523	-	0,368
	1,3	2,675	0,4523	-	0,789
sigmoidal Emax	0,8	2,675	0,9	3	0,840
	1,3	2,675	0,9	3	1,083

Um die Spline-Interpolationen mit den in Kapitel 2.3 vorgestellten Maximum-Likelihood Regressionsmethoden vergleichen zu können, sind zusätzliche Annahmen notwendig. Im Vergleich zu den

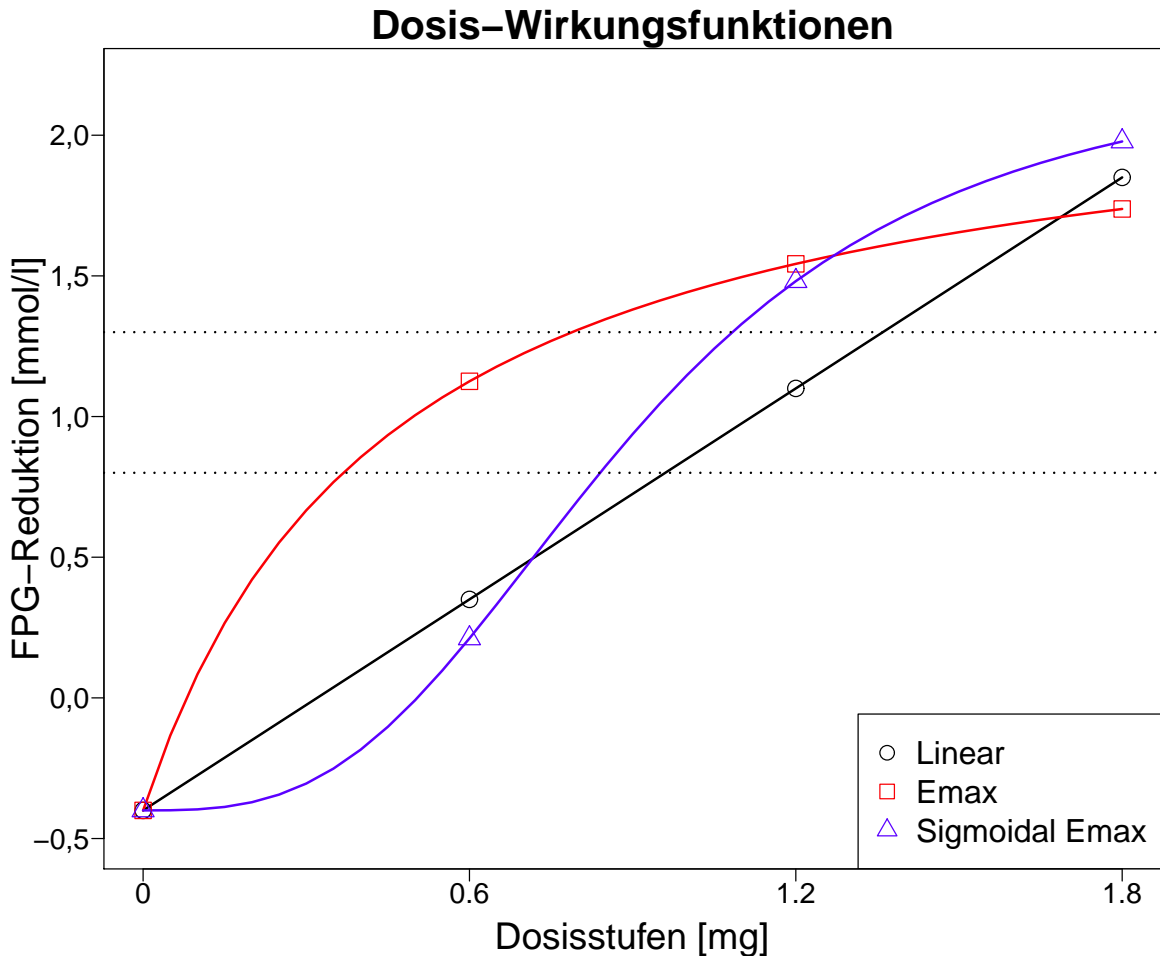


Abbildung 4.1.: Simulationsszenarien: Darstellung der verwendeten linearen, Emax und sigmoidal Emax-Dosis-Wirkungsfunktion (Tab. 4.1).

präsentierten Interpolationsverfahren benötigen die parametrischen Regressionen Informationen über die Gestalt der erwarteten Wirkungsfunktion $f_{\theta}(d)$. Es werden daher zwei Regressionen basierend auf einer Emax-Funktion bzw. auf einer sigmoidalen Emax-Funktion zur Schätzung der Zieldosis verwendet. Die zugehörigen Konfidenzintervalle werden mittels Normalapproximation durch die Δ -Regel aus Korollar 2.5 bestimmt. Die Regressionen und die Konfidenzintervallberechnungen können in der Statistiksoftware SAS 9.3 mit Hilfe der Prozedur NLMIXED auf Grundlage des Newton-Raphson Verfahrens durchgeführt werden (für eine detaillierte Beschreibung siehe SAS-Institute, 2008b, S. 4343-4442). Um numerische Probleme bezüglich der parametrischen Regressionen, speziell für die flexible sigmoidale Emax-Regression zu vermeiden, wird eine dreistufige Anpassungsprozedur verwendet, die in Kirby et al. (2011) und Jones et al. (2011) beschrieben wird. In den Simulationsläufen wird abgefragt, ob ein sigmoidales Emax-Modell auf die Daten angepasst werden kann, falls dies numerisch nicht möglich ist oder die verwendete NLMIXED Prozedur nach 50 Iterationen nicht konvergiert, wird eine Emax-Funktion angepasst und falls dies ebenfalls nicht möglich ist, wird im letzten Schritt eine lineare Funktion angepasst. Eine Übersicht der verschiedenen Konvergenzbedingungen der Prozedur NLMIXED ist in

SAS-Institute (2008b, S. 4355-4363) gegeben.

4.5.1. Eigenschaften der Schätzer der Zieldosis

In diesem Abschnitt wird die Verzerrung der Spline-basierten Zieldosis-Schätzer, die in Abschnitt 4.3 vorgestellt wurde, in Abhängigkeit der Wirkung der aktiven Kontrolle untersucht. Es werden dabei nur Wirkungsgrade der aktiven Kontrolle μ betrachtet, die zu einer Zieldosis innerhalb des Dosisbereichs $D = [0; 1,8]$ führen und somit einen Schnittpunkt mit der Dosis-Wirkungsfunktion f_{θ} aufweisen ($\mu \in [f_{\theta}(0), f_{\theta}(1,8)]$). In Abbildung 4.2 sind die lineare und die kubische Spline-Interpolation sowie die Verzerrung im Falle der Emax-Wirkungsfunktion dargestellt. Die Verzerrungen für die verwendeten Erwartungswerte der aktiven Kontrolle ($\mu = 0,8; 1,3$) liegen bei 0,066 und $-0,056$ für den kubischen und bei 0,104 und 0,063 für den linearen Spline. Die maximale absolute Verzerrung beträgt für den kubischen Spline 0,087 und für den linearen Spline 0,125.

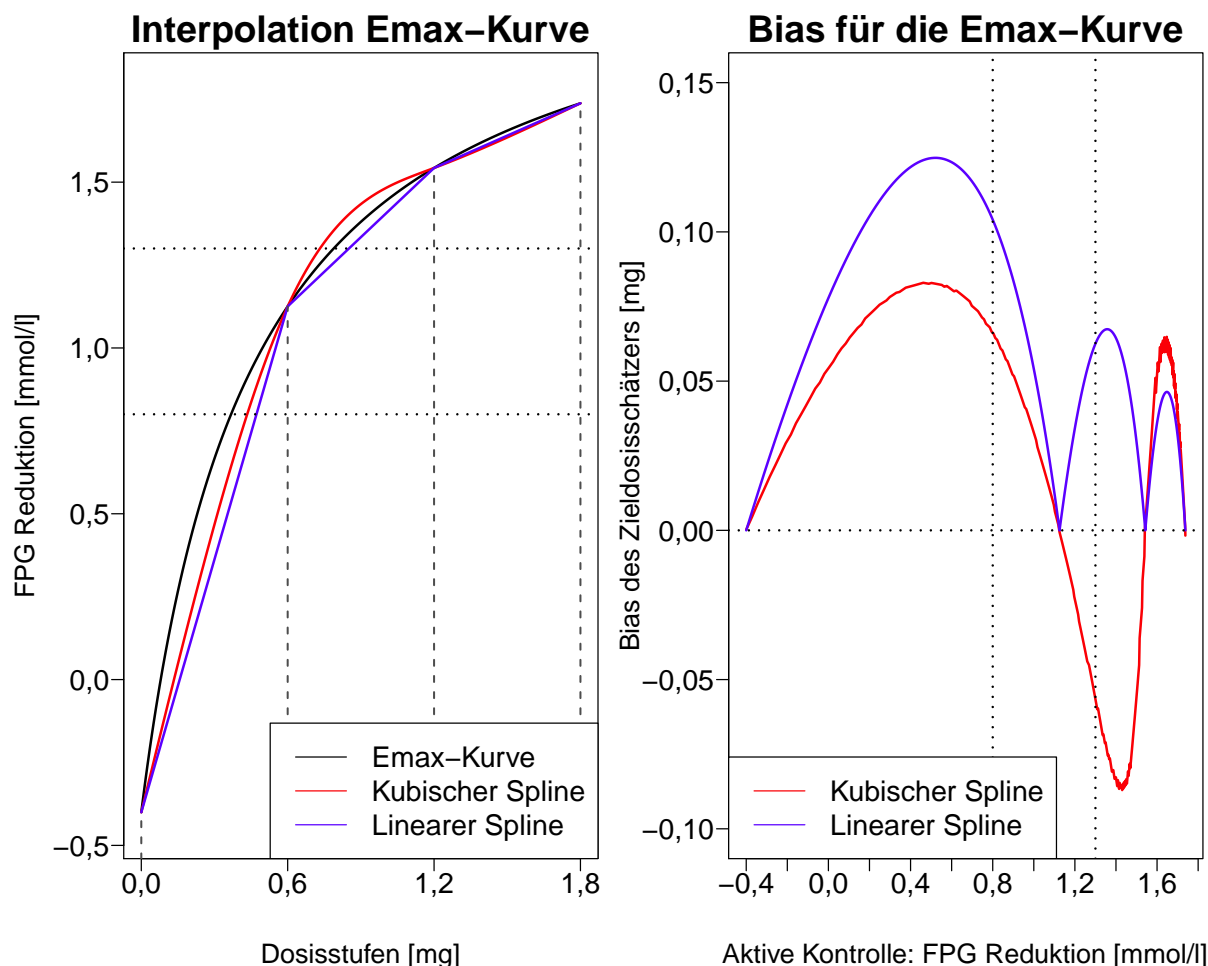


Abbildung 4.2.: Lineare und kubische Spline-Interpolation der Emax-Wirkungsfunktion aus Tabelle 4.1 mit entsprechender Verzerrung von \hat{d}^* für den linearen und den kubischen Spline.

4.5.2. Eigenschaften der Konfidenzintervalle

In diesem Abschnitt wird die Überdeckungswahrscheinlichkeit der verschiedenen Methoden untersucht sowie das Verhalten der medianen Konfidenzintervalllänge diskutiert. Die Überdeckungswahrscheinlichkeit zu einem nominalen 95% Niveau wird für die vier verwendeten Verfahren, linear Spline, kubischer Spline, Δ -Regel Emax-Fit und Δ -Regel sigmoidaler Emax-Fit, in den beschriebenen Simulationsszenarien (siehe Tab. 4.1) berechnet. Um die Güte der verschiedenen Verfahren zu beschreiben, wird erneut das 99%-Intervall des Simulationsfehlers aus Gl. (3.17) verwendet und als gepunktete Linien in Abbildung 4.3 dargestellt. Simulierte Überdeckungswahrscheinlichkeiten außerhalb des Intervalls werden als systematisch verschieden vom nominalen 95% Niveau angesehen. Aus Abbildung 4.3 ist ersichtlich, dass im Falle einer linearen Wirkungskurve beide Spline-Verfahren innerhalb des Simulationsstreifens liegen. Demgegenüber liefern die Emax-Regression und die sigmoidale Emax-Regression in einigen bzw. allen Situationen liberale Überdeckungswahrscheinlichkeiten. Im Falle der sigmoidalen Emax-Dosis-Wirkungsfunktion ist ersichtlich, dass die sigmoidale Regression sowie die Spline-Verfahren die Überdeckungswahrscheinlichkeit gut einhalten und sich nur für größere Fallzahlen ($n > 60$) bei den Spline-Verfahren ein liberaler Trend abzeichnet. Die Emax-Regression hingegen verliert im Fall von $\mu = 1,3$ mit wachsender Fallzahl an Überdeckungswahrscheinlichkeit (bis auf 88%). Für den Fall $\mu = 0,8$ hält sie jedoch das Niveau gut ein. Im Falle der Emax-Dosis-Wirkungsfunktion und $\mu = 1,3$ ist die Überdeckungswahrscheinlichkeit von allen Verfahren innerhalb des Intervalls des Simulationsfehlers, wobei die sigmoidale Emax-Regression, die eine leicht liberale Überdeckungswahrscheinlichkeit aufweist, eine Ausnahme bildet. Am interessantesten ist das Szenario Emax-Wirkungskurve mit $\mu = 0,8$, für welches keines der Verfahren die Überdeckungswahrscheinlichkeit einhält. Mit wachsender Fallzahl werden alle Methoden - mit Ausnahme der Emax-Regression - liberaler; besonders ist dieses Verhalten beim linearen Spline und bei der sigmoidalen Emax-Regression zu sehen. Für den kubischen Spline fällt der liberale Trend etwas schwächer aus, so dass für $n = 100$ die Überdeckungswahrscheinlichkeit bei ca. 90% liegt. Dass die sigmoidale Emax-Regression im Falle einer Emax-Wirkungsfunktion die Zieldosis nicht gut schätzt, ist wegen der Abweichung von den Modellannahmen nicht sehr verwunderlich. Für den kubischen und speziell den linearen Spline kann das sehr liberale Verhalten zu großen Teilen durch die starke Verzerrung des Zieldosissschätzers erklärt werden, die in Abbildung 4.2 dargestellt wurde.

Neben der Überdeckungswahrscheinlichkeit wird die mediane Konfidenzintervalllänge untersucht. Dies führt zu einigen Problemen, denn es gibt Simulationsszenarien in denen die Konfidenzintervallgrenzen außerhalb des untersuchten Dosisbereichs liegen. Parametrische Verfahren haben damit keine Schwierigkeiten. Anders ist es für die vorgestellten Spline-Interpolationen, die keine Extrapolation über die Grenzen des untersuchten Dosisbereichs hinaus vorsehen. In diesen Fällen führen die Spline-Verfahren zu offenen oder halboffenen Konfidenzintervallen. In allen Fällen mit endlichen Intervalllängen gibt es keine erkennbaren Unterschiede zwischen den vier Methoden und aus diesen Gründen wird auf die Darstellung der medianen Intervalllänge verzichtet.

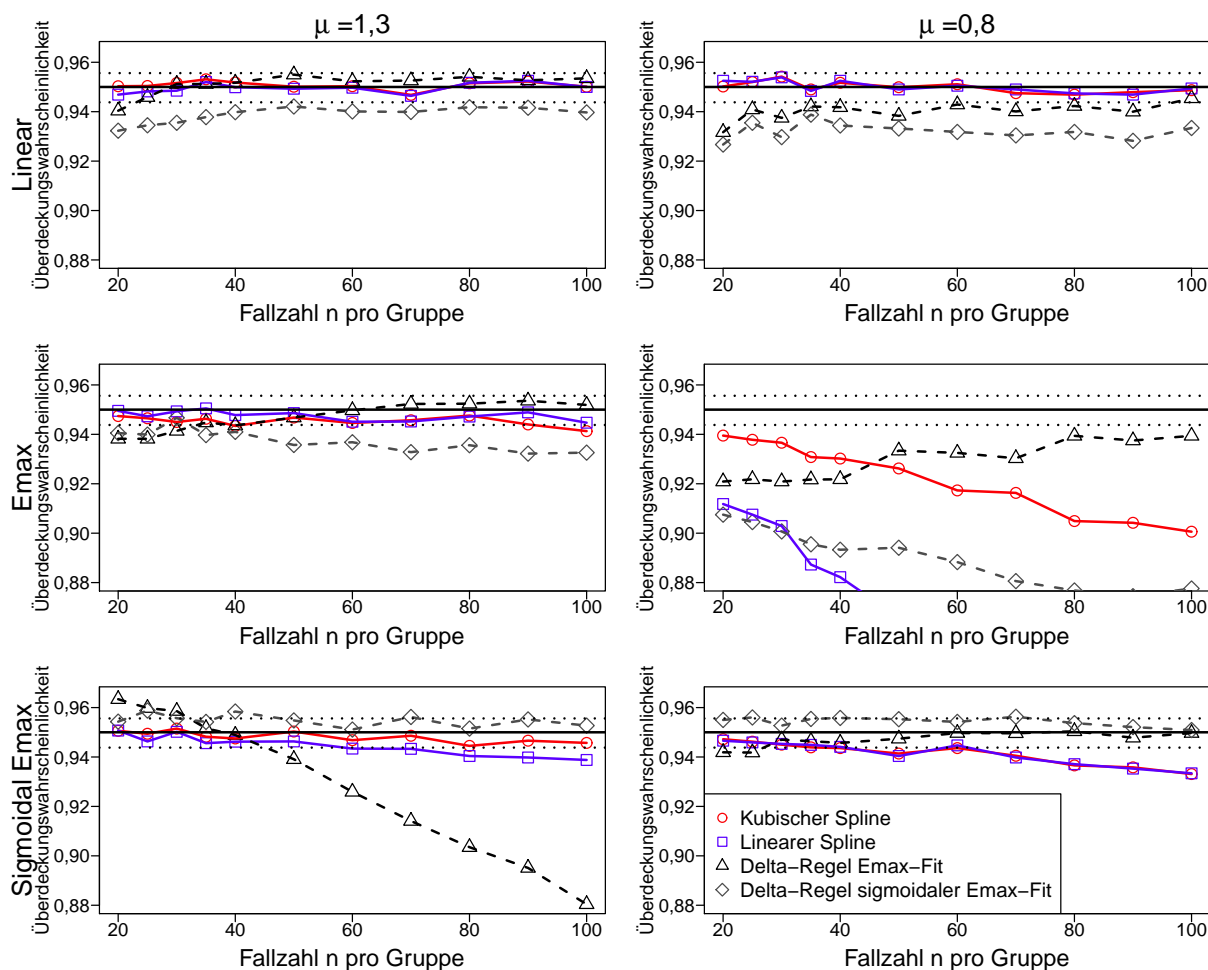


Abbildung 4.3.: Überdeckungswahrscheinlichkeit des kubischen Splines, des linearen Splines, der Δ -Regel Emax-Regression (Δ -Regel Emax-Fit) und der Δ -Regel sigmoidale Emax-Regression (Δ -Regel sigmoidaler Emax-Fit) aufgetragen gegen die Fallzahlen n pro Gruppe für lineare, Emax- und sigmoidalen Emax-Wirkungskurven und verschiedene μ . Die gepunkteten Linien bezeichnen das 99%-Intervall des Simulationsfehlers. Für alle Szenarien betrug die Anzahl der Simulationen $n_{sim} = 10.000$ mit jeweils $n_{boot} = 5.000$ Bootstraps pro Simulation.

Zusätzlich zu den vorgestellten Ergebnissen mit $k = 4$ Dosisstufen werden die Simulationsszenarien aus Tabelle 4.1 auch für $k \in \{5, 6, 8\}$ äquidistante Dosisstufen auf dem Dosisbereich $[0; 1,8]$ untersucht. Ziel dieser Simulationen ist es zu überprüfen, ob sich die Eigenschaften der Spline-Interpolationen mit wachsender Anzahl an Dosisstufen (Stützstellen) verbessern oder ob eine große Anzahl an Dosisstufen möglicherweise zu Problemen führt.

Die Überdeckungswahrscheinlichkeiten sind exemplarisch für die Szenarien mit einer Emax-Dosis-Wirkungskurve in Abbildung 4.4 dargestellt. Mit wachsender Anzahl an Dosisstufen kann eine deutliche Verbesserung der Überdeckungswahrscheinlichkeit der kubischen und linearen Spline-Interpolation beobachtet werden. Der kubische Spline hält bereits für $k = 5$ und der lineare Spline ab $k = 6$ Dosisstufen die nominale Überdeckungswahrscheinlichkeit über alle untersuchten Fallzahlen sehr gut ein und liegt innerhalb des 99%-Intervalls des Simulationsfehlers. Andererseits

führt eine Erhöhung der Anzahl der Dosisstufen, mit welcher auch eine Erhöhung der Gesamtfallzahl N einhergeht, zu keiner beobachtbaren Verbesserung der Überdeckungswahrscheinlichkeit Emax-Regression. Die Überdeckungswahrscheinlichkeiten sind liberal, insbesondere für kleinere Fallzahlen n pro Gruppe und nahezu identisch zu den Ergebnissen für $k = 4$ Dosisstufen (siehe Abb. 4.3). Die sigmoidale Emax-Regression schwankt in Abhängigkeit der Fallzahl n zwischen liberalen und konservativen Überdeckungswahrscheinlichkeiten.

In den Simulationsszenarien mit linearer und sigmoidaler Emax-Dosis-Wirkungskurve halten die Spline-Interpolationen die nominale Überdeckungswahrscheinlichkeit auch für mehrere Dosisstufen sehr gut ein (vgl. Abbildung 4.3). Die Ergebnisse der Emax-Regression sind ebenfalls nahezu unverändert zu den Überdeckungswahrscheinlichkeiten der Szenarien mit $k = 4$ Dosisstufen. Nur die sigmoidale Emax-Regression führt mit wachsender Anzahl an Dosisstufen zu liberaleren Überdeckungswahrscheinlichkeiten. Diese Simulationsergebnisse werden in der vorliegenden Arbeit nicht dargestellt und es wird auf Helms et al. (2014b) verwiesen.

4.5.3. Penalisierter kubischer Glättungsspline

Das in Abschnitt 4.5.2 vorgestellte Bootstrap-Konfidenzintervall kann für jede Art von Schätzung der Dosis-Wirkungsfunktion verwendet werden. Daher wird in diesem Abschnitt der mögliche Vorteil des komplexeren penalisierten kubischen Glättungssplines gegenüber der vorgestellten kubischen Spline-Interpolation untersucht. Die Methode der Spline-Glättung zur Schätzung der Dosis-Wirkungskurve in Dosis-Findungsstudien ohne aktive Kontrolle wurde von Kirby et al. (2009) untersucht. Ein möglicher Vorteil des Glättungssplines ist eine robustere Schätzung der Wirkungskurve und eine daraus resultierende mögliche Verbesserung der Schätzung der Zieldosis. Analog zu Kirby et al. (2009) wird daher ein kubischer Spline in Kombination mit einem quadratischen Strafterm (*penalty term*) verwendet. Nun wird die folgende Gleichung optimiert:

$$A = \sum_{i=1}^k (\bar{Y}_{i\cdot} - s_3(d_i))^2 + \alpha_{pen} \int_{d_1}^{d_k} (s_3''(t))^2 dt,$$

wobei $\alpha_{pen} \in \mathbb{R}_0^+$ der sogenannte Glättungsparameter ist und $s_3''(\cdot)$ die zweite Ableitung der kubischen Spline-Funktion. Der Parameter α_{pen} kontrolliert die Gewichtung des Strafterms und beeinflusst damit die Balance zwischen Anpassungsgüte (*goodness of fit*) und Glätte (*smoothness*) der Spline-Schätzung. Größere Werte von α_{pen} führen zu einer glatten Spline-Schätzung der Dosis-Wirkungsfunktion, wohingegen kleinere Werte zu einer Spline-Funktion führen, die näher am Verlauf der Spline-Interpolation liegt. Für $\alpha_{pen} = 0$ wären die kubische Spline-Interpolation und der Glättungsspline identisch. Demgegenüber wird für $\alpha_{pen} \rightarrow \infty$ der Strafterm dominieren und somit ein einziges quadratisches Polynom angepasst werden. Daher ist die Wahl des Glättungsparameters von großer Bedeutung und sollte mit Bedacht durchgeführt werden. Hinweise diesbezüglich finden sich zum Beispiel in SAS-Institute (2008b, S. 7068).

