

Monitoring von klinischen Studien mit rekurrierenden Ereignissen

20-wöchige Abschlussarbeit an der wirtschaftswissenschaftlichen Fakultät
im Rahmen der Prüfung im Studiengang Angewandte Statistik
an der Georg-August-Universität Göttingen

Masterarbeit zur Erlangung des akademischen Grades
Master of Science (M. Sc.)

vorgelegt von Susanna Salem
aus Berlin (20250670)

- 1. Betreuer:** Prof. Dr. Tim Friede
- 2. Betreuer:** PD Dr. Norbert Benda

vorgelegt am: 23.07.2018

Inhaltsverzeichnis

Abbildungsverzeichnis	ii
Tabellenverzeichnis	iii
1 Einleitung	1
1.1 Verblindetes kontinuierliches Monitorieren	2
1.2 Zeitliche Trends in den Schubraten bei RRMS	4
1.3 Aufgabenstellung	4
1.4 Struktur der Masterarbeit	5
2 Statistische Modelle, Hypothesentest und Fallzahlplanung	6
2.1 Inhomogene Poisson Prozesse	6
2.1.1 Simulation eines inhomogenen Poisson Prozesses	7
2.2 Modell	7
2.2.1 Zeittrend	7
2.2.2 Behandlungseffekt	8
2.2.3 Überdispersion	9
2.2.4 <i>Likelihood Funktion</i> und <i>Log-Likelihood Funktion</i>	9
2.3 Hypothesentests	10
2.4 Fallzahlplanung	11
3 Verblindetes kontinuierliches Monitorieren	12
3.1 Stop-Kriterium	12
3.2 Information \mathcal{I}_{ML}	12
3.3 Verblindete Schätzung der Information	13
4 Implementierung und Szenarien der Simulationsstudie	16
4.1 Datengenerierung	16
4.1.1 Aufbau der Simulationsfunktion	17
4.2 Verblindetes kontinuierliches Monitorieren	18
4.3 Parameterschätzung und Wald Test	20
4.4 Szenarien	21
5 Ergebnisse der Simulationsstudie	23
5.1 Wahrscheinlichkeit des Fehlers 1. Art	23
5.2 Überdeckungswahrscheinlichkeiten der 95%-Konfidenzintervalle für den Effektschätzer	25

5.3	Bias	27
5.4	Power, Studienlaufzeiten und Fallzahlen	33
6	Diskussion	42
	Literaturverzeichnis	46
	Anhang	48

Abbildungsverzeichnis

1	Graphische Darstellung der Intensitätsfunktion	8
2	Schematischer Ablauf des verblindeten kontinuierlichen Monitorierens . .	22
3	Verteilung der Wahrscheinlichkeit des Fehlers 1. Art	25
4	Wahrscheinlichkeit des Fehlers 1. Art in den einzelnen Szenarien	26
5	Überdeckungswahrscheinlichkeiten der 95%-Konfidenzintervalle für den Effektschätzer	28
6	Überdeckungswahrscheinlichkeiten der 95%-Konfidenzintervalle für den Effektschätzer für einzelne Szenarien	29
7	Bias des Effektschätzers	30
8	Bias des Schätzers für α_0 im Trendmodell	31
9	Bias des Schätzers für den Trendparameter α_1 im Trendmodell	32
10	Power für BCM-CONST und BCM-TREND	36
11	Bias des Schätzers für φ im Trendmodell	38
12	Quantile der Studienlaufzeiten für kleine Zeittrends	39
13	Quantile der Studienlaufzeiten für große Zeittrends	40
14	Quantile der Fallzahlen für BCM-CONST und BCM-TREND.	41

Tabellenverzeichnis

1	Prozentsatz der Fälle in denen wegen fehlender Überdispersion ein Poisson Modell in BCM-CONST angepasst wurde oder $\hat{\varphi} < 1e^{-6}$ bei BCM-TREND-	19
2	Prozentsatz der Fälle in denen die Schätzungen in BCM-TREND <code>optim()</code> mit den ersten Startwerten nicht konvergierten und neue Startwerte geschätzt wurden.	20
3	Prozentsatz der Fälle in denen der geschätzte Parameter für den Zeittrend in BCM-TREND sehr nahe null lag.	20
4	Simulierte Szenarien	21
5	Wahrscheinlichkeit des Fehlers 1. Art für BCM-CONST und BCM-TREND	24
6	Überdeckungswahrscheinlichkeiten der 95%-Konfidenzintervalle für BCM-CONST und BCM-TREND unter H_0	27
7	Überdeckungswahrscheinlichkeiten der 95%-Konfidenzintervalle für BCM-CONST und BCM-TREND unter H_1	27
8	Mittelwert, Standardabweichung, Minimum und Maximum des Bias für $\hat{\beta}_1$	33
9	Power im Planungsszenario für kleine Zeittrends	34
10	Power im Planungsszenario für große Zeittrends	34
11	Power für Abweichungen vom Planungsszenario für große Zeittrends. . .	37

1 Einleitung

Ein Ziel der Planung klinischer Studien ist, die Entscheidung darüber wie viele Patienten rekrutiert und über welchen Zeitraum sie gegebenenfalls nachbeobachtet werden, so optimal wie möglich zu treffen. Es sollen genügend Patienten in die Studie aufgenommen werden, um eine bestimmte Aussagekraft der Ergebnisse zu erzielen, aber auch nicht zu viele Patienten zu lange in der Studie gehalten werden, um Patienten keinem unnötigen Risiko auszusetzen. Der zweite Punkt, also die geringste nötige Fallzahl zu ermitteln, spielt gerade bei klinischen Studien mit besonders zu schützenden Studienteilnehmern, wie zum Beispiel Kindern, oder im Fall von seltenen Erkrankungen eine wichtige Rolle.

Um eine in diesem Sinne optimale Fallzahl zu planen, ist es notwendig einige sogenannte Störparameter vor Beginn der Studie abzuschätzen. Im Fall von rekurrierenden Ereignisdaten sind das unter anderem die erwartete Basisrate und, falls man eine solche erwartet, die Höhe der Überdispersion. Die Basisrate ist die Ereignisrate, die man in einer Wirksamkeitsstudie für die Vergleichsbehandlung, also die Behandlung mit der beispielsweise eine neue Behandlung verglichen werden soll, erwartet. Der Begriff Überdispersion besagt, dass die Varianz der Ereignisraten größer ist als deren Erwartungswert (Cullagh u. Nelder, 1983, S.131). Überdispersion kann auf verschiedene Arten entstehen, zum Beispiel durch positive Korrelationen zwischen den Ereignissen der untersuchten Individuen (zum Beispiel Angehörigkeit zum selben Haushalt) oder durch Heterogenität, die man in Bezug auf die Basisraten zwischen den Patienten erwartet (Fahrmeir u. a., 2009, S.197). Wenn Überdispersion in den rekurrierenden Ereignissen vorliegt, können diese durch eine negative Binomialverteilung modelliert werden, bei der die Anzahl der Ereignisse für jeden einzelnen Patienten einer Poissonverteilung folgt und die Verteilung der Ereignisraten zwischen den Patienten als Gamma-verteilt angenommen wird (Cullagh u. Nelder, 1983, S.132). Eine genaue Abschätzung dieser Parameter gestaltet sich in der Praxis häufig als schwierig, weil in der Regel keine ausreichenden Informationen aus Vorstudien vorliegen beziehungsweise sich diese nicht einfach auf die untersuchte Population oder das verwendete Design übertragen lassen. Deshalb kann es sinnvoll sein Interimsanalysen durchzuführen, um die Fallzahl nach Beginn der Studie an den spezifischen Charakteristika des untersuchten Patientenkollektivs auszurichten.

1.1 Verblindetes kontinuierliches Monitorieren

Ein adaptives Design umfasst alle Studiendesigns, bei denen in mehreren Stufen bereits im Studienverlauf gewonnene Daten verwendet werden, um Aspekte der Studie zu verändern (zum Beispiel Fallzahl, Randomisierungsverhältnis, Anzahl der Behandlungsgruppen), ohne dass dabei die Integrität und Interpretierbarkeit der Studienergebnisse auf irgendeine Weise negativ beeinträchtigt werden, siehe Bauer u. a. (2016).

Die Möglichkeit zur Fallzahlanpassung im Verlauf einer Studie stellte eine wichtige Motivation für die Einführung von adaptiven Designs dar (Bauer u. a., 2016). Es existieren verschiedene Ansätze, um die Fallzahl oder Laufzeit einer Studie durch eine Interimsanalyse, in der die Störparameter aus den bereits vorhandenen Daten geschätzt werden, anzupassen. Ziel einer solchen Interimsanalyse ist es, die Laufzeit oder Fallzahl bei Abweichungen der in der Studie beobachteten Störparameter von den im Planungsstadium der Studie angenommenen Störparametern so anzupassen, dass die vorgegebene Power erreicht wird. Wie bei allen adaptiven Designs muss sichergestellt werden, dass dadurch keine Inflation des Fehlers 1. Art eintritt.

Die Wahrscheinlichkeit eines Fehlers 1. Art α ist definiert als die Wahrscheinlichkeit, die Nullhypothese in einem Experiment abzulehnen, obwohl sie tatsächlich wahr ist. Ein Hypothesentest hat eine bestimmte Wahrscheinlichkeit α für den Fehler 1. Art, und die Nullhypothese wird abgelehnt, sobald der p-Wert kleiner ist als α (Held u. Sabanés Bové, 2014). Im betrachteten Beispiel, einer klinischen Studie zur Wirksamkeit eines Medikaments, bedeutet ein Fehler 1. Art, sich dafür zu entscheiden ein Medikament als wirksam zu erachten, obwohl das Medikament in Wahrheit keinen Effekt hat (oder keinen einer bereits etablierten Behandlung überlegenen Effekt). Es ist offensichtlich, dass es gerade bei der Auswertung von klinischen Studien besonders wichtig ist, sicherzustellen, dass die Wahrscheinlichkeit des Fehlers 1. Art eingehalten wird, um Patienten nicht zu gefährden. Die Power eines statistischen Tests ist die Wahrscheinlichkeit, die Nullhypothese (auf einem bestimmten Signifikanzniveau α) zu verwerfen, falls ein Effekt einer bestimmten Größenordnung tatsächlich vorliegt. Auf das Beispiel einer klinischen Studie bezogen bedeutet es, die Wahrscheinlichkeit die Wirksamkeit eines Medikaments zu erkennen, falls es tatsächlich wirksam ist (oder eine höhere Wirksamkeit gegenüber einer Vergleichsbehandlung hat).

Eine Frage, die sich bei der Durchführung von Interimsanalysen stellt, ist, ob die Zuordnung der Patienten zu den Behandlungsgruppen für die Interimsanalyse „entblindet“ werden sollte. Durch eine verblindete Interimsanalyse kann die Gefahr eines Fehlers 1. Art reduziert werden (Friede u. Miller, 2012; Chuang-Stein u. a., 2006). Daher ist ei-

ne verblindete Schätzung der Störparameter für viele Verfahren zur Fallzahlanpassung empfehlenswert. Auch aus logistischer Sicht betrachtet, stellt eine verblindete Schätzung einen geringeren Aufwand für die Durchführenden der Studie dar, da kein unabhängiges Datenmonitorierungskomitee, eine unabhängige Expertengruppe, welche die Patientensicherheit und Daten zur Wirksamkeit einer Behandlung im Verlauf einer klinischen Studie überwachen, eingesetzt werden muss, um die Zuordnung der Patienten zu den Behandlungsgruppe für die Fallzahlanpassung zu entblinden. Randomisierte, kontrollierte klinische Studien sind in der Regel Blindstudien. In einer Doppelblindstudie wissen weder die Patienten, noch das klinische Personal welcher Patient welche Behandlung erhält. Dies soll die Objektivität der an der Studie beteiligten Personen gewährleisten, so dass keine Möglichkeit besteht, dass die Ergebnisse durch das Wissen um die Zugehörigkeit der Patienten zu einer bestimmte Behandlung beeinflusst werden.

Führt man im Verlauf der Studie zu einem einzigen Zeitpunkt eine Schätzung der Störparameter anhand der bis zu diesem Zeitpunkt erhobenen Daten durch, und berechnet davon ausgehend eine neue Fallzahl ist das ein Verfahren zur Fallzahlanpassung. Eine Erweiterung auf mehrere Schätzzeitpunkte, also eine wiederholte Fallzahlanpassung, ist ebenfalls möglich. Führt man im Zeitverlauf der Studie zu kontinuierlich aufeinanderfolgenden Zeitpunkten eine verblindete Schätzung der Information des Effektschätzers durch, spricht man vom verblindeten kontinuierlichen Monitorieren (*blinded continuous monitoring*, kurz BCM). Beim kontinuierlichen Monitorieren wird die Studie gestoppt, sobald ein Stop-Kriterium für die Information erreicht wird. Ein Vorteil der wiederholten, kontinuierlichen Schätzung gegenüber Verfahren mit Fallzahlanpassung ist eine geringere Varianz in den resultierenden Fallzahlen bei etwa gleichen erwarteten Fallzahlen (Friede u. Miller, 2012).

Friede u. a. (2018) entwickelten eine Methode zum verblindeten kontinuierlichen Monitorieren klinischer Studien mit negativ binomialverteilten Ereignisdaten als Endpunkt. In einer Simulationsstudie zeigten sie, dass die Methode für die von ihnen untersuchten Szenarien unter Einhaltung der Wahrscheinlichkeit des Fehlers 1. Art und der Power im Mittel zu kürzeren Studienlaufzeiten führte. Wurde die Rate zu Anfang der Studie in der Planung korrekt eingeschätzt, verkürzte sich die Laufzeit im Mittel um ca. 7%. War die tatsächliche Rate doppelt so hoch wie bei der Planung vermutet, verkürzte sich die Laufzeit der Studien im Mittel um ca. 40%. Bei den simulierten Szenarien orientierten sich Friede u. a. (2018) am Beispiel einer klinischen Studie zur schubförmig remittierenden Multiplen Sklerose (RRMS) im Kindesalter mit jährlichen Schubraten als Endpunkt.

1.2 Zeitliche Trends in den Schubraten bei RRMS

Bei der von Friede u. a. (2018) entwickelten Methode wird davon ausgegangen, dass sich die jährlichen Raten der Ereignisse (d.h. Schübe) über die Zeit nicht verändern. Die Schubrate bei Multipler Sklerose bleibt jedoch im Verlauf der Erkrankung nicht konstant. Tremlett u. a. (2008) berichten, dass sich ab dem zwanzigsten Lebensjahr die Schubrate bei RRMS alle 5 Jahre um circa 17% verringert. In einer Metaanalyse über 13 placebo-kontrollierte klinische Studien zur schubförmigen Multiplen Sklerose (RMS) berichten Nicholas u. a. (2012), dass die jährlichen Schubraten vom ersten zum zweiten Jahr deutlich abnehmen. Dabei wurden aus 13 klinischen Studien jeweils die Raten des ersten und zweiten Jahres sowie der ersten beiden Halbjahre verglichen. Nicholas u. a. (2012) stellten eine signifikante Verringerung der Schubraten innerhalb der klinischen Studien vom ersten zum zweiten Jahr sowohl in der Placebogruppe als auch der Behandlungsgruppe fest. Dieser allgemeine negative Zeittrend wird damit erklärt, dass Patienten zu einem Zeitpunkt rekrutiert werden, an dem gerade sehr viele Schübe auftreten und die Schubraten im Laufe der Studie wieder auf ihr durchschnittliches Niveau zurückgehen (*“regression to the mean”*).

Eine Methode zur Fallzahlanpassung für rekurrende Ereignisdaten mit zeitlichem Trend in den Raten wurde von Schneider u. a. (2013) vorgestellt und in einer Simulationsstudie untersucht. Die berichteten Ergebnisse zeigen, dass eine Fallzahlanpassung mit Erreichen der Power und ohne Inflation der Wahrscheinlichkeit des Fehlers 1. Art erreicht wird. Das dort beschriebene Modell der Daten wurde als Grundlage für die vorliegende Simulationsstudie genommen, um das verblindete kontinuierliche Monitorieren von rekurrenden Ereignisdaten mit Zeittrend zu untersuchen.

1.3 Aufgabenstellung

Methoden zum verblindeten kontinuierlichen Monitorieren werden angewendet, um in Bezug auf Laufzeit und Fallzahl effiziente Studiendesigns zu erhalten. Es soll untersucht werden, ob das von Friede u. a. (2018) vorgeschlagene verblindete kontinuierliche Monitorieren von rekurrenden Ereignisdaten zuverlässige Ergebnisse liefert, selbst wenn ein negativer zeitlicher Trend in den Raten vorhanden ist, der beim Monitorieren und in der Auswertung unberücksichtigt bleibt. Weiterhin soll die Methode erweitert werden auf ein Modell für rekurrende Ereignisdaten mit negativem Zeittrend und Gamma-verteilten Ereignisraten. Beide Methoden sollen in R implementiert werden und in einer Simulationsstudie verglichen werden. Beide Methoden werden in Bezug auf die Power, die

Wahrscheinlichkeit des Fehlers 1. Art und die durchschnittlichen Laufzeiten der Studien nach Monitorieren verglichen.

1.4 Struktur der Masterarbeit

Es wird der Fall betrachtet, in dem in einer klinischen Studie die Anzahl von wiederholt auftretenden Ereignissen (z.B. Schüben bei multipler Sklerose) erfasst werden sowie die Zeitpunkte zu denen die Ereignisse auftreten. Der Schweregrad der Erkrankung variiert zwischen den Patienten. Zusätzlich dazu nimmt die Rate, mit der Ereignisse auftreten, im Verlauf der Studie für alle Patienten und beide Behandlungsgruppen in gleichem Maße ab. Solche Fälle wurden zum Beispiel für klinische Studien zur Multiplen Sklerose berichtet (Nicholas u. a., 2012). Es werden also Ereignisdaten mit heterogenen Raten und einem negativen zeitlichen Trend in den Raten betrachtet. Daten mit den beschriebenen Eigenschaften können durch einen inhomogenen Poisson Prozess modelliert werden (Lawless, 1987b; Schneider u. a., 2013).

Im folgenden Kapitel wird das statistische Modell, das der Simulation und Auswertung der Daten zugrunde gelegt wurde, erläutert, sowie der verwendete Hypothesentest und die Formel zur Fallzahlplanung. In Kapitel 3 wird die Methode zur Simulation der Daten und die Implementierung des kontinuierlichen verblindeten Monitorierens in R dargestellt. Im darauffolgenden Kapitel 4 werden die Ergebnisse beschrieben. Im letzten Kapitel wird eine kurze Zusammenfassung der Ergebnisse gegeben und Einschränkungen der durchgeführten Untersuchung und mögliche Fragestellungen für weiterführende Studien diskutiert.

2 Statistische Modelle, Hypothesentest und Fallzahlplanung

2.1 Inhomogene Poisson Prozesse

Mehrere Möglichkeiten existieren, um einen inhomogenen Poisson Prozess zu definieren. Eine gebräuchliche Definition legt den inhomogenen Poisson Prozess als stochastischen Prozess mit folgenden Eigenschaften fest, wobei $N(t)$ die Anzahl der Ereignisse zum Zeitpunkt t ist:

1. $N(0) = 0$
2. $\{N(t), t \geq 0\}$ hat unabhängige Zuwächse
3. $P\{N(t+h) - N(t) \geq 2\} = o(h)$
4. $P\{N(t+h) - N(t) = 1\} = \lambda(t)h + o(h)$

und der Term $o(h)$ eine Größe darstellt, die schneller gegen null strebt als h , wenn $h \rightarrow 0$. Das bedeutet eine Funktionen $f(h)$ ist $o(h)$, wenn $\lim_{h \rightarrow 0} \left\{ \frac{f(h)}{h} = 0 \right\}$ (Ross, 2008, Kap. 2).

Aus der Definition ergibt sich, dass die Anzahl der Ereignisse $N(t)$ aus einem solchen Prozess, die in einem Zeitintervall s auftreten, Poisson-verteilt ist mit :

$$P\{N(t+s) - N(t) = n\} = \exp(-(\Lambda(t+s) - \Lambda(t))) \frac{(\Lambda(t+s) - \Lambda(t))^n}{n!}, \quad n \geq 0 \quad (1)$$

wobei $\Lambda(T) = \int_0^T \lambda(t) dt$ die kumulative Rate bezeichnet (Ross, 2008, S.79).

Im Fall einer nach oben beschränkten Intensitätsfunktion $\lambda(t)$ kann der inhomogene Poisson Prozess auch als Zufallsauswahl aus einem homogenen Poisson Prozess verstanden werden (Ross, 2008, Kap. 2). Da bei einem negativen Trend die Raten im Verlauf der Zeit immer kleiner werden, trifft dies auch auf die betrachteten Daten zu und es konnte der *thinning* Algorithmus verwendet werden, um die Daten zu simulieren (Stoyan u. a., 1995, S.45-46).