Der Glättungsparameter α_{pen} kann auf vielfältige Weise bestimmt werden. Zum Beispiel kann ein

Wert für α_{pen} direkt vorgegeben, die Spline-Funktionen für verschiedene Ausprägungen von α_{pen} visualisiert oder der Parameter kann mit Hilfe einer Kreuzvalidierung (*Cross-Validation*) berechnet werden. Es stehen mehrere Kreuzvalidierungsverfahren zur Verfügung. Bei der gewöhnlichen Kreuzvalidierung (*leave one out Cross-Validation*) wird die Spline-Funktion k -mal berechnet und in jedem Durchlauf auf die Informationen der i -ten Dosis verzichtet. Analog zu Kirby et al. (2009) wird im Folgenden die generalisierte Kreuzvalidierung verwendet, um den Glättungsparameter zu bestimmen (siehe z. B. Wahba, 1990, Kap. 4). Um diesen komplexeren Spline-Ansatz mit dem kubischen Spline zu vergleichen, wird eine kleine Simulationsstudie durchgeführt, die die Szenarien der Emax-Wirkungskurve aus Tabelle 4.1 mit $k = 4$ bzw. 8 Dosisstufen zu den Fallzahlen $n = 25$ bzw. 50 pro Gruppe sowie $n_{sim} = 2.000$ und $n_{boot} = 1.000$ untersucht. Die geringe Anzahl an Simulationsdurchläufen ist der erhöhten numerischen Komplexität des penalisierten Splines und der damit verbundenen erhöhten Laufzeit geschuldet. Die Situationen mit $k = 8$ Dosisstufen werden in die Simulationsstudie aufgenommen, um die potentiellen Vorteile des penalisierten Glättungssplines aufzuzeigen, welche in der Glättung von verrauschten Daten liegen. Weiterhin werden neben dem Finden von α_{pen} durch Kreuzvalidierung (*CV*) zwei vorgegebene Glättungsparameter, die zu einer mittleren sowie einer hohen Glättung führen ($\log_{10}(k \cdot \alpha_{pen}) = -4, -2$) mit der kubischen Interpolation verglichen. Um den penalisierten Glättungsspline zu berechnen, wird die SAS 9.3 Prozedur TPSPLINE verwendet. Die Prozedur lässt vordefinierte Werte für α_{pen} zu und berechnet α_{pen} mittels generalisierter Kreuzvalidierung (SAS-Institute, 2008b, S. 7048-7068).

Die Ergebnisse dieser Simulationsstudie sind in Tabelle 4.2 zusammengefasst. Es ist ersichtlich, dass sich in den betrachteten Situationen kein substantieller Vorteil durch die Verwendung des penalisierten Splines gegenüber der kubischen Spline-Interpolation ergibt. Weiterhin ist ersichtlich, dass die Verwendung der Kreuzvalidierung für die Wahl von α_{pen} für $k = 4$ zu einem stärker verzerrten Zieldosis-Schätzer und zu einer sehr liberalen Überdeckungswahrscheinlichkeit (88%) führt. Dies kann durch die geringe Anzahl der Dosisstufen erklärt werden, auf dessen Grundlage eine Kreuzvalidierung zu sehr unzuverlässigen Ergebnissen führt. In diesen Fällen wäre es von Vorteil auf die Kreuzvalidierung zu verzichten und den Glättungsparameter vorgeben oder direkt die kubische Interpolation zu verwenden. Diese liefert für $k = 4$ die beste Überdeckungswahrscheinlichkeit und die geringste Verzerrung. Für $k = 8$ ist ersichtlich, dass alle Spline-Verfahren sehr gute Ergebnisse liefern und somit weder bei der Überdeckungswahrscheinlichkeit noch bei der Verzerrung ein wirklicher Vorteil des penalisierten Splines gezeigt werden kann. Da Dosis-Findungsstudien oft nur 4 bis 8 Dosisstufen untersuchen, kann im Falle von sehr wenigen Dosisstufen die Verwendung des penalisierten Splines mit Kreuzvalidierung zu unerwünschten Problemen führen. Dies ist einer der Hauptgründe, warum auf den simpleren Ansatz der kubischen Spline-Interpolation zurückgegriffen wird. Ist neben der Schätzung der Zieldosis auch die Identifikation der erwarteten Dosis-Wirkungskurve von Interesse, so ist es mittels penalisiertem Spline möglich, deutlich glattere Kurven als mit Hilfe der Spline-Interpolation anzupassen. Dies gilt speziell für eine hohe Anzahl an untersuchten Dosisstufen.

4.6. Zusammenfassung

In diesem Kapitel wurde die Schätzung der Zieldosis ohne Annahmen an die Klasse der Dosis-Wirkungsfunktion $f_{\theta}(d)$ mit Hilfe linearer und kubischer Spline-Interpolation vorgestellt. Zusätzlich wurde eine Methode zur Konstruktion eines Konfidenzintervalls der Zieldosis ohne Annahmen an die Wirkungsfunktion mit Hilfe eines parametrischen Bootstraps der Wirkungsmittelwerte pro Dosisstufe entwickelt, die für beliebige Berechnungsmethoden des Zieldosis-Schätzers verwendet werden kann. Speziell ist es möglich, für jede Art von Spline-basierter Schätzung der Zieldosis ein Konfidenzintervall zu berechnen. Die Eigenschaften dieser Interpolationsverfahren wurden mittels Simulationsstudien für verschiedene Wirkungsfunktionen untersucht und mit zwei parametrischen Regressionen verglichen.

Es konnte gezeigt werden, dass die Verzerrung der Interpolationsmethoden unabhängig von der untersuchten Fallzahl N ist und die zugehörigen Zieldosis-Schätzer somit nicht asymptotisch erwartungstreu sind. Im Gegensatz dazu folgt die asymptotische Erwartungstreue für die parametrischen Maximum-Likelihood Regressionen aus Korollar 2.5. Daher wurde die Verzerrung der Spline-Interpolationen gesondert betrachtet, wobei gezeigt werden konnte, dass die Verzerrung des Zieldosis-Schätzers stark vom Schnittpunkt der Interpolationsfunktion mit dem Wirkungsgrad der aktiven Kontrolle μ abhängt (Abb. 4.2).

Zusätzlich wurde die Überdeckungswahrscheinlichkeit und die mediane Konfidenzintervalllänge der linearen und kubischen Spline-Interpolation sowie der Emax und der sigmoidalen Emax-Regression untersucht. Es konnte gezeigt werden, dass die kubische Spline-Interpolation die nominale Überdeckungswahrscheinlichkeit von 95% für die lineare, die Emax und die sigmoidale Emax-Wirkungskurve in fast allen Szenarien gut einhält. Für den Fall einer Emax-Wirkungskurve mit erwarteter Wirkung der aktiven Kontrolle von $\mu = 0,8$ wurde ersichtlich, dass der kubische Spline liberal wird und sich dieser Trend für größere Fallzahlen n pro Gruppe noch verstärkt. Die Überdeckungswahrscheinlichkeit der linearen Spline-Interpolation ist nahezu identisch mit Überdeckungswahrscheinlichkeit der kubischen Interpolation. Es fällt allerdings auf, dass die Überdeckungswahrscheinlichkeit im beschriebenen Emax-Szenario mit $\mu = 0,8$ erheblich unter 95% sinkt und somit zu extrem liberalen Ergebnissen führt. Dieses Problem kann zu großen Teilen durch die starke Verzerrung der Spline-basierten Zieldosis-Schätzer in diesem Szenario erklärt werden, die für wachsende Fallzahlen immer dominanter wird (siehe Abb. 4.2).

Die Emax-Regression hält in den betrachteten Situationen die Überdeckungswahrscheinlichkeit gut ein, falls eine lineare oder Emax-Wirkungsfunktion vorliegt. Eine Ausnahme bildet das beschriebene Emax-Szenario für $\mu = 0,8$, in welchem auch die Emax-Regression liberal ist. Liegt eine sigmoidale Emax-Wirkungskurve vor, so fällt die Überdeckungswahrscheinlichkeit in einigen Situationen drastisch ab (bis auf 88%). Die sigmoidale Emax-Regression hat in nahezu allen betrachteten Situationen eine liberale Überdeckungswahrscheinlichkeit, außer die simulierte Wirkungsfunktion ist ebenfalls eine sigmoidale Emax-Funktion. Weiterhin war es in 30 – 38% der Simulationen nicht möglich die sigmoidale Emax-Regression mit Hilfe der SAS-Prozedur

NLMIXED zu bestimmen, falls die simulierte Dosis-Wirkungskurve einer Emax-Funktion entsprach. Wenn Unsicherheit über die zugrundeliegende Dosis-Wirkungsfunktion besteht, ist es nicht zwangsläufig empfehlenswert diese durch die Verwendung flexiblerer Regressionsmodelle auszugleichen. In solchen Fällen bieten sich zum Beispiel eher Regressionsverfahren unter Modellunsicherheit (Bornkamp et al., 2011) an. Neben den präsentierten Simulationsergebnissen mit $k = 4$ Dosisstufen wurden auch Studiendesigns mit $k \in \{5, 6, 8\}$ Dosisstufen simuliert und untersucht. Es konnte gezeigt werden, dass beide Spline-Verfahren mit wachsender Anzahl an Dosisstufen die nominale Überdeckungswahrscheinlichkeit für alle untersuchten Fallzahlen n gut einhalten und zu ähnlichen Ergebnissen führen.

Aufgrund der Konstruktion der Spline-Interpolationen ist eine Extrapolation über die Grenzen der Dosis-Range $D = [d_1, d_k]$ hinaus nicht sinnvoll und bei der Berechnung der Konfidenzintervalle besteht die Möglichkeit von halboffenen Intervallen. In diesen Fällen wäre die mediane Intervallbreite unendlich und ein Vergleich zwischen den Methoden nicht sinnvoll. In den anderen Fällen, in denen die mediane Intervallbreite für alle betrachteten Methoden endlich ist, konnten keine größeren Unterschiede zwischen den vier Methoden beobachtet werden.

In einem weiteren Abschnitt wurden mögliche Vorteile der Verwendung eines kubischen Glättungssplines mit quadratischem Strafterm und Kreuzvalidierung gegenüber der kubischen Spline-Interpolation durch eine kleine Simulationsstudie untersucht. Es konnte gezeigt werden, dass die Überdeckungswahrscheinlichkeit des Glättungssplines in Szenarien mit vier Dosisstufen deutlich liberaler ausfällt als für die Spline-Interpolation. Das Gleiche gilt für die geschätzte Verzerrung, die im Falle des Glättungssplines größer ausfällt. In allen betrachteten Simulationsszenarien war die Spline-Interpolation entweder überlegen oder es gab keinen nennenswerten Unterschied. Daher ist die Spline-Interpolation, speziell im Falle von wenigen Dosisstufen, dem komplexeren Verfahren vorzuziehen.

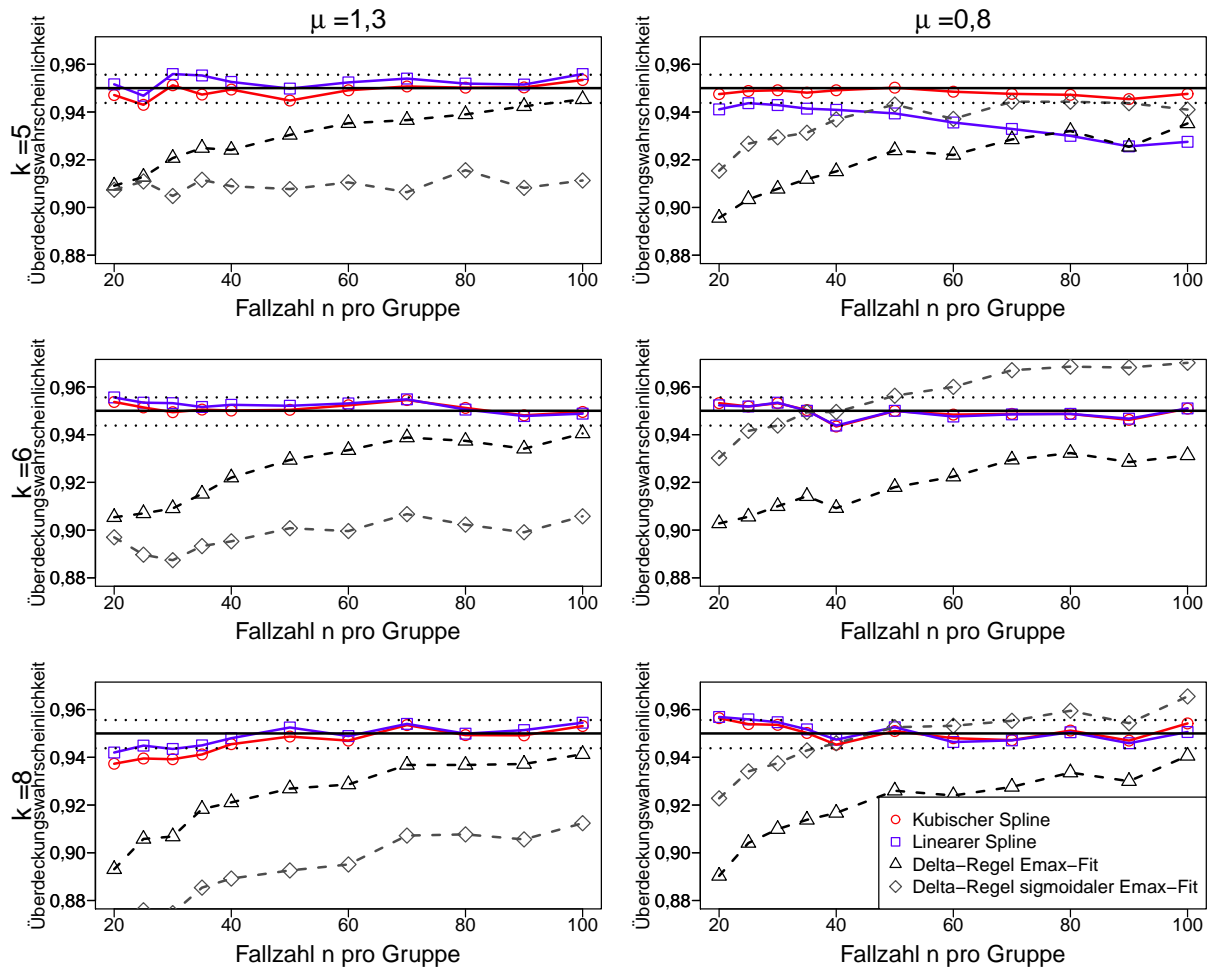


Abbildung 4.4.: Überdeckungswahrscheinlichkeit des kubischen Splines, des linearen Splines, der Δ -Regel Emax-Regression (Δ -Regel Emax-Fit) und der Δ -Regel sigmoidale Emax-Regression (Δ -Regel sigmoidaler Emax-Fit) aufgetragen gegen die Fallzahlen n pro Gruppe für Emax-Dosis-Wirkungskurven, verschiedene Anzahl an Dosisstufen $k \in \{5, 6, 8\}$ und Wirkungsgrade der aktiven Kontrolle μ . Die gepunkteten Linien bezeichnen das 99%-Intervall des Simulationsfehlers. Für alle Szenarien betrug die Anzahl der Simulationsdurchläufe $n_{sim} = 10.000$ mit jeweils $n_{boot} = 5.000$ Bootstraps pro Simulation.

Tabelle 4.2.: Der mittlere Zieldosissschätzer, die Überdeckungswahrscheinlichkeit des nominalen 95%-KI (Ü-Wkeit) und die mediane Intervalllänge für verschiedene Anzahlen an Dosisstufen (k), erwartete Wirkungsgrade der aktiven Kontrolle (μ) und Fallzahlen (n) pro Gruppe wurden für die kubische Spline-Interpolation ($\alpha_{pen} = \text{kein}$) und den Glättungspline mit verschiedenen Glättungsparametern ($\alpha_{pen} = \text{Kreuzvalidierung (CV), hoch, mittel}$) im Szenario der Emax-Dosis-Wirkungsfunktion (Tabelle 4.1) mit $n_{sim} = 2.000$ und $n_{boot} = 1.000$ angegeben.

k	μ	d^*	n	Penalty α_{pen}	Gemittelte \widehat{d}^*	Ü-Wkeit	Mediane Intervallbreite
4	0,8	0,3680	25	CV	0,5814	0,8875	1,2090
				hoch	0,5301	0,9285	1,1870
				mittel	0,5089	0,9305	1,1775
				kein	0,4931	0,9350	1,1680
	1,3	0,7886	50	CV	0,5267	0,8750	0,8510
				hoch	0,5023	0,8825	0,7360
				mittel	0,4651	0,9175	0,6975
				kein	0,4578	0,9260	0,7000
	0,8	0,3680	25	CV	0,9722	0,9490	1,3590
				hoch	0,9302	0,9425	1,4040
				mittel	0,8814	0,9465	1,4365
				kein	0,8753	0,9410	1,4350
1,3	0,7886	50	CV	0,9197	0,9505	1,2100	
			hoch	0,9032	0,9570	1,2525	
			mittel	0,8530	0,9465	1,3220	
			kein	0,8577	0,9435	1,3150	
8	0,8	0,3680	25	CV	0,4668	0,9560	1,1405
				hoch	0,4723	0,9295	1,0960
				mittel	0,4205	0,9485	0,9170
				kein	0,4278	0,9555	0,9040
	1,3	0,7886	50	CV	0,4388	0,9520	0,7900
				hoch	0,4546	0,9335	0,7030
				mittel	0,3962	0,9560	0,7070
				kein	0,4030	0,9455	0,6960
	0,8	0,3680	25	CV	0,9134	0,9525	1,4495
				hoch	0,9074	0,9490	1,4090
				mittel	0,7960	0,9385	1,4010
				kein	0,7734	0,9420	1,3860
1,3	0,7886	50	CV	0,8503	0,9705	1,2790	
			hoch	0,8751	0,9435	1,2695	
			mittel	0,8069	0,9545	1,2325	
			kein	0,7976	0,9455	1,2190	

5. Design- und Fallzahlplanung für Dosis-Findungsstudien mit aktiver Kontrolle

In den vorangegangenen Kapiteln lag der Fokus auf der Schätzung und anschließend auf der Konstruktion eines Konfidenzintervalls für die Zieldosis. Im folgenden Kapitel wird nun die Planung von Dosis-Findungsstudien mit aktiver Kontrolle in Bezug auf Dosisallokation und Fallzahlberechnung analysiert und diskutiert.

5.1. Verzerrungs-reduzierte Designs für die Verwendung Spline-basierter Auswertungen

In diesem Abschnitt werden zunächst Methoden für die in Kapitel 4 vorgestellten Spline-basierten Auswertungen entwickelt, die unter geringen Annahmen an die Wirkungsfunktion die Verzerrung (den Bias) der Zieldosissschätzer mittels Dosisallokation in der Planungsphase reduzieren. Dafür werden im Folgenden die Modellannahmen getroffen. Es werden Dosis-Findungsstudien mit k Dosisstufen $0 \leq d_1 < \dots < d_k$ und einer aktiven Kontrolle (ak) untersucht, wobei nur die minimale und maximale Dosisstufe d_1 und d_k festgelegt sein sollen. Diese können zum Beispiel durch Informationen über die Verträglichkeit des Medikamentes aus einer Phase I Studie berechnet oder aus präklinischen Experimenten abgeleitet werden. Weiterhin soll der Wirkungsgrad μ der aktiven Kontrolle unbekannt sein. Zusätzlich zu den verbleibenden Annahmen aus Abschnitt 2.2 sind in der Planungsphase Informationen oder Annahmen über die Gestalt der Wirkungsfunktion notwendig, auch wenn die Spline-Interpolationen für die Auswertung die Form der Wirkungsfunktion nicht benötigen. Daraus folgt, dass das Bias-reduzierte Design von der Klasse der Dosis-Wirkungsfunktion abhängt. Im Folgenden werden drei Möglichkeiten zur Dosisauswahl mit dem Ziel vorgestellt, die maximale absolute Verzerrung für alle in Frage kommenden Wirkungsgrade der aktiven Kontrolle μ zu reduzieren. Wirkungsgrade μ , die zu einer Zieldosis außerhalb des Dosisbereichs $D = [d_1, d_k]$ führen oder keinen Schnittpunkt mit der angenommenen Dosis-Wirkungsgeraden besitzen, sollen nicht betrachtet werden. Die daraus resultierenden Designs werden somit unabhängig vom Schnittpunkt der Interpolation mit der Wirkung der aktiven Kontrolle μ berechnet. Weiterhin wird für die allgemeinen Verfahren anhand verschiedener Dosis-Wirkungsfunktionen (linear, Emax und sigmoidale Emax-Funktion) illustriert, welche

Parameter zur Bestimmung des Designs benötigt werden.

5.1.1. Minimum des maximalen globalen Interpolationsfehlers

Unter Verwendung der Taylor-Approximation konnten Hall (1968), Hall und Meyer (1976) und de Boor (2001, Kap. 5) zeigen, dass die folgenden globalen Obergrenzen des Interpolationsfehlers für die lineare und kubische Spline-Interpolation existieren. Für den linearen Spline s_1 ist die globale obere Fehlergrenze gegeben durch

$$\|f_{\theta} - s_1\|_{\infty} \leq C_{s_1} \|f''_{\theta}\|_{\infty} h_{max}^2 = C_{s_1} \cdot \max_{d \in [d_1, d_k]} |f''_{\theta}(d)| \cdot h_{max}^2 \leq \frac{1}{8} \|f''_{\theta}\|_{\infty} h_{max}^2$$

und für die kubische Spline-Interpolation s_3 ist die globale obere Fehlergrenze gegeben durch

$$\|f_{\theta} - s_3\|_{\infty} \leq C_{s_3} \|f_{\theta}^{(4)}\|_{\infty} h_{max}^4 = C_{s_3} \cdot \max_{d \in [d_1, d_k]} |f_{\theta}^{(4)}(d)| \cdot h_{max}^4 \leq \frac{5}{384} \|f_{\theta}^{(4)}\|_{\infty} h_{max}^4$$

mit Konstanten C_{s_1} , C_{s_3} und $h_{max} = \max_i (d_{i+1} - d_i)$ sowie der Maximumnorm $\|g\|_{\infty} = \max_{x \in [a, b]} |g(x)|$. Um diese Obergrenzen des absoluten Interpolationsfehlers zu minimieren, muss h_{max} minimiert werden, da $\|f''_{\theta}\|_{\infty}$ und $\|f_{\theta}^{(4)}\|_{\infty}$ konstant auf $D = [d_1, d_k]$ sind.

Das Minimum von h_{max} wird bei einer äquidistanten Allokation der verbleibenden Dosisstufen zwischen d_1 und d_k erreicht; dann gilt: $h_{max} = d_{i+1} - d_i$ für $i = 1, \dots, k - 1$. Diese Art der Wahl der Dosisstufen (Knoten) wird hauptsächlich zum Nachweis asymptotischer Eigenschaften verwendet, indem die Anzahl der Dosisstufen gegen unendlich geht ($k \rightarrow \infty$) und somit aus den oben genannten Abschätzungen die Konvergenz der Spline-Interpolationen mit $h_{max} \rightarrow 0$ folgt (siehe auch Floater, 2007, Kap. 5). Für klinische Studien ist diese Annahme nicht realisierbar, so dass die Asymptotik mit Hilfe der Fallzahl N und einer festen Anzahl an Dosisstufen k erzielt werden soll. Diese Methode erzeugt unabhängig von der zugrundeliegenden Wirkungsfunktion $f_{\theta}(\cdot)$ eine äquidistante Dosisallokation zwischen d_1 und d_k und benötigt somit keine zusätzlichen Annahmen.

5.1.2. Dosisstufen mit abstandsgleichen Wirkungsgraden

In diesem Ansatz werden die verbleibenden Dosisstufen so gewählt, dass die Differenzen der Wirkungsgrade zweier adjazenter (benachbarter) Dosisstufen identisch für alle Dosislevel sind:

$$|f_{\theta}(d_{i+1}) - f_{\theta}(d_i)| \stackrel{!}{=} h_{equ} \text{ für } i = 1, \dots, k - 1.$$

Durch diese Bedingung werden Dosisstufen in den Abschnitten der Dosis-Wirkungsfunktion f_{θ} gewählt, welche starke Veränderungen im Wirkungsgrad aufweisen. Dieses Verfahren hängt von der gewählten Dosis-Wirkungsfunktion ab. Allerdings kann gezeigt werden, dass die Wahl des Designs nach diesem Verfahren für die lineare, die Emax und die sigmoidale Emax-Funktion unabhängig von Achsenabschnitt θ_0 und von der Steigung θ_1 ist (Beweis siehe Abs. B.2.1 im

Anhang). Daraus folgt für eine lineare Wirkungsfunktion, dass keinerlei Annahmen an die Funktionsparameter gemacht werden müssen und sich das optimale Design analog zu Abschnitt 5.1.1 durch äquidistante Allokation der verbleibenden Dosisstufen zwischen d_1 und d_k ergibt. Weiterhin muss zur Berechnung des optimalen Designs einer Emax-Wirkungskurve nur die ED_{50} θ_2 und für eine sigmoidale Emax-Funktion zusätzlich zu θ_2 der Grad der sigmoidalen Krümmung θ_3 vorgegeben werden. Unter diesem und dem vorangegangenen Verfahren sind die optimalen Designs für die lineare und die kubische Spline-Interpolation identisch.

5.1.3. Gleichmäßige Verteilung der Obergrenze des Interpolationsfehlers

In diesem Ansatz werden die Ideen der beiden vorangegangenen Methoden kombiniert. Dafür werden Abschätzungen der globalen Obergrenzen des Interpolationsfehlers aus Abschnitt 5.1.1 bedingt auf jedes Intervall zwischen zwei benachbarten Dosisstufen gleichzeitig berechnet. Die Optimalitätsbedingungen für die lineare Spline-Interpolation s_1 sind durch

$$\|f_{\boldsymbol{\theta}} - s_1\|_{\infty} |_{d \in [d_i, d_{i+1}]} \leq \|f''_{\boldsymbol{\theta}}\|_{\infty} |_{d \in [d_i, d_{i+1}]} (d_{i+1} - d_i)^2 \stackrel{!}{=} h_1 \text{ mit } i = 1, \dots, k-1, \quad (5.1)$$

und für die kubische Spline-Interpolation s_3 durch

$$\|f_{\boldsymbol{\theta}} - s_3\|_{\infty} |_{d \in [d_i, d_{i+1}]} \leq \|f_{\boldsymbol{\theta}}^{(4)}\|_{\infty} |_{d \in [d_i, d_{i+1}]} (d_{i+1} - d_i)^4 \stackrel{!}{=} h_3 \text{ mit } i = 1, \dots, k-1, \quad (5.2)$$

gegeben. Die verbleibenden Dosisstufen werden so gewählt, dass die Gleichungen des jeweiligen Splines erfüllt sind. Folglich ist der maximale Interpolationsfehler auf jedem Intervall zwischen zwei Dosisstufen gleich groß und sorgt für eine gleichmäßig gute Interpolation über den untersuchten Dosisbereich $D = [d_1, d_k]$. Analog zu den vorangegangenen Methoden kann gezeigt werden, dass dieses Verfahren für die lineare, die Emax und die sigmoidale Emax-Funktion unabhängig vom Achsenabschnitt θ_0 und der Steigung θ_1 ist (Beweis siehe Anhang B.2.5). Zur Konstruktion des optimalen Designs werden, wie auch für den Ansatz aus Abschnitt 5.1.2, numerische Optimierungsverfahren zum Lösen des jeweiligen Systems der Designbedingungen benötigt. Dabei kann wieder auf die in Abschnitt 3.2.3 beschriebenen Optimierungsroutinen zum Finden von Nullstellen in Gleichungssystemen zurückgegriffen werden. Um diese Art von Routinen auf die präsentierten Gleichungssysteme anwenden zu können, werden alle $k-1$ Bedingungen des entsprechenden Ansatzes in einem Vektor \mathbf{b} zusammengefasst (z. B. $\mathbf{b} = \mathbf{1}_{k-1}h$) und mit der Projektionsmatrix $\mathbf{P}_{k-1} = \mathbf{I}_{k-1} - \frac{1}{(k-1)}\mathbf{1}_{k-1}\mathbf{1}'_{k-1}$ multipliziert. Dadurch vereinfacht sich das Finden der Dosisallokation auf die Bestimmung der $k-1$ Nullstellen ($\mathbf{P}_{k-1}\mathbf{b} = \mathbf{0}$). Zur numerischen Berechnung der gesuchten Dosisstufen kann dann zum Beispiel die Levenberg-Marquardt Kleinste-Quadrate-Methode verwendet werden (Abs. 3.2.3), die in SAS/IML durch die Routine NLPLM implementiert ist.

5.1.4. Das Minimum der maximalen globalen absoluten Verzerrung des Zieldosischätzers

Im Folgenden werden die vorgestellten Verfahren aus den Abschnitten 5.1.1 bis 5.1.3 mit $m = 1, 2, 3$ durchnummeriert. Um für die gewünschte Spline-Interpolation (linear oder kubisch) entscheiden zu können, welches der drei vorgestellten Verfahren hinsichtlich der maximalen absoluten Verzerrung des Zieldosischätzers überlegen ist, muss ein Kriterium definiert werden. Dieses ist durch

$$\text{Bias}_{opt} = \min_{m \in \{1, 2, 3\}} \max_{\mu \in [f_{\theta}(d_1) + c_u, f_{\theta}(d_k) - c_o]} |\text{Bias}_m(\mu)|, \quad (5.3)$$

wobei $\text{Bias}_m(\mu)$ die Verzerrung von Methode m in Abhängigkeit von μ widerspiegelt. Mit Hilfe dieser Kennzahl wird das Verfahren m zur Berechnung der optimalen Dosisallokation verwendet, welches die kleinste maximale absolute Verzerrung der Zieldosis über alle Wirkungsgrade der aktiven Kontrolle μ ($c_u, c_o \in [0, (f_{\theta}(d_k) - f_{\theta}(d_1))/2]$) aufweist. Es ist somit möglich, ein Design in Abhängigkeit der zugrundeliegenden Wirkungsfunktion zu bestimmen, das unabhängig vom Wirkungsgrad der aktiven Kontrolle μ ist und das Minimum der maximalen globalen absoluten Verzerrung der Zieldosis der untersuchten Methoden aufweist. Enthält ein Studiendesign eine Placebo-Kontrolle ($d_1 = 0$), dann sollte das aktive Vergleichspräparat zu einem höheren Wirkungsgrad als Placebo führen. Somit könnte der Wirkungsbereich der aktiven Kontrolle mit Hilfe von $c_u > 0$ sinnvoll eingeschränkt werden. In allen präsentierten Beispielen wurden $c_u = c_o = 0$ gewählt.

Exemplarisch wird in Abbildung 5.1 das Bias minimale Design des kubischen Splines für die Emax-Dosis-Wirkungskurve, die in den Simulationsszenarien (siehe Tabelle 4.1) verwendet wurde, dargestellt, um einen direkten Vergleich mit der Verzerrung des ursprünglich äquidistanten Designs herzustellen (siehe Abb. 4.2). Anstelle der ursprünglichen Dosisstufen $(d_1, \dots, d_4)' = (0; 0,6; 1,2; 1,8)'$ ergibt die optimale Dosisallokation (mittels des Ansatzes abstandsgleicher Wirkungsgrade $m = 2$) die Dosisstufen $(0; 0,165; 0,516; 1,8)'$. Es ist ersichtlich, dass im Vergleich zu Abbildung 4.2 die maximale absolute Verzerrung des Zieldosischätzers mit 0,023 anstelle von 0,087 deutlich geringer ausfällt als zuvor. Allerdings ist auch erkennbar, dass das optimale Design für die kubische Spline-Interpolation nicht automatisch zum optimalen Design der linearen Spline-Interpolation führt, da sich die maximale absolute Verzerrung von 0,125 auf 0,267 erhöht. In solchen Situationen sollte auf die Verwendung der linearen Spline-Interpolation in optimalen Designs des kubischen Splines verzichtet werden. Im betrachteten Szenario ist die äquidistante Dosisallokation das optimale Design für den linearen Spline (siehe Abs. 4.5.1).

Werden mehr als vier Dosisstufen untersucht, verringert sich die maximale absolute Verzerrung für beide Interpolationsansätze. Dies folgt aus den globalen Obergrenzen des Interpolationsfehlers (Abs. 5.1.1) und der Bedingung aus Gl. (5.3), dass die minimale maximale absolute Verzerrung Bias_{opt} kleiner oder gleich der maximalen absoluten Verzerrung der äquidistanten Dosisallokation

($m=1$) ist. Werden zum Beispiel im betrachteten Emax-Szenario Dosisallokationen mit $k = 6$ oder $k = 10$ Dosisstufen untersucht, dann ergibt sich für die optimale Dosisallokationen der kubischen Spline-Interpolation eine maximale absolute Verzerrung des kubischen Splines von 0,0103 für $k = 6$ und 0,0094 für $k = 10$ Dosisstufen. Für die optimale Dosisallokation der linearen Spline-Interpolation ergibt sich für den linearen Spline eine absolute maximale Verzerrung von 0,0523 für $k = 6$ und 0,0182 für $k = 10$ Dosisstufen.

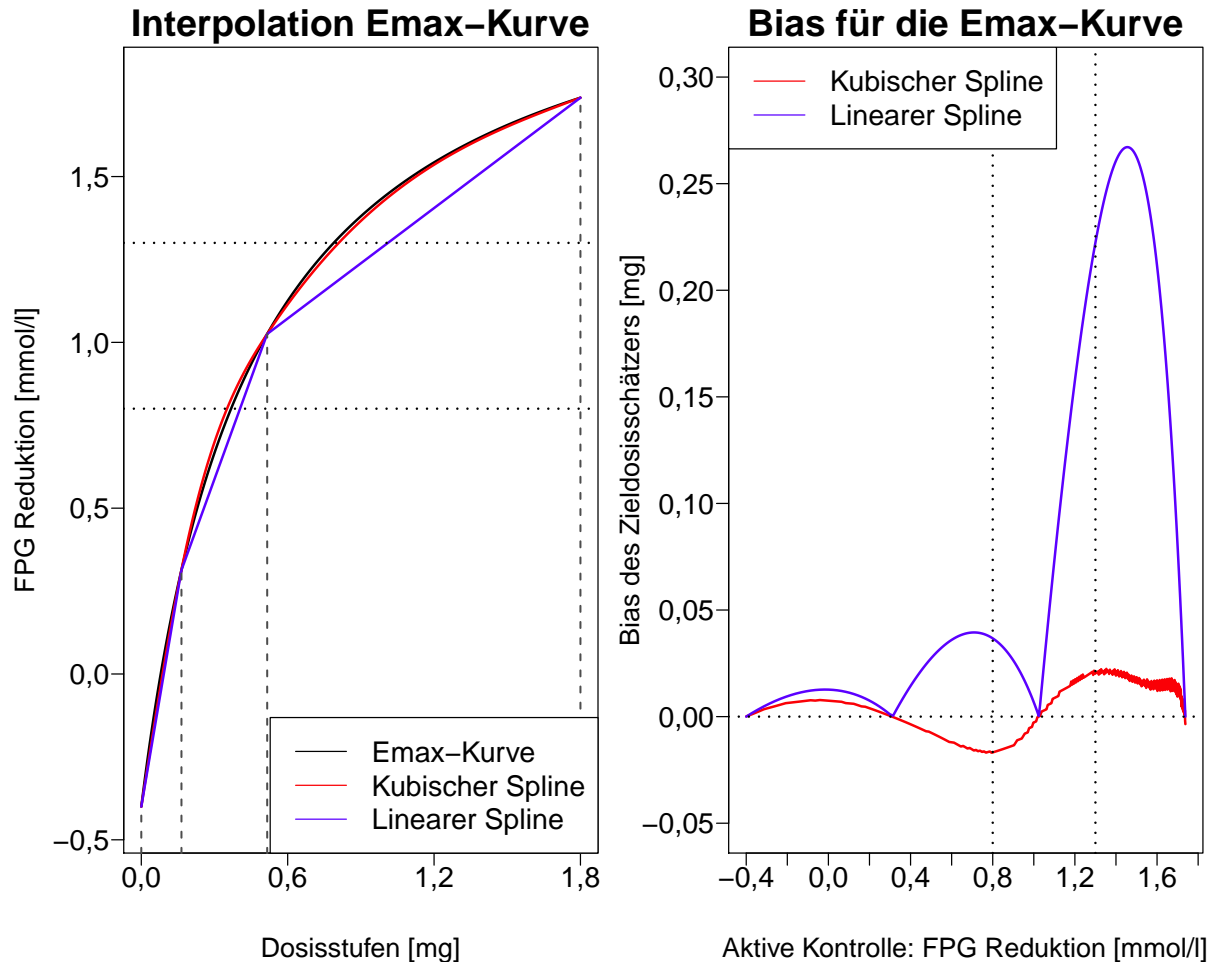


Abbildung 5.1.: Optimale Dosisauswahl ($k = 4$): Lineare und kubische Spline-Interpolation der Emax-Wirkungsfunktion aus Tabelle 4.1 mit entsprechender Verzerrung von \hat{d}^* für den linearen und den kubischen Spline im optimalen Design des kubischen Splines.

5.1.5. Verzerrungs-minimales Design: Simulationsstudie

Um den Einfluss des Bias reduzierten Designs auf die Überdeckungswahrscheinlichkeit zu untersuchen, wird eine Simulationsstudie basierend auf dem optimalen Emax-Design der kubischen Spline-Interpolation für die Simulationsszenarien in Tabelle 4.1 durchgeführt. Die Emax-Wirkungskurve wird für die neuen Dosisstufen $(d_1, \dots, d_4)' = (0; 0,165; 0,516; 1,8)'$ und die Fallzahlen $n = 20, 25, 30, 40, \dots, 100$ pro Gruppe mit $n_{sim} = 10.000$ und $n_{boot} = 5.000$ simuliert.

Abbildung 5.2 zeigt eine Verbesserung der Überdeckungswahrscheinlichkeit für $\mu = 0,8$ für die kubische Spline-Interpolation im Vergleich zu Abbildung 4.3. Fast alle Werte liegen innerhalb des 99%-Intervalls des Simulationsfehlers. Im Falle von $\mu = 1,3$ ist zu erkennen, dass die Überdeckungswahrscheinlichkeit des kubischen Splines leicht konservativ wird und für größere Fallzahlen bei 96% liegt. Das Bias-reduzierte Design führt somit nicht nur zu einer geringen Verzerrung der Zieldosis, sondern auch zu einer Verbesserung der Überdeckungswahrscheinlichkeit in den untersuchten Szenarien. Dies gilt allerdings nur für die kubische Spline-Interpolation. Die Überdeckungswahrscheinlichkeit der linearen Spline-Interpolation verbessert sich zwar für $\mu = 0,8$, wird aber für $\mu = 1,3$ mit größeren Fallzahlen n pro Gruppe liberaler (92%). Dieses Verhalten ist allerdings nicht verwunderlich und kann auf die Verzerrung des linearen Splines unter dem optimalen Design des kubischen Splines zurückgeführt werden (Abb. 5.1). Das optimale Design für die lineare Spline-Interpolation ergibt im beschriebenen Emax-Szenario eine äquidistante Allokation der Dosisstufen und ist somit gleich dem ursprünglichen Design, das bereits in den vorangegangenen Simulationsstudien (siehe Abbildungen 4.2 und 4.3) verwendet wurde. Daher ist es nicht möglich die maximale absolute Verzerrung des Zieldosischätzers für den linearen Spline durch die Verwendung einer anderen Dosisallokation zu verringern. Es ist weiterhin interessant zu sehen, dass die beiden parametrischen Regressionen in allen betrachteten Situationen stark liberale Überdeckungswahrscheinlichkeiten aufweisen, speziell für kleine und mittlere Fallzahlen ($n < 70$) pro Gruppe und somit ebenfalls sehr sensitiv auf die Wahl der Dosisstufen reagieren.

5.1.6. Zusammenfassung

Mit den vorgestellten Methoden ist es möglich, eine Dosisauswahl zu bestimmen, welche die Verzerrung des Zieldosischätzers kontrolliert und das unabhängig vom Wirkungsgrad der aktiven Kontrolle μ . Die Berechnung des optimalen Designs oder der Dosisallokation zur Planung einer Studie bedarf nur weniger Annahmen, da lediglich die Gestalt der Wirkungsfunktion $f_{\theta}(d)$ und einige zugehörige Parameter prespezifiziert werden müssen. Zum Beispiel werden zur Berechnung des vorgestellten optimalen Designs der Emax-Dosis-Wirkungsfunktion nur zusätzliche Informationen zur ED_{50} (θ_2) benötigt. Ein weiterer Baustein in der Planungsphase einer klinischen Studie stellt die Fallzahlberechnung dar. Der folgende Abschnitt beschäftigt sich mit Fallzahlberechnung von Dosis-Findungsstudien mit aktiver Kontrolle.

5.2. Fallzahlplanung

In diesem Abschnitt werden verschiedene Ansätze zur Kalkulation der Fallzahl in Dosis-Findungsstudien mit aktiver Kontrolle basierend auf den Ergebnissen der Maximum-Likelihood Methode aus Abschnitt 2.3 vorgestellt und entwickelt. Die grundlegende Idee der Fallzahlplanung beschäftigt sich mit der möglichst genauen Berechnung der Stichprobengröße einer Studie. Dabei muss ein gewähltes Gütekriterium erfüllt werden; zum Beispiel soll eine vorgegebene mittlere Differenz zwischen zwei Gruppen mit einer festen Wahrscheinlichkeit (Power) aufgedeckt werden. Aus

ethischen und wirtschaftlichen Gesichtspunkten ist es von großem Interesse nicht zu viele oder zu wenige Patienten in eine Studie einzuschließen. In Dosis-Findungsstudien mit aktiver Kontrolle liegt der Fokus auf der Schätzung der kleinsten Dosis d^* , welche zum gleichen Wirkungsgrad wie die aktive Kontrolle führt sowie der Berechnung eines zugehörigen $(1 - \alpha)$ -Konfidenzintervalls. Die Fallzahlplanung wird dazu verwendet, die Unsicherheit des Zieldosischätzers \hat{d}^* zu kontrollieren, indem die Länge des zugehörigen Konfidenzintervalls beschränkt wird. In der Literatur existiert bereits eine Vielzahl an Methoden zur Fallzahlkalkulation, die sich auf die Länge des Konfidenzintervalls beziehen.

Beal (1989) untersucht Fallzahlkalkulationen in Studiendesigns mit zwei Behandlungsgruppen und überträgt dabei das Konzept der Power eines statistischen Tests auf die Länge des Konfidenzintervalls der mittleren Gruppendifferenz. Die Power $\Phi(c)$ wird dann in Abhängigkeit der geforderten halben Intervallbreite c definiert, so dass das Konfidenzintervall mit der Wahrscheinlichkeit $\Phi(c)$ kleiner oder gleich $2c > 0$ ist. Allerdings schränkt Beal die Aussage der Power in der Form ein, dass eine ermittelte Fallzahl die Länge des KI mit der Wahrscheinlichkeit $\Phi(c)$ beschränkt, falls der wahre Parameter enthalten ist. Weiterhin berücksichtigt Beal (1989) nicht die Unsicherheit der Konfidenzintervallschätzung. Moore et al. (1989) hingegen untersucht Fallzahlplanungen, die die Länge eines Konfidenzintervalls in Erwartung durch eine gewählte Konstante beschränken ($E(KI_{\text{Länge}}) \leq 2c$) und kann zeigen, dass das daraus resultierende Intervall nur mit einer Wahrscheinlichkeit von ca. 50% kleiner oder gleich der gewählten Konstanten ist. Grieve und Beal (1991) stellten, basierend auf Beal (1989), einen Ansatz vor, der die Unsicherheit des Konfidenzintervalls berücksichtigt, indem anstelle der Varianz ein Quantil der Verteilung des Varianzschätzers verwendet wird. Die berechnete Fallzahl kontrolliert die Länge des Konfidenzintervalls, so dass die Breite des KI mit einer Wahrscheinlichkeit von mindestens γ kleiner als eine vorgegebene Konstante ist. Lehmann (1959) argumentiert gegen die Verwendung der erwarteten Konfidenzintervallbreite als wünschenswertes Kriterium um das Konfidenzintervall zu beschränken: „short intervals are desirable when they cover the true parameter value but not necessarily otherwise“.

Im Folgenden werden Fallzahlberechnungen für Dosis-Findungsstudien mit aktiver Kontrolle vorgestellt, die die Länge des $(1 - \alpha)$ -Konfidenzintervalls entweder in Erwartung ($E(KI_{\text{Länge}}) \leq 2 \cdot c$) oder in Wahrscheinlichkeit mit dem Kriterium $P(KI_{\text{Länge}} \leq 2 \cdot c) \leq \gamma \in]0,1[$ kontrollieren (Graybill, 1958). Dafür werden im folgenden Abschnitt die Resultate der Normalapproximation aus Abschnitt 2.4.3 zusammengefasst und auf die Problematik der Fallzahlplanung übertragen.

5.2.1. Normalapproximation mittels Δ -Regel

Es gelten die Annahmen aus Abschnitt 2.2 mit der Annahme gleicher Varianzen der Fehlerterme $\sigma_i = \sigma_{ak} = \sigma > 0$. Da es das Ziel ist, die Gesamtfallzahl N für eine vorgegebene Wirkungsfunktion $f_{\theta}(d)$ zu berechnen, sind anstelle der Fallzahlen n_i , $i = 1, \dots, k$, und n_{ak} die Gewichte w_i , $w_d = \sum_{i=1}^k w_i$ und w_{ak} ($w_d + w_{ak} = 1$) vorgegeben, so dass sich die Fallzahlen pro Gruppe aus der Gesamtfallzahl N berechnen lassen (z. B. $n_i = \lceil w_i \cdot N \rceil$). Die verschiedenen Ansätze zur Fallzahl-

berechnung werden für allgemeine Dosis-Wirkungsfunktionen über das $(1 - \alpha)$ -Konfidenzintervall der Normalapproximation mittels Δ -Regel hergeleitet, das bereits in Abschnitt 2.4.3 in Gl. (2.11) eingeführt wurde:

$$KI_{\Delta} = \left[\hat{d}^* - \frac{u_{1-\frac{\alpha}{2}}}{\sqrt{N}} \sqrt{\hat{\tau}^2}, \hat{d}^* + \frac{u_{1-\frac{\alpha}{2}}}{\sqrt{N}} \sqrt{\hat{\tau}^2} \right].$$

Definition 5.1 Die Länge des $(1 - \alpha)$ -Konfidenzintervalls kann mit Hilfe des Plug-in Schätzers $\hat{\tau}^2$ unter Verwendung von Gl. (2.11) als Differenz zwischen oberer und unterer Intervallgrenze definiert werden:

$$KI_{Länge} = 2 \frac{u_{1-\frac{\alpha}{2}}}{\sqrt{N}} \sqrt{\hat{\tau}^2}.$$

Basierend auf der Zufallsvariable der Länge des Konfidenzintervalls werden in den folgenden Abschnitten die verschiedenen Formeln zur Fallzahlberechnung hergeleitet.

5.2.2. Fallzahl zur Kontrolle der erwarteten Länge des Konfidenzintervalls

In einem ersten Schritt soll eine Fallzahlformel bestimmt werden, die die Länge des Konfidenzintervalls (Def. 5.1) in Erwartung beschränkt. Die dazu benötigte Fallzahl N_E kann durch Lösen der folgenden Gleichung berechnet werden:

$$E(KI_{Länge}) \approx 2 \frac{u_{1-\frac{\alpha}{2}}}{\sqrt{N}} \sqrt{\tau^2} \leq 2c.$$

Die benötigte Fallzahl lässt sich durch:

$$N_E \gtrsim \tau^2 \left(\frac{u_{1-\frac{\alpha}{2}}}{c} \right)^2 \tag{5.4}$$

mit $\tau^2 = \sigma^2 \Delta g(\boldsymbol{\theta}_{d^*}) \left(\tilde{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}} \oplus \frac{1}{w_{ak}} \right) \Delta g(\boldsymbol{\theta}_{d^*})'$ (siehe Gl. 2.10) beschreiben. Um diese Fallzahl berechnen zu können, müssen neben den gemachten Modellannahmen auch die Wirkungsfunktion $f_{\boldsymbol{\theta}}(d)$ sowie der Parametervektor der Zieldosis $\boldsymbol{\theta}_{d^*} = (\boldsymbol{\theta}', \mu)'$ und die Varianz der Fehlerterme σ^2 vorgegeben oder auf Grundlage vorangegangener Studien bestimmt werden. Um sicherzustellen, dass die berechnete Fallzahl ganzzahlig ist, wird für N_E die kleinste ganze Zahl gewählt, die die Ungleichung (5.4) erfüllt. Dies gilt analog für die anschließende Fallzahlallokation von N_E auf die Fallzahlen pro Dosisstufe n_i und die aktive Kontrolle n_{ak} . Im Folgenden wird die Fallzahl N_E exemplarisch für eine lineare Wirkungsfunktion der Form $f_{\boldsymbol{\theta}}(d) = \theta_0 + \theta_1 d$ bestimmt. Dazu können die Resultate aus Abschnitt 3.2.1 in Gl. (5.4) eingesetzt werden. Dann ergibt sich für die Fallzahl:

$$N_E \gtrsim \frac{\sigma^2}{\theta_1^2} \left[\frac{\left(\sum_{i=1}^k w_i [d_i - d^*]^2 \right)}{\left(\left(w_d \cdot \sum_{i=1}^k w_i d_i^2 \right) - \left(\sum_{i=1}^k w_i d_i \right)^2 \right)} + \frac{1}{w_{ak}} \right] \left(\frac{u_{1-\frac{\alpha}{2}}}{c} \right)^2 \quad (5.5)$$

mit den bekannten Gewichten der Fallzahlen pro Gruppe w_i , w_d und der aktiven Kontrolle w_{ak} (siehe auch Helms et al., 2012). Im folgenden Abschnitt wird die Fallzahlformel zur Kontrolle der Länge des Konfidenzintervalls mit Wahrscheinlichkeit γ hergeleitet.