2.1.1 Simulation eines inhomogenen Poisson Prozesses

In der Literatur finden sich unterschiedliche Algorithmen, um Daten aus einem inhomogenen Poisson Prozess mit Intensitätsfunktion $\lambda(t)$ zu erzeugen, siehe zum Beispiel Pasupathy (2010). Beim sog. *thinning* werden im ersten Schritt Ereignisse T_1^*, \dots, T_n^* aus einem homogenen Poisson Prozess mit konstanter Rate λ^* erzeugt, wobei λ^* eine obere Schranke für $\lambda(t)$ ist. Im nächsten Schritt wird ein geeigneter Anteil dieser Ereignisse wieder „gelöscht“, gerade so, dass die Auftretenshäufigkeit der Ereignisse der gewünschten Rate $\lambda(t)$ entspricht. Das Löschen geschieht für jedes Ereignis unabhängig und mit der Wahrscheinlichkeit $1 - \frac{\lambda(t)}{\lambda^*}$. Ein Beweis dafür, dass dieses Vorgehen tatsächlich Daten gemäß der gewünschten Intensitätsfunktion $\lambda(t)$ erzeugt, findet man bei Lewis u. Shedler (1979).

Das Ausdünnen kann dabei praktisch mit Hilfe einer Folge von unabhängigen Zufallszahlen $[z_1, z_2, \dots]$ geschehen, welche aus einer Gleichverteilung über dem Intervall $[0, 1]$ stammen. Ein Ereignis T_k^* wird gelöscht, falls $z_k > \frac{\lambda(T_k^*)}{\lambda^*}$ (Stoyan u. a., 1995, S.45-46).

2.2 Modell

Für die Simulation und Auswertung der Daten wurde das in Schneider u. a. (2013) beschriebene statistische Modell übernommen, welches im Folgenden beschrieben wird.

2.2.1 Zeittrend

Die Basisintensitätsfunktion $\lambda_0(t, \theta)$ ist die Intensitätsfunktion der Kontrollgruppe. Sie hängt von der Zeit t und einem Parametervektor $\theta = (\alpha_0, \alpha_1)$, $\theta \in \mathbb{R}^d$ ab. Der Zusammenhang zwischen Zeit und Intensität wurde, wie bei Schneider u. a. (2013), mit der exponentiellen Funktion

$$\lambda_0(t, \theta) = \exp(\alpha_0 + \alpha_1 t) \quad (2)$$

festgelegt.

Bei einem negativen Zeittrend ($\alpha_1 < 0$) fallen die Raten also im Laufe der Zeit t exponentiell ab (siehe Abbildung 1). Dieses log-lineare Modell bietet den Vorteil, dass die Rate $\lambda_0(t, \theta)$ für alle α_0 und α_1 positiv ist (Cox u. Lewis, 1966, S. 45).

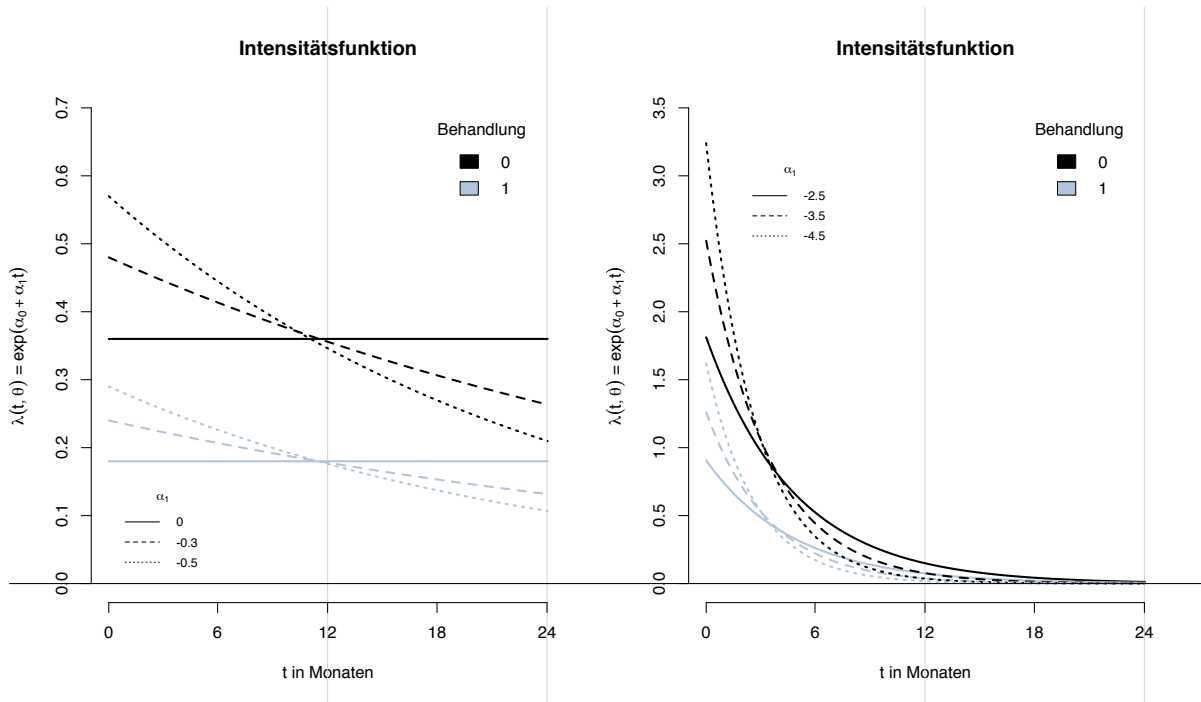


Abbildung 1: Intensitätsfunktion im Zeitverlauf für beide Behandlungsgruppen und verschiedene Werte von α_1 . Der Behandlungseffekt ist $\beta_1 = \log(0.5)$. Links: kleine Zeittrends, rechts: große Zeittrends.

2.2.2 Behandlungseffekt

Als Parametrisierung für den Behandlungseffekt β_1 wurde das logarithmierte Verhältnis der Raten aus den beiden Behandlungsgruppen verwendet: $\beta_1 = \log\left(\frac{\lambda_1(t)}{\lambda_0(t)}\right)$. Es wurde davon ausgegangen, dass der Behandlungseffekt über die Zeit konstant bleibt.

Daraus ergibt sich die proportionale Intensitätsfunktion mit $x_0 = 0$ und $x_{i1} = 1$ für $i = 0, 1$

$$\lambda_i(t, \theta, \beta_1) = \lambda_0(t, \theta) \exp(\beta_1 x_i) \quad (3)$$

und die kumulative Intensitätsfunktion

$$\Lambda_i(t, \theta, \beta_1) = \Lambda_0(t, \theta) \exp(\beta_1 x_i) = \frac{1}{\alpha_1} \exp(\alpha_0) (\exp(\alpha_1 t) - 1) \exp(\beta_1 x_i) \quad (4)$$

mit $\Lambda_0(t, \theta) = \int_0^t \lambda_0(u, \theta) du$.

2.2.3 Überdispersion

Die interindividuelle Heterogenität in den Raten wurde durch eine Gamma-verteilte Variable U_{ij} mit Erwartungswert 1 und Varianz φ modelliert ($U_{ij} \sim \Gamma(1, \varphi^{-1})$).

Die individuelle Intensitätsfunktion, bedingt auf die Variable U_{ij} , ist

$$\lambda_{ij}(t, \theta, \beta_1 | U_{ij} = u_{ij}) = u_{ij} \lambda_0(t, \theta) \exp(\beta_1 x_i) \quad (5)$$

und die bedingte kumulative Intensitätsfunktion

$$\Lambda_{ij}(t, \alpha_0, \alpha_1 | U_{ij} = u_{ij}) = u_{ij} \frac{1}{\alpha_1} \exp(\alpha_0) (\exp(\alpha_1 t) - 1) \exp(\beta_1 x_i) \quad (6)$$

(Schneider u. a., 2013).

2.2.4 Likelihood Funktion und Log-Likelihood Funktion

Es werden N Patienten zufällig und im Verhältnis $k : 1$ den Behandlungsgruppen ($i = 0, 1$) zugeordnet. Jeder Patient j ($j = 1, \dots, n_i$) wird über den Zeitraum $[0, T_{ij}]$ beobachtet. Im Beobachtungszeitraum treten wiederholt Ereignisse zu den Zeitpunkten $t_{ij1} < \dots < t_{ijY_{ij}}$ auf, wobei Y_{ij} die Gesamtzahl der Ereignisse für einen Patienten j in Gruppe i darstellt.

Die Likelihood $\mathcal{L}(\varphi, \theta, \beta_1)$ ist gegeben als

$$\mathcal{L}(\varphi, \theta, \beta_1) = \prod_{i=0}^1 \prod_{j=1}^{n_i} \left(\prod_{k=1}^{Y_{ij}} \frac{\lambda_0(t_{ijk}, \theta) \exp(x_i \beta_1)}{\Lambda_0(T_{ij}, \theta) \exp(x_i \beta_1)} \right) \times \frac{\Gamma(\varphi^{-1} + Y_{ij})}{\Gamma(\varphi^{-1}) Y_{ij}!} \frac{(\varphi \Lambda_0(T_{ij}, \theta) \exp(x_i \beta_1))^{Y_{ij}}}{(1 + \varphi \Lambda_0(T_{ij}, \theta) \exp(x_i \beta_1))^{Y_{ij} + \varphi - 1}} \quad (7)$$

wobei der zweite Teil die Likelihood eines Negativ-binomial Modells ist Lawless (1987b).

Die Log-Likelihood ist

$$\begin{aligned}
l(\alpha_0, \alpha_1, \beta_1, \varphi) &= \sum_{i=0}^1 \sum_{j=1}^{n_i} \sum_{k=1}^{Y_{ij}} [\alpha_0 + \alpha_1 t_{ijk}] + \sum_{i=0}^1 \sum_{j=1}^{n_i} Y_{ij} (\log(\varphi) + x_i \beta_1) \\
&+ \sum_{i=0}^1 \sum_{j=1}^{n_i} (\log(\Gamma(Y_{ij} + \varphi^{-1})) - \log(\Gamma(\varphi^{-1}))) \\
&- \sum_{i=0}^1 \sum_{j=1}^{n_i} [(Y_{ij} + \varphi^{-1}) \log(1 + \varphi \alpha_1^{-1} \exp(\alpha_0) \exp(x_i \beta_1) (\exp(\alpha_1 T_{ij}) - 1))] \quad (8)
\end{aligned}$$

(Schneider u. a., 2013).

2.3 Hypothesentests

Im betrachteten Modell lässt sich die Wirksamkeit einer Behandlung über den Vergleich der Raten $\lambda_0(t)$ und $\lambda_1(t)$ erfassen. Je wirksamer eine Behandlung, desto weniger Ereignisse (zum Beispiel Schübe) treten auf. Es soll getestet werden, ob eine neue Behandlung wirksamer ist als eine Kontrollbehandlung, also, ob das Verhältnis der Raten $\frac{\lambda_1(t)}{\lambda_0(t)} < 1$ beziehungsweise das logarithmierte Verhältnis $\beta_1 = \log\left(\frac{\lambda_1(t)}{\lambda_0(t)}\right) < 0$.

Um die Nullhypothese $H_0 : \beta_1 \geq 0$ gegen die Alternative $H_1 : \beta_1 < 0$ zu testen, kann ein Wald-Test verwendet werden. Die Teststatistik des Wald-Tests ist unter Annahme der Nullhypothese asymptotisch standardnormalverteilt, also $\frac{\hat{\beta}_1}{\sigma_{\hat{\beta}_1}} \stackrel{a}{\sim} N(0, 1)$. Die Nullhypothese wird abgelehnt, falls $\frac{\hat{\beta}_1}{\sigma_{\hat{\beta}_1}} < z_{(\alpha/2)}$.

Um den Effektschätzer $\hat{\beta}_1$ zu erhalten wurde die Log-Likelihood aus (8) über alle Parameter im Trend-Modell $\gamma = (\beta_1, \alpha_0, \alpha_1, \varphi)$ maximiert. Der ML-Schätzer des Behandlungseffekts ist asymptotisch verteilt als $\hat{\beta}_{1ML} \stackrel{a}{\sim} \mathcal{N}(\beta_1, [\mathbf{I}(\gamma)^{-1}]_{\beta_1\beta_1})$, wobei $\mathbf{I}(\gamma)$ die Fisher-Informationsmatrix, das heißt der Erwartungswert der negativen zweiten partiellen Ableitung der Likelihood aus (7) nach den Parametern, ist, siehe z.B. Held u. Sabanés Bové (2014). Der Standardfehler des Effektschätzers entspricht dann $\sqrt{[\mathbf{I}(\gamma)^{-1}]_{\beta_1\beta_1}}$. Für das Trend-Modell wurde $\sqrt{[\mathbf{I}(\gamma)^{-1}]_{\beta_1\beta_1}}$ basierend auf der von Schneider u. a. (2013) berichteten Log-likelihood (8) berechnet (siehe Anhang A). Die Formeln zur Berechnung der Varianz von $\hat{\beta}_1$ sind in Kapitel 3 sind im Abschnitt 3.2 Information für beide Modelle dargestellt. Die Formel zur Berechnung der Varianz des ML-Schätzers von β_1 im Negativ-Binomial-Modell wurde aus Friede u. a. (2018) übernommen und ist ebenfalls in Lawless (1987b) beschrieben.

2.4 Fallzahlplanung

Vor Beginn einer klinischen Studie wird die Fallzahl berechnet, die unter Annahme eines bestimmten Effekts benötigt wird, um, unter festgelegter Wahrscheinlichkeit für einen Fehler 1. Art α , die erwünschte Power $(1 - \beta)$ zu erreichen. Im Planungsstadium wird davon ausgegangen, dass jeder in die Studie aufgenommene Patient über die gleiche Zeitspanne T nachbeobachtet wird (das heißt für die Nachbeobachtungszeiten wird geplant $T_{ij} = T$ für alle T_{ij}). Das entspricht der von Schneider u. a. (2013) untersuchten Situation und die benötigte Fallzahl der Kontrollgruppe kann mit der in Schneider u. a. (2013) berichteten Formel berechnet werden:

$$n_0 = \frac{(z_{1-\alpha} + z_{1-\beta})^2}{\beta_1^{*2}} \left(\frac{1}{\Lambda_0(T, \theta^*)} \left(1 + \frac{1}{k \exp(\beta_1^*)} \right) + \varphi^*(1 + 1/k) \right). \quad (9)$$

Die Parameter, die für die Bestimmung der Fallzahl abgeschätzt werden müssen, sind die Überdispersion φ , die kumulative Rate $\Lambda_0(T, \theta^*)$ sowie der erwartete, bzw. kleinste klinisch relevante Effekt, β^* . Die Parameter $\theta = (\alpha_0, \alpha_1)$ müssen im Stadium der Fallzahlplanung nicht festgelegt werden, da von gleichen Nachbeobachtungszeiten, und somit der gleichen kumulativen Rate $\Lambda_0(T, \theta^*)$ für alle Patienten ausgegangen wird (Schneider u. a., 2013). Das bedeutet, dass es nicht notwendig ist, vor Beginn der Studie den Parameter für den zeitlichen Trend festzulegen, wenn zunächst einmal davon ausgegangen wird, dass alle Patienten die gleiche Nachbeobachtungszeit haben.

Im Fall ohne zeitlichen Trend kann diesselbe Formel verwendet werden. Hier sind die kumulativen Raten gleich $\lambda_0 T$ (Friede u. Schmidli, 2010; Keene u. a., 2007).

3 Verblindetes kontinuierliches Monitorieren

3.1 Stop-Kriterium

Beim verblindeten kontinuierlichen Monitorieren wird vor dem Start der Studie ein Stop-Kriterium festgelegt. Sobald die beobachtete Information in den Daten das Kriterium erreicht oder überschreitet, wird die Studie beendet. Als Stop-Kriterium dient die Information des Effektschätzers. Das Stop-Kriterium ist der Wert der Information, der erreicht werden muss, um die vorher festgelegte Power zu erreichen bei gegebenem α -Fehler-Niveau. Die Information des Effektschätzers $\hat{\beta}_1$ ist definiert als die Inverse der Varianz des Effektschätzers $\mathcal{I}(\hat{\beta}_1) = \text{var}(\hat{\beta}_1)^{-1}$ (Friede u. Miller, 2012). Die kritische Information \mathcal{I}^* bei einem Wald-Test lässt sich mit folgender Formel berechnen:

$$\mathcal{I}^* = \frac{(z_{1-\alpha} + z_{1-\beta})^2}{\beta_1^{*2}} \quad (10)$$

siehe zum Beispiel Jennison u. Turnbull (2000, S.53).

3.2 Information \mathcal{I}_{ML}

In dieser Arbeit sollen zwei Verfahren zum verblindeten kontinuierlichen Monitorieren von klinischen Studien mit rekurrierenden Ereignissen als Endpunkt miteinander verglichen werden. Das erste Verfahren, im folgenden BCM-CONST, wurde von Friede u. a. (2018) vorgestellt und ist eine Methode zum verblindeten kontinuierlichen Monitorieren für negativ binomialverteilte Daten. Die Information \mathcal{I}_{ML} in diesem Modell wird in folgendem Abschnitt beschrieben.

Negativ-Binomial Modell: Die Verteilung des Effektschätzers im Negativ-Binomial Modell wird in Lawless (1987a) beschrieben. Der ML-Schätzer des Effekts $\hat{\beta}_{1ML}$ ist approximativ normalverteilt mit

$$\hat{\beta}_{1ML} \stackrel{a}{\sim} \mathcal{N}(\beta_1, 1/I_0 + 1/I_1) \quad (11)$$

und mit

$$I_0 = \sum_{j=1}^{n_0} \frac{T_{0j} \lambda_0}{1 + \varphi T_{0j} \lambda_0} \quad (12)$$

und

$$I_1 = \sum_{j=1}^{n_1} \frac{T_{1j}\lambda_1}{1 + \varphi T_{1j}\lambda_1} \quad (13)$$

Die Information des Effektschätzers ist die Inverse seiner Varianz

$$\mathcal{I}_{ML} = \frac{1}{1/I_0 + 1/I_1} = \frac{I_0 I_1}{I_0 + I_1}. \quad (14)$$

Das zweite Verfahren, im folgenden BCM-TREND, ist eine Erweiterung von BCM-CONST für rekurrende Ereignisdaten mit negativem zeitlichen Trend in den Raten und der Likelihood aus (7).

Trend-Modell: Für das Trend-Modell wurde die Information $\mathcal{I}(\hat{\beta}_{1ML})$ aus der Fisher-Informationsmatrix der Log-Likelihood Funktion berechnet. Der ML-Schätzer des Effekts $\hat{\beta}_{1ML}$ im Trend-Modell ist verteilt mit

$$\hat{\beta}_{1ML} \overset{a}{\sim} \mathcal{N}(\beta_1, [\mathbf{I}(\boldsymbol{\gamma})^{-1}]_{ii}). \quad (15)$$

Die Information des Effektschätzers ist die Inverse seiner Varianz

$$\mathcal{I}_{ML} = \frac{1}{[\mathbf{I}(\boldsymbol{\gamma})^{-1}]_{ii}} \quad (16)$$

Die Fisher-Informationsmatrix $\mathbf{I}(\boldsymbol{\gamma})$, deren Inverse und der Eintrag $[\mathbf{I}(\boldsymbol{\gamma})^{-1}]_{ii}$ wurden aus den partiellen zweiten Ableitungen der Log-Likelihood (7) hergeleitet, wobei der Eintrag $[\mathbf{I}(\boldsymbol{\gamma})^{-1}]_{ii}$ aus einer 3x3 Matrix stammt, da alle partiellen gemischten zweiten Ableitungen nach φ null sind (siehe Anhang A). Der Schätzer der Information $\hat{\mathcal{I}}_{ML}$ wird berechnet indem man für die Parameter deren ML-Schätzer einsetzt.

3.3 Verblindete Schätzung der Information

Die verblindete Information $\mathcal{I}_{ML}^{(b)}$, welche in jedem Monitorierungsschritt mit der kritischen Information \mathcal{I}^* verglichen wird, muss aus den Daten ohne die Zuordnung der Patienten zu den Behandlungsgruppen zu kennen, geschätzt werden. Die verblindete Schätzung wurde mit Hilfe des sog. *lumping approaches* umgesetzt (Friede u. a., 2018). Hierbei wird zuerst eine Gesamtrate aus den Daten aller Patienten geschätzt und dann, mit Hilfe des vor Studienbeginn erwarteten Effekts β_1^* , Schätzer für die Raten in beiden Behandlungsgruppen berechnet.

Negativ-Binomial-Modell ML-Schätzer der Gesamtrate $\hat{\lambda}$ und der Überdispersion $\hat{\varphi}$ können mit der Funktion `glm.nb()` in R aus dem Paket MASS (Venables u. Ripley, 2002) berechnet werden. Aus dem Schätzer der Gesamtrate wird mit Hilfe des angenommenen Behandlungseffekts β_1^* eine Rate für die jeweilige Behandlungsgruppe verblindet geschätzt.

$$\hat{\lambda}_{0(b)} = 2\hat{\lambda}/(1 + \exp(\beta_1^*)) \quad (17)$$

$$\hat{\lambda}_{1(b)} = 2\hat{\lambda} \exp(\beta_1^*)/(1 + \exp(\beta_1^*)) = \hat{\lambda}_{0(b)} \exp(\beta_1^*) \quad (18)$$

(Friede u. a., 2018)

Verblindete Schätzung der Information:

$$I_1^{(b)} = \frac{1}{1+k} \sum_{m=1}^N \left(T_m \hat{\lambda}_{1(b)} / (1 + \hat{\varphi} T_m \hat{\lambda}_{1(b)}) \right) \quad (19)$$

$$I_0^{(b)} = \frac{1}{1+k} \sum_{m=1}^N \left(T_m \hat{\lambda}_{0(b)} / (1 + \hat{\varphi} T_m \hat{\lambda}_{0(b)}) \right) \quad (20)$$

mit m in T_m als Index der über alle Patienten $m = 1, \dots, N$ läuft. Für λ_i wird der verblindete Schätzer $\hat{\lambda}_i^{(b)}$ und für φ der Schätzer $\hat{\varphi}$ eingesetzt. Dabei wird angenommen, dass die Beobachtungszeiten in beiden Gruppen gleich verteilt sind. $\frac{1}{1+k}$ beträgt für gleich große Patientenzahlen in Behandlungs- und Kontrollgruppe 0.5 mit $k = 1$. Der verblindete Schätzer der Information des ML-Schätzers für den Behandlungseffekt setzt sich zusammen als

$$\mathcal{I}_{ML}^{(b)} = \frac{I_0^{(b)} I_1^{(b)}}{I_0^{(b)} + I_1^{(b)}} \quad (21)$$

(Friede u. a., 2018).