5.2.3. Fallzahlformel zur Kontrolle der Länge des Konfidenzintervalls mit Wahrscheinlichkeit γ

Wie zu Beginn des Abschnitts beschrieben, wäre es wünschenswert, neben der Fallzahl N_E eine Methode zu finden, die die Kontrolle der Intervalllänge nicht in Erwartung, sondern mit einer vorgegebenen Wahrscheinlichkeit ermöglicht. Sei dafür N_γ die Fallzahl, die zu einer Konfidenzintervallbreite kleiner $2c > 0$ mit einer Wahrscheinlichkeit von mindestens $\gamma \in]0,1[$ führt (Graybill, 1958). Diese Wahrscheinlichkeit kann wie folgt dargestellt werden:

$$P(KI_{\text{Länge}} \leq 2c) = P\left(N \geq \hat{\tau}^2 \left(\frac{u_{1-\frac{\alpha}{2}}}{c}\right)^2\right) \geq \gamma, \quad (5.6)$$

wobei $\hat{\tau}^2$ (Kor. 2.6) den konsistenten Plug-in Schätzer der Varianz τ^2 beschreibt. Da τ^2 im Allgemeinen nicht in der Planungsphase einer Studie bekannt ist, wird anstelle des Erwartungswertes des Varianzschätzers das γ -Quantil der Verteilung von $\hat{\tau}^2$ zur Berechnung der Fallzahl N_γ verwendet. Dann gilt für die Fallzahl:

$$N_\gamma \geq \psi(\tau^2, \gamma) \cdot \left(\frac{u_{1-\frac{\alpha}{2}}}{c}\right)^2, \quad (5.7)$$

wobei $\psi(\tau^2, \gamma)$ das γ -Quantil der Verteilungsfunktion von $\hat{\tau}^2$ mit Erwartungswert τ^2 beschreibt. N_γ ist das γ -Quantil der Verteilung der Fallzahl, dessen Erwartungswert gerade durch die Fallzahl N_E definiert wird. Für $\gamma > 0,5$ ist $N_\gamma > N_E$. Die Verteilung von $\hat{\tau}^2$ ist im Allgemeinen unbekannt, da die benötigten Parameterschätzer häufig nur numerisch bestimmt werden können. Die Fallzahl N_γ kann in diesen Fällen nicht analytisch angegeben werden. Allerdings ist es möglich N_γ mittels Simulationen und numerischer Optimierung zu bestimmen, solange die Annahmen aus Abschnitt 2.2 erfüllt und die in Abschnitt 5.2.2 beschriebenen Informationen bezüglich der Dosis-Wirkungsfunktion $f_\theta(d)$ bekannt sind. Im Folgenden wird ein Algorithmus zur Berechnung der Fallzahl N_γ vorgestellt, der die Gleichung (5.6) erfüllt.

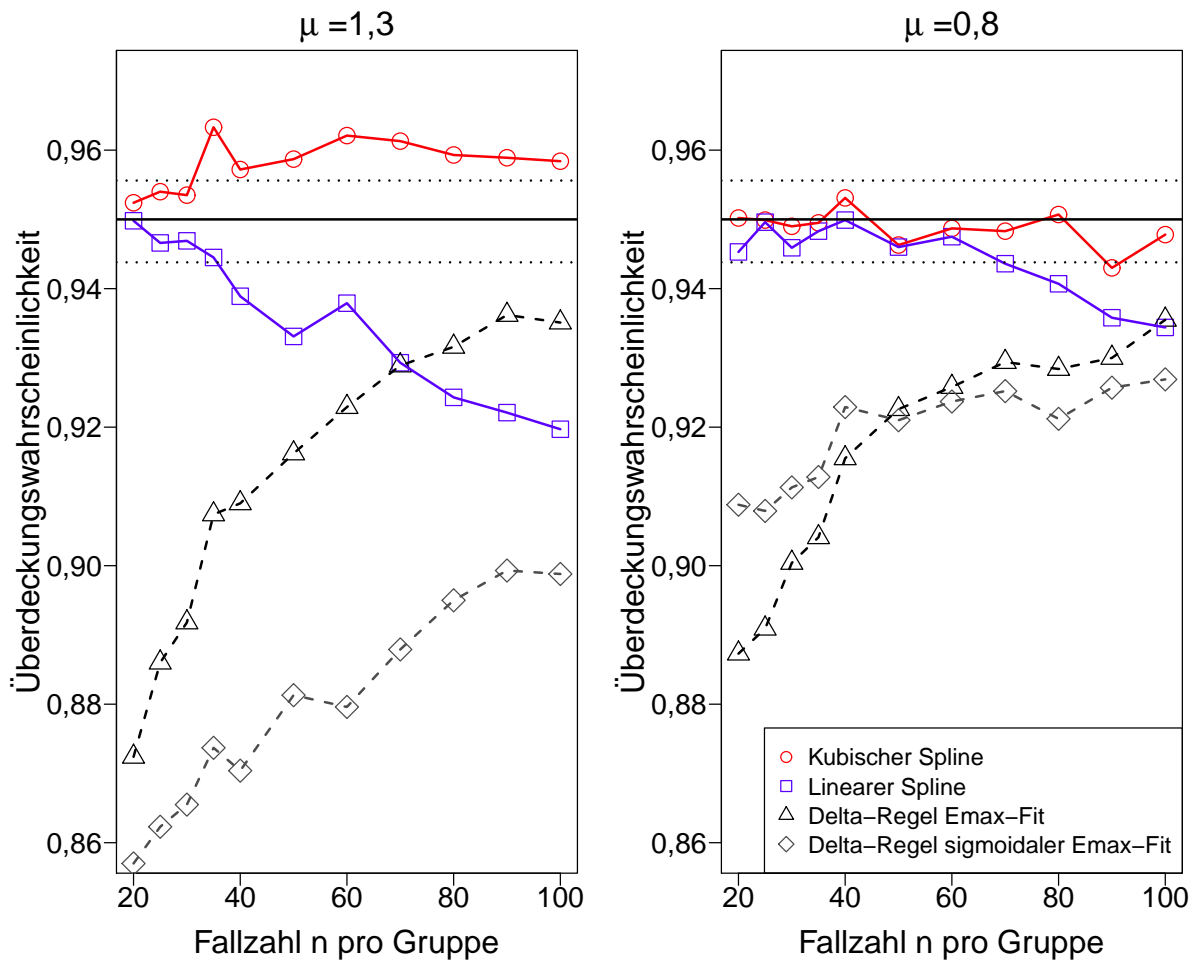


Abbildung 5.2.: Optimale Dosisauswahl: Überdeckungswahrscheinlichkeit des kubischen und linearen Splines, der Δ -Regel Emax-Regression (Δ -Regel Emax-Fit) und der Δ -Regel sigmoidalen Emax-Regression (Δ -Regel sigmoidaler Emax-Fit) aufgetragen gegen die Fallzahlen n pro Gruppe für lineare, Emax- und sigmoidale Emax-Wirkungskurven sowie verschiedene μ . Die gepunkteten Linien bezeichnen das 99%-Intervall des Simulationsfehlers. Für alle Szenarien betrug die Anzahl der Simulationsdurchläufe $n_{sim} = 10.000$ mit jeweils $n_{boot} = 5.000$ Bootstraps pro Simulation.

Algorithmus zur Berechnung der Fallzahl N_γ

Input: Fallzahl-Startwert N sowie n_{sim} , f_θ , θ , μ , d_1, \dots, d_k , γ , die Gewichte w_{ak}, w_i , $i = 1, \dots, k$, und die halbe Intervallbreite c .

Falls $\gamma > 0,5$ und $N < N_E \rightarrow N = N_E$.

Sei $s = 0$ und $t = 0$

1. Berechnung der Fallzahlen pro Gruppe: $n_i = \lceil w_i N \rceil, i = 1, \dots, k$, und $n_{ak} = \lceil w_{ak} N \rceil$ sowie $N = \sum_{i=1}^k n_i + n_{ak}$.
2. Simulationsschleife mit n_{sim} Durchläufen:
 - Generierung der N Realisationen \mathbf{Y} .
 - Berechnung von \hat{d}^* , KI_Δ und $KI_{L\ddot{a}nge}$.
 - Falls $KI_{L\ddot{a}nge} \leq 2c \rightarrow s = s + 1$.
3. Berechnung des Schätzers für $P(KI_{L\ddot{a}nge} \leq 2c)$: $p_\gamma = s/n_{sim}$.
4. Abbruchbedingungen des Algorithmus:
 - Falls $t = 1$ und $p_\gamma \geq \gamma \rightarrow$ Abbruch des Algorithmus und $N_\gamma = N$.
 - Falls $t = 2$ und $p_\gamma < \gamma \rightarrow$ Abbruch des Algorithmus und $N_\gamma = N_{alt}$.
5. Iterationsschritt
 - Falls $p_\gamma < \gamma \rightarrow t = 1, N_{alt} = N$ und $N = N + 1$.
 - Falls $p_\gamma \geq \gamma \rightarrow t = 2, N_{alt} = N$ und $N = N - 1$.
6. Wiederhole die Schritte 1 bis 5 bis 4. erfüllt ist.

Output: Fallzahl N_γ .

Da im Algorithmus die Genauigkeit des Schätzers p_γ (Zielfunktion) der Wahrscheinlichkeit $P(KI_{L\ddot{a}nge} \leq 2c)$ maßgeblich von der Anzahl der Simulationen n_{sim} abhängt und selbst für $n_{sim} = 100.000$ Schwankungen unterworfen ist, kann es bei Anwendung des Algorithmus zur Intervallhalbierung aus Abschnitt 3.2.3 oder des Newton-Raphson Verfahrens zu Konvergenzproblemen kommen. Die simulationsbedingte Schwankung des Schätzers macht es unter Umständen nötig, in der Suche einen Schritt zurück zu gehen, was in der Intervallhalbierung (Abs. 3.2.3) nicht möglich ist. Der hier vorgestellte Algorithmus ist sehr simpel und findet die gesuchte Fallzahl, jedoch hängt die Genauigkeit der Fallzahlbestimmung von der verwendeten Anzahl an Simulationen n_{sim} und die Anzahl der Iterationen von der Wahl eines geeigneten Startwertes N ab. Zur Bestimmung geeigneter Startwerte siehe Abschnitt 5.2.5.

Für den Fall einer linearen Dosis-Wirkungsfunktion wird im Folgenden eine Approximation für die Fallzahl N_γ hergeleitet. Mit den Ergebnissen aus Abschnitt 3.2.1 ergibt sich der konsistente

Plug-in Schätzer $\hat{\tau}^2$ (Gl. 3.12) für lineare Wirkungsfunktionen zu:

$$\hat{\tau}^2 = \frac{\hat{\sigma}^2}{\hat{\theta}_1^2} \left[\frac{\left(\sum_{i=1}^k w_i [d_i - \hat{d}^*]^2 \right)}{\left(\left(w_d \cdot \sum_{i=1}^k w_i d_i^2 \right) - \left(\sum_{i=1}^k w_i d_i \right)^2 \right)} + \frac{1}{w_{ak}} \right] \left(\frac{u_{1-\frac{\alpha}{2}}}{c} \right)^2.$$

Der Schätzer $\hat{\tau}^2$ verwendet die Schätzer $\hat{\sigma}^2$, $\hat{\theta}_1$ und \hat{d}^* , welche bereits in Kapitel 3 definiert wurden. Da die Zufallsvariablen $\hat{\theta}_1$ und \hat{d}^* per Konstruktion voneinander abhängen, kann selbst im Fall einer linearen Wirkungskurve die Verteilung von $\hat{\tau}^2$ nicht analytisch angegeben werden. Um dieses Problem zu umgehen, wird auf einen Teil der Unsicherheit verzichtet und die daraus resultierende approximative Verteilung verwendet, um die gewünschte Fallzahl zu berechnen. Dann vereinfacht sich die Fallzahlformel zu

$$\begin{aligned} N_\gamma &\geq \psi(\tau^2, \gamma) \left(\frac{u_{1-\frac{\alpha}{2}}}{c} \right)^2 \\ &\approx \varphi\left(\frac{\sigma^2}{\theta_1^2}, \gamma\right) \left[\frac{\left(\sum_{i=1}^k w_i [d_i - d^*]^2 \right)}{\left(\left(w_d \cdot \sum_{i=1}^k w_i d_i^2 \right) - \left(\sum_{i=1}^k w_i d_i \right)^2 \right)} + \frac{1}{w_{ak}} \right] \left(\frac{u_{1-\frac{\alpha}{2}}}{c} \right)^2, \end{aligned} \quad (5.8)$$

wobei $\varphi(\sigma^2/\theta_1^2, \gamma)$ das γ -Quantil der Verteilung des Quotienten der Zufallsvariablen $Q = \hat{\sigma}^2/\hat{\theta}_1^2$ mit Erwartungswert σ^2/θ_1^2 widerspiegelt. Anstatt die Verteilung des Varianzschätzers $\hat{\tau}^2$ zu bestimmen, wird der Zieldosischätzer \hat{d}^* durch seinen Erwartungswert d^* ersetzt und nur die Verteilung des Quotienten Q berechnet. Die daraus resultierende Approximation der Fallzahl wird im folgenden Satz zusammengefasst.

Satz 5.1 *Seien die Modellannahmen aus Abschnitt 2.2 sowie Abschnitt 5.2 erfüllt und sei $f_\theta(d) = \theta_0 + \theta_1 \cdot d$ eine lineare Dosis-Wirkungsfunktion. Weiter sei N_E die in Gl. (5.5) definierte Fallzahl der erwarteten KI-Breite und $\gamma \in]0, 1[$, dann folgt mit Hilfe der Approximation aus Gl. (5.8):*

$$N_\gamma \gtrsim \frac{1}{\Sigma_{\theta_1} \cdot F(1, N_E - 3, \lambda, 1 - \gamma)} \cdot \left[\frac{\left(\sum_{i=1}^k w_i [d_i - d^*]^2 \right)}{\left(\left(w_d \cdot \sum_{i=1}^k w_i d_i^2 \right) - \left(\sum_{i=1}^k w_i d_i \right)^2 \right)} + \frac{1}{w_{ak}} \right] \left(\frac{u_{1-\frac{\alpha}{2}}}{c} \right)^2, \quad (5.9)$$

wobei $F(1, N_E - 3, \lambda, 1 - \gamma)$ das $(1 - \gamma)$ -Quantil der F -Verteilung mit Freiheitsgraden 1 und $N_E - 3$ sowie Nichtzentralitätsparameter $\lambda = \theta_1^2/(\sigma^2 \Sigma_{\theta_1})$ beschreibt. Die Varianz Σ_{θ_1} von $\hat{\theta}_1$ ist durch $\Sigma_{\theta_1} = \mathbf{H}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{H}'$ mit $\mathbf{H} = (0, 1)$ gegeben.

BEWEIS. Aus Gl. (5.8) folgt zunächst, dass nur die Verteilung des Quotienten $Q = \hat{\sigma}^2/\hat{\theta}_1^2$ bestimmt werden muss. Um die Berechnungen zu vereinfachen, wird zuerst die Verteilung der

monotonen Transformation $1/Q$ berechnet. Dafür werden als erstes die Verteilungen des Varianzschätzers $\hat{\sigma}^2$ und des quadrierten Schätzers der Steigung $\hat{\theta}_1^2$ benötigt. In Abschnitt 3.1.3 wurde für die lineare Wirkungskurve bereits die Verteilung des Varianzschätzers der Fehlerterme in Gl. (3.6) durch:

$$\hat{\sigma}^2 = \frac{1}{N-3} \mathbf{Y}' \left(\mathbf{I}_N - \mathbf{X}_t (\mathbf{X}'_t \mathbf{X}_t)^{-1} \mathbf{X}'_t \right) \mathbf{Y} \sim \frac{\sigma^2}{N-3} \cdot \chi_{N-3}^2 \quad (5.10)$$

mit entsprechender Designmatrix $\mathbf{X}_t = (\mathbf{X} \oplus \mathbf{1}_{n_{ak}})$ angegeben. Weiterhin wurde im selben Abschnitt die Verteilung des Vektors der Parameterschätzer mit $\hat{\boldsymbol{\theta}} \sim \mathcal{N}(\boldsymbol{\theta}, \sigma^2 (\mathbf{X}'\mathbf{X})^{-1})$ hergeleitet. Daraus folgt für die Verteilung von $\hat{\theta}_1$:

$$\hat{\theta}_1 = \mathbf{H}\hat{\boldsymbol{\theta}} \sim \mathcal{N}(\theta_1, \sigma^2 \Sigma_{\theta_1})$$

mit $\mathbf{H} = (0, 1)$ und $\Sigma_{\theta_1} = \mathbf{H} (\mathbf{X}'\mathbf{X})^{-1} \mathbf{H}'$. Weiterhin folgt für das normalisierte Quadrat der normalverteilten Zufallsvariable des Schätzers (Brunner, 2005, S. 49-52):

$$\frac{\hat{\theta}_1^2}{\text{Var}(\hat{\theta}_1)} = \frac{\hat{\boldsymbol{\theta}}' \mathbf{H}' \mathbf{H} \hat{\boldsymbol{\theta}}}{\sigma^2 \Sigma_{\theta_1}} = \frac{1}{\sigma^2} \hat{\boldsymbol{\theta}}' \mathbf{H}' \left(\mathbf{H} (\mathbf{X}'\mathbf{X})^{-1} \mathbf{H}' \right)^{-1} \mathbf{H} \hat{\boldsymbol{\theta}} \sim \chi_1^2(\lambda)$$

mit $\lambda = \theta_1^2 / (\sigma^2 \Sigma_{\theta_1})$. Folglich sind Zähler und Nenner des Quotienten Q χ^2 -verteilt mit Freiheitsgraden $\nu = 1, N-3$ und nach Satz A.5 stochastisch unabhängig. Speziell ist der Parameterschätzer $\hat{\boldsymbol{\theta}}$ unabhängig vom Varianzschätzer $\hat{\sigma}^2$. Daraus folgt für den Quotienten von zwei unabhängig χ^2 -verteilten Zufallsvariablen:

$$\frac{\frac{\hat{\theta}_1^2}{\sigma^2}}{\Sigma_{\theta_1} \cdot \frac{\hat{\sigma}^2}{\sigma^2}} \sim \frac{\frac{\chi_1^2(\lambda)}{1}}{\frac{\chi_{N-3}^2}{N-3}} \sim F(1, N-3, \lambda). \quad (5.11)$$

Mit Hilfe der Verteilung von $1/Q$ ist es möglich das γ -Quantil der Verteilung von Q durch das $(1-\gamma)$ -Quantil der F-Verteilung $F(1, N-3, \lambda, 1-\gamma)$ aus Gl. (5.11) zu berechnen:

$$\text{P} \left(\frac{\hat{\theta}_1^2}{\Sigma_{\theta_1} \hat{\sigma}^2} \leq F(1, N-3, \lambda, 1-\gamma) \right) = 1-\gamma \iff \text{P} \left(\frac{\hat{\sigma}^2}{\hat{\theta}_1^2} \leq \frac{1}{\Sigma_{\theta_1} F(1, N-3, \lambda, 1-\gamma)} \right) = \gamma. \quad (5.12)$$

Allerdings tritt dabei nun ein weiteres Problem auf, denn das benötigte Quantil der F-Verteilung $F(1, N-3, \lambda, 1-\gamma)$ sowie die Varianz Σ_{θ_1} hängen von der unbekanntem Gesamtfallzahl N ab, die berechnet werden soll. Dieses Problem kann zu Beispiel durch numerische Optimierung von

Gl. (5.8) mit $\varphi(\sigma^2/\theta_1^2, \gamma) = 1/\Sigma_{\theta_1} F(1, N - 3, \lambda, 1 - \gamma)$ gelöst werden. Um die Notwendigkeit von numerischen Verfahren zu umgehen, wird im Folgenden die unbekannte Fallzahl N im Ausdruck $1/\Sigma_{\theta_1} F(1, N - 3, \lambda, 1 - \gamma)$ durch die in Gl. (5.5) bestimmte Fallzahl N_E ersetzt, um die Fallzahl N_γ , mit $F(1, N_E - 3, \lambda, 1 - \gamma)$ zu bestimmen. Damit ergibt sich für die Approximation der Fallzahl:

$$N_\gamma \gtrsim \frac{1}{\Sigma_{\theta_1} \cdot F(1, N_E - 3, \lambda, 1 - \gamma)} \cdot \left[\frac{\left(\sum_{i=1}^k w_i [d_i - d^*]^2 \right)}{\left(\left(w_d \cdot \sum_{i=1}^k w_i d_i^2 \right) - \left(\sum_{i=1}^k w_i d_i \right)^2 \right)} + \frac{1}{w_{ak}} \right] \left(\frac{u_{1-\frac{\alpha}{2}}}{c} \right)^2,$$

was zu zeigen war. □

Das Ersetzen der unbekanntes Fallzahl durch eine leichter zu berechnende Variante im benötigten Quantil kommt in Fallzahlkalkulationen häufiger vor. Zum Beispiel wird die Fallzahl im Zwei-Stichprobenfall basierend auf einem Quantil der Normalverteilung als Annäherung oder Startwert berechnet, um die Fallzahl basierend auf dem Quantil der t-Verteilung zu bestimmen (Julious und Owen, 2006). Nachdem für beide Kriterien Formeln zur Fallzahlberechnung hergeleitet wurden, werden diese im folgenden Abschnitt mittels Simulationsstudien untersucht.

5.2.4. Simulationen

In diesem Abschnitt wird die Qualität der präsentierten Ansätze zur Fallzahlberechnung durch umfangreiche Simulationsstudien untersucht. Das Planungsdesign der Dosis-Findungsstudie ist analog zu Abschnitt 3.3 motiviert durch die Studie von Chapple et al. (2004) (Abs. 1.2.1). Es wird daher eine sechsarmige Dosis-Findungsstudie ($k=5$) mit den Dosisstufen $(d_1, \dots, d_5)' = (0; 0,25; 0,5; 0,75; 1)'$ und einem aktiven Komparator untersucht. Der erwartete Wirkungsgrad der aktiven Kontrolle ist $\mu = 1$. Die Dosisstufen wurden auf dem standardisiertem Intervall $[0, 1]$ gewählt, um die Vergleichbarkeit zu erhöhen. Dies ist keine Einschränkung der Allgemeinheit, da sich jeder Dosisbereich $D = [d_1, d_k]$ dividiert durch die maximale Dosisstufe $d_k > 0$ auf das standardisierte Intervall $[\tilde{d}_1, 1]$ überführen lässt. Zur Vereinfachung werden zudem die Fallzahlen pro Dosisstufe auf $n_i = n, i = 1, \dots, k$, und die Fallzahl der aktiven Kontrolle auf $n_{ak} = r \cdot n$, mit $r = 2$ gesetzt. Dies führt zu identischen Dosisgewichten $w_i = 1/(k+r) = 1/7$ und zu $w_{ak} = r/(k+r) = 2/7$ für die aktive Kontrolle, so dass die Gesamtfallzahl als $N = (k+r)n$ beschrieben werden kann. Zwei verschiedene Dosis-Wirkungskurven werden in den Simulationsstudien verwendet: Die (log)-lineare Dosis-Wirkungsfunktion (Kap. 3), um die entsprechenden Fallzahlformeln untersuchen zu können und die Emax-Dosis-Wirkungsfunktion (Abs. 4.5) als ein Beispiel für eine nichtlineare Dosis-Wirkungskurve. Weiterhin werden alle Fallzahlen für das nominale 95%-Konfidenzintervall (KI) ($\alpha = 0,05$) und die halben Konfidenzintervallbreiten $c \in \{0,1; 0,15; 0,2\}$ berechnet. Zusätzlich variiert das Quantil der Fallzahl N_γ mit

$\gamma \in \{0,7; 0,8; 0,9\}$. Auf dem gewählten Dosisbereich $D = [0, 1]$ beschreibt $2c$ die relative Breite des Konfidenzintervalls bezogen auf die Länge des Dosisbereichs, da $d_k - d_1 = 1$. Es werden die Fallzahlen N_E (Gl. 5.5) und N_γ (Gl. 5.9) berechnet und zusätzlich wird die exakt benötigte Fallzahl $N_{\gamma, ex}$ für das γ -Quantil mittels numerischer Optimierung in den Simulationsszenarien bestimmt (siehe Algorithmus in Abs. 5.2.3). Dafür wird die Fallzahl ausgehend von N_E oder einem entsprechenden Startwert so lange verändert, bis die geschätzte Konfidenzintervallbreite in mindestens $\gamma\%$ der n_{sim} Simulationsdurchläufe kleiner als $2c$ ist. Die Anzahl der Simulationsdurchläufe ist im Falle von Szenarien mit linearer Wirkungskurve durch $n_{sim} = 100.000$ und für Emax-Wirkungskurven durch $n_{sim} = 10.000$ festgelegt. Alle Szenarien sind in Tabelle 5.1 zusammengefasst. Die Generierung der Zufallszahlen erfolgte nach der Fallzahlberechnung analog zu den vorangegangenen Simulationskapiteln (Abs. 3.3 und Abs. 4.5).

Tabelle 5.1.: Simulationsszenarien für die lineare und die Emax-Wirkungskurve mit $\theta_0 = 0$, residualer Varianz $\sigma^2 = 1$, erwarteter Wirkung der aktiven Kontrolle $\mu = 1$, halben Intervallbreiten $c \in \{0,1; 0,15; 0,2\}$ und Quantilen $\gamma \in \{0,7; 0,8; 0,9\}$.

Dosis-Wirkungskurve	θ_1	θ_2	d^*
Linear	1,5	-	0,667
	2	-	0,5
	2,5	-	0,4
Emax		0,2	0,2
	2	0,4	0,4
		0,6	0,6
		0,2	0,133
	2,5	0,4	0,26
		0,6	0,4

Ergebnisse der Fallzahlberechnungen für lineare Dosis-Wirkungsfunktionen

Die Ergebnisse der Fallzahlberechnungen für lineare Dosis-Wirkungsfunktionen sind exemplarisch für $\theta_1 = 2$ und $\theta_1 = 2,5$ in Tabelle 5.2 dargestellt. Es wird die benötigte Fallzahl pro Gruppe n , die Überdeckungswahrscheinlichkeit des nominalen 95%-KI, die mittlere Konfidenzintervallbreite und die Wahrscheinlichkeit $P(KI_{Länge} \leq 2c)$ für N_E , N_γ und die exakte Fallzahl $N_{\gamma, ex}$ angegeben. Weiterhin wird der Quotient $N_\gamma/N_{\gamma, ex}$ verwendet, um die Genauigkeit der approximierten Fallzahl N_γ zu bestimmen. Aus Tabelle 5.2 ist ersichtlich, dass für die Fallzahl N_E über alle Szenarien hinweg die mittlere Konfidenzintervallbreite von $2c$ sehr gut eingehalten wird. Das entsprechende Konfidenzintervall ist aber nur mit einer Wahrscheinlichkeit von ca. 50% kleiner als die geforderte Breite von $2c$. Diese Eigenschaft ist nicht überraschend, da sie bereits von Moore et al. (1989) im Rahmen anderer Fallzahlplanungssituationen beobachtet wurde. Weiterhin ist ersichtlich, dass für die approximative Fallzahl N_γ die geforderte Bedingung $P(KI_{Länge} \leq 2c) \geq \gamma$ über alle untersuchten Situationen hinweg gut eingehalten wird und die Approximation sehr dicht an der exakten Fallzahl $N_{\gamma, ex}$ liegt. Der Mittelwert des Quotienten

$N_\gamma/N_{\gamma,ex}$ beträgt über alle Szenarien 1. Das Minimum liegt bei 0,92 und das Maximum bei 1,08. Somit ergibt sich eine maximale Abweichung der approximierten Fallzahl von der wirklich benötigten Fallzahl von $\pm 8\%$. Die Überdeckungswahrscheinlichkeit ist leicht konservativ und sehr dicht am gewünschten nominalen 95%-Level für alle untersuchten Szenarien und Fallzahlmethoden. Um die Korrelation zwischen den Fallzahlen N_E und $N_{\gamma,ex}$ zu untersuchen, werden die entsprechenden Fallzahlen aller 27 Simulationsszenarien in Abbildung 5.3 gegeneinander aufgetragen. Es ist ein starker linearer Zusammenhang zu erkennen, der vom gewählten γ -Quantil abhängt.

Ergebnisse der Fallzahlberechnungen für Emax-Dosis-Wirkungskurven

Da im Falle der Emax-Dosis-Wirkungskurve als nichtlineare Wirkungskurve keine explizite Approximation bestimmt werden konnte, werden im Folgenden nur die Fallzahl pro Gruppe n , die Überdeckungswahrscheinlichkeit des 95%-KI, die mittlere Konfidenzintervallbreite ($KI_{Länge}$), die Wahrscheinlichkeit $P(KI_{Länge} \leq 2c)$ für die Fallzahlformel N_E (Gl. 5.4) und die exakte Fallzahl $N_{\gamma,ex}$ (Gl. 5.7) angegeben. In Tabelle 5.3 sind die Simulationsergebnisse exemplarisch für $\theta_1 = 2$ und $\theta_2 = 0,2$ sowie $\theta_2 = 0,4$ dargestellt. Es ist ersichtlich, dass auch im Falle einer Emax-Wirkungskurve die Fallzahl N_E zu einer mittleren Intervallbreite von $2c$ führt und das zugehörige Konfidenzintervall mit einer Wahrscheinlichkeit von ca. 50% kleiner als die geforderte Breite von $2c$ ist. Die numerisch bestimmte exakte Fallzahl $N_{\gamma,ex}$ führt in allen untersuchten Szenarien zu einer Konfidenzintervallbreite, die die Bedingung $P(KI_{Länge} \leq 2c) \geq \gamma$ erfüllt. Weiterhin ist ersichtlich, dass mit Ausnahme der Szenarien mit einer Intervallbreite von $2c = 0,4$, die simulierten Überdeckungswahrscheinlichkeiten in fast allen untersuchten Situationen das nominale 95%-Niveau sehr gut einhalten. Analog zur linearen Wirkungskurve wird der Zusammenhang zwischen N_E und $N_{\gamma,ex}$ bezüglich der 54 verwendeten Simulationsszenarien untersucht und in Abbildung 5.3 dargestellt. Der Zusammenhang der beiden Fallzahlformeln ist für die untersuchten Emax-Wirkungskurven ebenfalls linear und von der Wahl des γ -Quantils abhängig. Motiviert durch diese Ergebnisse werden im nachfolgenden Abschnitt Verfahren zur Approximation der Fallzahl N_γ vorgestellt.

5.2.5. Empirische Approximationsansätze für N_γ im Falle nichtlinearer Wirkungskurven

Wie bereits in Abschnitt 5.2.3 beschrieben, ist die analytische Berechnung oder Approximation der Fallzahl N_γ im Falle einer nichtlinearen Dosis-Wirkungsfunktion kompliziert. Um zur Bestimmung von N_γ nicht ausschließlich auf Simulationen und numerische Optimierung angewiesen zu sein, werden drei empirische Ansätze am Beispiel der Emax-Funktion als nichtlineare Wirkungskurve vorgestellt. Die Emax-Dosis-Wirkungskurve wurde bereits in Abschnitt 4.5 als

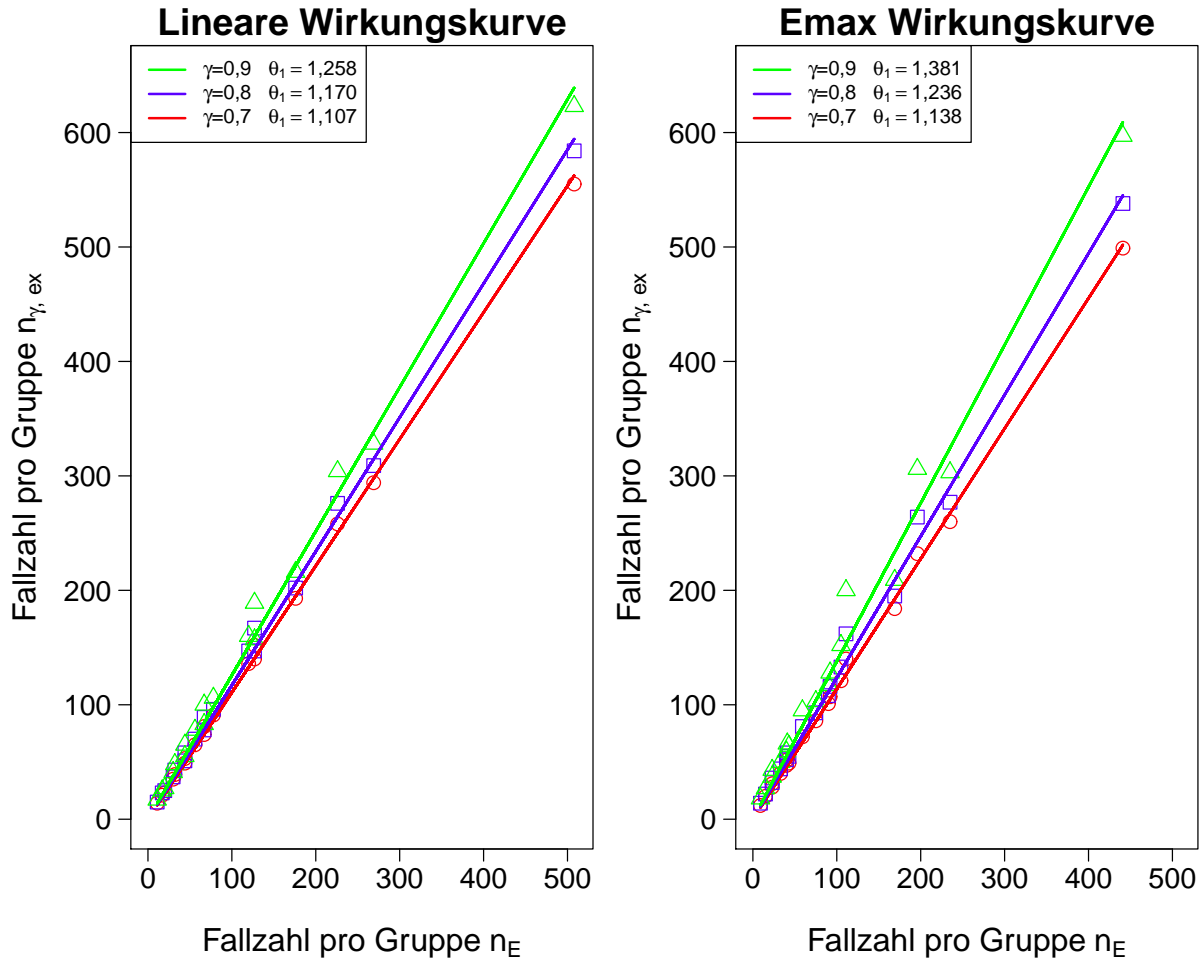


Abbildung 5.3.: Lineare Regression ohne Achsenabschnitt zwischen der Fallzahl $n_E = N_E/(k+r)$ pro Gruppe und der exakten benötigten Fallzahl zum γ -Quantil $n_{\gamma, ex} = N_{\gamma, ex}/(k+r)$ pro Gruppe für verschiedene γ -Quantile im Falle linearer und Emax-Dosis-Wirkungsfunktionen basierend auf den Szenarien aus Tabelle 5.1.

$$f_{\theta}(d) = \theta_0 + \frac{\theta_1 d}{(d + \theta_2)}$$

mit Achsenabschnitt θ_0 , Steigung θ_1 und der ED_{50} θ_2 , definiert. Wie in Abschnitt 5.2.2 gezeigt, lässt sich die Fallzahl N_E der erwarteten Konfidenzintervallbreite auch für nichtlineare Funktionen angeben und unterscheidet sich von der Fallzahl N_{γ} lediglich in der Behandlung des Varianzterms $\hat{\tau}^2$ (Def. 5.1). Daher wird in Abbildung 5.3 das Verhältnis von N_E zur exakt benötigten Fallzahl $N_{\gamma, ex}$ für die lineare und Emax-Dosis-Wirkungskurve dargestellt, indem die entsprechenden Fallzahlen für alle untersuchten Simulationsszenarien aus Tabelle 5.1 zusammengefasst werden. Die Idee ist es, einen Multiplikator $a(\gamma)$ in Abhängigkeit des Quantils γ zu finden, der in Kombination mit der Fallzahl N_E des betrachteten Szenarios die Fallzahl N_{γ} durch $N_{\gamma} \approx a(\gamma) \cdot N_E$ annähert. Aus Abbildung 5.3 ist ersichtlich, dass der Zusammenhang zwischen den Fallzahlen N_E

und $N_{\gamma, ex}$ für die betrachteten linearen und Emax-Dosis-Wirkungsfunktionen linear ist. Weiterhin konnte in Abschnitt 5.2.4 gezeigt werden, dass die Approximation der Fallzahl N_{γ} für lineare Wirkungsfunktionen sehr nahe an der benötigten Fallzahl liegt und analytisch angegeben werden kann (Gl. 5.9). Somit ist der erste Approximationsansatz eine Taylor-Approximation ersten Grades (Linearisierung) der Emax-Funktion an der angenommenen Zieldosis d^* , was zu einer Tangente an der Emax-Funktion mit Steigung $\tilde{\theta}_1 = \theta_1\theta_2/(\theta_2 + d^*)^2$ und Achsenabschnitt $\tilde{\theta}_0 = \mu - \tilde{\theta}_1 d^*$ führt. Diese Funktion ist eine Linearisierung der Emax-Wirkungsfunktion mit der Eigenschaft $\tilde{f}_{\tilde{\theta}}(d^*) = \tilde{\theta}_0 + \tilde{\theta}_1 d^* = \mu$. Werden $\tilde{\theta}_0$ und $\tilde{\theta}_1$ zusammen mit den übrigen Annahmen des untersuchten Szenarios (α, c, γ, μ und σ^2) in die Fallzahlformeln N_E^{lin} (Gl. 5.5) und N_{γ}^{lin} (Gl. 5.9) der linearen Wirkungsfunktion eingesetzt, so kann das Verhältnis dieser beiden Fallzahlen $a(\gamma) = N_{\gamma}^{lin}/N_E^{lin}$ für die Approximation der Fallzahl N_{γ} der Emax-Dosis-Wirkungskurve verwendet werden:

$$N_{\gamma, tan} \approx \frac{N_{\gamma}^{lin}}{N_E^{lin}} \cdot N_E. \quad (5.13)$$

Alternativ zu dieser Approximation kann das Verhältnis zwischen den beiden Fallzahlen auch empirisch mittels linearer Regression ohne Achsenabschnitt auf Grundlage der Fallzahlen der Simulationsszenarien (siehe Abb. 5.3) bestimmt werden ($N_{\gamma, reg}$). Dieser empirische Regressionsansatz kann zu:

$$N_{\gamma, emp} \approx N_E \cdot (1 + (\gamma - 0,5)). \quad (5.14)$$

vereinfacht werden. In Abbildung 5.4 ist die vereinfachte Regression für die untersuchten Szenarien mit einer Emax-Dosis-Wirkungskurve als gepunktete Linien dargestellt.

Mit Hilfe dieser Ansätze ist es möglich, für nichtlineare Wirkungsfunktionen die Fallzahl N_{γ} auf Grundlage der Fallzahl N_E zu approximieren. Dabei kann der Linearisierungsansatz (Gl. 5.13) für alle nichtlinearen Wirkungskurven mit existierender erster Ableitung berechnet werden, ohne auf empirische Ergebnisse zurückgreifen zu müssen. Diese Approximationsmethoden können entweder zum Finden von Startwerten für die numerische Berechnung von $N_{\gamma, ex}$ verwendet werden (siehe Abs. 5.2.3) oder um schnell abschätzen zu können, in welchem Rahmen sich die Fallzahlen bewegen oder verändern. In Abbildung 5.5 ist die prozentuale Abweichung der Fallzahlen der Approximationsverfahren gegenüber der exakt benötigten Fallzahl $N_{\gamma, ex}$ für alle 54 untersuchten Simulationsszenarien der Emax-Wirkungskurve aus Tabelle 5.1 dargestellt. Es ist ersichtlich, dass für den Linearisierungsansatz (Gl. 5.13) der Quotient $Q_{tan} = N_{\gamma, tan}/N_{\gamma, ex}$ in nahezu 75% aller untersuchten Szenarien eine Abweichung von weniger als $\pm 10\%$ aufweist aber im Allgemeinen die exakte Fallzahl ein wenig unterschätzt. Für die grobe Approximation aus (5.14) ist die angenäherte Fallzahl im Median dichter an der exakten Fallzahl, allerdings ist die Streuung deutlich größer als im Falle der Linearisierung. Die Fallzahlen des empirischen Regressionsansatzes liegen zwischen den anderen Methoden, wobei der vorgestellte Ansatz die Fallzahlen unterschätzt.

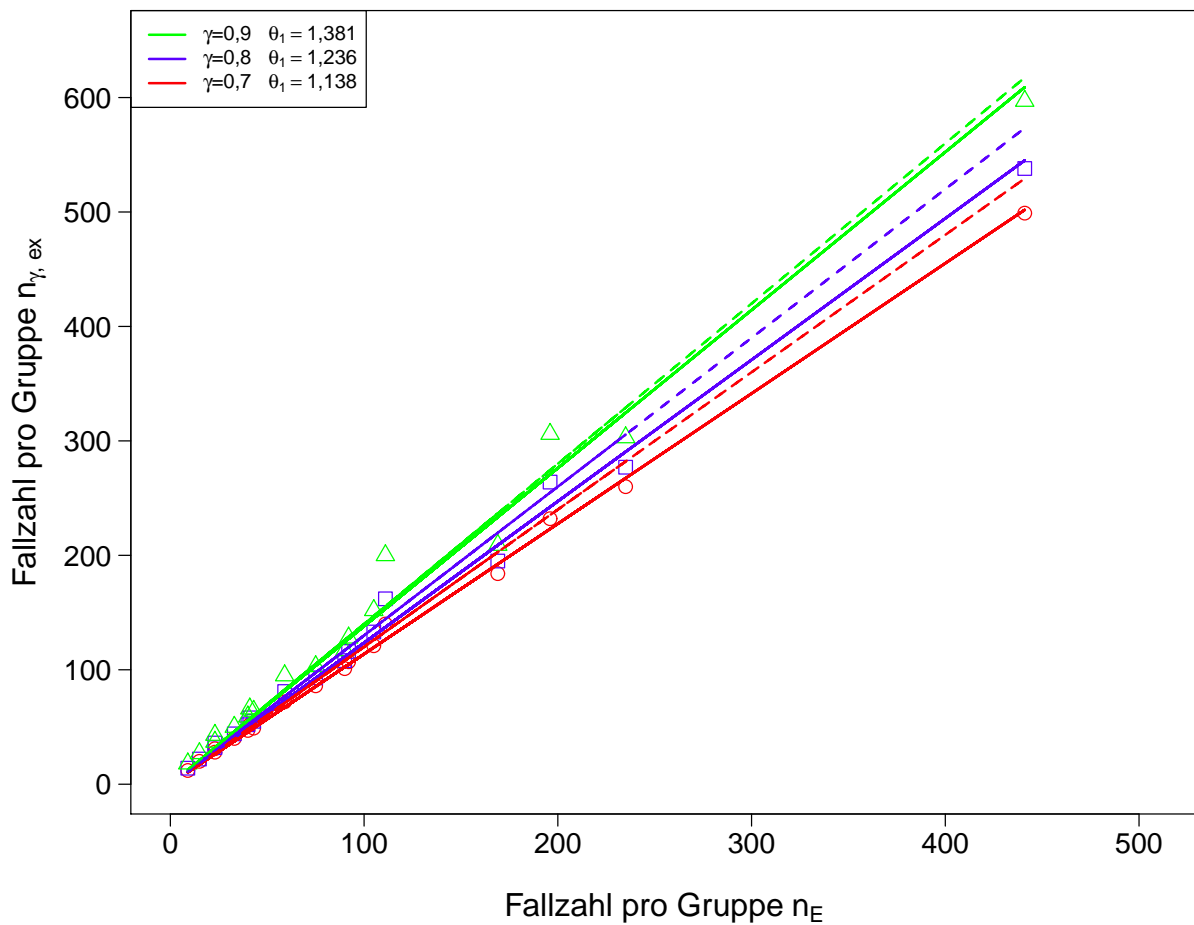


Abbildung 5.4.: Lineare Regression ohne Achsenabschnitt (—) zwischen der Fallzahl $n_E = N_E/(k+r)$ pro Gruppe und der exakten benötigten Fallzahl zum γ -Quantil $n_{\gamma, ex} = N_{\gamma, ex}/(k+r)$ pro Gruppe für verschiedene γ -Quantile und die vereinfachte Regression (---) $n_{\gamma, emp} = N_{\gamma, emp}/(k+r)$ im Falle der Emax-Dosis-Wirkungsfunktion basierend auf den Szenarien aus Tabelle 5.1.

Beispiel einer Fallzahlplanung

Nachdem die Eigenschaften der verschiedenen Fallzahlberechnungen vorgestellt und verglichen wurden, wird in diesem Abschnitt die Fallzahlplanung einer fiktiven Dosis-Findungsstudie mit aktiver Kontrolle vorgestellt. Das Studiendesign basiert auf der Studie von Chapple et al. (2004), die in Abschnitt 1.2.1 vorgestellt wurde. Daher werden eine Placebo-Kontrolle und vier verschiedene Dosisstufen ($k=5$) des Wirkstoffs Solifenacin $(d_1, \dots, d_5)' = (0; 2,5; 10; 20)'$ (mg) sowie eine Dosis Tolterodine als aktive Kontrolle untersucht. Der primäre Endpunkt in dieser Studie soll die mittlere Differenz des ausgeschiedenen Urin-Volumens (ml) zum Ende der Studie sein (Chapple et al., 2004). Es wird eine lineare Dosis-Wirkungskurve mit Achsenabschnitt $\theta_0 = 0$ und Steigung $\theta_1 = 1$ sowie eine erwartete Wirkung der aktiven Kontrolle von $\mu = 10 ml$ angenommen. Weiterhin wird für die normalverteilten Fehlerterme eine Standardabweichung von $\sigma = 10$ gewählt. Die Gewichte der Fallzahlen pro Gruppe seien wie in Abschnitt 5.2.4 mit

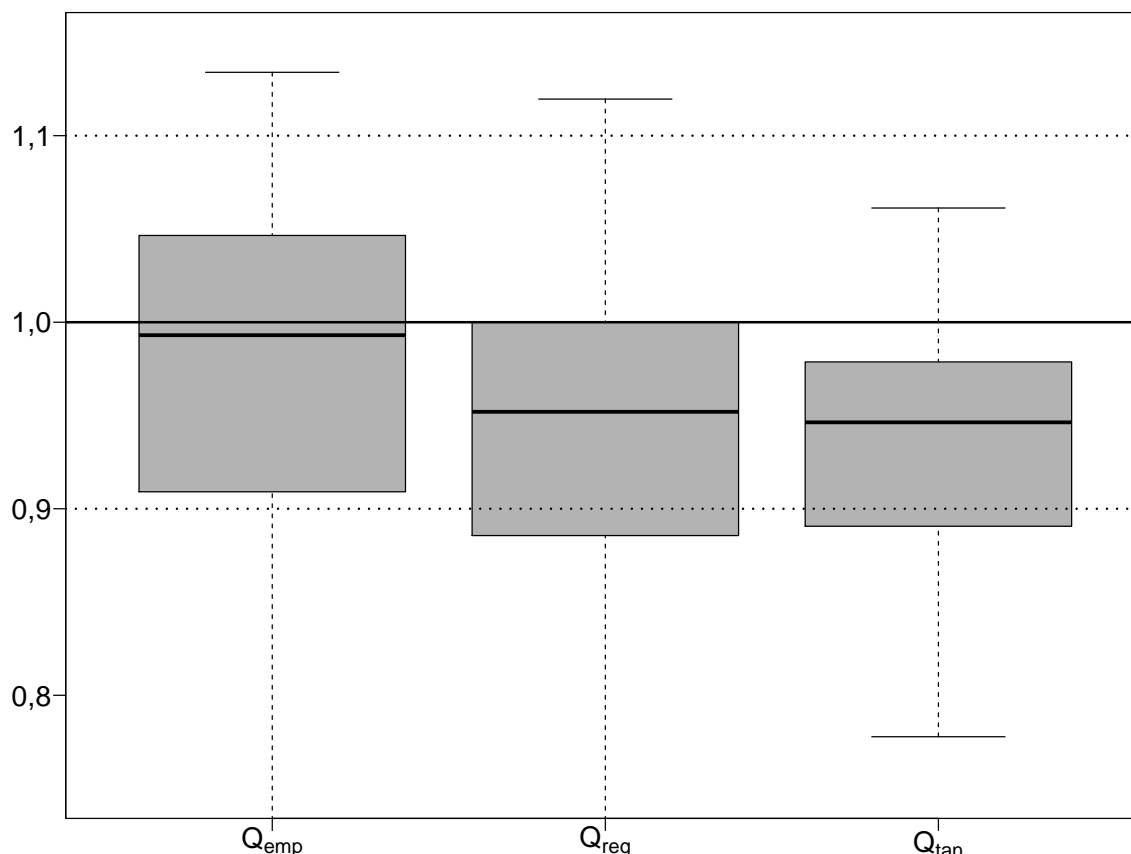


Abbildung 5.5.: Boxplot der Fallzahl-Quotienten des empirischen Multiplikators ($Q_{emp} = N_{\gamma, emp}/N_{\gamma, ex}$), der empirischen Regression ($Q_{reg} = N_{\gamma, reg}/N_{\gamma, ex}$) und des Linearisierungsansatzes ($Q_{tan} = N_{\gamma, tan}/N_{\gamma, ex}$) in Relation zur exakt benötigten Fallzahl $N_{\gamma, ex}$ der untersuchten Szenarien für die Emax-Wirkungskurve.