Trend-Modell ML-Schätzer der Gesamtrate $\hat{\alpha}_0$ zum Zeitpunkt $t = 0$, des zeitlichen Trends $\hat{\alpha}_1$ und der Überdispersion $\hat{\varphi}$ können mittels numerischer Maximierung der Log-Likelihood aus 8 mit Hilfe der Funktion `optim()` in R geschätzt.

Der Parameter β_1 wird in der Formel zur Schätzung der Information durch den angenommenen Behandlungseffekt β_1^* ersetzt, um eine verblindete Schätzung der Information durchzuführen. Es werden, wie oben für das Negativ-Binomial-Modell beschrieben, alle

Beobachtungszeiten $T_m = T_1, \dots, T_N$ bei der Berechnung der Terme der jeweiligen Behandlungsgruppe verwendet und das Randomisierungsverhältnis $\frac{1}{1+k}$, genutzt, um eine verblindete Schätzung der jeweiligen, für eine Behandlungsgruppe spezifischen Terme zu erhalten. Dabei wird wieder angenommen, dass die Beobachtungszeiten in beiden Gruppen gleich verteilt sind. Im Fall von gleich großen Gruppen ist der verblindete Schätzer der Information hier:

$$[I(\gamma)^{-1}]_{ii}^{(b)} = \frac{0.5i^{(b)}(0.5^2e^{(b)}i^{(b)} - (0.5f^{(b)})^2)}{(0.5^2a^{(b)}i^{(b)} - (0.5a^{(b)})^2)(0.5^2e^{(b)}i^{(b)} - (0.5f^{(b)})^2) - (0.5^2i^{(b)}b^{(b)} - 0.5^2a^{(b)}f^{(b)})^2} \quad (22)$$

mit

$$i^{(b)} = E \left\{ -\frac{\partial^2 l}{\partial \alpha_0^2} \right\}^{(b)}$$

$$e^{(b)} = E \left\{ -\frac{\partial^2 l}{\partial \beta_1 \partial \alpha_1} \right\}^{(b)}$$

$$f^{(b)} = E \left\{ -\frac{\partial^2 l}{\partial \alpha_1 \partial \alpha_0} \right\}^{(b)}$$

$$a^{(b)} = E \left\{ -\frac{\partial^2 l}{\partial \beta_1^2} \right\}^{(b)}$$

$$b^{(b)} = E \left\{ -\frac{\partial^2 l}{\partial \alpha_0^2} \right\}^{(b)}$$

(siehe Anhang A).

4 Implementierung und Szenarien der Simulationsstudie

Die Fragestellung, die in Kapitel 1.3 beschrieben wurde, wurde mit einer Simulationsstudie untersucht. Dazu wurden Datensätze in R (Version 3.3.1) R Development Core Team (2008) für verschiedene Parameterkombinationen simuliert und beide Verfahren zum verblindeten kontinuierlichen Monitorieren (BCM-CONST und BCM-TREND) auf die Datensätze angewendet. Da die Methode zum verblindeten kontinuierlichen Monitorieren aus Friede u. a. (2018) für Daten mit negativen Zeittrends in den Ereignisraten untersucht werden sollte, wurden, zur besseren Vergleichbarkeit der Ergebnisse, die dort beschriebenen Szenarien übernommen und zusätzlich unterschiedliche Ausprägungen des Zeittrends hinzugefügt. Die simulierten Szenarien sind durch das Beispiel einer klinischen Studie zu Multiplen Sklerose im Kindesalter motiviert.

Die Studienlaufzeit betrug in allen Szenarien maximal 48 Monate, mit einer 24-monatigen Rekrutierungsphase, über die hinweg Patienten kontinuierlich und gleichmäßig rekrutiert werden. Die geplante Nachbeobachtungszeit im Design mit fester Studienlaufzeit betrug 24 Monate. Das Monitorieren wurden mit zwei möglichen Startzeitpunkten durchgeführt: dem Start vor Ende der Rekrutierungsphase (nach 13 Monaten) und dem Start einen Monat nach Ende der Rekrutierungsphase (nach 25 Monaten). Das Studiendesign mit geplanter fester Studienlaufzeit von 48 Monaten wurde zum Vergleich ebenfalls simuliert.

Die Anzahl der Ereignisse in allen simulierten Datensätzen sollte der in (7) beschriebenen Verteilung folgen. Wie die Methode zur Simulation der Daten eines inhomogenen Poisson-Prozesses in R implementiert wurde, wird im folgenden Abschnitt beschrieben. Im Anschluss wird die Umsetzung des verblindeten kontinuierlichen Monitorierens, der Umgang mit aufgetretenen Schwierigkeiten bei der Schätzung der Parameter und der verblindeten Information sowie die Datenauswertung dargestellt.

4.1 Datengenerierung

Für die Simulation der Daten wurde die Funktion `simPP()` in R geschrieben. Mit der Funktion kann ein Datensatz generiert werden, der Ereignisdaten (Anzahl und Zeiten von Ereignissen) aus einer fiktiven klinischen Studie mit N Patienten aus zwei Behandlungsgruppen enthält. Die simulierten Ereignisse folgen einem inhomogenen Poisson Prozess mit Intensitätsfunktion $\lambda_i(t) = \exp(\alpha_0 + \alpha_1 t)$ und gamma-verteilten Raten. Das Randomisierungsverhältnis der zwei Gruppen und der Behandlungseffekt kann über die Eingabe von `k` und `effect` spezifiziert werden.

```
simPP(n0, theta, effect, k, phi, fllwYRS, recrYRS, interimYRS)
```

Als Werte können übergeben werden:

- `n0` Anzahl der Patienten in der Kontrollgruppe
- `theta` Parameter $\theta = (\alpha_0, \alpha_1)$
- `effect` $\exp(\beta_1)$
- `k` Randomisierungsverhältnis
- `phi` Varianz der Gammaverteilung
- `fllwYRS` Nachbeobachtungszeit T in Jahren im Planungsstadium (Eingabe als Zahl, für alle Individuen gleich)
- `recrYRS` Länge der Rekrutierungsperiode in Jahren
- `interimYRS` Auswertungszeitpunkt (zur Simulation von Datensätzen mit interindividuell variablen Nachbeobachtungszeiten)

4.1.1 Aufbau der Simulationsfunktion

Rekrutierungszeiten Die Rekrutierungszeiten werden als kontinuierliche Zeiten gleichmässig auf dem Intervall von $T=0$ bis zum Ende der Rekrutierungszeit (hier $T=2$) in Jahren generiert.

Heterogene Raten Es werden N gamma-verteilte Variablen $U_{ij} \sim \Gamma(1, \varphi)$ in \mathbb{R} mit Hilfe der Funktion `rgamma()` generiert. Die heterogenen Raten für den Zeitpunkt $T = 0$ werden mit Hilfe dieser Variablen erzeugt als: $\lambda_{ij}(0) = u_{ij} \exp(\alpha_0) \exp(x_i \beta_1)$.

Zeittrend Es wird zuerst die Anzahl der Ereignisse pro Patient Y_{ij} als Variable eines stationären Poisson Prozesses mit der Rate $\lambda_{ij}T$ simuliert ($\lambda_{ij}T = u_{ij} \exp(\alpha_0) \exp(x_i \beta_1)T$, T = Gesamtbeobachtungszeit für einen Patienten). Dies geschieht mit Hilfe der R-Funktion zur Simulation von Poisson-verteilten Zufallszahlen `rpois()`. Für jeden Patienten werden dann Y_{ij} exponentialverteilte Ereigniszeiten simuliert. Die Ereigniszeiten werden mittels Y_{ij} gleichverteilter Zufallsvariablen auf dem Intervall $[0, T]$ mit der R-Funktion `runif()` zur Simulation gleichverteilter Zufallszahlen generiert. Der Zeittrend wird dann mit Hilfe der Ausdünnung der Daten erreicht, siehe Lewis u. Shedler (1979) oder Stoyan u. a. (1995) und Kapitel 2.1.1. Dazu wird zu jeder Ereigniszeit die entsprechende Rate $\lambda(t)$ berechnet. Das Verhältnis dieser Rate zur Anfangsrate $\lambda_{ij}T$ wird mit einer Zufallszahl,

die gleichverteilt zwischen 0 und 1 ist, verglichen. Ist dieses Verhältnis kleiner als die Zufallszahl wird das Ereignis verworfen, andernfalls wird es behalten.

4.2 Verblindetes kontinuierliches Monitorieren

Für das Monitorieren wurde zunächst ein Datensatz erzeugt, der die Daten einer Studie mit fester Studienlaufzeit (48 Monate) und gleichen Nachbeobachtungszeiten für alle Individuen (24 Monate) enthält. Zu jedem Monitorierungszeitpunkt wurde eine Unter-
menge der Daten betrachtet, die nur Ereignisse enthielt, welche vor dem Monitorierungszeitpunkt aufgetreten sind. Für das Monitorieren wurden zwei mögliche Startzeitpunkte gewählt: Monat 13 oder Monat 25. Aus diesen Daten wurde die verblindete Information $\mathcal{I}_{ML}^{(b)}$ geschätzt.

Das Monitorieren wurde in monatlichen Zeitabständen durchgeführt. Sobald die verblindet geschätzte Information $\hat{\mathcal{I}}_{ML}^{(b)}$ gleich oder größer der kritischen Information \mathcal{I}^* war, wurden die Daten unverblindet ausgewertet und die Ergebnisse gespeichert. Für diese Simulationsstudie wurde die Power auf 0.8 ($\beta = 0.2$) und das Fehlerniveau auf $\alpha = 0.025$ (einseitig) festgelegt. Daraus ergibt sich eine kritische Information von $\mathcal{I}^* = 16.34$ gemäß (9). Eine schematische Übersicht über den Ablauf des verblindeten kontinuierlichen Monitorierens wird in Abbildung 2 gegeben.

Bei der Umsetzung des kontinuierlichen Monitorierens mit den simulierten Daten in R traten einige numerische Schwierigkeiten auf, die wie folgt gelöst wurden. In einigen Datensätzen (abhängig vom Parameter φ) war die Überdispersion in den simulierten Daten nahe bei null. In solchen Fällen wurde bei BCM-CONST kein Negativ-Binomial-Modell an die Daten angepasst, sondern auf ein einfacheres Modell zurückgegriffen, das Poisson-Modell und der Schätzer für die Überdispersion $\hat{\varphi}$ auf null gesetzt.

Für BCM-TREND ergab sich ein ähnliches Problem beim Schätzen der kumulativen Raten. In Fällen, in denen kein negativer Trend in den Daten vorhanden ist, lassen sich die kumulativen Raten nicht mit Formel (4) berechnen. Trat ein solcher Fall ein, wurden die kumulativen Rate unter Auslassung des Trendparameters (entspricht $\hat{\alpha}_1 = 1$) geschätzt. Bei der Maximierung einer Funktion mit dem Nelder-Mead Algorithmus über die Funktion `optim()` in R können die Wertebereiche der zu schätzenden Funktionswerte nicht auf einfache Weise auf bestimmte Bereiche restringiert werden. Dadurch resultierten bei der Maximierung der Log-Likelihood im Trendmodell auch negative Werte für den Schätzer

der Überdispersion. In allen Fällen, in denen dies auftrat, wurde, nach Abschluss der Optimierung der Schätzer der Überdispersion auf einen Wert nahe bei null (10^{-6}) festgelegt.

Um die Konvergenz der ML-Schätzung mittels numerischen Optimierung zu verbessern, wurden der Funktion `optim()` Start-Werte übergeben. Im ersten Schritt wurden die folgenden Startwerte für die Parameter gewählt:

- Als Startwert für $\hat{\alpha}_0$ wurden die Rate der Ereignisse aus dem ersten halben Jahr von jedem bis zum Monitorierungszeitpunkt rekrutierten Patienten berechnet.
- Startwert für $\hat{\alpha}_1$ wurde auf den Wert -0.1 gesetzt, da hier keine sinnvolle Berechnung möglich schien.
- Als Orientierung für den Startwert für $\hat{\varphi}$ diene die Formel für die Varianz der Anzahl von Ereignissen in einem Negativ-Binomial Modell $Y_m \sim \text{NegBin}(\lambda T_m, \varphi)$, $\text{Var}(Y_m) = \mu_m(1 + \varphi\mu_m)$, mit $m = 1, \dots, N$. Die Formel wurde nach φ umgestellt und für μ_m der Gesamtmittelwert der Anzahl der Ereignisse und für $\hat{\sigma}_{Y_m}^2$ die Varianz der Anzahl der Ereignisse bis zum Monitorierungszeitpunkt eingesetzt.
- Bei unverblindeter Auswertung: Startwert für $\hat{\beta}_1 = \log(Y_1./Y_0.)$
 Y_i ist der Mittelwert der Anzahl der Ereignisse in Gruppe i .

Konvergierte die Schätzung mit diesen Startwerte nicht, wurden in einem zweiten Schritt die ML-Schätzer aus einem Negativ-Binomial-Modell (berechnet mit `glm.nb()`) als Startwerte verwendet. Als Startwert für den zu schätzenden Parameter α_0 wurde $\log(\hat{\lambda}_{0_{ML}})$ und als Startwert für den zu schätzenden Parameter φ , $\hat{\varphi}_{ML}$ verwendet.

Tabelle 1: Prozentsatz der Fälle in denen wegen fehlender Überdispersion ein Poisson Modell in BCM-CONST angepasst wurde oder $\hat{\varphi} < 1e^{-6}$ bei BCM-TREND-Iterationslimit wegen fehlender Überdispersion in `glm.nb()` erreicht bei BCM-CONST [% der Fälle] oder $\hat{\varphi} < 1e^{-6}$ bei BCM-TREND [% der Fälle]

Monitorierungsstart:	In finaler Analyse			Während des Monitorierens		
	13	25	Feste Laufzeit	13	25	Feste Laufzeit
Methode: CONST	4.33%	2.65%	1.91%	18.55%	2.71%	0.84%
Methode: TREND	8.28%	5.38%	3.35%	22.44%	3.75%	1.35%

Sowohl in BCM-TREND also auch in BCM-CONST konvergierten nach diesen Schritten alle Parameterschätzungen.

Tabelle 2: Prozentsatz der Fälle in denen die Schätzungen in BCM-TREND `optim()` mit den ersten Startwerten nicht konvergierten und neue Startwerte geschätzt wurden.

	Alternative Schätzung der Startwerte für <code>optim()</code> [% der Fälle]					
	In finaler Analyse			Während des Monitorierens		
Monitorierungsstart:	13	25	Feste Laufzeit	13	25	Feste Laufzeit
Methode: TREND	0.0577%	0.0679%	0.0908%	0.1224%	0.0301%	0.0429%

Tabelle 3: Prozentsatz der Fälle in denen der geschätzte Parameter für den Zeittrend in BCM-TREND sehr nahe null lag.

	$-1e^{-6} < \hat{\alpha}_1 < 1e^{-6}$ [% der Fälle]					
	In finaler Analyse			Während des Monitorierens		
Monitorierungsstart:	13	25	Feste Laufzeit	13	25	Feste Laufzeit
Methode: TREND	0.0020%	0.0020%	0.0026%	0.0168%	0.0026%	0.0005%

4.3 Parameterschätzung und Wald Test

Beim BCM-TREND wurde die Log-Likelihood Funktion in R über `optim()` mit dem Standardwert für den Algorithmus (Nelder-Mead) maximiert um ML-Schätzer der Parameter zu erhalten. Beim BCM-CONST wurden die ML-Schätzer der Parameter mit `glm.nb()` ermittelt. Der Wald-Test wurde für beide Verfahren mit dem jeweiligen Schätzer des Behandlungseffekts und dessen Standardfehler durchgeführt. Bei BCM-TREND wurde der Standardfehler von $\hat{\beta}_{1MLTREND}$ über die Inverse der Fisher-Informationsmatrix der Log-Likelihoodfunktion geschätzt (siehe Anhang A). Bei BCM-CONST wurde der Standardfehler von $\hat{\beta}_{1MLCONST}$ über `glm.nb()` ermittelt.

4.4 Szenarien

Das Planungsszenario entsprach dem in Friede u. a. (2018) verwendeten Planungsfall ($\lambda_0 T = 0.36T$, $\beta^* = \log(0.5)$, $\varphi = 0.82$). Mit der Fallzahlplanungsformel (9), einer festen Studienlaufzeit von 4 Jahren und einer Nachbeobachtungszeit von 2 Jahren für jeden Patient, wurde eine Fallzahl von insgesamt $N = 190$ Patienten (entspricht 95 Patienten pro Behandlungsgruppe) ermittelt. Simulationen wurden für folgende Szenarien durchgeführt, die, bis auf die Spezifikation des Zeittrends, analog zu den in Friede u. a. (2018) simulierten Szenarien sind. Die Anfangsraten für das Trendmodell wurden jeweils aus den kumulativen Raten und der Beobachtungszeit berechnet und der Parameter α_0 entsprechend angepasst ($\exp(\alpha_0) = \frac{\Lambda_0(T)\alpha_1}{(\exp(\alpha_1 T) - 1)}$). Es wurden 2000 Simulationen pro Szenario durchgeführt. Bei 2000 Simulationen liegt der Monte Carlo Fehler bei ≈ 0.003 für die Wahrscheinlichkeit des Fehlers 1. Art ($\alpha = 0.025$) und ≈ 0.009 für die Power ($1 - \beta = 0.8$).

Tabelle 4: Simulierte Szenarien

		Parameter unter H_0
Kumulative Basisrate	$\Lambda_0(1)$	0.36, 0.54, 0.72
Effekt	$\exp(\beta_1)$	1
Zeittrend	α_1	0, -0.3, -0.5, -0.7, -2.5, -3.5, -4.5
Heterogenität	φ	0.35, 0.82
		Parameter unter H_1
Kumulative Basisrate	$\Lambda_0(1)$	0.36, 0.72
Effekt	$\exp(\beta_1)$	0.5, 0.375
Zeittrend	α_1	0, -0.3, -0.5, -0.7, -2.5, -3.5, -4.5
Heterogenität	φ	0.35, 0.82
Rekrutierungsphase	T_{rp}	24 Monate
Startzeitpunkts des Monitorierens	T_{mn}	13. Monat, 25. Monat
Nachbeobachtungszeit	T	24 Monate
Studienlänge	T_{sl}	48 Monate
Gesamtfallzahl	N	190
Anzahl der Simulationen pro Szenario		2000

Simuliere Datensatz

Die Anzahl der Ereignisse Y eines Individuums j ($j = 1, \dots, n_i$) aus Gruppe i ($i = 0, 1$) mit Ereigniszeiten $t_{ij1} < t_{ij2} < \dots < t_{ijk}$ im Beobachtungszeitraum $[0, T_{ij}]$ hat die Wahrscheinlichkeitsdichte