$w_i = 1/(k+r) = 1/7$ und $w_{ak} = r/(k+r) = 2/7$ mit $r = 2$ definiert. Das 95%-Konfidenzintervall ($\alpha = 0,05$) der Zieldosis soll in Erwartung bzw. mit einer Wahrscheinlichkeit von $\gamma = 0,8$ nicht breiter sein als $2c = 4\text{ mg}$. Um die Länge des Konfidenzintervalls in Erwartung zu kontrollieren, wird eine Gesamtfallzahl von $N_E = 490$ Probanden, $n = 70$ pro Dosisstufe und $n_{ak} = 140$ für die aktive Kontrolle, benötigt. Soll die Fallzahl das Intervall nicht nur in Erwartung sondern mit einer Wahrscheinlichkeit von $\gamma = 0,8$ kontrollieren, so wird stattdessen eine Gesamtfallzahl von $N_\gamma = 567$ Probanden, $n = 81$ pro Gruppe und $n_{ak} = 162$ für die aktive Kontrolle, benötigt.

5.2.6. Alternative Verfahren zur Fallzahlplanung

Im vorangegangenen Abschnitt wurden Fallzahlkalkulationen auf Grundlage der ML-basierten Regression und der Konfidenzintervallberechnung mittels Δ -Regel (Abs. 2.3) hergeleitet. Der Vorteil dieser Verfahren ist die Möglichkeit das $(1 - \alpha)$ -Konfidenzintervall explizit angeben zu können, was eine direkte Berechnung der Fallzahlformeln ermöglicht. Allerdings muss zur Be-

rechnung der Fallzahl N_γ in vielen Fällen auf numerische Verfahren zurückgegriffen werden. Soll ein anderes Verfahren zur Bestimmung des Konfidenzintervalls oder zur Schätzung der Zieldosis verwendet werden, kann die benötigte Fallzahl ebenfalls numerisch mit Hilfe von Simulationen durch den in Abschnitt 5.2.3 vorgestellten Algorithmus bestimmt werden. Dazu muss das gewünschte Verfahren zur Berechnung des Konfidenzintervall oder zum Schätzen der Zieldosis im 2. Schritt des Algorithmus eingesetzt werden. Damit ist es zum Beispiel möglich im Falle einer linearen Wirkungskurve anstelle der Δ -Regel auf die Methode von Fisch und Strehlau (1993) (Abs. 3.2.5) zurückzugreifen.

Dies wird im Folgenden anhand des Beispiels einer Fallzahlplanung aus Abschnitt 5.2.5 demonstriert. Unter Verwendung der Δ -Regel ergaben sich die Fallzahlen $N_E^\Delta = 490$ mit $n = 70$ und $n_{ak} = 140$ zur Kontrolle der erwarteten Konfidenzintervallbreite und $N_\gamma^\Delta = 567$ mit $n = 81$ und $n_{ak} = 162$ zur Beschränkung des KI mit einer Wahrscheinlichkeit von mindestens $\gamma = 0,8$. Die Fallzahl N_E muss für die Methode von Fisch und Strehlau ebenfalls numerisch bestimmt werden. Dazu kann entweder im 3. Schritt des Algorithmus die Abfrage $1/n_{sim} \sum_{s=1}^{n_{sim}} KI_{Länge} \leq 2c$ verwendet werden oder es wird $\gamma = 0,5$ gewählt. Letztere Bedingung soll hier verwendet werden und stützt sich auf die Resultate aus Tabelle 5.2, in welcher ersichtlich wird, dass die Fallzahl N_E zu einer Wahrscheinlichkeit von $P(KI_{Länge} \leq 2c) \approx 50\%$ führt.

Als Startwert für die numerische Suche von N_E wird die Fallzahl der Δ -Regel N_E^Δ verwendet. Der Algorithmus führt zu der Fallzahl $N_E = 504$ mit $n = 72$ und $n_{ak} = 144$. Analog kann die Fallzahl N_γ für $\gamma = 0,8$ mit Hilfe des Startwertes N_γ^Δ berechnet werden und führt zu der Fallzahl $N_\gamma = 581$ mit $n = 83$ und $n_{ak} = 166$. Zu den berechneten Fallzahlen wird im Folgenden für das vorgestellte Beispiel eine kleine Simulationsstudie mit $n_{sim} = 100.000$ durchgeführt und die Δ -Regel mit der Methode von Fisch und Strehlau verglichen. Die Ergebnisse sind in Tabelle 5.4 vorgestellt. Es ist ersichtlich, dass die Methode von Fisch und Strehlau die Überdeckungswahrscheinlichkeit gut einhält und die mittlere Intervallbreite im Fall von N_E kontrolliert. Für N_γ ist das Konfidenzintervall in mindestens $\gamma\%$ der Simulationen kleiner als $2c$.

Analog kann für nichtlineare Wirkungsfunktionen anstelle einer parametrischen Regression mittels ML-Schätzern auch die kubische Spline-Interpolation (Abs. 4.2.1) verwendet werden. Für die Fallzahlplanung mit kubischer Spline-Interpolation muss analog zu den parametrischen Verfahren die angenommene Wirkungsfunktion, die Wirkung der aktiven Kontrolle und die Varianz der Fehlerterme spezifiziert werden. Unter Verwendung dieser Bedingungen ist es möglich, für eine angenommene Wirkungskurve zuerst die optimale Dosisallokation des kubischen Splines zu bestimmen (Abs. 5.1) und anschließend wie oben beschrieben die benötigte Fallzahl numerisch zu berechnen.

5.3. Zusammenfassung

In diesem Kapitel wurden für die Spline-basierte Schätzung der Zieldosis verschiedene Ansätze zur Wahl der Dosisstufen in der Planungsphase einer Dosis-Findungsstudie hergeleitet. Dabei

lag der Fokus auf der Reduktion der interpolationsbedingten Verzerrung des Zieldosischätzers. Es wurden drei verschiedene Verfahren zur Dosiswahl vorgestellt und in einem neuen Kriterium kombiniert. In Simulationsstudien konnte gezeigt werden, dass sich nicht nur die Verzerrung reduziert, sondern auch die Überdeckungswahrscheinlichkeit des kubischen Splines durch die Wahl des optimalen Designs verbessert.

In einem zweiten Teil wurden Methoden zur Fallzahlkalkulation für lineare und nichtlineare Wirkungskurven in Dosis-Findungsstudien mit aktiver Kontrolle hergeleitet. Dabei wurde eine parametrische Regression mittels ML-Schätzern und die Δ -Regel (Abs. 2.3) zur Konstruktion des Konfidenzintervalls verwendet. Es konnte gezeigt werden, dass die Fallzahlen N_E und N_γ ihre entsprechenden Bedingungen an die Länge des Konfidenzintervalls sehr gut erfüllen. Im Falle einer linearen Wirkungskurve konnte eine Approximation für die Fallzahl N_γ bestimmt werden, die sehr nahe an der wirklich benötigten Fallzahl liegt. Für komplexere Wirkungsfunktionen kann die Fallzahl N_γ numerisch mit Hilfe von Simulationen bestimmt werden. Dieses Verfahren kann auch auf andere Methoden zur Konfidenzintervallberechnung übertragen werden. Allerdings müssen die entsprechenden Fallzahlen dann durch Simulationen und numerische Optimierung bestimmt werden. Im Falle einer nichtlinearen Wirkungsfunktion kann daher eine Fallzahlkalkulation für eine kubische Spline-Interpolation (Abs. 4.2.1) durchgeführt und zusätzlich mit den Methoden zur optimalen Dosisallokation kombiniert werden.

Tabelle 5.2.: Die Fallzahl pro Gruppe n , die Überdeckungswahrscheinlichkeit (Ü-Wkeit), die Wahrscheinlichkeit $P(KI_{Länge} \leq 2c)$ und die mittlere Konfidenzintervallbreite (mittlere $KI_{Länge}$) sind für $\sigma^2=1$, verschiedene Steigungen θ_1 , γ -Quantile und halbe Intervallbreiten c für die Fallzahlformeln N_E , N_γ und die exakte Fallzahl $N_{\gamma, ex}$ mittels numerischer Optimierung für lineare Wirkungskurven mit $n_{sim} = 100.000$ angegeben.

θ_1	d^*	c	γ	Methode	n	Ü-Wkeit	$P(KI_{Länge} \leq 2c)$	mittlere $KI_{Länge}$	
2	0,5	0,1	-	N_E	67	0,9515	0,4807	0,2020	
			0,7	N_γ	73	0,9517	0,6813	0,1935	
				$N_{\gamma, ex}$	74	0,9515	0,7145	0,1921	
			0,8	N_γ	78	0,9516	0,8166	0,1871	
				$N_{\gamma, ex}$	78	0,9524	0,8171	0,1870	
			0,9	N_γ	84	0,9517	0,9178	0,1802	
			$N_{\gamma, ex}$	83	0,9520	0,9059	0,1812		
		0,15	-	N_E	30	0,9543	0,4910	0,3052	
			0,7	N_γ	34	0,9546	0,6899	0,2860	
				$N_{\gamma, ex}$	35	0,9540	0,7290	0,2820	
			0,8	N_γ	37	0,9535	0,8061	0,2738	
				$N_{\gamma, ex}$	37	0,9543	0,8058	0,2740	
	0,9		N_γ	42	0,9542	0,9290	0,2566		
		$N_{\gamma, ex}$	41	0,9525	0,9111	0,2598			
	0,2	-	N_E	17	0,9584	0,4930	0,4122		
		0,7	N_γ	20	0,9569	0,6887	0,3779		
			$N_{\gamma, ex}$	21	0,9563	0,7445	0,3680		
		0,8	N_γ	23	0,9566	0,8354	0,3510		
			$N_{\gamma, ex}$	23	0,9567	0,8346	0,3508		
		0,9	N_γ	27	0,9559	0,9427	0,3226		
		$N_{\gamma, ex}$	25	0,9544	0,8993	0,3359			
	2,5	0,4	0,1	-	N_E	44	0,9523	0,4923	0,2017
				0,7	N_γ	48	0,9512	0,6894	0,1928
					$N_{\gamma, ex}$	49	0,9513	0,7314	0,1909
0,8				N_γ	51	0,9518	0,8054	0,1871	
				$N_{\gamma, ex}$	51	0,9519	0,8042	0,1871	
0,9				N_γ	55	0,9513	0,9093	0,1800	
		$N_{\gamma, ex}$	55	0,9513	0,9089	0,1800			
0,15		-	N_E	20	0,9538	0,5214	0,3022		
		0,7	N_γ	22	0,9523	0,6683	0,2877		
			$N_{\gamma, ex}$	23	0,9534	0,7316	0,2811		
		0,8	N_γ	24	0,9526	0,7861	0,2751		
			$N_{\gamma, ex}$	25	0,9527	0,8348	0,2693		
	0,9	N_γ	28	0,9519	0,9308	0,2540			
	$N_{\gamma, ex}$	27	0,9523	0,9057	0,2588				
0,2	-	N_E	11	0,9550	0,4868	0,4142			
	0,7	N_γ	13	0,9557	0,6788	0,3789			
		$N_{\gamma, ex}$	14	0,9552	0,7572	0,3647			
	0,8	N_γ	15	0,9546	0,8240	0,3510			
		$N_{\gamma, ex}$	15	0,9543	0,8246	0,3512			
	0,9	N_γ	18	0,9547	0,9439	0,3193			
	$N_{\gamma, ex}$	17	0,9545	0,9150	0,3290				

Tabelle 5.3.: Die Fallzahl pro Gruppe n , die Überdeckungswahrscheinlichkeit (Ü-Wkeit), die Wahrscheinlichkeit $P(KI_{Länge} \leq 2c)$ und die mittlere Konfidenzintervallbreite (mittlere $KI_{Länge}$) sind für verschiedene γ -Quantile, halbe Intervallbreiten c und ED_{50} θ_2 für die Fallzahlformel N_E und die exakte Fallzahl $N_{\gamma,ex}$ mittels numerischer Optimierung angegeben. Weiterhin sind der Achsenabschnitt $\theta_0 = 0$ und die Steigung $\theta_1 = 2$ und die Varianz des Fehlerterms $\sigma^2 = 1$ für Emax-Wirkungskurven gegeben. Es wurden $n_{sim} = 10.000$ Simulationen durchgeführt.

θ_2	d^*	c	γ	Methode	n	Ü-Wkeit	$P(KI_{Länge} \leq 2c)$	mittlere $KI_{Länge}$
0,2	0,1	0,1	-	N_E	92	0,9490	0,5073	0,2029
			0,7	$N_{\gamma,ex}$	107	0,9448	0,7103	0,1878
			0,8	$N_{\gamma,ex}$	116	0,9460	0,8077	0,1805
			0,9	$N_{\gamma,ex}$	128	0,9464	0,9083	0,1711
	0,2	0,15	-	N_E	41	0,9384	0,5051	0,3133
			0,7	$N_{\gamma,ex}$	52	0,9401	0,7299	0,2747
			0,8	$N_{\gamma,ex}$	58	0,9457	0,8313	0,2577
			0,9	$N_{\gamma,ex}$	66	0,9436	0,8969	0,2426
		0,2	-	N_E	23	0,9351	0,5072	0,4650
			0,7	$N_{\gamma,ex}$	31	0,9384	0,7237	0,3896
			0,8	$N_{\gamma,ex}$	36	0,9398	0,8191	0,3368
			0,9	$N_{\gamma,ex}$	43	0,9422	0,9112	0,3048
0,4	0,1	0,15	-	N_E	235	0,9490	0,5162	0,2019
			0,7	$N_{\gamma,ex}$	260	0,9459	0,7039	0,1919
			0,8	$N_{\gamma,ex}$	277	0,9492	0,8001	0,1858
			0,9	$N_{\gamma,ex}$	303	0,9480	0,9034	0,1774
	0,4	0,15	-	N_E	105	0,9469	0,5361	0,3071
			0,7	$N_{\gamma,ex}$	121	0,9457	0,7022	0,2854
			0,8	$N_{\gamma,ex}$	133	0,9473	0,8080	0,2708
			0,9	$N_{\gamma,ex}$	152	0,9437	0,9080	0,2521
		0,2	-	N_E	59	0,9464	0,5367	0,4339
			0,7	$N_{\gamma,ex}$	72	0,9425	0,7207	0,3785
			0,8	$N_{\gamma,ex}$	81	0,9437	0,8099	0,3541
			0,9	$N_{\gamma,ex}$	95	0,9473	0,9011	0,3251

Tabelle 5.4.: Die Überdeckungswahrscheinlichkeit (Ü-Wkeit), die Wahrscheinlichkeit $P(KI_{Länge} \leq 2c)$ und die mittlere Konfidenzintervallbreite (mittlere $KI_{Länge}$) sind für die Methode von Fisch und Strehlau ($F\&S$) und die Δ -Regel im Szenario aus Abschnitt 5.2.5 mit $\sigma^2=10$, Steigung $\theta_1 = 1$ und $\gamma = 0,8$ für die Fallzahlen N_E und N_γ mit $n_{sim} = 100.000$ angegeben.

	$F\&S$	Δ -Regel
$N_E = 504, n = 72$		
Ü-Wkeit	0,9486	0,9519
$P(KI_{Länge} \leq 2c)$	0,5091	0,5665
mittlere $KI_{Länge}$	4,0200	3,9680
$N_\gamma = 581, n = 83$		
Ü-Wkeit	0,9501	0,9517
$P(KI_{Länge} \leq 2c)$	0,8175	0,8532
mittlere $KI_{Länge}$	3,7340	3,6940

6. Zusammenfassung, Diskussion und Ausblick

In der vorliegenden Arbeit wurde die Auswertung und Planung von Dosis-Findungsstudien mit aktiver Kontrolle mit Schwerpunkt auf der Schätzung der Zieldosis und der Berechnung von zugehörigen Konfidenzintervallen untersucht. Es wurden beliebige Fallzahlen pro Dosisstufe n_i und aktiver Kontrolle n_{ak} sowie eine nicht-äquidistante Allokation der Dosisstufen $d_i \in D$, $i = 1, \dots, k$, zugelassen. Für allgemeine Dosis-Wirkungskurven $f_{\theta}(d)$ wurde eine Maximum-Likelihood Methode (Abs. 2.3) zur Schätzung der Zieldosis und zur Berechnung eines asymptotischen Konfidenzintervalls vorgestellt. Daraufhin folgte die Betrachtung von (log-)linearen Wirkungskurven als Spezialfall (Kap. 3). Die Herleitung des Zieldosis-Schätzers basierte auf Resultaten der linearen Modelle Theorie (Ravishanker und Dey, 2002, Kap. 4) mittels der Methode der kleinsten Quadrate. Unter den gemachten Modellannahmen ist dieser Schätzer identisch mit dem Schätzer der Maximum-Likelihood Methode. Zur Bestimmung eines Konfidenzintervalls der Zieldosis wurden vier verschiedene Ansätze vorgestellt und erweitert. Neben der Normalapproximation mittels Δ -Regel wurde ein parametrischer Bootstrap, ein Profile-Likelihood Ansatz und eine Erweiterung der Methode von Fisch und Strehlau (1993) untersucht. Auch wenn die Δ -Regel im Vergleich zu den anderen Methoden in den Simulationsstudien zu konservativeren Überdeckungswahrscheinlichkeiten führte, kann das resultierende Konfidenzintervall der Δ -Regel als einziges explizit angegeben werden. Diese Eigenschaft spielt für die Herleitung der Fallzahlformel eine große Rolle. Zusätzlich wurde untersucht, inwieweit eine Abweichung vom normalverteilten Fehlerterm im Falle einer linearen Wirkungskurve zu Problemen führt.

Um nicht von der Wahl der zugrundeliegenden Wirkungsfunktion abhängig zu sein, wurde in Kapitel 4 die Schätzung der Zieldosis für unbekannte Dosis-Wirkungsfunktionen mit Hilfe linearer und kubischer Spline-Interpolationen hergeleitet und untersucht. Zur Konstruktion eines Konfidenzintervalls der Zieldosis wurde ein parametrischer Bootstrap in Kombination mit den Interpolationsverfahren vorgestellt. Die Eigenschaften der Spline-Verfahren wurden in Simulationsstudien mit parametrischen Regressionen, deren Schätzer auf der Maximum-Likelihood Methode basieren (Abs. 2.3), verglichen. Die zum Teil substantielle Verzerrung der Spline-basierten Zieldosis-Schätzer, die nicht asymptotisch gegen null geht, kann zu Problemen führen und die Überdeckungswahrscheinlichkeit der Interpolationsverfahren beeinflussen. Um dieses Problem lösen zu können, wurden in Abschnitt 5.1 die Aspekte der Designplanung von Dosis-Findungsstudien mit dem Fokus auf Dosisallokation untersucht. Es wurden verschiedene Bedingungen zur Be-

rechnung einer günstigen Dosisallokation vorgestellt und in einem neuen Kriterium kombiniert. Dieses Kriterium bestimmt für die gewünschte Spline-Interpolation die Dosisallokation, die die maximale absolute Verzerrung des Zieldosischätzers minimiert. Die Funktionsweise der Bias-Reduktion konnte eindrucksvoll anhand einer Emax-Dosis-Wirkungskurve demonstriert werden (Abs. 5.1.5). In den Simulationsstudien konnte gezeigt werden, dass die reduzierte Verzerrung des Zieldosischätzers für die optimale Dosisallokation zu einer Verbesserung der Überdeckungswahrscheinlichkeit führt. Entgegen der Spline-Interpolation, die keine Annahmen an die Gestalt der Wirkungsfunktion benötigt, ist es mit der vorgestellten Methode nicht möglich die optimale Dosisallokation ohne Annahmen an die zugrundeliegende Wirkungsfunktion $f_{\theta}(d)$ zu berechnen.

Um die Präzision des Zieldosischätzers in der Planungsphase kontrollieren zu können, wurden in Abschnitt 5.2 zwei verschiedene Kriterien zur Fallzahlplanung untersucht, die die Breite des Konfidenzintervalls der Zieldosis beschränken. Im Falle der Fallzahl N_E sollte die Intervallbreite in Erwartung nicht breiter als die vordefinierte Konstante $2c$ sein. Für die Fallzahl N_{γ} sollte die Intervallbreite mit einer Wahrscheinlichkeit von γ nicht breiter als die Konstante $2c$ sein. Basierend auf der Regression mittels Maximum-Likelihood Methode und dem Konfidenzintervall der Δ -Regel wurden die Fallzahlen N_E und N_{γ} für beliebige Dosis-Wirkungskurven hergeleitet und für lineare Dosis-Wirkungsfunktionen explizit berechnet. In Simulationsstudien konnte gezeigt werden, dass die Fallzahl N_E die mittlere Breite des Konfidenzintervalls sehr gut kontrolliert. Ebenfalls führt die Fallzahl N_{γ} in $100 \cdot \gamma\%$ der Simulationen zu Intervallbreiten $\leq 2c$. Beide Fallzahlformeln erfüllen ihre jeweiligen Bedingungen und sind somit geeignet die Güte der Zieldosischätzung in der Planungsphase zu kontrollieren. Es ist möglich diese Planungskriterien auch auf andere Verfahren zur Schätzung der Zieldosis und Konfidenzintervallbestimmung (z. B. die Methode von Fisch und Strehlau, 1993) zu übertragen, indem die entsprechenden Fallzahlen numerisch berechnet werden (Abs. 5.2.6). Speziell ist es möglich, die Dosisallokation für eine kubische Spline-Interpolation zu berechnen und darauf basierend die benötigte Fallzahl zur Kontrolle der Konfidenzintervalllänge zu bestimmen.

Das Planen und Auswerten von Dosis-Findungsstudien findet großes Interesse in Forschung und Wirtschaft, da es einen essentiellen Teil zur Entwicklung neuer Medikamente beiträgt. Aus diesem Grund befindet sich die Erforschung und Weiterentwicklung neuer Verfahren im Bereich der Dosis-Findung im Fokus des wissenschaftlichen Interesses (Pinheiro et al., 2014; Dette et al., 2013, 2014). Dies wird auch an der breiten Palette zur Verfügung stehenden Methodik deutlich (Bornkamp et al., 2007). Die Statistiksoftware ADDPLAN 6 fasst nahezu alle verwendeten Methoden zusammen und ermöglicht neben der Simulation und Auswertung von Dosis-Findungsstudien auch die entsprechende Fallzahlplanung und Wahl der Dosisstufen. Trotz der Vielzahl an Möglichkeiten ist die Planung und Auswertung von Dosis-Findungsstudien mit aktiver Kontrolle bisher nicht in dieser Software implementiert. Weiterhin sind die Verwendung einer kubischen

Interpolation und die Bestimmung einer entsprechenden Spline-bedingten Dosisallokation nicht möglich.

In Abschnitt 3.4 wurden zwei SAS/IML-Makros vorgestellt, die die Auswertung von Dosis-Findungsstudien mit aktiver Kontrolle und linearer Dosis-Wirkungsfunktion ermöglichen und neben einem Zieldosis-schätzer auch die in Kapitel 3.2 vorgestellten Methoden zur Konfidenzintervallberechnung beinhalten (Helms et al., 2014a). Dabei ermöglicht die Matrizen-Sprache in SAS/IML eine direkte Umsetzung der vorgestellten Theorie und Algorithmen (Kap. 3). Neben den erwähnten Makros zur Auswertung von Studien mit linearer Dosis-Wirkungsfunktion sind im Zuge der vorliegenden Arbeit SAS-Programme zur Berechnung und zur Simulation Spline-basierter Auswertungen, optimaler Dosisallokationen und verschiedener Fallzahlformeln entstanden.

Mit Ausnahme von Kapitel 4 wurden in der gesamten Arbeit identische Varianzen der Fehlerterme pro Dosisstufe und aktiver Kontrolle angenommen. Diese Annahmen an die Fehlerterme sind in der Literatur der Auswertung von Dosis-Findungsstudien nicht unüblich (siehe z. B. Kirby et al., 2011; Bornkamp et al., 2007). Dies gilt speziell für die Planung von Dosis-Findungsstudien (siehe z. B. Dette et al., 2008, 2014). Allerdings lässt sich ein großer Teil der Methodik auch ungleiche Varianzen zu und kann dahingehend erweitert werden. Die Erweiterung der vorgestellten Verfahren zur Auswertung von Dosis-Findungsstudien mit aktiver Kontrolle auf ungleiche Varianzen der Fehlerterme wird einen Teil der weiteren Forschung darstellen. Dabei ist zu beachten, dass der Ansatz der Spline-Interpolation (Kap. 4) bereits für ungleiche Varianzen hergeleitet aber noch nicht untersucht wurde (Helms et al., 2014b). Weiterhin können die Methoden zur Fallzahlkalkulation verallgemeinert werden, indem ungleiche Varianzen auch in der Planungsphase von Dosis-Wirkungsstudien mit aktiver Kontrolle berücksichtigt werden. Zusätzlich wäre die explizite Analyse der präsentierten Fallzahlkalkulationen für weitere Wirkungskurven (z. B. quadratisch oder sigmoidal Emax) interessant; vor allem in Bezug auf empirische Approximationen der Fallzahl N_γ . Außerdem ist die Veröffentlichung weiterer SAS oder R-Makros für die vorgestellten Methoden inklusive der angesprochenen Verallgemeinerungen erstrebenswert.