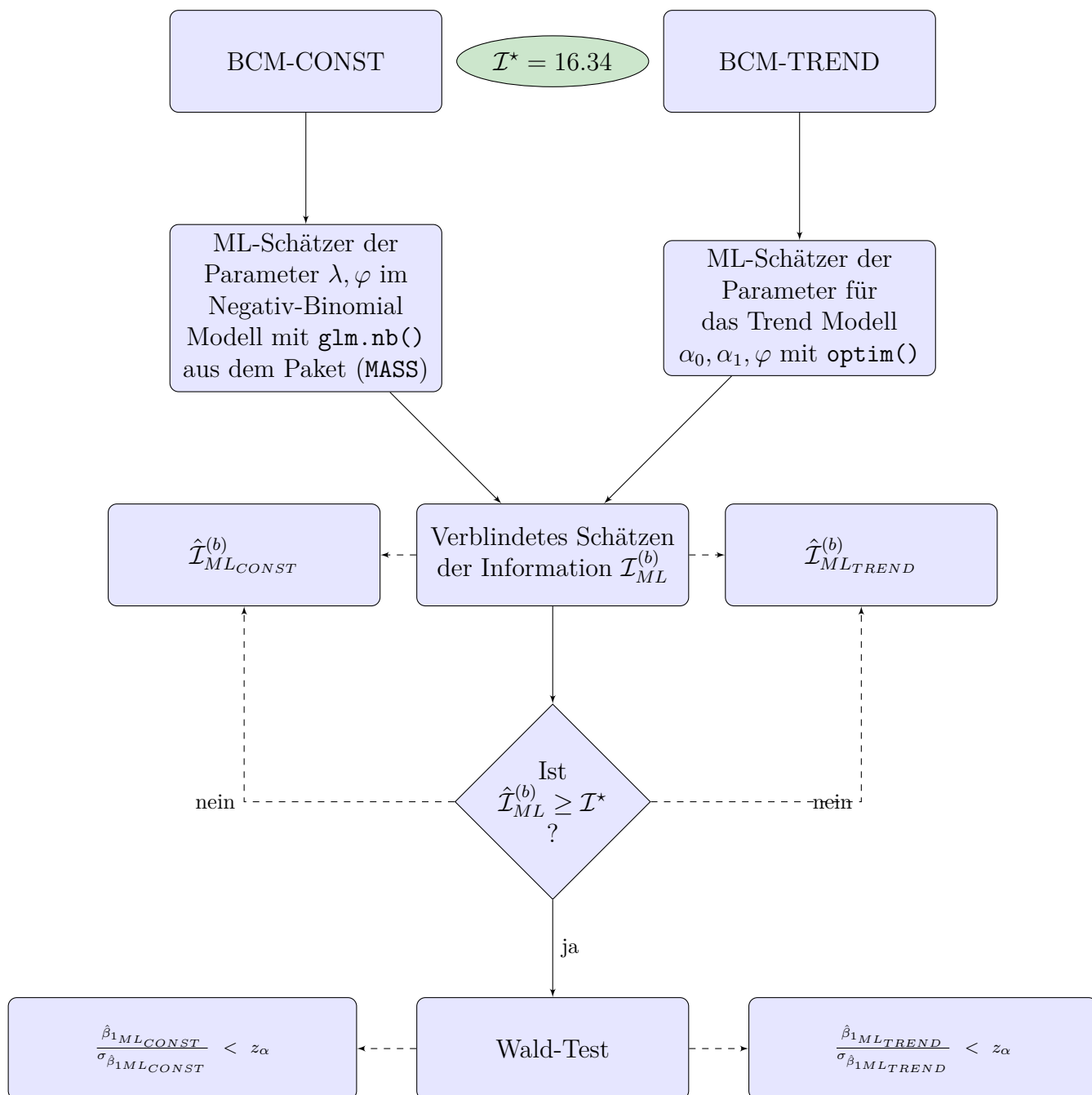
$$f(y|\alpha_0, \alpha_1, \varphi, \beta_1, t, T) = \left(\frac{\lambda_0(t_{ijk}, \theta) \exp(x_i \beta_1)}{\Lambda_0(T_{ij}, \theta) \exp(x_i \beta_1)} \right) \times \frac{\Gamma(\varphi^{-1} + y_{ij})}{\Gamma(\varphi^{-1}) y_{ij}!} \frac{(\varphi \Lambda_0(T_{ij}, \theta) \exp(x_i \beta_1))^{y_{ij}}}{(1 + \varphi \Lambda_0(T_{ij}, \theta) \exp(x_i \beta_1))^{y_{ij} + \varphi - 1}}$$


Abbildung 2: Schematischer Ablauf des verblindeten kontinuierlichen Monitorierens

5 Ergebnisse der Simulationsstudie

Methoden zum verblindeten kontinuierlichen Monitorieren werden angewendet, um in Bezug auf Laufzeit und Fallzahl effiziente Studiendesigns zu erhalten. Damit beurteilt werden kann, ob die Methode die regulatorischen Anforderungen erfüllt, muss festgestellt werden, ob die vorher festgelegte Wahrscheinlichkeit des Fehlers 1. Art (α) eingehalten wird. Aus Sicht des Sponsors einer Studie ist es von hoher Bedeutung, ob die vorher festgelegte Power erreicht wird und ob die Studienlaufzeiten und/oder Fallzahlen, sobald dies möglich ist, optimal angepasst werden. Mit Hilfe von Simulationsstudien für verschiedene Datentypen und zu praxisnahen Beispielszenarien können die beschriebenen Eigenschaften untersucht werden (Chuang-Stein u. a., 2006). In Friede u. a. (2018) wurde das verblindete kontinuierliche Monitorieren für negativ binomialverteilte Endpunkte einer klinischen Studie mit einer Monte Carlo Simulationsstudie untersucht und es konnte gezeigt werden, dass in Fällen, in denen die Studienlaufzeit im Planungsstadium der Studie überschätzt wird, nach Monitorieren geringere Studienlaufzeiten resultieren, unter Beibehaltung der geplanten Power und ohne Inflation der Wahrscheinlichkeit des Fehlers 1. Art.

In diesem Abschnitt werden die Ergebnisse einer Simulationsstudie berichtet, in der die Robustheit der von Friede u. a. (2018) entwickelten Methode (BCM-CONST) gegenüber der Verletzung der Annahme überprüft wurde, dass die Ereignisraten über die Zeit konstant bleiben. Zusätzlich werden Ergebnisse zu einer Erweiterung des verblindeten kontinuierlichen Monitorierens für rekurrierende Ereignisdaten mit gamma-verteilen Raten und negativem Zeittrend in den Raten (BCM-TREND) vorgestellt. Es wurden Ereignisdaten aus einem inhomogenen Poisson Prozess mit über die Zeit abnehmenden Ereignisraten für verschiedene Szenarien simuliert und beide Methoden zum verblindeten Monitorieren, BCM-CONST und BCM-TREND darauf angewendet. Es werden für beide Verfahren die Wahrscheinlichkeit des Fehlers 1. Art, die Power ($1 - \beta$), Überdeckungswahrscheinlichkeiten der Konfidenzintervalle und Ergebnisse zur Verzerrung der Schätzer (Bias) berichtet. Beide Methoden werden in Bezug auf Studienlaufzeit und Fallzahl für zwei mögliche Startzeitpunkte für das Monitorieren verglichen (Beginn vor oder nach Ende der Rekrutierungsphase).

5.1 Wahrscheinlichkeit des Fehlers 1. Art

Um die Rate des Fehlers 1. Art für die beiden verglichenen Verfahren zu bewerten, werden die Szenarien betrachtet, die unter Gültigkeit der Nullhypothese simuliert wurden. Für

Tabelle 5: Mittelwerte, Minimum, Maximum und Prozentsatz der Fälle außerhalb des Bereichs von ± 2 Monte Carlo Fehlern für die Wahrscheinlichkeit des Fehlers 1. Art bei BCM-CONST, BCM-TREND und das Design mit fester Studienlaufzeit (= ohne Monitorieren) für verschiedene Monitorierungsstartzeitpunkte.

Methode	Beginn des Monitorierens nach	MW	Min.	Max.	% der Fälle innerhalb von $0.025 \pm 2MCSE$
BCM-CONST	Monat 13	0.026	0.019	0.032	95.24%
	Monat 25	0.028	0.021	0.037	90.48%
BCM-TREND	Monat 13	0.027	0.022	0.032	92.86%
	Monat 25	0.028	0.021	0.036	90.48%
<i>Gesamt</i>					92.26 %
BCM-CONST	Feste Laufzeit	0.027	0.022	0.033	88.1%
BCM-TREND	Feste Laufzeit	0.027	0.022	0.033	88.1%

BCM-CONST und BCM-TREND wird die Wahrscheinlichkeit eines Fehlers 1. Art gemittelt über die Szenarien, in denen monitoriert wurde, betrachtet. Insgesamt weisen beide Verfahren ähnliche Wahrscheinlichkeiten für den Fehler 1. Art auf. Für BCM-CONST liegt die mittlere Wahrscheinlichkeit des Fehlers 1. Art im Durchschnitt bei ≈ 0.027 (Spannweite: 0.020 bis 0.037) und für BCM-TREND bei ≈ 0.027 (Spannweite: 0.022 bis 0.036). Die Ergebnisse unterschieden sich nur unwesentlich zwischen einem frühen und späten Start des Monitorierens (siehe Tabelle 5). Die mittleren Häufigkeiten eines Fehlers 1. Art liegen somit für beide Verfahren nahe am vorgegebenen Niveau von $\alpha = 0.025$. Unter Berücksichtigung des Monte Carlo Fehlers ($MCSE$) für 2000 Simulationen von ≈ 0.003 liegen die Abweichungen für 92.26% der Szenarien im Bereich von $\pm \approx 2MCSE$ (siehe Tabelle 5). Für die Szenarien außerhalb dieses Bereichs liegen die höchsten Raten bei 0.036 (BCM-TREND) und 0.037 (BCM-CONST). Für das Design mit fester Studienlaufzeit sind die Wahrscheinlichkeiten des Fehlers 1. Art ähnlich verteilt. Hier liegen 88% der Wahrscheinlichkeiten für den Fehler 1. Art im Bereich von $\pm \approx 2MCSE$, die maximale Fehlerwahrscheinlichkeit fällt mit 0.033 ähnlich aus wie beim Monitorierungsstart nach 13 Monaten, und etwas kleiner aus als beim Monitorierungsstart nach 25 Monaten.

Die Fehlerwahrscheinlichkeiten für beide Verfahren unterscheiden sich nicht bedeutsam zwischen den Methoden oder Startzeitpunkten des Monitorierens. Ebenso unterschieden sich die Fehlerwahrscheinlichkeiten nach Monitorieren im Mittel nicht wesentlich von denen im Design mit fester Laufzeit (siehe Abbildung 3). Es gibt keine Anzeichen dafür, dass die Fehlerwahrscheinlichkeit bei der Anwendung von BCM-CONST durch einen negativen Zeittrend in den Ereignisraten erhöht wird (siehe Abbildung 4). Die Abweichungen von der nominalen Fehlerwahrscheinlichkeit $\alpha = 0.025$ liegen, unter Betrachtung des $MCSE$, im Mittel nur geringfügig über dem erfordernten Wert. Die Ergebnisse zeigen,

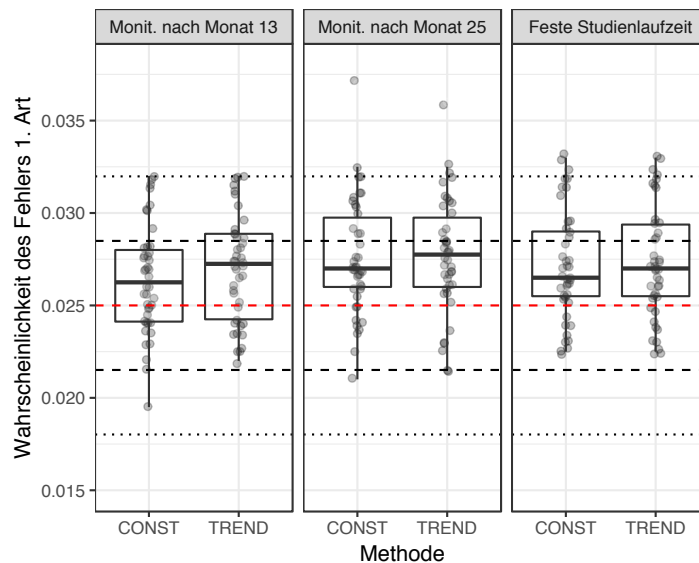


Abbildung 3: Verteilung der Wahrscheinlichkeit des Fehlers 1. Art für BCM-CONST und BCM-TREND für verschiedene Monitorierungsstartzeiten. Rote Linie: festgelegtes Niveau α . Gestrichelte Linie: Monte Carlo Fehler ($MCSE$). Gepunktete Linie: $2MCSE$.

unter Berücksichtigung, dass pro Szenario nur 2000 Simulationen durchgeführt wurden und im Vergleich mit den Fehlerwahrscheinlichkeiten im Design mit fester Laufzeit, dass durch das Monitorieren bei beiden Methoden von keiner praktisch relevanten Inflation des Fehlers 1. Art auszugehen ist.

5.2 Überdeckungswahrscheinlichkeiten der 95%-Konfidenzintervalle für den Effektschätzer

Für alle Szenarien in denen unter der H_0 oder der H_1 monitoriert wurde, liegt die Überdeckungswahrscheinlichkeit der 95%-Konfidenzintervalle ($\hat{\beta}_1 \pm z_{0.975} \sigma_{\hat{\beta}_1}$) im Durchschnitt bei 94.78% für BCM-CONST und 94.76% für BCM-TREND, und somit nahe an den geplanten 95% (siehe MW in Tabelle 6 und 7). Die Spannweite der 95%-Konfidenzintervalle beträgt für BCM-CONST 93.45% - 96.20% und für BCM-TREND 93.50% - 95.80%. Die Überdeckungswahrscheinlichkeiten unterscheiden sich nicht wesentlich zwischen BCM-CONST und BCM-TREND (siehe Abbildung 5). Für eine Abbildung aufgeschlüsselt nach allen Szenarien siehe Abbildung 6.

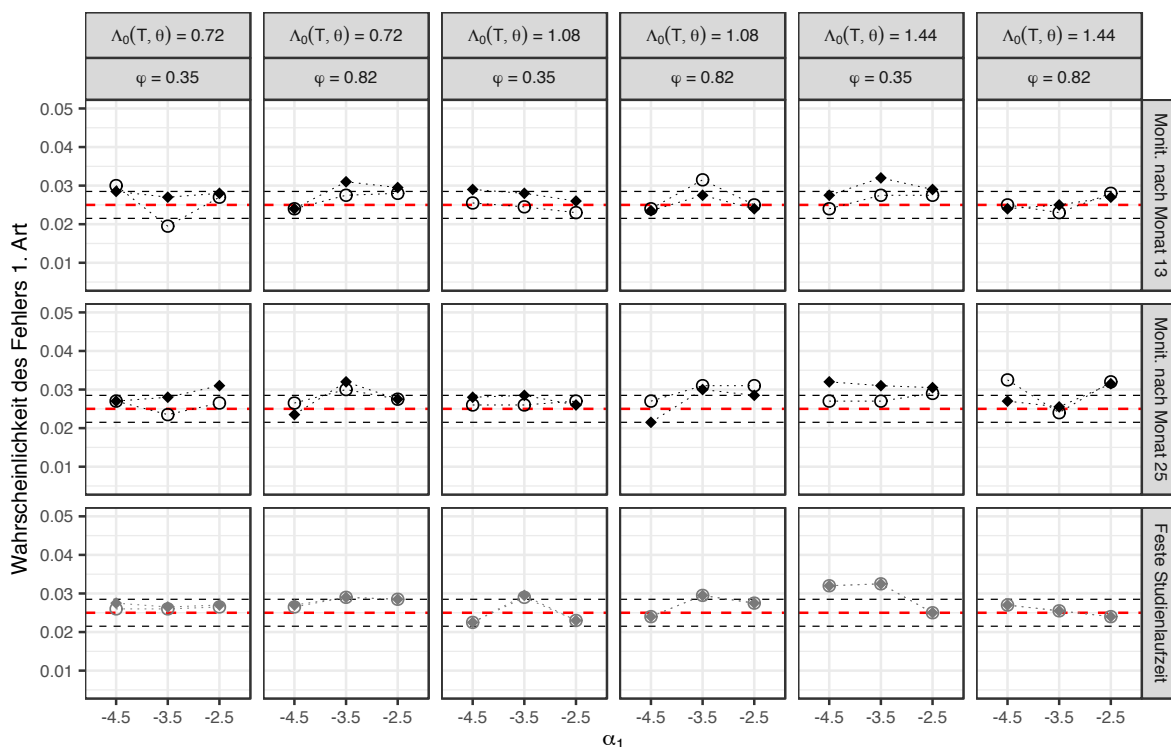
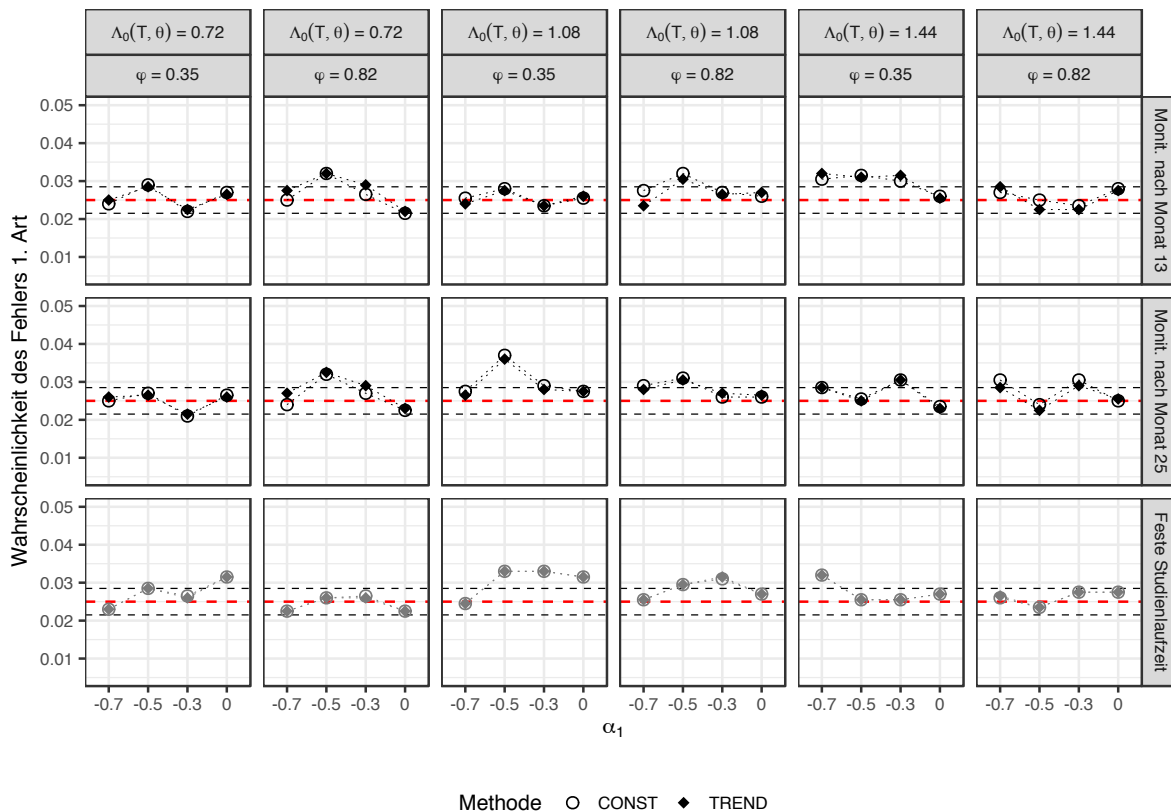


Abbildung 4: Wahrscheinlichkeit des Fehlers 1. Art für BCM-CONST, BCM-TREND und das Design mit fester Laufzeit (= ohne Monitorieren). Die rote gestrichelte Linie zeigt das vorgegebene Niveau für den Fehler 1. Art ($\alpha = 0.025$). Die schwarze gestrichelte Linie zeigt den Monte Carlo Fehler für 2000 Wiederholungen.

Tabelle 6: Mittelwerte, Minimum und Maximum für BCM-CONST und BCM-TREND und das Design mit fester Studienlaufzeit (= ohne Monitorieren) für verschiedene Monitorierungsstartzeitpunkte unter H_0 .

Methode	Beginn des			
	Monitorierens nach	MW	Min.	Max.
BCM-CONST	Monat 13	94.73%	93.70%	95.65%
	Monat 25	94.62%	93.45%	95.70%
BCM-TREND	Monat 13	94.61%	93.70%	95.60%
	Monat 25	94.62%	93.70%	95.80%
CONST	Feste Laufzeit	94.68%	93.75%	96.20%
TREND	Feste Laufzeit	94.67%	93.75%	96.20%

Tabelle 7: Mittelwerte, Minimum und Maximum für BCM-CONST und BCM-TREND und das Design mit fester Studienlaufzeit (= ohne Monitorieren) für verschiedene Monitorierungsstartzeitpunkte unter H_1 .

Methode	Beginn des			
	Monitorierens nach	MW	Min	Max
BCM-CONST	Monat 13	94.84%	93.85%	96.20%
	Monat 25	94.88%	93.60%	96.10%
BCM-TREND	Monat 13	94.87%	93.65%	95.70%
	Monat 25	94.87%	93.50%	95.70%
CONST	Feste Laufzeit	94.83%	93.70%	96.10%
TREND	Feste Laufzeit	94.82%	93.70%	96.10%

5.3 Bias

Nach dem Monitorieren mit BCM-TREND wurden die ML-Schätzer $\hat{\alpha}_0$, $\hat{\alpha}_1$, $\hat{\varphi}$ und $\hat{\beta}_1$ berechnet, indem die Log-Likelihood Funktion numerisch mit der R-Funktion `optim()` maximiert wurde. Im Verfahren BCM-CONST wurden die ML-Schätzer $\hat{\lambda}_0$, $\hat{\varphi}$ und $\hat{\beta}_1$ mit der R-Funktion `glm.nb()` ermittelt. Der Bias der Schätzer für BCM-TREND wurde berechnet, indem diese innerhalb jedes Szenarios gemittelt (über 2000 Simulationen) und die Differenz zum jeweiligen wahren Parameter genommen wurde.

Der Bias fällt bei BCM-TREND für alle ML-Schätzer gering aus (siehe Tabelle 8). Betrachtet man den Bias bei BCM-TREND über alle Szenarien hinweg, in denen ein großer Behandlungseffekt ($\beta_1 = -0.981$) simuliert wurde, liegt der Bias im Mittel bei -0.0069 . Das entspricht einer geringen Abweichung vom wahren Parameter (um ca. 1%). Für einen etwas geringeren Behandlungseffekt ($\beta_1 = -0.693$) beträgt der Bias -0.0019 , und in Fällen, in denen kein Behandlungseffekt simuliert wurde ($\beta_1 = 0$), liegt der über die jeweiligen Szenarien gemittelte Bias von $\hat{\beta}_1$ bei ≈ 0.0001 . Der Behandlungseffekt wird bei BCM-TREND im Mittel um eine Größe, die maximal 3% des wahren Werts von

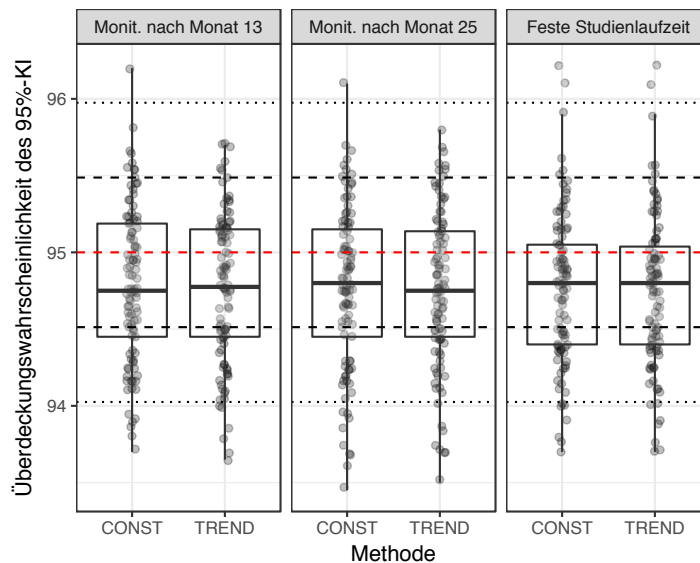


Abbildung 5: Verteilung der Überdeckungswahrscheinlichkeiten der 95%-Konfidenzintervalle für die Schätzer des Behandlungseffekts. Rote Linie: vorgegebenes Niveau von 95%. Gestrichelte Linie: Monte Carlo Fehler ($MCSE = 0.49\%$). Gepunktete Linie: $2MCSE$.

β_1 beträgt, überschätzt und um maximal 2% unterschätzt. Auch für die Schätzer der anderen Parameter nach BCM-TREND findet sich kein Hinweis auf einen bedeutsamen Bias (siehe Abbildungen 8 und 9).

BCM-CONST liefert ebenfalls kaum verzerrte Effektschätzer $\hat{\beta}_1$. Die Abweichung beträgt im Mittel -0.019 für große Effekte ($\beta_1 = -0.981$), -0.013 für mittlere Effekte ($\beta_1 = -0.693$) und 0.0002 für Szenarien ohne Behandlungseffekt. Der Behandlungseffekt wird bei BCM-CONST im Mittel um eine Größe, die maximal 8% des wahren Werts von β_1 beträgt, überschätzt und um maximal 1.4% unterschätzt. Bei der Schätzung von Überdispersion zeigt sich nach BCM-CONST für große Zeittrends ein deutlicher Bias der Schätzer (siehe Abbildung 11). In einigen Fällen ist der Bias von $\hat{\varphi}$ bei BCM-CONST ungewöhnlich hoch. Das kommt daher, dass vereinzelte Schätzer von $\hat{\varphi}$ sehr hoch sind ($\hat{\varphi} > 4000$). Dies tritt ausschließlich bei BCM-CONST mit Monitorierungsstart nach 13 Monaten in sieben aller unter der Alternativhypothese simulierten Datensätze ($\approx 0.006\%$) und in einem von allen unter der Nullhypothese simulierten Datensätze auf ($\approx 0.001\%$). Die hohen Werte entstehen, wenn bei einem simulierten Datensatz nach Monitorieren mehrere Ereignisse innerhalb einer sehr kurzen Nachbeobachtungszeit auftreten (beispielweise 5 Ereignisse innerhalb eines Tages). In der Praxis treten Schübe bei MS nicht in so kurzen Zeitabständen auf, weshalb solche Fälle bei einer klinischen Studie zur MS nicht zu erwarten sind.

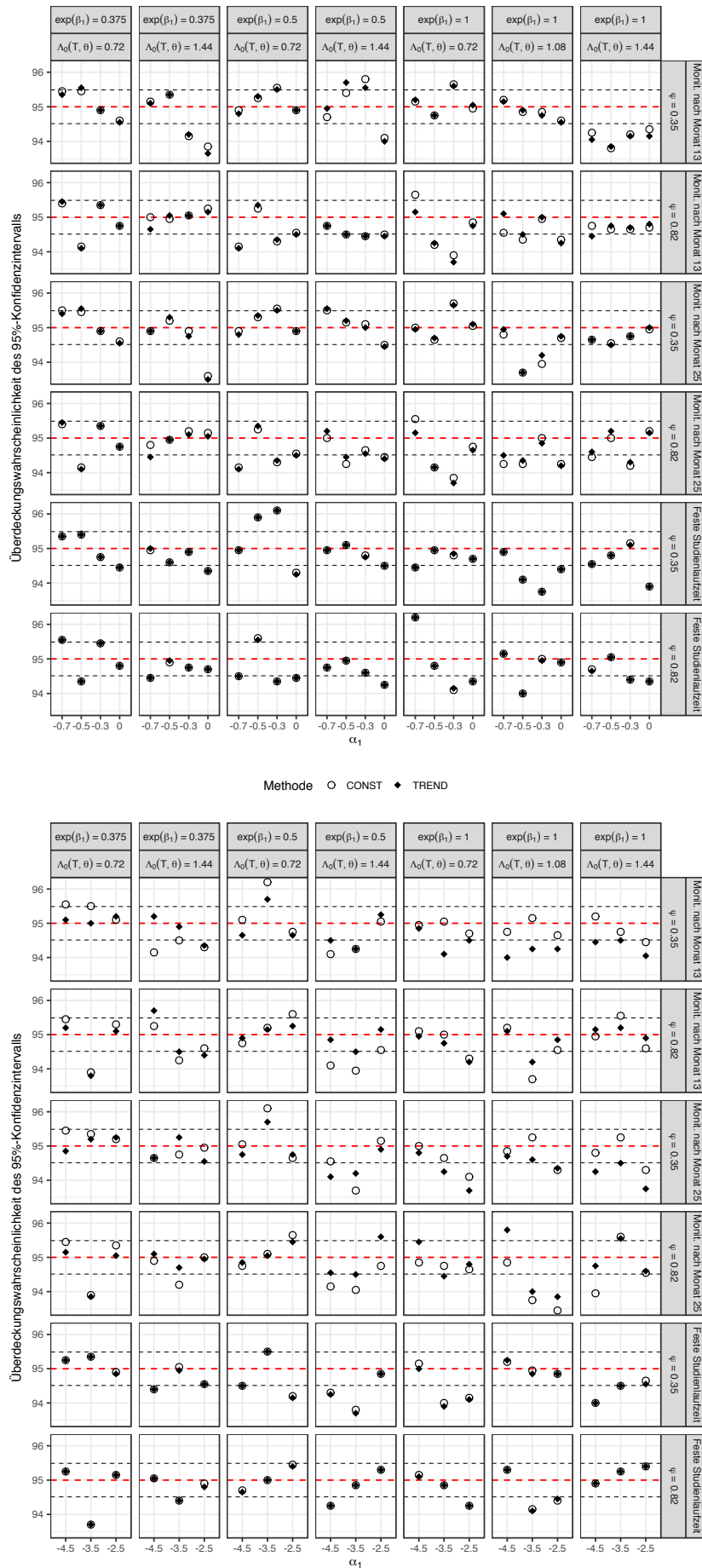


Abbildung 6: Überdeckungswahrscheinlichkeiten der 95%-Konfidenzintervalle für die einzelnen Szenarien Rote Linie: vorgegebenes Niveau von 95%. Gestrichelte Linie: Monte Carlo Fehler ($MCSE = 0.49\%$)

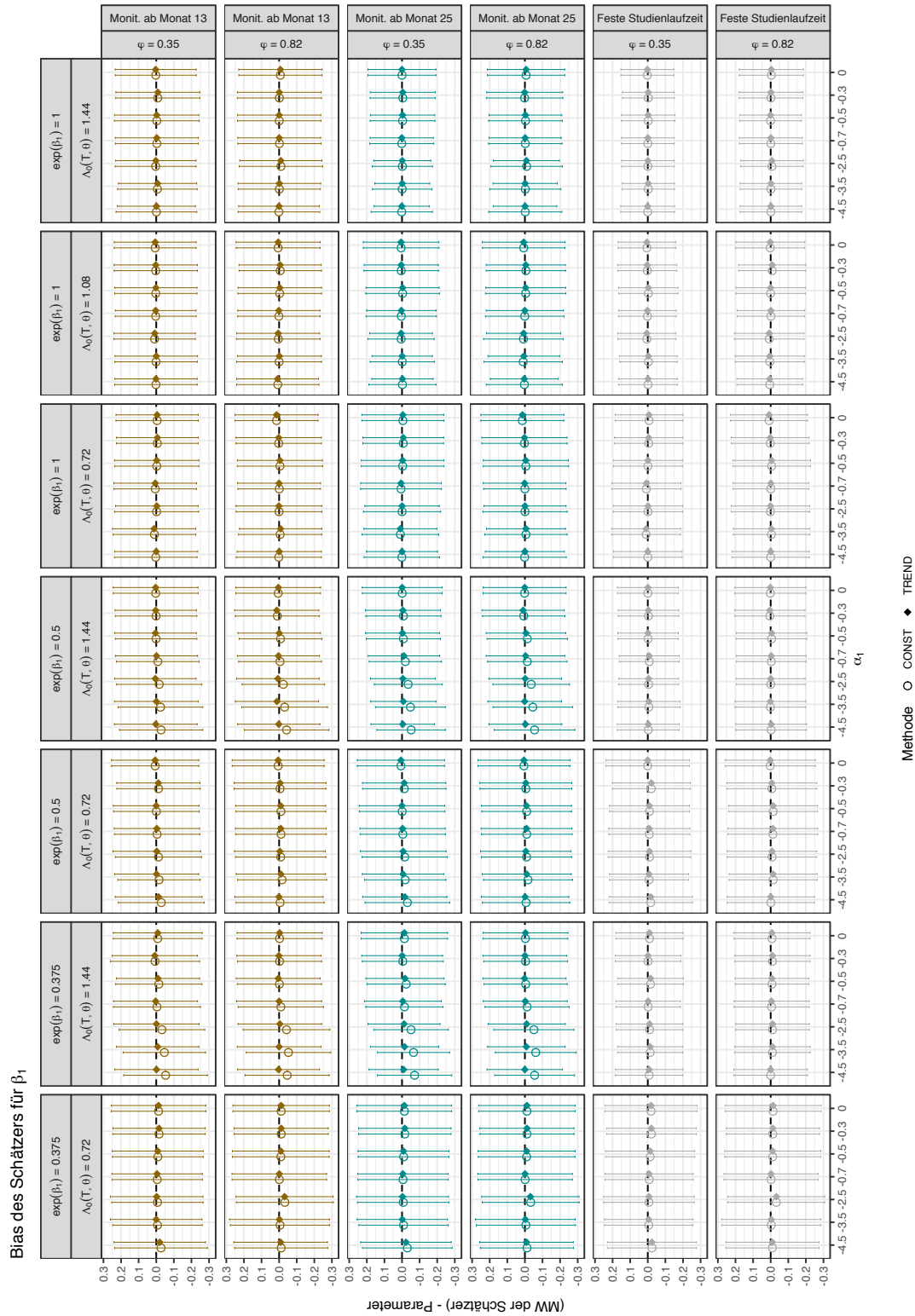


Abbildung 7: Bias des Effektschätzers

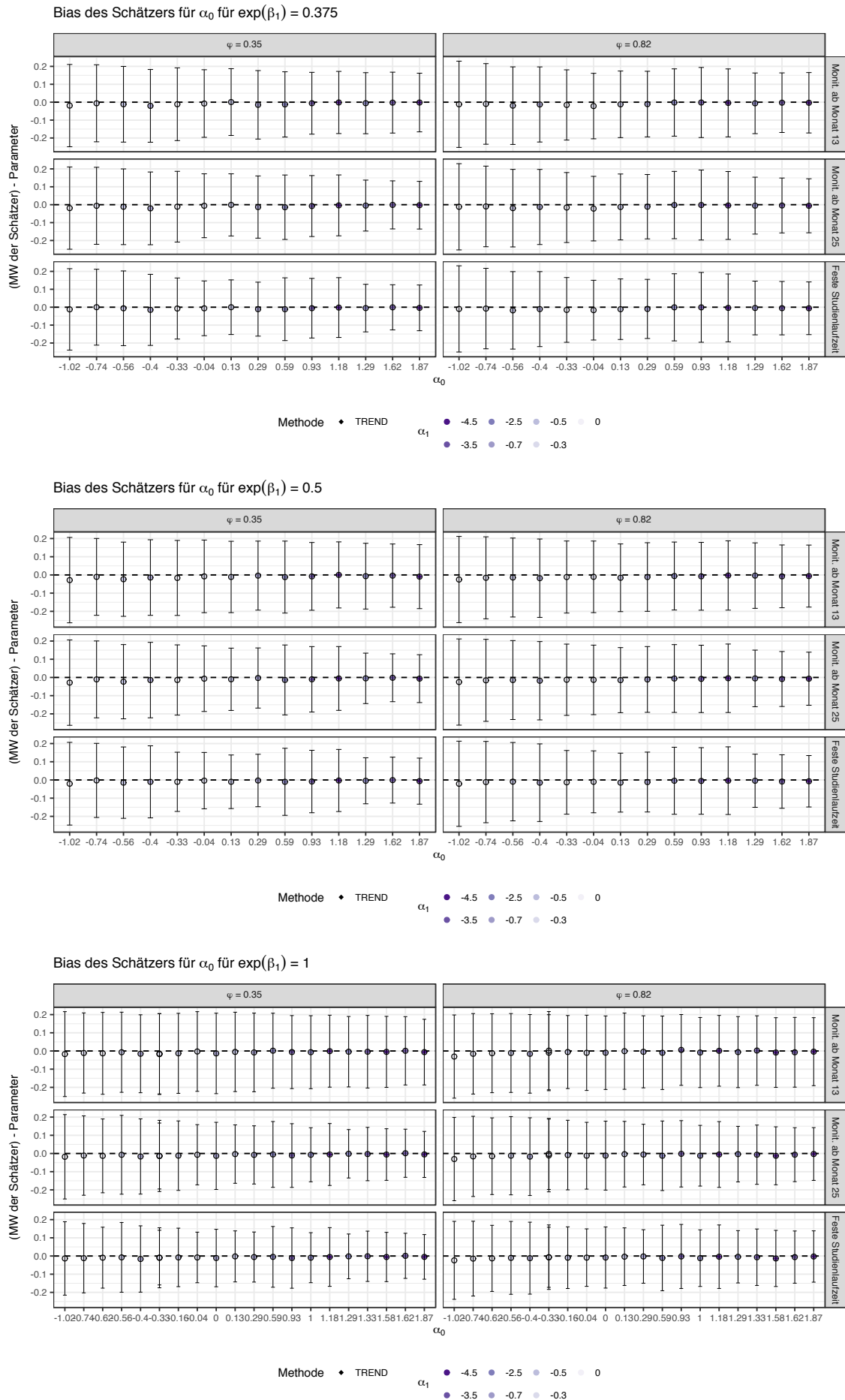


Abbildung 8: Bias des Schätzers für α_0 im Trendmodell

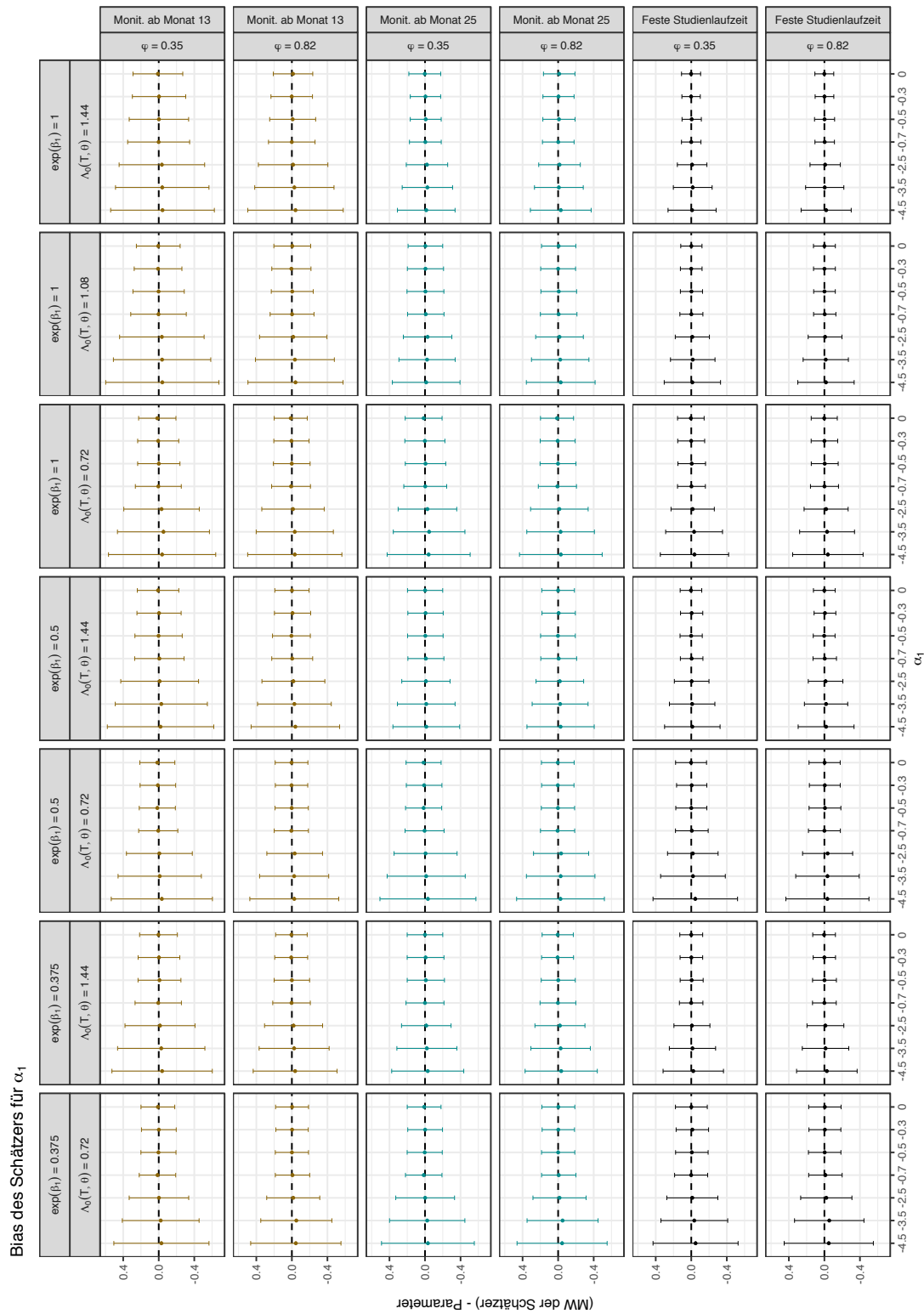


Abbildung 9: Bias des Schätzers für den Trendparameter α_1 im Trendmodell

Tabelle 8: Mittelwert, Standardabweichung, Minimum und Maximum des Bias für $\hat{\beta}_1$ für verschiedene Startzeitpunkte des Monitorierens bei BCM-TREND, BCM-CONST und dem Design mit fester Laufzeit.

Methode	Parameter	13 Monate		25 Monate		Feste Laufzeit	
		β_1	MW (SD) [Min;Max]	MW (SD) [Min;Max]	MW (SD) [Min;Max]	MW (SD) [Min;Max]	
TREND	-0.981	-0.0054 (0.0083)	[-0.0308;0.0085]	-0.0084 (0.0075)	[-0.0308;0.0024]	-0.0104 (0.0071)	[-0.0320;0.0004]
	-0.693	0.0001 (0.0068)	[-0.0135;0.0133]	-0.0039 (0.0058)	[-0.0161;0.0102]	-0.0055 (0.0061)	[-0.0201;0.0049]
	0	0.0001 (0.0055)	[-0.0097;0.0144]	0.0002 (0.0049)	[-0.0084;0.0145]	0.0001 (0.0044)	[-0.0087;0.0094]
CONST	-0.981	-0.0171 (0.0171)	[-0.0532;0.0069]	-0.0217 (0.0215)	[-0.0712;-0.0012]	-0.0104 (0.0071)	[-0.0321;0.0004]
	-0.693	-0.0104 (0.0128)	[-0.0424;0.0099]	-0.0163 (0.0177)	[-0.0547; 0.0067]	-0.0056 (0.0061)	[-0.0201;0.0048]
	0	0.0002 (0.0053)	[-0.0103;0.0147]	0.0001 (0.0051)	[-0.0100; 0.0148]	0.0001 (0.0044)	[-0.0087;0.0094]

5.4 Power, Studienlaufzeiten und Fallzahlen

Für ein effizientes Design ist es wünschenswert, dass die Power nahe beim vorher festgelegten Wert liegt. Wird fortlaufend eine höhere Power erreicht, bedeutet das, dass mehr Patienten oder eine längere Studienlaufzeit realisiert wurden als nötig gewesen wäre, um die Vorgaben zur Power aus dem Planungsstadium zu erfüllen. Durch das verblindete Monitorieren soll in solchen Fällen eine kürzere Laufzeit erreicht werden. Die Laufzeit muss so abgeschätzt werden, dass noch genug Daten vorhanden sind, um die erwünschte Power zu erreichen.