Ein anderer interessanter Aspekt wäre, die vorgestellten Resultate für Dosis-Findungsstudien mit aktiver Kontrolle auf die adaptive Design-Planung (siehe z. B. Jones et al., 2011) zu übertragen. Dadurch kann die Möglichkeit geschaffen werden, die Fallzahlen nach einer Zwischenauswertung neu zu bestimmen und gegebenenfalls anzupassen. Weiterhin ist die Variation der Dosisstufen in der Zwischenauswertung von Interesse. Dabei können zum Beispiel während der Studie Dosisstufen zusätzlich in das Versuchsdesign aufgenommen oder daraus entfernt werden (siehe z. B. Bornkamp et al., 2011; Dragalin et al., 2007). Dieser Ansatz könnte auf die vorgestellten parametrischen Verfahren zur Auswertung von Dosis-Findungsstudien mit aktiver Kontrolle erweitert oder in Kombination mit der Spline-orientierten Dosisallokation verwendet werden (siehe z. B. Dragalin et al., 2010).

Ebenfalls von Interesse wäre der Vergleich und die Kombination der optimalen Spline-Designs

mit der optimalen Design Theorie parametrischer Verfahren (siehe z. B. Dette et al., 2008, 2014; Bretz et al., 2010). Dabei könnte untersucht werden, welche Unterschiede die jeweiligen Designs aufweisen und ob sich Resultate der parametrischen Verfahren auf die Spline-orientierte Allokation übertragen lassen. Speziell wäre die optimale Bestimmung der Fallzahlallokation beziehungsweise Gewichtung w_i pro Dosisstufe interessant, um die Güte der Interpolation weiter zu verbessern. Zusätzlich könnte die Berechnung des optimalen Spline-Designs unter Modellunsicherheit mit Hilfe eines Min-Max-Ansatzes betrachtet werden. Dann wäre es möglich ein Design nicht nur für eine vorgegebene Dosis-Wirkungskurve, sondern auch für verschiedene Parameterkonstellationen anzugeben. Die resultierende Dosis-Allokation würde die maximale Verzerrung über alle betrachteten Modellvariationen der Dosis-Wirkungskurve minimieren. Zusammenfassend kann im Hinblick auf die weitere Forschung im Bereich der Dosis-Findungsstudien mit aktiver Kontrolle gesagt werden, dass noch viele Aspekte der Planung und Auswertung, speziell im adaptiven und optimalen Design bearbeitet werden können.

A. Definitionen, Lemmata und Sätze

A.1. ML-Schätzer und deren asymptotische Eigenschaften

Definition A.1 (Konsistenz)

Eine Folge von Schätzern $\hat{\theta}_n$ heißt konsistent für θ , falls $\hat{\theta}_n - \theta \xrightarrow{P} 0$ für festes θ , d.h. $\lim_{n \rightarrow \infty} P(|\hat{\theta}_n - \theta| > \epsilon) = 0 \forall \epsilon > 0$.

Definition A.2 (Starke Konsistenz)

Eine Folge von Schätzern $\hat{\theta}_n$ heißt stark konsistent für θ , falls $\hat{\theta}_n - \theta \xrightarrow{f.s.} 0$ für festes θ , d.h. $\lim_{n \rightarrow \infty} P(\omega \in \Omega : \sup_{m \geq n} |\hat{\theta}_m - \theta| > \epsilon) = 0 \forall \epsilon > 0$.

Die Konsistenz ist eine Minimaleigenschaft, welche Schätzer mindestens erfüllen sollten. Sie impliziert, dass sich der Schätzer mit wachsendem Stichprobenumfang n dem zu schätzenden Parameter immer besser annähert. Da fast sichere Konvergenz ($\xrightarrow{f.s.}$) die Konvergenz in Wahrscheinlichkeit (\xrightarrow{P}) impliziert, gilt selbiges auch für die starke Konsistenz und Konsistenz.

Satz A.1 (Starke Konsistenz der Maximum-Likelihood Schätzer) Seien Y_1, Y_2, \dots, Y_n unabhängig und identisch verteilte Zufallsvariablen mit Wahrscheinlichkeitsdichte $\phi(y|\theta)$, $\theta \in \Theta$ und bezeichne θ^0 den wahren Wert von θ . Weiterhin sei

1. Θ kompakt,
2. $\phi(y|\theta)$ stetig in θ für alle y ,
3. es existiert eine Funktion $K(y)$, so dass $E_{\theta^0}|K(y)| < \infty$ und

$$U(y, \theta) = \ln(\phi(y|\theta)) - \ln(\phi(y|\theta^0)) \leq K(y), \forall y \text{ und } \theta,$$

4. für alle $\theta \in \Theta$ und genügend kleines $\rho > 0$, $\sup_{|\theta^* - \theta| < \rho} \phi(y|\theta^*)$ messbar in y ,
5. (Identifizierbarkeit) $\phi(y|\theta) = \phi(y|\theta^0)$ (σ -endlich messbar) $\Rightarrow \theta = \theta^0$,

dann folgt für jede Folge von Maximum-Likelihood Schätzern $\hat{\theta}_n$ von θ :

$$\hat{\theta}_n \xrightarrow{f.s.} \theta^0.$$

BEWEIS. Für den Beweis siehe zum Beispiel Ferguson (1996, S. 112-115). □

Satz A.2 (Cramér-Asymptotische Normalität von Maximum-Likelihood Schätzern)

Seien Y_1, Y_2, \dots, Y_n unabhängig und identisch verteilte Zufallsvariablen mit Wahrscheinlichkeitsdichte $\phi(y|\boldsymbol{\theta})$, $\boldsymbol{\theta} \in \Theta \subset \mathbb{R}^{b+1}$ und bezeichne $\boldsymbol{\theta}^0$ den wahren Wert von $\boldsymbol{\theta}$. Weiterhin gelte

1. Θ ist eine offene Teilmenge des \mathbb{R}^{b+1} ,
2. die zweiten Ableitungen von $\phi(y|\boldsymbol{\theta})$ bezüglich $\boldsymbol{\theta}$ existieren, sind stetig für alle y und können in das Integral hineingezogen werden,
3. es existiert eine Funktion $K(y)$, so dass $E_{\boldsymbol{\theta}^0}|K(y)| < \infty$ und jede Komponente der Matrix $\mathbf{H}(y, \boldsymbol{\theta}) = \frac{\partial^2}{\partial \boldsymbol{\theta}^2} \ln(\phi(y|\boldsymbol{\theta}))$ ist absolut gleichmäßig beschränkt durch $K(y)$ in einer Umgebung um $\boldsymbol{\theta}^0$,
4. $\mathcal{F}(\boldsymbol{\theta}) = -E_{\boldsymbol{\theta}^0}(\mathbf{H}(y, \boldsymbol{\theta}^0))$ ist die Fisher-Informationsmatrix und positiv definit,
5. (Identifizierbarkeit) $\phi(y|\boldsymbol{\theta}) = \phi(y|\boldsymbol{\theta}^0)$ (dv σ -endlich messbar) $\Rightarrow \boldsymbol{\theta} = \boldsymbol{\theta}^0$.

dann existiert eine stark konsistente Folge von Maximum-Likelihood Schätzern $\widehat{\boldsymbol{\theta}}_n$, so dass

$$\sqrt{n}(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}^0) \xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}, \mathcal{F}(\boldsymbol{\theta}^0)^{-1}).$$

BEWEIS. Für den Beweis siehe Ferguson (1996, S. 121-122). □

A.2. Lineare Modelle Theorie

Satz A.3 Sei $\mathbf{X} = (X_1, \dots, X_n)'$ ein Zufallsvektor mit Erwartungswertvektor $\boldsymbol{\mu}$ und Kovarianzmatrix $\boldsymbol{\Sigma}$ und seien $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{b} \in \mathbb{R}^m$, dann gelten folgende Aussagen

1. $E(\mathbf{A}\mathbf{X} + \mathbf{b}) = \mathbf{A}\boldsymbol{\mu} + \mathbf{b}$
2. $E(\mathbf{A}\mathbf{X}\mathbf{b}') = \mathbf{A}\mathbf{X}\mathbf{b}'$
3. $\text{Var}(\mathbf{A}\mathbf{X} + \mathbf{b}) = \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}'$.

BEWEIS. Der Beweis der von Aussage 1 folgt direkt durch

$$E(\mathbf{Y}) = \mathbf{A}E(\mathbf{X}) + \mathbf{b} = \mathbf{A}\boldsymbol{\mu} + \mathbf{b}.$$

Für den Beweis von Aussage 2. und 3. siehe Mattai und Provost (1992, S. 53) oder Brunner (2005, S. 32-33). □

Definition A.3 (Multivariate Normalverteilung) 1. Die Zufallsvariablen Z_i seien unabhängig standardnormalverteilt ($Z_i \sim \mathcal{N}(0, 1)$). Dann heißt die Verteilung des Zufallsvektors $\mathbf{Z} = (Z_1, \dots, Z_n)'$ n -dimensionale Standardnormalverteilung mit Erwartungswertvektor $E(\mathbf{Z}) = \mathbf{0}$ und Kovarianzmatrix $\text{Var}(\mathbf{Z}) = \mathbf{I}_n$. Abkürzend schreibt man $\mathbf{Z} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_n)$.

2. Sei $\mathbf{A}_{m \times n}$, $m \leq n$ und $\mathbf{Y} = \mathbf{AZ} + \boldsymbol{\mu}$ mit $\boldsymbol{\mu} \in \mathbb{R}^m$. Dann heißt \mathbf{Y} m -dimensional normalverteilt mit $E(\mathbf{Y}) = \boldsymbol{\mu}$ und $\text{Var}(\mathbf{Y}) = \boldsymbol{\Sigma} = \mathbf{A}\mathbf{I}_n\mathbf{A}' = \mathbf{A}\mathbf{A}'$, $\mathbf{Y} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

Satz A.4 (Lancaster) Sei $\mathbf{Y} = (Y_1, \dots, Y_n)'$ ein Zufallsvektor mit $E(\mathbf{Y}) = \boldsymbol{\mu} = (\mu_1, \dots, \mu_n)'$ und $\text{Var}(\mathbf{Y}) = \boldsymbol{\Sigma}$. Ferner sei $\mathbf{A} = \mathbf{A}'$. Dann gilt

$$E(\mathbf{Y}'\mathbf{A}\mathbf{Y}) = \text{Sp}(\mathbf{A}\boldsymbol{\Sigma}) + \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}.$$

BEWEIS. Der Beweis folgt durch Ausnutzung der Invarianz der Spur unter zyklischer Vertauschung:

$$\begin{aligned} E(\mathbf{Y}'\mathbf{A}\mathbf{Y}) &= E(\text{Sp}(\mathbf{Y}'\mathbf{A}\mathbf{Y})) = E(\text{Sp}(\mathbf{A}\mathbf{Y}\mathbf{Y}')) \\ &= \text{Sp}(\mathbf{A} \cdot E(\mathbf{Y}\mathbf{Y}')) = \text{Sp}(\mathbf{A}\boldsymbol{\Sigma}) + \text{Sp}(\mathbf{A}\boldsymbol{\mu}\boldsymbol{\mu}') \\ &= \text{Sp}(\mathbf{A}\boldsymbol{\Sigma}) + \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}. \end{aligned}$$

Siehe auch Brunner (2005, S. 34) □

Satz A.5 (Craig und Sakamoto) Sei $\mathbf{Y} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\mathbf{A} = \mathbf{A}'$ positiv semidefinit (p.s.d.), $\mathbf{B} = \mathbf{B}'$ p.s.d. und \mathbf{b} ein konstanter Vektor. Dann gilt

1. $\mathbf{Y}'\mathbf{A}\mathbf{Y}$ und $\mathbf{Y}'\mathbf{B}\mathbf{Y}$ bzw. $\mathbf{Y}'\mathbf{A}\mathbf{Y}$ und $\mathbf{B}\mathbf{Y}$ sind stochastisch unabhängig, falls $\mathbf{B}\boldsymbol{\Sigma}\mathbf{A} = \mathbf{0}$ ist,
2. $\mathbf{A}\mathbf{Y}$ und $\mathbf{b}'\mathbf{Y}$ sind stochastisch unabhängig, falls $\mathbf{b}'\boldsymbol{\Sigma}\mathbf{A} = \mathbf{0}$ ist.

BEWEIS. Für den Beweis siehe Craig (1943) oder Brunner (2005, S. 41-42). □

Definition A.4 (Mahalanobis-Abstand) Sei $\mathbf{Z} \in \mathbb{R}^l$ mit $E(\mathbf{Z}) = \boldsymbol{\mu}$ und $\text{Var}(\mathbf{Z}) = \boldsymbol{\Sigma}$ positiv definit (p. d.), dann heißt

$$D_{\boldsymbol{\Sigma}}(\mathbf{Z}, \boldsymbol{\mu}) = (\mathbf{Z} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{Z} - \boldsymbol{\mu})$$

Mahalanobis-Abstand von \mathbf{Z} zu $\boldsymbol{\mu}$.

Satz A.6 (Gauss-Markov-Aitken) Sei $\mathbf{Y} = \mathbf{X}\boldsymbol{\theta} + \boldsymbol{\epsilon} \in \mathbb{R}^n$ mit konstantem, unbekanntem Parametervektor $\boldsymbol{\theta}$ und $\det(\mathbf{X}'\mathbf{X}) \neq 0$ sowie $E(\boldsymbol{\epsilon}) = \mathbf{0}$ und $\text{Var}(\boldsymbol{\epsilon}) = \mathbf{S}$ positiv definit (p. d.). Dann gilt

1. Der Schätzer $\hat{\boldsymbol{\theta}} = (\mathbf{X}'\mathbf{S}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{S}^{-1}\mathbf{Y}$ minimiert den Mahalanobis-Abstand (Def. A.4) von $\mathbf{X}\hat{\boldsymbol{\theta}}$ zu \mathbf{Y} . Dieser Schätzer wird als verallgemeinerter Kleinste-Quadrate-Schätzer (generalized least squares (GLS) estimator) für $\boldsymbol{\theta}$ bezeichnet.
2. Für jede positiv definite Kovarianzmatrix $\boldsymbol{\Sigma}$ ist $\tilde{\boldsymbol{\theta}} = (\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{Y}$ erwartungstreu für $\boldsymbol{\theta}$. Insbesondere ist $\hat{\boldsymbol{\theta}}_0 = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$ erwartungstreu für $\boldsymbol{\theta}$. Dieser Schätzer heißt gewöhnlicher Kleinste-Quadrate-Schätzer (ordinary least squares (OLS) estimator) für $\boldsymbol{\theta}$.
3. Unter allen Schätzern $\boldsymbol{\theta}^* = \mathbf{A}\mathbf{Y}$, für die $E(\boldsymbol{\theta}^*) = \boldsymbol{\theta}$ gilt, hat $\hat{\boldsymbol{\theta}}$ minimale Varianz (genauer: jedes Diagonalelement von $\text{Var}(\boldsymbol{\theta}^*)$ ist minimal, „best linear unbiased estimator“ (BLUE)) wobei $\mathbf{A} \neq \mathbf{0}$ eine beliebige $(l \times n)$ -Matrix ist.
4. Es ist $\hat{\boldsymbol{\theta}} = \hat{\boldsymbol{\theta}}_0$ genau dann, wenn

$$\mathbf{X}'\mathbf{S} \left[\mathbf{I}_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' \right] = \mathbf{0}$$

gilt.

BEWEIS. Für den Beweis der Punkte 1 bis 3 siehe Brunner (2005, S. 44-46) oder (Ravishanker und Dey, 2002, S.119-121) und für 4. siehe Kruskal (1968). \square

Satz A.7 (Varianz-Schätzer) Sei $\mathbf{Y} = \mathbf{X}\boldsymbol{\theta} + \boldsymbol{\epsilon} \in \mathbb{R}^n$ mit konstantem und unbekanntem Parametervektor $\boldsymbol{\theta}$ und $r(\mathbf{X})$ der Rang der Designmatrix \mathbf{X} sowie $\det(\mathbf{X}'\mathbf{X}) \neq 0$, $E(\boldsymbol{\epsilon}) = \mathbf{0}$ und $\text{Var}(\boldsymbol{\epsilon}) = \sigma^2\mathbf{I}_n$.

Dann ist

$$\hat{\sigma}^2 = \frac{1}{n - r(\mathbf{X})} \cdot \mathbf{Y}' \left[\mathbf{I}_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' \right] \mathbf{Y}$$

erwartungstreu für σ^2 und falls $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I}_n)$ ist, dann gilt

$$\frac{n - r(\mathbf{X})}{\sigma^2} \cdot \hat{\sigma}^2 \sim \chi_{n-r(\mathbf{X})}^2(0).$$

BEWEIS.

Die Erwartungstreue folgt aus dem Satz von Lancaster (Satz A.4) mit $\mathbf{S} = \sigma^2 \mathbf{I}_n = \text{Var}(\mathbf{Y})$ und $\mathbf{A} = \mathbf{I}_n - \mathbf{P}_v = \mathbf{I}_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$:

$$\begin{aligned} E([n - r(\mathbf{X})] \hat{\sigma}^2) &= \text{Sp}(\mathbf{A}\mathbf{S}) + \boldsymbol{\theta}'\mathbf{X}'[\mathbf{I}_n - \mathbf{P}_v]\mathbf{X}\boldsymbol{\theta} \\ &= \sigma^2 \text{Sp}(\mathbf{A}) = \sigma^2 \left[\text{Sp}(\mathbf{I}_n) - \text{Sp}\left(\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\right) \right] \\ &= \sigma^2 \left[n - \text{Sp}\left((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}\right) \right] \\ &= \sigma^2 [n - r(\mathbf{X})], \end{aligned}$$

da $\mathbf{X}'[\mathbf{I}_n - \mathbf{P}_v] = \mathbf{0}$ und die Spur invariant unter zyklischer Vertauschung ist. Für den Beweis der Verteilung des Varianzschätzers siehe zum Beispiel Brunner (2005, S. 38-39). \square

B. Beweise und Berechnungen

B.1. Eigenschaften des Schätzers der Zioldosis

Satz B.1 (Taylor-Approximation des Zioldosischätzers) Seien $Y_d = f_{\boldsymbol{\theta}}(\mathbf{d}) + \epsilon_d$ und $Y_{ak} = \mu \mathbf{1}_{ak} + \epsilon_{ak}$ die in Abschnitt 2.2 definierten Zufallsvariablen und sei die Gestalt der Wirkungsfunktion $f_{\boldsymbol{\theta}}(\mathbf{d})$ bekannt mit unbekanntem Parametervektor $\boldsymbol{\theta}$, speziell existiere die Umkehrfunktion $f_{\boldsymbol{\theta}}^{-1}(\mu)$. Dann kann die Zioldosis sowie der zugehörige Plug-in Schätzer analog zu Abschnitt 2.4.3 als $d^* = f_{\boldsymbol{\theta}}^{-1}(\mu) = g(\boldsymbol{\theta}_{d^*})$ und $\hat{d}^* = g(\hat{\boldsymbol{\theta}}_{d^*})$ geschrieben werden. Zusätzlich wird gefordert, dass die Umkehrfunktion bezüglich $\boldsymbol{\theta}_{d^*}$ zweimal stetig differenzierbar ist. Der Erwartungswert des Zioldosischätzers kann daraufhin mit Hilfe der Taylor-Approximation zweiten Grades wie folgt angenähert werden:

$$E\left(g\left(\hat{\boldsymbol{\theta}}_{d^*}\right)\right) \approx d^* + \frac{1}{2} \text{Sp}[\mathbf{H}(\boldsymbol{\theta}_{d^*}) \boldsymbol{\Sigma}],$$

wobei $\boldsymbol{\Sigma}$ die Kovarianzmatrix der asymptotischen Verteilung des Zufallsvektors $\hat{\boldsymbol{\theta}}_{d^*}$ aus Gl. (2.8) beschreibt und $\mathbf{H}(\boldsymbol{\theta}_{d^*}) = \frac{\partial^2 g}{\partial \boldsymbol{\theta}_{d^*} \partial \boldsymbol{\theta}_{d^*}'}$ die Hessematrix an der Stelle $\boldsymbol{\theta}_{d^*}$ von $g(\cdot)$.

BEWEIS. Mit Hilfe der Taylor-Approximation kann zuerst der Zioldosischätzer durch

$$\hat{d}^* = g\left(\hat{\boldsymbol{\theta}}_{d^*}\right) \approx g\left(\boldsymbol{\theta}_{d^*}\right) + \Delta g\left(\boldsymbol{\theta}_{d^*}\right)' \left(\hat{\boldsymbol{\theta}}_{d^*} - \boldsymbol{\theta}_{d^*}\right) + \frac{1}{2} \left(\hat{\boldsymbol{\theta}}_{d^*} - \boldsymbol{\theta}_{d^*}\right)' \mathbf{H}\left(\boldsymbol{\theta}_{d^*}\right) \left(\hat{\boldsymbol{\theta}}_{d^*} - \boldsymbol{\theta}_{d^*}\right)$$

approximiert werden, wobei $\Delta g(\boldsymbol{\theta}_{d^*})$ (Gl. 2.9) den Gradienten von $g(\cdot)$ an der Stelle $\boldsymbol{\theta}_{d^*}$ widerspiegelt. Dann kann der Erwartungswert des Zioldosischätzers durch den Erwartungswert der Approximation ersetzt werden:

$$\begin{aligned} E\left(g\left(\hat{\boldsymbol{\theta}}_{d^*}\right)\right) &\approx E\left[g\left(\boldsymbol{\theta}_{d^*}\right) + \Delta g\left(\boldsymbol{\theta}_{d^*}\right)' \left(\hat{\boldsymbol{\theta}}_{d^*} - \boldsymbol{\theta}_{d^*}\right) + \frac{1}{2} \left(\hat{\boldsymbol{\theta}}_{d^*} - \boldsymbol{\theta}_{d^*}\right)' \mathbf{H}\left(\boldsymbol{\theta}_{d^*}\right) \left(\hat{\boldsymbol{\theta}}_{d^*} - \boldsymbol{\theta}_{d^*}\right)\right] \\ &= g\left(\boldsymbol{\theta}_{d^*}\right) + \Delta g\left(\boldsymbol{\theta}_{d^*}\right)' E\left[\left(\hat{\boldsymbol{\theta}}_{d^*} - \boldsymbol{\theta}_{d^*}\right)\right] \\ &\quad + \frac{1}{2} E\left[\text{Sp}\left(\left(\hat{\boldsymbol{\theta}}_{d^*} - \boldsymbol{\theta}_{d^*}\right)' \mathbf{H}\left(\boldsymbol{\theta}_{d^*}\right) \left(\hat{\boldsymbol{\theta}}_{d^*} - \boldsymbol{\theta}_{d^*}\right)\right)\right] \\ &= d^* + \frac{1}{2} \text{Sp}\left(\mathbf{H}\left(\boldsymbol{\theta}_{d^*}\right) E\left[\left(\hat{\boldsymbol{\theta}}_{d^*} - \boldsymbol{\theta}_{d^*}\right) \left(\hat{\boldsymbol{\theta}}_{d^*} - \boldsymbol{\theta}_{d^*}\right)'\right]\right) \\ &= d^* + \frac{1}{2} \text{Sp}\left(\mathbf{H}\left(\boldsymbol{\theta}_{d^*}\right) \boldsymbol{\Sigma}\right), \end{aligned}$$

was zu zeigen war. Siehe für die Beweisführung auch Preacher et al. (2007). □

B.2. Verzerrungs-reduzierte Designs für die Verwendung Spline-basierter Auswertungen

B.2.1. Dosisstufen mit abstandsgleichen Wirkungsgraden

Es kann gezeigt werden, dass die in Abschnitt 5.1.2 vorgestellte Methode zur Bestimmung der Dosisallokation für lineare, Emax und sigmoidale Emax-Wirkungsfunktionen nicht von der Steigung θ_1 und dem Achsenabschnitt θ_0 der jeweiligen Funktion abhängt. Dafür wird in den folgenden Berechnungen angenommen, dass die Parameter $\theta_1, \theta_2, \theta_3$ von null verschieden sind.

B.2.2. Die lineare Wirkungskurve

Wird eine lineare Wirkungsfunktion $f_{\theta}(d) = \theta_0 + \theta_1 d$ angenommen, so vereinfachen sich die Bedingungen aus Abschnitt 5.1.2 zu:

$$|f_{\theta}(d_{i+1}) - f_{\theta}(d_i)| = |(\theta_0 + \theta_1 d_{i+1}) - (\theta_0 + \theta_1 d_i)| = |\theta_1| |d_{i+1} - d_i| \stackrel{!}{=} h_{equ} \text{ mit } i = 1, \dots, k-1.$$

Somit hängt das Gleichungssystem mit $k-1$ Bedingungen nicht vom Achsenabschnitt θ_0 und der Steigung θ_1 ab. Damit ergibt sich die optimale Dosisallokation analog zum globalen Interpolationsfehler in Abschnitt 5.1.1 mit $d_{i+1} - d_i = \tilde{h}_{equ} = h_{equ}/\theta_1$ für $i = 1, \dots, k-1$. Infolgedessen sind die Dosisstufen äquidistant auf dem Dosisbereich verteilt.