Die Power nach dem Monitorieren variiert über beide Methoden und alle unter der H_1 simulierten Szenarien hinweg zwischen 78.75 % und 99.95%. Der Monte Carlo Simulationsfehler ($MCSE$) liegt bei ≈ 0.009 . Für keines der beiden Verfahren unterschreitet die Power nach Monitorieren in den unter der H_1 simulierten Szenarien das vorgegebene Niveau von 0.8 um mehr als zwei $MCSE$.

Planungsszenario Im Planungsszenario ($\Lambda_0 T = 0.72$, $\alpha_1 = 0$, $\varphi = 0.82$, $\exp(\beta_1) = 0.5$) wird für BCM-CONST und für BCM-TREND nach beiden Monitorierungsstarts eine Power von $\approx 79\%$ erreicht. Die durchschnittliche Laufzeit liegt bei ≈ 44 Monaten und ca. 63% der simulierten Studien haben die maximale Laufzeit von 48 Monaten. Es kommt unter dem Planungsszenario auch bei einem frühem Monitorierungsstart vor Ende der Rekrutierungsphase zu keiner Reduktion der Fallzahl. Die Ergebnisse zeigen, dass beim Monitorieren in beiden Verfahren die Power eingehalten wird, wenn die Daten der Planungssituation entsprechen. Das ist wünschenswert, denn die Laufzeiten sollten hier für beide Verfahren nahe am geplanten Zeitrahmen bleiben. Auch wenn ein zeitlicher Trend mit ansonsten richtig geplanten Parametern hinzukommt, wird die Power eingehalten (siehe Tabellen 9 und 10). Bei starken Zeittrends sind die Anfangsraten sehr hoch, was erklärt, warum hier für beide Verfahren kürzere Laufzeiten erreicht werden. Ansonsten

führt eine Fehlspezifikation des zeitlichen Trends unterschiedlicher Größenordnung und ansonsten richtig geplanten Parametern, in den simulierten Szenarien zu keiner Unterschreitung der Zielpower bei BCM-CONST.

Tabelle 9: Power und Laufzeiten im Planungsszenario für kleine Zeittrends. Für die Laufzeiten ist in jeder Spalte unter dem Wert für die Power angegeben: Median der Laufzeit, die Interquartilsspanne und der Prozentsatz der Studien mit maximaler Laufzeit (Median [IQS] und % maximale Laufzeit).

Methode	Start nach		Zeittrend			
			$\alpha_1 = 0$	$\alpha_1 = -0.3$	$\alpha_1 = -0.5$	$\alpha_1 = -0.7$
CONST	Monat 13	Power	0.79	0.79	0.81	0.81
		Laufzeit	48 [40-48] 63.10%	48 [38-48] 62.40%	48 [38-48] 61.70%	48 [38-48] 63.50%
TREND	Monat 13	Power	0.79	0.80	0.81	0.81
		Laufzeit	48 [40-48] 63.10%	48 [38-48] 62.35%	48 [38-48] 61.60%	48 [37-48] 63.30%
CONST	Monat 25	Power	0.79	0.79	0.81	0.81
		Laufzeit	48 [40-48] 63.10%	48 [38-48] 62.40%	48 [38-48] 61.70%	48 [38-48] 63.50%
TREND	Monat 25	Power	0.79	0.80	0.81	0.81
		Laufzeit	48 [40-48] 63.10%	48 [38-48] 62.35%	48 [38-48] 61.60%	48 [37-48] 63.30%

Tabelle 10: Power im Planungsszenario für große Zeittrends. Für die Laufzeiten ist in jeder Spalte unter dem Wert für die Power angegeben: Median der Laufzeit, die Interquartilsspanne und der Prozentsatz der Studien mit maximaler Laufzeit (Median [IQS] und % maximale Laufzeit).

Methode	Start nach		Zeittrend		
			$\alpha_1 = -2.5$	$\alpha_1 = -3.5$	$\alpha_1 = -4.5$
CONST	Monat 13	Power	0.80	0.82	0.80
		Laufzeit	48 [34-48] 62.75%	48 [33-48] 63.95%	48 [32-48] 60.85%
TREND	Monat 13	Power	0.79	0.81	0.79
		Laufzeit	48 [30-48] 62.15%	48 [27-48] 63.15%	48 [26-48] 59.70%
CONST	Monat 25	Power	0.80	0.82	0.80
		Laufzeit	48 [34-48] 62.75%	48 [33-48] 63.95%	48 [32-48] 60.85%
TREND	Monat 25	Power	0.79	0.81	0.80
		Laufzeit	48 [30-48] 62.15%	48 [27.75-48] 63.20%	48 [26-48] 59.70%

Abweichungen vom Planungsszenario Unterschiede zwischen den Verfahren ergeben sich, wenn die Raten höher oder die Überdispersion geringer ist als die Parameter, die im Planungsstadium für die Fallzahlplanung der Studie verwendet wurden. In solchen Fällen sollten auch weniger Daten als geplant zur erwünschten Power führen. Hier ergibt sich zunächst einmal für beide Verfahren eine Verkürzung der Laufzeit (siehe Abbildung 12). Im Fall eines frühen Monitorierungsstarts und großen Zeittrends resultieren jedoch bei BCM-TREND effizientere Designs als bei BCM-CONST. Für BCM-TREND ergibt sich für den stärksten negativen Trend eine Differenz in den Medianen der Laufzeiten von 2 - 4 Monaten. Vor allem wenn in dieser Situation (früher Monitorierungsstart und starker

Zeittrend $\alpha_1 = -4.5$) nur einer der „Störparameter“ anders ausfällt als erwartet, also die Überdispersion bei der Planung überschätzt wurde oder die Basisrate unterschätzt wurde, zeigen sich die Vorteile von BCM-TREND. Bei einer Überschätzung der Überdispersion im Planungsstadium ($\varphi^* = 0.82$, $\varphi = 0.35$, $\Lambda_0 T = 0.72$) liegt das 50%-Quantil der Studienlaufzeiten bei BCM-TREND 4 Monate unter dem 50%-Quantil der Studienlaufzeit bei BCM-CONST. Das 75% Quantil der Studienlaufzeiten bei BCM-TREND ist 7 Monate geringer als das 75%-Quantil der Studienlaufzeiten bei BCM-CONST. In Bezug auf die Fallzahl ist für BCM-TREND $Q_{0.25_{TREND}} = 174$ Patienten und für BCM-CONST $Q_{0.25_{CONST}} = 188$, was einen Unterschied von 14 Patienten darstellt. Bei einer Unterschätzung der Basisrate ($\Lambda_0^* T = 0.72$, $\varphi = 0.82$, $\Lambda_0 T = 1.44$) zeigt sich im Median und im 75%-Quantil eine Laufzeitdifferenz von 3 Monaten, mit der geringeren Laufzeit unter BCM-TREND. Die der Median der Fallzahl ist dabei bei BCM-TREND um 24 Patienten geringer als bei BCM-CONST ($Q_{0.5_{CONST}} = 174$ Patienten, $Q_{0.5_{TREND}} = 150$ Patienten).

In Fällen mit größerem Behandlungseffekt kann man mit dem verblindeten Monitorieren die Power nicht steuern. Da beim verblindeten Monitorieren die Information verblindet unter Zuhilfenahme des Planungseffekts geschätzt wird, zielt die Methode nicht auf die Korrektur von Misspezifikationen der Effektgröße ab und ist diesbezüglich auch nicht robust. Daher resultiert hier für beide Verfahren eine sehr hohe Power (Minimum 95.3%). Wenn der Behandlungseffekt anders als erwartet ausfällt, findet man Unterschiede in den Laufzeiten zwischen den Verfahren bei großen Zeittrends und frühem Monitorierungsstart (siehe Abbildung 13).

Insgesamt weist BCM-CONST eine ähnliche Power wie BCM-TREND auf. Das deutet darauf hin, dass die verblindet geschätzte Information $\hat{\mathcal{I}}_{ML}^{(b)}$ in BCM-CONST ebenfalls ähnlich zu der in BCM-TREND ist. Da der Schätzer für φ bei BCM-CONST bei größer werdenden Zeittrends eine Verzerrung aufweist (siehe Abbildung 11), in dem Sinne, dass die Überdispersion überschätzt wird, wird vermutet, unter Betrachtung der Terme (12) und (13) die zur Berechnung der Information (14) verwendet werden, dass die kumulativen Basisraten bei BCM-CONST bei der verblindeten Schätzung überschätzt werden. Eine Verzerrung in beiden Schätzern, dem der kumulativen Basisrate und dem der Überdispersion, würde bei der Berechnung der geschätzten Information dazu führen, dass sich die Verzerrung bei der Schätzung, falls der Bias für beide Werte in die gleiche Richtung geht, gegenseitig in einem gewissen Umfang aufheben könnte.

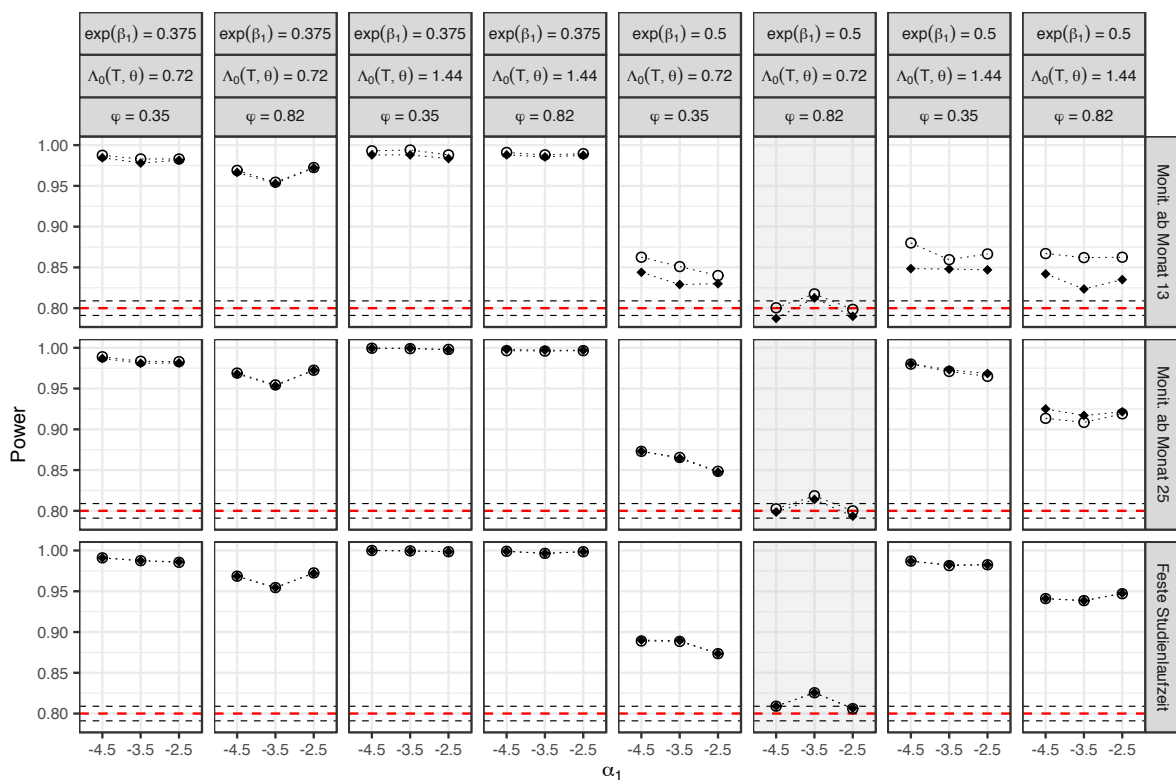
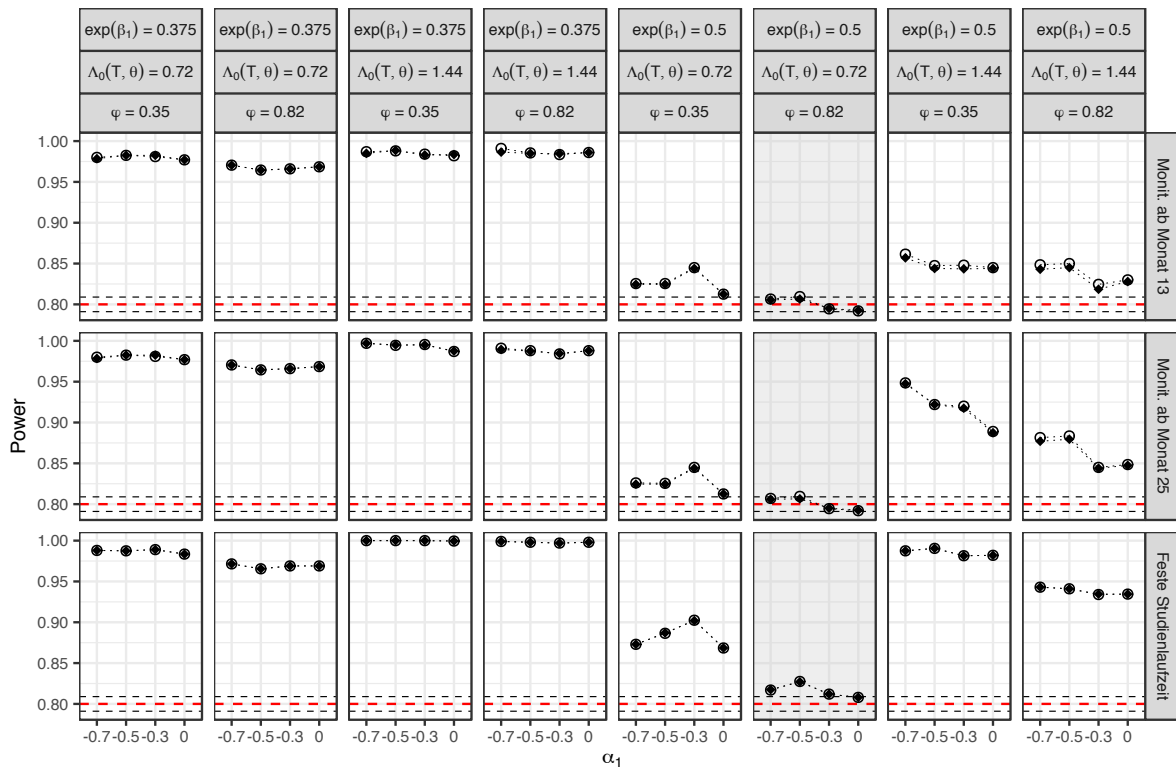


Abbildung 10: Power für BCM-CONST und BCM-TREND und Designs mit fester Studienlaufzeit (= ohne Monitorieren). Das Planungsszenario ist grau unterlegt. Die rote Linie zeigt das vorgegebene Ziel für die Power ($1 - \beta = 0.8$).

Tabelle 11: Power für Abweichungen vom Planungsszenario für große Zeittrends und einen Behandlungseffekt von $\exp(\beta_1) = 0.5$. Für die Laufzeiten ist in jeder Spalte unter dem Wert für die Power angegeben: Median der Laufzeit, die Interquartilsspanne und der Prozentsatz der Studien mit maximaler Laufzeit (Median [IQS] und % maximale Laufzeit).

Methode	Start nach	φ	$\Lambda_0 T$		$\alpha_1 = -2.5$	$\alpha_1 = -3.5$	$\alpha_1 = -4.5$
CONST	Monat 13	0.82	1.44	Power	0.86	0.86	0.87
				Median [IQS] %Max.	22 [20-24] 0.15%	22 [20-24] 0.30%	22 [20-24] 0.35%
TREND	Monat 13	0.82	1.44	Power	0.83	0.82	0.84
				Median [IQS] %Max.	20 [19-22] 0.15%	20 [18-21] 0.25%	19 [17-21] 0.30%
CONST	Monat 13	0.35	0.72	Power	0.84	0.85	0.86
				Median [IQS] %Max.	29 [25-37] 19%	28 [25-34] 17.5%	28 [24-35] 19.45%
TREND	Monat 13	0.35	0.72	Power	0.83	0.83	0.84
				Median [IQS] %Max.	26 [24-33] 18.7%	25 [23-29] 16.7%	24 [22-28] 18.75%
CONST	Monat 13	0.35	1.44	Power	0.87	0.86	0.88
				Median [IQS] %Max.	18 [16-19] 0%	17 [15-19] 0%	17 [15-19] 0%
TREND	Monat 13	0.35	1.44	Power	0.85	0.85	0.85
				Median [IQS] %Max.	17 [15-18] 0%	16 [14-17] 0%	15 [14-17] 0%
CONST	Monat 25	0.82	1.44	Power	0.92	0.91	0.91
				Median [IQS] %Max.	25 [25-25] 0.15%	25 [25-25] 0.3%	25 [25-25] 0.35%
TREND	Monat 25	0.82	1.44	Power	0.92	0.92	0.93
				Median [IQS] %Max.	25 [25-25] 0.15%	25 [25-25] 0.25%	25 [25-25] 0.30%
CONST	Monat 25	0.35	0.72	Power	0.85	0.87	0.87
				Median [IQS] %Max.	29 [25-37] 19%	28 [25-34] 17.5%	28 [25-35] 19.45%
TREND	Monat 25	0.35	0.72	Power	0.85	0.86	0.87
				Median [IQS] %Max.	26 [25-33] 18.7%	25 [25-29] 16.75%	25 [25-28] 18.75%
CONST	Monat 25	0.35	1.44	Power	0.96	0.97	0.98
				Median [IQS] %Max.	25 [25-25] 0%	25 [25-25] 0%	25 [25-25] 0%
TREND	Monat 25	0.35	1.44	Power	0.97	0.97	0.98
				Median [IQS] %Max.	25 [25-25] 0%	25 [25-25] 0%	25 [25-25] 0%

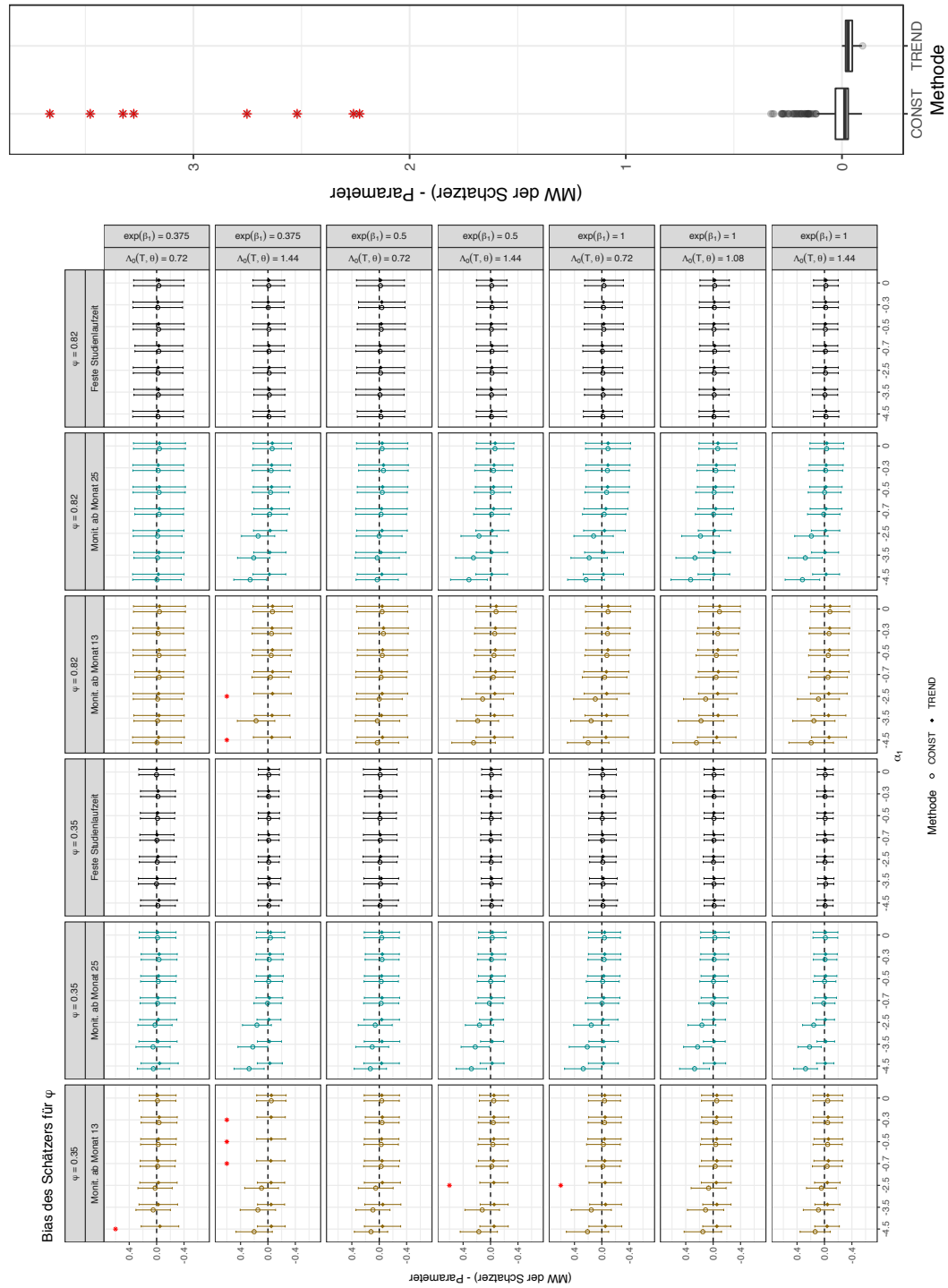


Abbildung 11: Bias des Schätzers für φ im Trendmodell. Mit rotem Stern markiert sind Werte des Bias > 2 .

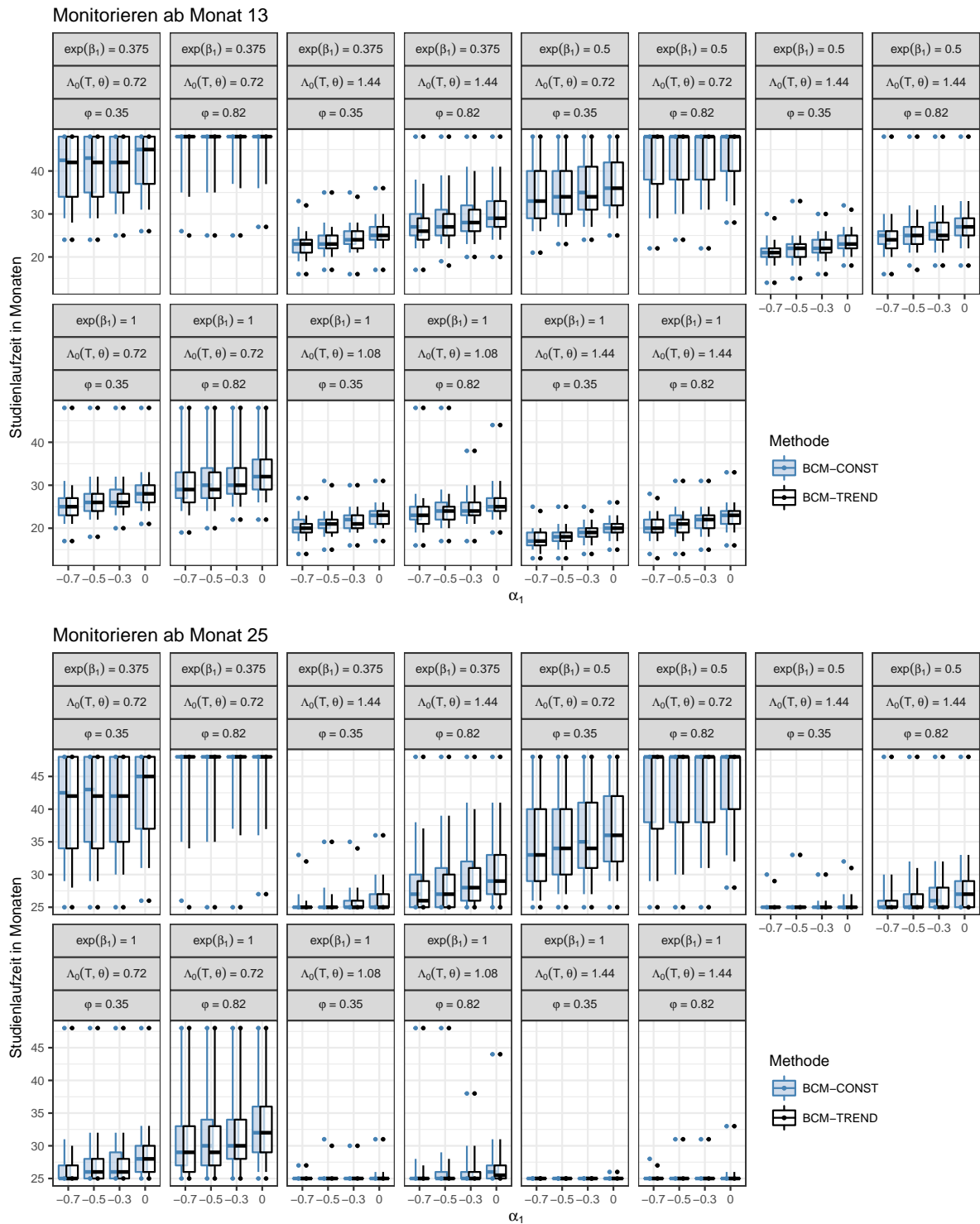


Abbildung 12: Quantile der Studienlaufzeiten (in Monaten) für BCM-CONST und BCM-TREND.

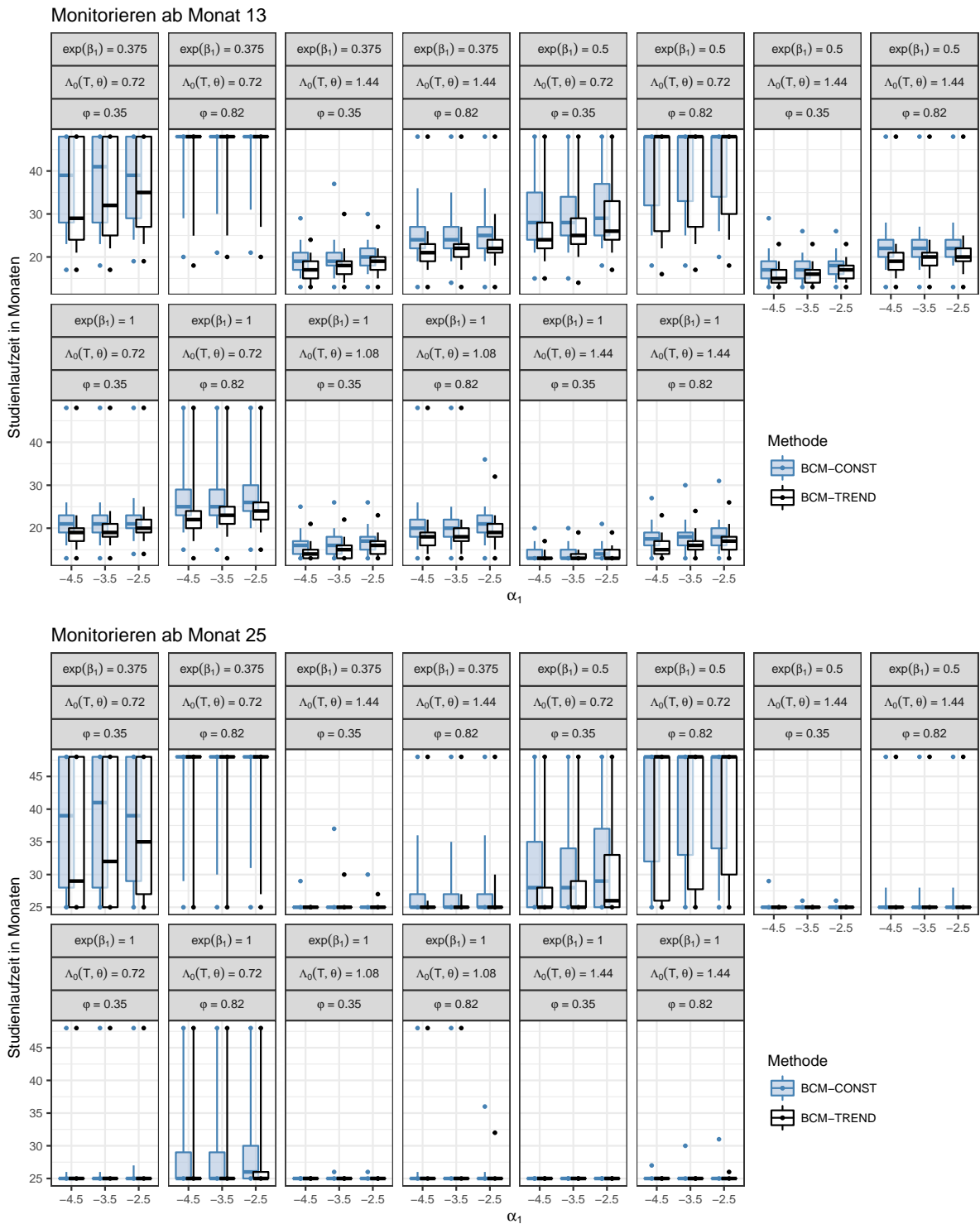


Abbildung 13: Quantile der Studienlaufzeiten (in Monaten) für BCM-CONST und BCM-TREND.

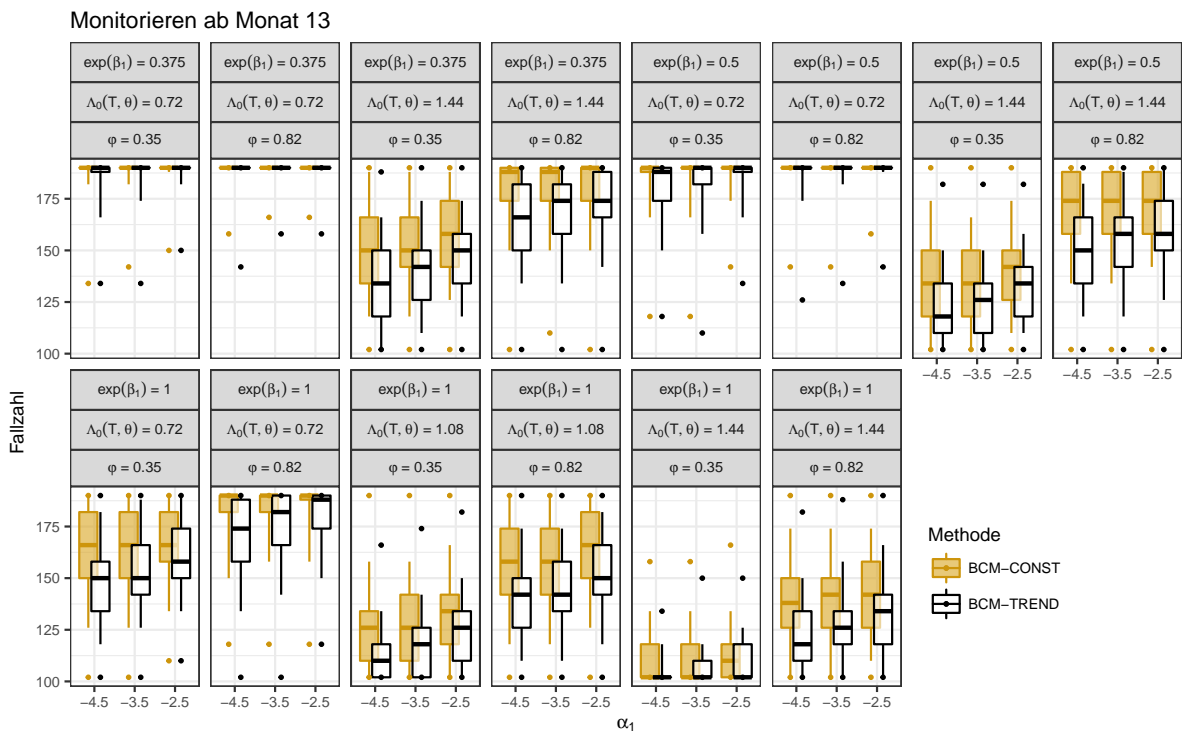
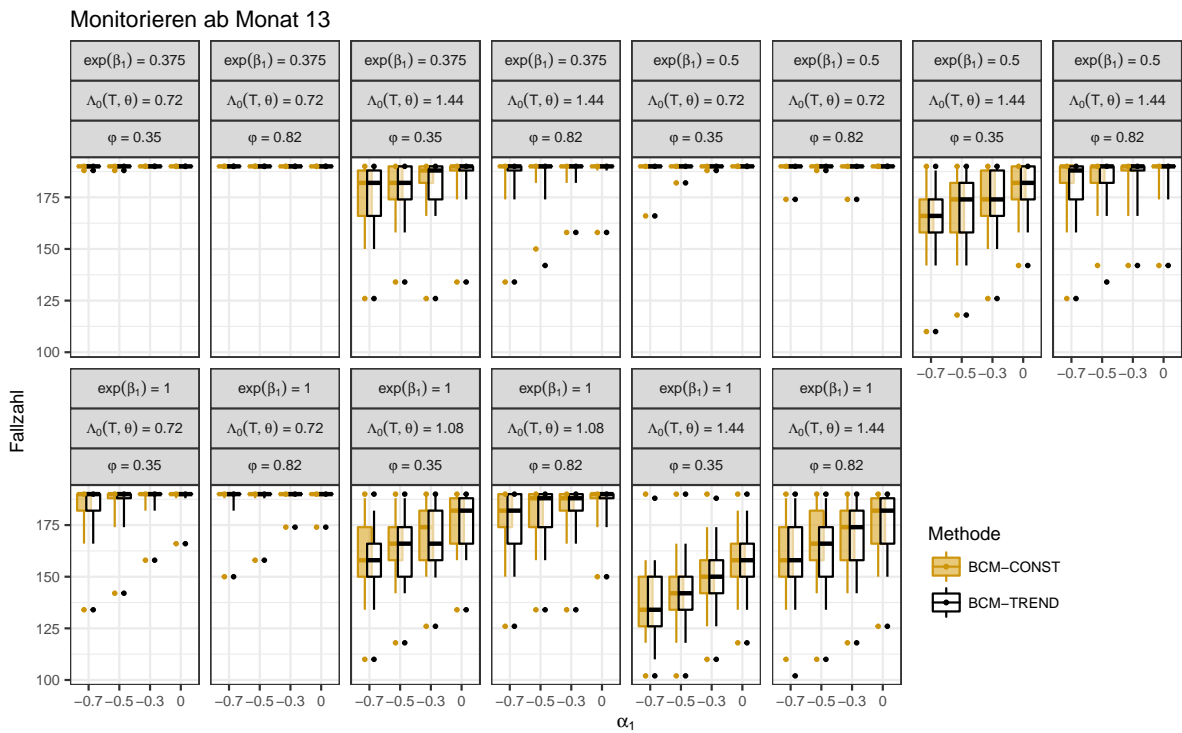


Abbildung 14: Quantile der Fallzahlen für BCM-CONST und BCM-TREND.

6 Diskussion

In dieser Arbeit wurde eine bereits bestehende Methode zum verblindeten kontinuierlichen Monitorieren für rekurrierende Ereignisdaten auf Fälle erweitert, in denen ein negativer Zeittrend in den Ereignisraten vorliegt. Das Verfahren wurde mit Hilfe einer Simulationsstudie auf seine Charakteristika, wie die Wahrscheinlichkeit des Fehlers 1. Art, den Bias der geschätzten Parameter und die Power, untersucht. Weiterhin wurden die Laufzeiten der Studien und die Fallzahlen nach verblindetem Monitorieren betrachtet. Als mögliche Startzeiten für das verblindete kontinuierliche Monitorieren wurden zwei Zeitpunkte untersucht, ein Zeitpunkt der circa ein Jahr vor Ende der Rekrutierungsphase liegt und ein zweiter Zeitpunkt, einen Monat nach Ende der Rekrutierungsphase. Zusätzlich wurde die bereits bestehende Methode zum verblindeten Monitorieren für rekurrierende Ereignisdaten ohne zeitlichem Trend in den Raten (Friede u. a., 2018) auf rekurrierende Ereignisdaten mit negativem Zeittrend in den Raten angewendet, um zu untersuchen, wie sich Abweichungen von der Annahme von über die Zeit konstant bleibenden Raten auf die Ergebnisse, die mit dem Verfahren erzielt werden können, auswirken.

Das verblindete kontinuierliche Monitorieren dient in klinischen Studien dazu, Studienlaufzeiten oder Fallzahlen im Verlauf einer Studie so anzupassen, dass die vorher für eine Studie festgelegte Power erreicht wird. Das ist vor allem in solchen Situationen nützlich, in denen es besonders schwierig ist die Störparameter, wie Basisrate oder Überdispersion vor Studienbeginn abzuschätzen, zum Beispiel weil nicht genug Daten aus früheren Studien vorliegen. Weiterhin ist es in Fällen hilfreich, in denen es eine besondere Herausforderung darstellt, Patienten zu rekrutieren oder besonders zu schützende Patienten getestet werden, beispielsweise bei klinischen Studien zu seltenen Erkrankungen, wie der Multiplen Sklerose im Kindesalter (Friede u. a., 2018). Um eine optimale Laufzeitanpassung beim verblindeten kontinuierlichen Monitorieren zu erreichen wird die Information des Effektschätzers im Studienverlauf verblindet geschätzt und die Studie gestoppt sobald die Information ein bestimmtes, vorher festgelegtes Niveau erreicht hat. Wie bei jedem adaptiven Design, ist es wichtig dabei sicherzustellen, dass es durch die Veränderung am Studiendesign auf Basis der Interimsschätzung nicht zu einer Inflation des Fehlers 1. Art kommt.

Bei negativen Zeittrends in den Ereignisraten, in Höhe wie sie beispielsweise für Schubraten in einer MS-Studie erwartet werden können, resultiert für das verblindete kontinuierliche Monitorieren in den untersuchten Fällen weder eine praktisch relevante Inflation der Wahrscheinlichkeit des Fehlers 1. Art, noch ein Verlust an Power. Auch eine Abwei-

chung von der Voraussetzung konstanter Rate bei den untersuchten Szenarien in Form eines negativen Zeittrends stellt keine Gefahr im Sinne eines möglichen verfrühten Studienendes, und somit einem Verlust an Power oder einer Inflation der Wahrscheinlichkeit des Fehlers 1. Art dar. Das gilt sowohl für Startzeitpunkte des Monitorierens nach ungefähr der Hälfte der Rekrutierungszeit als auch für Startzeitpunkte des Monitorierens einen Monat nach Ende der Rekrutierungsphase. Bei größeren negativen Zeittrends und einem frühen Start des Monitorierens liefert die Berücksichtigung des zeitlichen Trends im Trendmodell bei der verblindeten Schätzung der Information Studiendesigns mit kürzeren Laufzeiten und geringeren Fallzahlen unter Beibehaltung der Power und beinahe ohne Inflation des Fehlers 1. Art. Die Unterschiede der Laufzeitmediane betrug für solche Szenarien in dieser Simulationsstudie circa 2 bis 3 Monate.

Die Schätzer des Behandlungseffekts weisen für beide Methoden nur geringfügige Verzerrungen auf. Auch die Überdeckungswahrscheinlichkeiten der 95%-Konfidenzintervalle liegen im erwarteten Bereich. Die Schätzer von anfänglicher Basisrate, Zeittrend und Überdispersion aus dem Trendmodell weisen ebenfalls keine bedeutsame Verzerrung auf. Wird der Zeittrend bei der Auswertung nach Monitorieren nicht berücksichtigt (an die Daten wird ein Negativ-binomial Modell angepasst) ist der Schätzer der Überdispersion für Fälle, in denen die Studie aufgrund des Monitorierens früher gestoppt wurde, verzerrt.

Die Unterschiede in der Power beider Methoden sind für die meisten untersuchten Szenarien gering. Erst bei frühem Monitorierungsstart und großen zeitlichen Trends, die stark über dem Ausmaß eines Zeittrends liegen, der beispielsweise in einer klinischen Studie zu MS erwartet werden kann (Nicholas u. a., 2012), zeigt sich in den simulierten Daten, dass die Power im Trendmodell bei Abweichung der Parameter von der Planungssituation (im Sinne von höheren Basisraten oder niedrigerer Überdispersion) näher am vorgegebenen Niveau liegt. In solchen Fällen scheint im Trendmodell eine genauere Abschätzung der tatsächlichen Information in den Daten gewährleistet zu sein als im Negativ-Binomial Modell, das den zeitlichen Trend nicht berücksichtigt. Es stellt sich die Frage, warum die Unterschiede zwischen beiden Verfahren nicht deutlicher ausfallen. Die Power hängt von der Größe des Effekts, von der Menge der Daten und der Fähigkeit des Tests ab, einen signifikanten Effekt zu erkennen, falls dieser tatsächlich vorhanden ist. Die Effektschätzer beider Verfahren weisen keine deutliche Verzerrung auf und auch die Überdeckungswahrscheinlichkeiten der 95%-Konfidenzintervalle befinden sich für beide Verfahren im erwarteten Bereich. Auch unter großen Zeittrends sind die Effektschätzer bei der fälschlichen Annahme konstanter Raten nicht deutlich verzerrt. Die Studienlaufzeit hängt von der Höhe der verblindet geschätzten Information ab, da die Studie erst dann gestoppt

wird, wenn die Information ein bestimmtes kritisches Level erreicht. Bei großen Zeittrends scheint es zu einer leichten Unterschätzung der verblindet geschätzten Information zu kommen, falls man den Zeittrend im Negativ-Binomial Modell unberücksichtigt lässt, was in diesen Fällen zu etwas längeren Laufzeiten und höherer Power führt. In den meisten untersuchten Fällen scheint sich die verblindet geschätzte Information zwischen den Verfahren jedoch nicht stark zu unterscheiden. Im Negativ-Binomial-Modell ist der Schätzer der Überdispersion verzerrt, in Richtung einer Überschätzung der Überdispersion und die Verzerrung vergrößert sich, je größer der Zeittrend ist. Daher wird angenommen, dass in den Szenarien, in denen sich die Verfahren kaum unterscheiden, die Basisraten im Negativ-Binomial überschätzt werden. Eine Überschätzung der Überdispersion und der Raten würden sich, betrachtet man die Formel zur Berechnung der Information des Effektschätzers im Negativ-Binomial-Modell, sozusagen gegenseitig ausgleichen können. Das könnte erklären, warum die Nichtberücksichtigung des Trends in vielen Fällen weder zu einer starken Erhöhung, noch zu einem Verlust an Power im Vergleich zum Verfahren mit Trendmodell führt. Hier wäre es für Folgestudien interessant, die verblindet geschätzten Information für beide Verfahren systematisch zu vergleichen. Auch die Erwartung, dass bei der Schätzung der kumulativen Basisraten bei Auslassung des zeitlichen Trends bei der Datenauswertung eine Verzerrung auftritt, im Sinne, dass die Basisraten hier überschätzt werden und in welcher Höhe dies geschieht sollte untersucht werden.