B.2.3. Die Emax-Wirkungskurve

Wird eine Emax-Wirkungsfunktion $f_{\theta}(d) = \theta_0 + \theta_1 d/(d + \theta_2)$ angenommen, so können die Bedingungen aus Abschnitt 5.1.2 wie folgt beschrieben werden:

$$\begin{aligned} |f_{\theta}(d_{i+1}) - f_{\theta}(d_i)| &= \left| \left(\theta_0 + \theta_1 \frac{d_{i+1}}{(d_{i+1} + \theta_2)} \right) - \left(\theta_0 + \theta_1 \frac{d_i}{(d_i + \theta_2)} \right) \right| \\ &= |\theta_1| \left| \frac{d_{i+1}}{(d_{i+1} + \theta_2)} - \frac{d_i}{(d_i + \theta_2)} \right| \stackrel{!}{=} h_{equ} \text{ mit } i = 1, \dots, k-1. \end{aligned}$$

Es ist ersichtlich, dass auch im Falle einer Emax-Wirkungskurve der Achsenabschnitt θ_0 und die Steigung θ_1 keinen Einfluss auf das Gleichungssystem der optimalen Dosiswahl haben. Lediglich die ED_{50} (θ_2) muss bekannt sein, um das Gleichungssystem lösen und die optimale Dosisallokation berechnen zu können.

B.2.4. Die sigmoidale Emax-Wirkungskurve

Wird eine sigmoidale Emax-Wirkungsfunktion $f_{\theta}(d) = \theta_0 + \theta_1/(1 + (\theta_2/d)^{\theta_3})$ angenommen, so können die Bedingungen aus Abschnitt 5.1.2 wie folgt beschrieben werden:

$$\begin{aligned}
 |f_{\theta}(d_{i+1}) - f_{\theta}(d_i)| &= \left| \theta_0 + \theta_1 \frac{1}{\left(1 + (\theta_2/d_{i+1})^{\theta_3}\right)} - \theta_0 - \theta_1 \frac{1}{\left(1 + (\theta_2/d_i)^{\theta_3}\right)} \right| \\
 &= |\theta_1| \left| \frac{1}{\left(1 + (\theta_2/d_{i+1})^{\theta_3}\right)} - \frac{1}{\left(1 + (\theta_2/d_i)^{\theta_3}\right)} \right| \stackrel{!}{=} h_{equ} \text{ mit } i = 1, \dots, k-1.
 \end{aligned}$$

Weiterhin bleibt ersichtlich, dass die Optimierung der Dosisallokation unabhängig vom Achsenabschnitt θ_0 und der Steigung θ_1 durchgeführt werden kann. Allerdings werden im Falle einer sigmoidalen Emax-Funktion zusätzlich zur ED_{50} Informationen zum Grad der sigmoidalen Krümmung θ_3 für die Bestimmung der optimalen Dosisallokation benötigt.

B.2.5. Gleichmäßige Verteilung der Obergrenze des Interpolationsfehlers

Es ist ebenfalls möglich zu zeigen, dass die in Abschnitt 5.1.3 vorgestellte Methode zur Bestimmung der Dosisallokation für die lineare, die Emax und die sigmoidale Emax-Wirkungsfunktion nicht von der Steigung θ_1 und dem Achsenabschnitt θ_0 der jeweiligen Funktion abhängt. Da die Parameter der jeweiligen Wirkungsfunktion nur die zweite bzw. vierte Ableitung der Dosis-Wirkungsfunktion beeinflussen, wird die Differenz der Dosisstufen $|d_{i+1} - d_i|$ zum Nachweis der Unabhängigkeit vernachlässigt.

B.2.6. Die lineare Wirkungskurve

Die zweite und vierte Ableitung der linearen Wirkungsfunktion $f_{\theta}(d) = \theta_0 + \theta_1 d$, die für die Obergrenze der linearen und kubischen Spline-Interpolation benötigt werden, können wie folgt beschrieben werden:

$$\frac{\partial^2 f_{\theta}}{\partial d^2} = 0 \text{ und } \frac{\partial^4 f_{\theta}}{\partial d^4} = 0.$$

Daraus folgt nicht nur, dass das optimale Dosisdesign unabhängig von der Wahl des Achsenabschnitts θ_0 und der Steigung θ_1 ist, sondern auch jedes beliebige Design mit $d_1 < \dots < d_k$ optimal ist. Für jede betrachtete Wirkung der aktiven Kontrolle μ ist die Verzerrung gleich null. Dies kann durch die Tatsache erklärt werden, dass lineare Spline-Interpolationen Polynome bis zum Grad eins und kubische Spline-Interpolationen bis zum Grad drei exakt interpolieren (siehe z. B. Floater, 2007, S. 58-61).

B.2.7. Die Emax-Wirkungskurve

Die zweite und vierte Ableitung der Emax Wirkungsfunktion $f_{\theta}(d) = \theta_0 + \theta_1 d / (d + \theta_2)$, die für die Obergrenze der linearen und kubischen Spline-Interpolation benötigt werden, können wie folgt beschrieben werden:

$$\frac{\partial^2 f_{\theta}}{\partial d^2} = -2\theta_1 \frac{\theta_2}{(d + \theta_2)^3} = \tilde{c}_2 \frac{1}{(d + \theta_2)^3} \quad \text{mit } \tilde{c}_2 = -2\theta_1\theta_2$$

und

$$\frac{\partial^4 f_{\theta}}{\partial d^4} = -24\theta_1 \frac{\theta_2}{(d + \theta_2)^5} = \tilde{c}_4 \frac{1}{(d + \theta_2)^5} \quad \text{mit } \tilde{c}_4 = -24\theta_1\theta_2.$$

Das Einsetzen der Ausdrücke in die Gleichungen (5.1) bzw. (5.1) zeigt, dass die zweite und vierte Ableitung und somit die optimale Dosisallokation der linearen und kubischen Spline-Interpolation nur von der ED_{50} (θ_2) abhängen. Dies wird beispielhaft für Gleichung (5.1) demonstriert:

$$\begin{aligned} \|f_{\theta} - s_1\|_{\infty} |_{d \in [d_i, d_{i+1}]} &\leq \|f''_{\theta}\|_{\infty} |_{d \in [d_i, d_{i+1}]} (d_{i+1} - d_i)^2 \\ &= \tilde{c}_2 \frac{1}{(d + \theta_2)^3} (d_{i+1} - d_i)^2 \\ &= \frac{1}{(d + \theta_2)^3} (d_{i+1} - d_i)^2 \stackrel{!}{=} h_1 / \tilde{c}_2 = \tilde{h}_1 \quad \text{mit } i = 1, \dots, k - 1. \end{aligned}$$

B.2.8. Die sigmoidale Emax-Wirkungskurve

Die für die Obergrenze der linearen und kubischen Spline-Interpolation benötigten zweiten und vierten Ableitungen der sigmoidalen Emax-Wirkungsfunktion $f_{\theta}(d) = \theta_0 + \theta_1 / (1 + (\theta_2/d)^{\theta_3})$ können wie folgt beschrieben werden:

$$\frac{\partial^2 f_{\theta}}{\partial d^2} = \theta_1 \theta_3 \left[\frac{(2\theta_3) \left(\frac{\theta_2}{d}\right)^{2\theta_3}}{\left(\left(1 + \left(\frac{\theta_2}{d}\right)^{\theta_3}\right)^3\right) \cdot d^2} - \frac{(\theta_3 + 1) \left(\frac{\theta_2}{d}\right)^{\theta_3}}{\left(\left(1 + \left(\frac{\theta_2}{d}\right)^{\theta_3}\right)^2\right) \cdot d^2} \right]$$

und

$$\begin{aligned} \frac{\partial^4 f_{\theta}}{\partial d^4} &= \frac{\theta_1 \theta_3}{d^4} \cdot [(\theta_2/d)^{\theta_3} (-\theta_3^3 - 6\theta_3^2 - 11\theta_3 - 6) / ((\theta_2/d)^{\theta_3} + 1)^2 \\ &\quad + (\theta_2/d)^{2\theta_3} (14\theta_3^3 + 36\theta_3^2 + 22\theta_3) / ((\theta_2/d)^{\theta_3} + 1)^3 \\ &\quad + (\theta_2/d)^{3\theta_3} (-36\theta_3^3 - 66\theta_3^2) / ((\theta_2/d)^{\theta_3} + 1)^4 \\ &\quad + (\theta_2/d)^{4\theta_3} (25\theta_3^3) / ((\theta_2/d)^{\theta_3} + 1)^5]. \end{aligned}$$

Selbst wenn die Ableitungen deutlich komplexer als im Falle einer Emax-Wirkungsfunktion sind, ist es immer noch leicht ersichtlich, dass die Wahl der Dosisstufen für die lineare und kubische Spline-Interpolation nicht von der Steigung θ_1 und den Achsenabschnitt θ_0 abhängen. Allerdings werden Annahmen an die ED_{50} (θ_2) und an der Grad der sigmoidalen Krümmung θ_3 für die Bestimmung der optimalen Dosisallokation der beiden Spline-Interpolationen benötigt.

Literaturverzeichnis

- Beal, S. L. (1989). Sample size determination for confidence intervals on the population mean and on the difference between two population means. *Biometrics*, 45(3):969–977.
- Berkson, J. (1969). Estimation of a linear function for a calibration line; consideration of a recent proposal. *Technometrics*, 11:649–660.
- Bornkamp, B., Bretz, F., Dette, H., und Pinheiro, J. (2011). Response-adaptive dose-finding under model uncertainty. *The Annals of Applied Statistics*, 5(2B):1611–1631.
- Bornkamp, B., Bretz, F., Dmitrienko, A., Enas, G., Gaydos, B., Hsu, C.-H., König, F., Krams, M., Liu, Q., Neuenschwander, B., Parke, T., Pinheiro, J., Roy, A., Sax, R., und Shen, F. (2007). Innovative approaches for designing and analyzing adaptive dose-ranging trials. *Journal of Biopharmaceutical Statistics*, 17(6):965–995.
- Bretz, F., Dette, H., und Pinheiro, J. C. (2010). Practical considerations for optimal designs in clinical dose finding studies. *Statistics in Medicine*, 29(7-8):731–742.
- Bretz, F., Hsu, J., Pinheiro, J., und Liu, Y. (2008). Dose finding - a challenge in statistics. *Biometrical Journal*, 50(4):480–504.
- Bretz, F., Pinheiro, J. C., und Branson, M. (2005). Combining multiple comparisons and modeling techniques in dose-response studies. *Biometrics*, 61(3):738–748.
- Brunner, E. (2005). Angewandte Statistik II. Vorlesungsskript. Georg-August-Universität Göttingen.
- Budtz-Jørgensen, E. (2007). Estimation of the benchmark dose by structural equation models. *Biostatistics*, 8(4):675–688.
- Chapple, C., Araño, P., Bosch, J., De Ridder, D., Kramer, A., und Ridder, A. (2004). Solifenacin appears effective and well tolerated in patients with symptomatic idiopathic detrusor overactivity in a placebo- and tolterodine-controlled phase 2 dose-finding study. *BJU International*, 93(1):71–77.
- Chow, S.-C. und Liu, J.-P. (2004). *Design and Analysis of Clinical Trials: Concepts and Methodologies, Second Edition*. Wiley Series in Probability and Statistics.

- Craig, A. T. (1943). Note on the independence of certain quadratic forms. *The Annals of Mathematical Statistics*, 14(2):195–197.
- de Boor, C. (2001). *A Practical Guide to Splines*. Number 27 in Applied Mathematical Sciences. Springer, New York.
- Dehling, H. und Haupt, B. (2004). *Einführung in die Wahrscheinlichkeitstheorie und Statistik*. Springer, Berlin.
- Dette, H., Bornkamp, B., und Bretz, F. (2013). On the efficiency of two-stage response-adaptive designs. *Statistics in Medicine*, 32(10):1646–1660.
- Dette, H., Bretz, F., Pepelyshev, A., und Pinheiro, J. (2008). Optimal designs for dose-finding studies. *Journal of the American Statistical Association*, 103(483):1225–1237.
- Dette, H., Kiss, C., Benda, N., und Bretz, F. (2014). Optimal designs for dose finding studies with an active control. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 76(1):265–295.
- Dilleen, M., Heimann, G., und Hirsch, I. (2003). Non-parametric estimators of a monotonic dose-response curve and bootstrap confidence intervals. *Statistics in Medicine*, 22:869–882.
- Dragalin, V., Bornkamp, B., Bretz, F., Miller, F., Padmanabhan, S. K., Patel, N., Perevozskaya, I., Pinheiro, J., und Smith, J. R. (2010). A simulation study to compare new adaptive dose-ranging designs. *Statistics in Biopharmaceutical Research*, 2(4):487–512.
- Dragalin, V., Hsuan, F., und Padmanabhan, S. K. (2007). Adaptive designs for dose-finding studies based on sigmoid Emax model. *Journal of Biopharmaceutical Statistics*, 17(6):1051–1070.
- Dunnett, C. W. (1955). A multiple comparison procedure for comparing several treatments with a control. *Journal of the American Statistical Association*, 50(272):1096–1121.
- Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *The Annals of Statistics*, 7(1):1–26.
- Efron, B. (1987). Better bootstrap confidence intervals. *Journal of the American Statistical Association*, 82(397):171–185.
- Eisenhart, C. (1939). The interpretation of certain regression methods and their use in biological and industrial research. *The Annals of Mathematical Statistics*, 10:162–186.
- EMA (2014). Qualification opinion of MCP-Mod as an efficient statistical methodology for model-based design and analysis of phase II dose finding studies under model uncertainty. http://www.ema.europa.eu/docs/en_GB/document_library/Regulatory_and_procedural_guideline/2014/02/WC500161027.pdf.

- Ferguson, T. (1996). *A Course in Large Sample Theory*. Chapman & Hall, London.
- Fisch, R. D. und Strehlau, G. A. (1993). A simplified approach to calibration confidence sets. *The American Statistician*, 47(3):168–171.
- Floater, M. (2007). Inf-mat5340 - Spline methods. Lecture notes. University of Oslo.
<http://www.uio.no/studier/emner/matnat/ifi/INF-MAT5340/v07/undervisningsmateriale/>.
- Friede, T., Mitchell, C., und Müller-Velten, G. (2007). Blinded sample size reestimation in non-inferiority trials with binary endpoints. *Biometrical Journal*, 49(6):903–916.
- Graybill, F. A. (1958). Determining sample size for a specified width confidence interval. *The Annals of Mathematical Statistics*, 29(1):282–287.
- Grieve, A. P. und Beal, S. L. (1991). Confidence intervals and sample sizes. *Biometrics*, 47(4):1597–1603.
- Hall, C. A. (1968). On error bounds for spline interpolation. *Journal of Approximation Theory*, 1:209–218.
- Hall, C. A. und Meyer, W. W. (1976). Optimal error bounds for cubic spline interpolation. *Journal of Approximation Theory*, 16:105–122.
- Helms, H.-J. (2011). Robuste Verfahren für strukturierte hochdimensionale Repeated-Measures-Designs unter Nicht-Normalverteilung. Master’s thesis, Universität Göttingen.
- Helms, H.-J., Benda, N., und Friede, T. (2012). Auswertung und Fallzahlberechnung in Dosis-Findungs-Studien mit aktiver Kontrolle mit SAS/STAT und SAS/IML. In *Proceedingsband 16. KSFE*. http://saswiki.org/images/8/83/16._KSFE_2012_-_Helms_Benda_Friede_-_Auswertung_und_Fallzahlberechnung_in_Dosis-Findungs-Studien_mit_aktiver_Kontrolle_mit_SAS_STAT_und_SAS_IML.pdf.
- Helms, H.-J., Benda, N., und Friede, T. (2014a). Point and interval estimators of the target dose in clinical dose-finding studies with active control. *Journal of Biopharmaceutical Statistics*. in press.
- Helms, H.-J., Benda, N., Zinserling, J., und Friede, T. (2013). Auswertung von Dosis-Findungs-Studien mit aktiver Kontrolle in SAS 9.3. In *Proceedingsband 17. KSFE*. http://saswiki.org/images/a/ab/17._KSFE_2013_Helms_-_Auswertung_von_Dosis-Findungs-Studien.pdf.
- Helms, H.-J., Benda, N., Zinserling, J., Kneib, T., und Friede, T. (2014b). Spline-based procedures for dose-finding studies with active control. submitted.

- Hsu, J. C. und Berger, R. L. (1999). Stepwise confidence intervals without multiplicity adjustment for dose-response and toxicity studies. *Journal of the American Statistical Association*, 94(446):468–482.
- Jacobson, N. (2009). *Basic algebra 1 (2nd ed.)*. Dover Publications.
- Janssen, A. und Pauls, T. (2003). How do bootstrap and permutation tests work? *The Annals of Statistics*, 31(3):768–806.
- Jones, B., Layton, G., Richardson, H., und Thomas, N. (2011). Model-based bayesian adaptive dose-finding designs for a phase 2 trial. *Statistics in Biopharmaceutical Research*, 3(2):276–287.
- Julious, S. A. und Owen, R. J. (2006). Sample size calculations for clinical studies allowing for uncertainty about the variance. *Pharmaceutical Statistics*, 5(1):29–37.
- Källén, A. und Larsson, P. (1999). Dose response studies: How do we make them conclusive? *Statistics in Medicine*, 18(6):629–641.
- Kannan, N., Keating, J. P., und Mason, R. L. (2007). A comparison of classical and inverse estimators in the calibration problem. *Communications in Statistics - Theory and Methods*, 36:83–95.
- Kirby, S., Brain, P., und Jones, B. (2011). Fitting Emax models to clinical trial dose-response data. *Pharmaceutical Statistics*, 10(2):143–149.
- Kirby, S., Colman, P., und Morris, M. (2009). Adaptive modelling of dose-response relationships using smoothing splines. *Pharmaceutical Statistics*, 8(4):346–355.
- Kruskal, W. (1968). When are Gauss-Markov and least squares estimators identical? A coordinate-free approach. *The Annals of Mathematical Statistics*, 39(1):70–75.
- Krutchkoff, R. G. (1967). Classical and inverse regression methods of calibration. *Technometrics*, 9(3):425–439.
- Lehmann, E. (1959). *Testing Statistical Hypotheses*. Wiley, New York.
- Mattai, A. M. und Provost, S. B. (1992). *Quadratic Forms in Random Variables: Theory and Applications*. Marcel Dekker, New York.
- Moore, N., Senant, J., Richard, M., Grieve, A., und Day, S. (1989). Confidence intervals and trail sizes. *The Lancet*, 333(8633):337–338.
- Moré, J. J. (1978). The Levenberg-Marquardt algorithm: Implementation and theory. In Watson, G., editor, *Numerical Analysis*, volume 630 of *Lecture Notes in Mathematics*, pages 105–116. Springer, Berlin Heidelberg.

- Morales, K., Ibrahim, J., Chen, C., und Ryan, L. (2006). Bayesian model averaging with applications to benchmark dose estimation for arsenic in drinking water. *Journal of the American Statistical Association*, 101(473):9–17.
- Nauck, M., Frid, A., Hermansen, K., Shah, N., Tankova, T., Mitha, I., Zdravkovic, M., Düring, M., und Matthews, D. (2009). Efficacy and safety comparison of liraglutide, glimepiride, and placebo, all in combination with metformin, in type 2 diabetes: the LEAD (liraglutide effect and action in diabetes)-2 study. *Diabetes Care*, 32:84–90.
- Pawitan, Y. (2001). *In all Likelihood: Statistical Modelling and Inference Using Likelihood*. Oxford University Press.
- Piegl, L. und Tiller, W. (1997). *The NURBS Book*. Springer, Berlin Heidelberg.
- Pinheiro, J., Bornkamp, B., und Bretz, F. (2006a). Design and analysis of dose-finding studies combining multiple comparisons and modeling procedures. *Journal of Biopharmaceutical Statistics*, 16(5):639–656.
- Pinheiro, J., Bornkamp, B., Glimm, E., und Bretz, F. (2014). Model-based dose finding under model uncertainty using general parametric models. *Statistics in Medicine*, 33:1646–1661. online version.
- Pinheiro, J., Bretz, F., und Branson, M. (2006b). Analysis of dose-response studies—modeling approaches. In *Dose Finding in Drug Development*, Statistics for Biology and Health, pages 146–171. Springer, New York.
- Preacher, K. J., Rucker, D. D., und Hayes, A. F. (2007). Addressing moderated mediation hypotheses: Theory, methods, and prescriptions. *Multivariate Behavioral Research*, 42(1):185–227.
- Ravishanker, N. und Dey, D. K. (2002). *A First Course in Linear Model Theory*. Chapman & Hall/CRC, Boca Raton, FL.
- Ruberg, S. J. (1995a). Dose response studies I. Some design considerations. *Journal of Biopharmaceutical Statistics*, 5(1):1–14.
- Ruberg, S. J. (1995b). Dose response studies II. Analysis and interpretation. *Journal of Biopharmaceutical Statistics*, 5(1):15–42.
- SAS-Institute (2008a). *SAS/IML 9.2 User’s Guide*. Cary, NC: SAS Institute, USA. <http://support.sas.com/documentation/cdl/en/imlug/59656/PDF/default/imlug.pdf>.
- SAS-Institute (2008b). *SAS/STAT 9.2 User’s Guide*. Cary, NC: SAS Institute, USA. <http://support.sas.com/documentation/cdl/en/statug/63033/PDF/default/statug.pdf>.

- Schöbel, A. (2007). Numerik I. Vorlesungsskript. Georg-August-Universität Göttingen. <http://optimierung.math.uni-goettingen.de/skripte/numerik1.pdf>.
- Seber, G. A. F. und Wild, C. J. (1989). *Nonlinear Regression*. Wiley Series in Probability and Mathematical Statistics. Wiley, New York.
- Sheiner, L. (1997). Learning versus confirming in clinical drug development. *Clinical Pharmacology and Therapeutics*, 61:275–291.
- Smajic, A. (2009). Optimale Box-Einschließungen von NURBS Kurven. Master’s thesis, Universität Wien. http://othes.univie.ac.at/6783/1/2009-10-05_0300959.pdf.
- Stoer, J. (1993). *Numerische Mathematik 1*. Springer Verlag, Berlin.
- Tamhane, A. C. und Logan, B. R. (2002). Multiple test procedures for identifying the minimum effective and maximum safe doses of a drug. *Journal of the American Statistical Association*, 97(457):293–301.
- Thonnard, M. (2006). Confidence intervals in inverse regression. Master’s thesis, Eindhoven University of Technology. <http://alexandria.tue.nl/extra2/afstvers1/wsk-i/thonnard2006.pdf>.
- Wahba, G. (1990). *Spline Models for Observational Data*. Society for Industrial and Applied Mathematics, Philadelphia.
- Williams, E. J. (1969). Regression methods in calibration problems. *The International Statistical Institute*, 43:17–28.

Publikationen, Vorträge und Poster zum Thema der Dissertation

Publikationen

Ein großer Teil der Resultate der vorliegenden Arbeit sind veröffentlicht oder eingereicht in:

- Helms, H.-J. und Friede, T.. Sample size calculation for dose-finding studies with active control. Biometrical Journal (in preparation).
- Helms, H.-J., Benda, N., Zinserling, J., Kneib, T. und Friede, T. (2014). Spline-based procedures for dose-finding studies with active control. Statistics in Medicine (submitted).
- Helms, H.-J., Benda, N., und Friede, T. (2014). Point and interval estimators of the target dose in clinical dose-finding studies with active control. Journal of Biopharmaceutical Statistics (in press).

Tagungsband-Beiträge

- Proceedingsband 17. KSFE, Ulm (2013): Auswertung von Dosis-Findungs-Studien mit aktiver Kontrolle in SAS 9.3.
- Proceedingsband 16. KSFE, Dresden (2012): Auswertung und Fallzahlberechnung in Dosis-Findungs-Studien mit aktiver Kontrolle mit SAS/NLMIXED und SAS/IML.

Vorträge

- 60. Biometrisches Kolloquium, Bremen (2014): Spline based analysis and design of dose-finding studies with active control.
- DAGStat, Freiburg (2013): Analysis of dose-finding studies with active control.
- KSFE, Ulm (2013): Auswertung von Dosis-Findungs-Studien mit aktiver Kontrolle in SAS 9.3. KSFE-Award 3. Platz.
- 58. Biometrisches Kolloquium, Berlin (2012): Auswertung und Fallzahlberechnung in Dosis-Findungs-Studien mit aktiver Kontrolle.
- KSFE, Dresden (2012): Auswertung und Fallzahlberechnung in Dosis-Findungs-Studien mit aktiver Kontrolle mit SAS/NLMIXED und SAS/IML.

Poster

- ISCB, München (2013): Spline-based procedures for dose-finding studies with active control.