Bei der Bewertung der Ergebnisse sind einige Einschränkungen zu beachten. Zum einen wurden nur Fälle untersucht in denen die Nachbeobachtungszeiten in beiden Gruppen gleich verteilt sind. Dies ist in der Praxis nicht immer der Fall. Durch Studienabbrüche kann es zwischen beiden Behandlungsgruppen zu unterschiedlichen Fallzahlen und Nachbeobachtungszeiten kommen. Hier muss die Robustheit der Methoden weiter untersucht werden. Auch sind Situationen vorstellbar, in denen der zeitliche Trend nicht, wie hier simuliert, in beiden Gruppen gleich ist, sondern in einer der Behandlungsgruppen stärker ausfällt, beispielsweise aufgrund der Wirkweise des Medikaments. In zukünftigen Studien könnte überprüft werden, wie gut das Verfahren zum verblindeten kontinuierlichen Monitorieren mit Trendmodell bei unterschiedlichen Zeittrends für Behandlungs- und Kontrollgruppe funktionieren. Hier wäre interessant systematisch zu untersuchen wie sich eine Verzerrung bei der verblindeten Schätzung des Zeittrends auf die Information und somit auf die Power und Laufzeiten der Studien auswirkt. Interessant wäre es in zukünftigen Studien das Trendmodell für saisonale Zeittrends (beispielsweise bei der chronisch-obstruktiven Lungenerkrankung) zu untersuchen.

Das verblindete kontinuierliche Monitorieren bietet Vorteile in Bezug auf die Variabilität der Fallzahlen bei kontinuierlichen Daten gegenüber Fallzahlenanpassungsverfahren mit einer Fallzahlneuschätzung zu nur einem Zeitpunkt während der Studie (BSSR). Ein Verfahren zur BSSR bei rekurrierenden Ereignisdaten mit Zeittrend wurde von Schneider u. a. (2013) vorgestellt. Interessant wäre es das verblindete kontinuierliche Monitorieren zu erweitern, so dass auch eine Verlängerung der Studie möglich wird, sobald dies als nötig erachtet wird und die von Schneider u. a. (2013) vorgestellte Methode mit dem verblindeten kontinuierlichen Monitorieren für Trenddaten vergleichen. In dieser Studie wurden monatliche Monitorierungsschritte untersucht. Es wäre interessant zu untersuchen, was der beste Abstand der Monitorierungsschritte ist, der zu einem guten Ergebnis in Bezug auf Power und Wahrscheinlichkeit des Fehlers 1. Art und Variabilität der Laufzeiten führt, um den logistischen Aufwand des kontinuierlichen Monitorierens weniger aufwendig zu machen.

Insgesamt weist das verblindete kontinuierliche Monitorieren mit Trendmodell gute Eigenschaften in Bezug auf die Einhaltung der Power, der Wahrscheinlichkeit eines Fehlers 1. Art und der Anpassung der Laufzeit der Studien auf. Das Trendmodell liefert unverzerrte Schätzer. Beim Verfahren ohne Berücksichtigung des Zeittrends stellen negative Zeittrends im untersuchten Ausmaß kein Risiko für eine Inflation der Wahrscheinlichkeit des Fehlers 1. Art oder eine zu niedrige Laufzeit dar. Bei relativ großen zeitlichen Trends ermöglicht das verblindete kontinuierliche Monitorieren mit Trendmodell etwas kürzere Laufzeiten. Das Verfahren sollte in weiteren Simulationsstudien mit anderen Arten und Mustern von zeitlichen Trends (zum Beispiel positive Zeittrends oder saisonale Zeittrends) untersucht werden.

Literaturverzeichnis

- [Bauer u. a. 2016] BAUER, Peter ; BRETZ, Frank ; DRAGALIN, Vladimir ; KÖNIG, Franz ; WASSMER, Gernot: Twenty-five years of confirmatory adaptive designs: opportunities and pitfalls. In: *Statistics in Medicine* 35 (2016), Nr. 3, S. 325–347
- [Chuang-Stein u. a. 2006] CHUANG-STEIN, Christy ; ANDERSON, Keaven ; GALLO, Paul ; COLLINS, Sylva: Sample size reestimation: a review and recommendations. In: *Drug Information Journal* 40 (2006), Nr. 4, S. 475–484
- [Cox u. Lewis 1966] COX, DR ; LEWIS, PAWL: *The statistical analysis of series of events*. John Wiley and Sons, 1966
- [Cullagh u. Nelder 1983] CULLAGH, MC ; NELDER, J: *Generalised Linear Models*. 1983
- [Fahrmeir u. a. 2009] FAHRMEIR, Ludwig ; KNEIB, Thomas ; LANG, Stefan: *Regression: Modelle, Methoden und Anwendungen*. Springer-Verlag, 2009
- [Friede u. a. 2018] FRIEDE, Tim ; HÄRING, Dieter ; SCHMIDLI, Heinz: Blinded continuous monitoring in clinical trials with recurrent event endpoints. In: *In Begutachtung und wieder eingereicht bei Pharmaceutical Statistics* (2018, eingereicht)
- [Friede u. Miller 2012] FRIEDE, Tim ; MILLER, Frank: Blinded continuous monitoring of nuisance parameters in clinical trials. In: *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 61 (2012), Nr. 4, S. 601–618
- [Friede u. Schmidli 2010] FRIEDE, Tim ; SCHMIDLI, Heinz: Blinded sample size reestimation with negative binomial counts in superiority and non-inferiority trials. In: *Methods of Information in Medicine* 49 (2010), Nr. 06, S. 618–624
- [Held u. Sabanés Bové 2014] HELD, Leonhard ; SABANÉS BOVÉ, D: *Applied statistical inference*. Bd. 10. Springer, 2014. – 978–3 S.
- [Jennison u. Turnbull 2000] JENNISON, Christopher ; TURNBULL, Bruce W.: Group sequential methods with applications to clinical trials, 2000. In: *Chapman-Hall/CRC, Boca Raton, FL* (2000)
- [Keene u. a. 2007] KEENE, Oliver N. ; JONES, Mark R. ; LANE, Peter W. ; ANDERSON, Julie: Analysis of exacerbation rates in asthma and chronic obstructive pulmonary disease: example from the TRISTAN study. In: *Pharmaceutical Statistics: The Journal of Applied Statistics in the Pharmaceutical Industry* 6 (2007), Nr. 2, S. 89–97
- [Lawless 1987a] LAWLESS, Jerald F.: Negative binomial and mixed Poisson regression. In: *Canadian Journal of Statistics* 15 (1987), Nr. 3, S. 209–225
- [Lawless 1987b] LAWLESS, Jerald F.: Regression methods for Poisson process data. In: *Journal of the American Statistical Association* 82 (1987), Nr. 399, S. 808–815
- [Lewis u. Shedler 1979] LEWIS, Peter A. ; SHEDLER, Gerald S.: Simulation of nonhomogeneous Poisson processes by thinning. In: *Naval Research Logistics (NRL)* 26 (1979), Nr. 3, S. 403–413

- [Nicholas u. a. 2012] NICHOLAS, Richard ; STRAUBE, Sebastian ; SCHMIDLI, Heinz ; PFEIFFER, Sebastian ; FRIEDE, Tim: Time-patterns of annualized relapse rates in randomized placebo-controlled clinical trials in relapsing multiple sclerosis: A systematic review and meta-analysis. In: *Multiple Sclerosis Journal* 18 (2012), Nr. 9, S. 1290–1296
- [Pasupathy 2010] PASUPATHY, Raghu: Generating homogeneous poisson processes. In: *Wiley Encyclopedia of Operations Research and Management Science* (2010)
- [R Development Core Team 2008] R DEVELOPMENT CORE TEAM: *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, 2008. <http://www.R-project.org>. – ISBN 3-900051-07-0
- [Ross 2008] ROSS, Sheldon M.: *STOCHASTIC PROCESSES, 2ND ED.* Wiley India Pvt. Limited, 2008 (Wiley series in probability and statistics). <https://books.google.de/books?id=HVHqPgAACAAJ>. – ISBN 9788126517572
- [Schneider u. a. 2013] SCHNEIDER, Simon ; SCHMIDLI, Heinz ; FRIEDE, Tim: Blinded sample size re-estimation for recurrent event data with time trends. In: *Statistics in medicine* 32 (2013), Nr. 30, S. 5448–5457
- [Stoyan u. a. 1995] STOYAN, Dietrich ; KENDALL, Wilfrid S. ; MECKE, Joseph: *Stochastic Geometry and its Applications*. Wiley, Chichester, 1995
- [Tremlett u. a. 2008] TREMLETT, Helen ; ZHAO, Yinshan ; JOSEPH, Jay ; DEVONSHIRE, Virginia u. a.: Relapses in multiple sclerosis are age-and time-dependent. In: *Journal of Neurology, Neurosurgery & Psychiatry* 79 (2008), Nr. 12, S. 1368–1374
- [Venables u. Ripley 2002] VENABLES, W. N. ; RIPLEY, B. D.: *Modern Applied Statistics with S*. Fourth. New York : Springer, 2002 <http://www.stats.ox.ac.uk/pub/MASS4>. – ISBN 0-387-95457-0

Anhang A

Fisher-Information $I(\beta_1, \boldsymbol{\theta}, \varphi)$ im Trend-Modell

Man erhält die Fisher-Informationsmatrix für das Trend-Modell, indem man den Erwartungswert der negativen zweiten partiellen Ableitung der Log-Likelihood-Funktion nach dem Parametervektor bildet. Die Standardfehler der Parameterschätzer können aus der Inversen der Fisher-Informationsmatrix ermittelt werden, da $\hat{\boldsymbol{\theta}}_{ML} \overset{a}{\sim} \mathcal{N}_p(\boldsymbol{\theta}, \mathbf{I}(\hat{\boldsymbol{\theta}}_{ML})^{-1})$ und $(\hat{\boldsymbol{\theta}}_{ML})_i \overset{a}{\sim} \mathcal{N}(\theta_i, [\mathbf{I}(\hat{\boldsymbol{\theta}}_{ML})^{-1}]_{ii})$ (Held u. Sabanés Bové, 2014).

Beim Monitorieren wird die vor Beginn der Studie berechnete Information I^* mit der während der laufenden Studie berechneten Information des Effektschätzers I_{ML} verglichen (Studie wird gestoppt falls $I_{ML} \geq I^*$), wobei

$I^* = \frac{(z_{1-\alpha} + z_{1-\beta})^2}{\beta_1^{*2}}$ und $I_{ML} = \frac{1}{[\mathbf{I}(\hat{\boldsymbol{\theta}}_{ML})^{-1}]_{\beta_1 \beta_1}}$ und $\beta_1 = \log\left(\frac{\lambda_1}{\lambda_0}\right)$ und β_1^* der Effekt ist, der vor der Studie angenommen wird.

Die Log-Likelihood-Funktion für das Trend-Modell ist gegeben durch

$$l(\beta_1, \alpha_0, \alpha_1, \varphi) = \sum_{i=0}^1 \sum_{j=1}^{n_i} \sum_{k=1}^{Y_{ij}} (\alpha_0 + \alpha_1 t_{ijk}) + \sum_{i=0}^1 \sum_{j=1}^{n_i} Y_{ij} (\log(\varphi) + x_i \beta_1) +$$

$$\sum_{i=0}^1 \sum_{j=1}^{n_i} (\log(\Gamma(Y_{ij} + \varphi^{-1})) - \log(\Gamma(\varphi^{-1}))) -$$

$$\sum_{i=0}^1 \sum_{j=1}^{n_i} (Y_{ij} + \varphi^{-1}) \log\left(1 + \varphi \alpha_1^{-1} \exp(\alpha_0) \exp(x_i \beta_1) (\exp(\alpha_1 T_{ij}) - 1)\right)$$

Schneider u. a. (2013).

In der Fisher-Informationsmatrix

$$E \left\{ - \left[\frac{\partial^2}{\partial \theta^2} \log f(x; \beta_1, \alpha_1, \alpha_0, \varphi | \beta_1, \alpha_1, \alpha_0, \varphi) \right] \right\} = E \left\{ - \begin{pmatrix} \frac{\partial^2 l}{\partial \beta_1^2} & \frac{\partial^2 l}{\partial \beta_1 \partial \alpha_1} & \frac{\partial^2 l}{\partial \beta_1 \partial \alpha_0} & \frac{\partial^2 l}{\partial \beta_1 \partial \varphi} \\ \cdots & \frac{\partial^2 l}{\partial \alpha_1^2} & \frac{\partial^2 l}{\partial \alpha_1 \partial \alpha_0} & \frac{\partial^2 l}{\partial \alpha_1 \partial \varphi} \\ \cdots & \cdots & \frac{\partial^2 l}{\partial \alpha_0^2} & \frac{\partial^2 l}{\partial \alpha_0 \partial \varphi} \\ \cdots & \cdots & \cdots & \frac{\partial^2 l}{\partial \varphi^2} \end{pmatrix} \right\}$$

sind die Erwartungswerte aller gemischten partiellen Ableitungen zweiter Ordnung nach φ null.

Die ersten partiellen Ableitungen nach β_1 , α_0 und α_1 sind

$$\frac{\partial l}{\partial \beta_1} = \sum_{j=1}^{n_1} Y_{1j} - \sum_{j=1}^{n_1} \frac{\Lambda_1(T_{1j}, \theta)(1 + \varphi Y_{1j})}{1 + \varphi \Lambda_1(T_{1j}, \theta)} \quad (23)$$

$$\frac{\partial l}{\partial \alpha_0} = \sum_{i=0}^1 \sum_{j=1}^{n_i} \sum_{k=1}^{Y_{ij}} 1 - \sum_{i=0}^1 \sum_{j=1}^{n_i} \frac{\Lambda_i(T_{ij}, \theta)(1 + \varphi Y_{ij})}{1 + \varphi \Lambda_i(T_{ij}, \theta)} \quad (24)$$

$$\frac{\partial l}{\partial \alpha_1} = \sum_{i=0}^1 \sum_{j=1}^{n_i} \sum_{k=1}^{Y_{ij}} t_{ijk} - \sum_{i=0}^1 \sum_{j=1}^{n_i} \Lambda_i(T_{ij}, \theta) \left(\frac{1 + \varphi Y_{ij}}{\alpha_1(1 + \varphi \Lambda_i(T_{ij}, \theta))} - \frac{(1 + \varphi Y_{ij}) \exp(\alpha_1 T_{ij}) T_{ij}}{\alpha_1(1 + \varphi \Lambda_i(T_{ij}, \theta))(\exp(\alpha_1 T_{ij}) - 1)} \right) \quad (25)$$

wobei $\Lambda_i(T_{ij}, \theta) = E \{ \Lambda_i(T_{ij}, \theta) | U_{ij} = u_{ij} \} = 1/\alpha_1 \exp(\alpha_0) (\exp(\alpha_1 T_{ij}) - 1) \exp(x_i \beta_1)$.

Die zweiten gemischten partiellen Ableitungen nach φ sind

$$\frac{\partial^2 l}{\partial \beta_1 \partial \varphi} = - \sum_{j=1}^{n_1} \frac{\Lambda_1(T_{1j}, \theta)(Y_{1j} - \Lambda_1(T_{1j}, \theta))}{(1 + \varphi \Lambda_1(T_{1j}, \theta))^2} \quad (26)$$

$$\frac{\partial^2 l}{\partial \alpha_0 \partial \varphi} = - \sum_{i=0}^1 \sum_{j=1}^{n_i} \frac{\Lambda_i(T_{ij}, \theta)(Y_{ij} - \Lambda_i(T_{ij}, \theta))}{(1 + \varphi \Lambda_i(T_{ij}, \theta))^2} \quad (27)$$

$$\frac{\partial^2 l}{\partial \alpha_1 \partial \varphi} = - \sum_{i=0}^1 \sum_{j=1}^{n_i} \Lambda_i(T_{ij}, \theta) \left(\frac{Y_{ij} - \Lambda_i(T_{ij}, \theta)}{\alpha_1(1 + \varphi \Lambda_i(T_{ij}, \theta))^2} - \frac{(Y_{ij} - \Lambda_i(T_{ij}, \theta)) \exp(\alpha_1 T_{ij}) T_{ij}}{\alpha_1(1 + \varphi \Lambda_i(T_{ij}, \theta))^2 (\exp(\alpha_1 T_{ij}) - 1)} \right) \quad (28)$$

Da $E\{Y_{ij}\} = \Lambda_i(T_{ij}, \theta)$ ergibt sich, dass $E \left\{ -\frac{\partial^2 l}{\partial \beta_1 \varphi} \right\} = \sum_{j=1}^{n_1} \frac{\Lambda_1(T_{1j}, \theta)(E\{Y_{1j}\} - \Lambda_1(T_{1j}, \theta))}{(1 + \varphi \Lambda_1(T_{1j}, \theta))^2} = 0$ und ebenfalls, dass $E \left\{ -\frac{\partial^2 l}{\partial \alpha_0 \varphi} \right\} = E \left\{ -\frac{\partial^2 l}{\partial \alpha_1 \varphi} \right\} = 0$.

Für die Berechnung der Varianz von $\hat{\beta}_1$ über die Inverse der Fisher-Informationsmatrix ist daher nur $\nu = (\beta_1, \alpha_1, \alpha_0)$ notwendig, denn

$$\begin{bmatrix} \mathcal{I}(\nu) & 0 \\ 0 & a \end{bmatrix}^{-1} = \begin{bmatrix} \mathcal{I}(\nu)^{-1} & 0 \\ 0 & \frac{1}{a} \end{bmatrix}$$

$$\text{wobei } \mathcal{I}(\nu) = \begin{bmatrix} E \left\{ -\frac{\partial^2 l}{\partial \beta_1^2} \right\} & E \left\{ -\frac{\partial^2 l}{\partial \beta_1 \alpha_1} \right\} & E \left\{ -\frac{\partial^2 l}{\partial \beta_1 \alpha_0} \right\} \\ E \left\{ -\frac{\partial^2 l}{\partial \beta_1 \alpha_1} \right\} & E \left\{ -\frac{\partial^2 l}{\partial \alpha_1^2} \right\} & E \left\{ -\frac{\partial^2 l}{\partial \alpha_0 \alpha_1} \right\} \\ E \left\{ -\frac{\partial^2 l}{\partial \beta_1 \alpha_0} \right\} & E \left\{ -\frac{\partial^2 l}{\partial \alpha_0 \alpha_1} \right\} & E \left\{ -\frac{\partial^2 l}{\partial \alpha_0^2} \right\} \end{bmatrix}$$

$$H_{ij} = \exp(\alpha_0) \exp(x_i \beta_1) \exp(\alpha_1 T_{ij}) T_{ij}$$

$$H_{1j} = \exp(\alpha_0) \exp(\beta_1) \exp(\alpha_1 T_{1j}) T_{1j}$$

$$H_{0j} = \exp(\alpha_0) \exp(\alpha_1 T_{0j}) T_{0j}$$

$$E \left\{ -\frac{\partial^2 l}{\partial \beta_1 \beta_1} \right\} = \sum_{j=1}^{n_1} \frac{\Lambda_1(T_{1j})}{(1 + \varphi \Lambda_1(T_{1j}, \theta))}$$

$$E \left\{ -\frac{\partial^2 l}{\partial \alpha_0^2} \right\} = \sum_{j=1}^{n_0} \frac{\Lambda_0(T_{0j}, \theta)}{(1 + \varphi \Lambda_0(T_{0j}, \theta))} + \sum_{j=1}^{n_1} \frac{\Lambda_1(T_{1j}, \theta)}{(1 + \varphi \Lambda_1(T_{1j}, \theta))}$$

$$E \left\{ -\frac{\partial^2 l}{\partial \beta_1 \alpha_0} \right\} = \sum_{j=1}^{n_1} \frac{\Lambda_1(T_{1j})}{(1 + \varphi \Lambda_1(T_{1j}, \theta))}$$

$$\begin{aligned} E \left\{ -\frac{\partial^2 l}{\partial \alpha_0 \alpha_1} \right\} &= \sum_{j=1}^{n_0} \frac{\exp(\alpha_0) \exp(\alpha_1 T_{0j}) T_{0j} - \Lambda_0(T_{0j}, \theta)}{\alpha_1 (1 + \varphi \Lambda_0(T_{0j}, \theta))} + \sum_{j=1}^{n_1} \frac{\exp(\alpha_0) \exp(\beta_1) \exp(\alpha_1 T_{1j}) T_{1j} - \Lambda_1(T_{1j}, \theta)}{\alpha_1 (1 + \varphi \Lambda_1(T_{1j}, \theta))} \\ &= \sum_{i=0}^1 \sum_{j=1}^{n_i} \frac{H_{ij} - \Lambda_i(T_{ij}, \theta)}{\alpha_1 (1 + \varphi \Lambda_i(T_{ij}, \theta))} \end{aligned}$$

$$E \left\{ -\frac{\partial^2 l}{\partial \beta_1 \alpha_1} \right\} = \sum_{j=1}^{n_1} \frac{H_{1j} - \Lambda_1(T_{1j}, \theta)}{\alpha_1 (1 + \varphi \Lambda_1(T_{1j}, \theta))}$$

$$E \left\{ -\frac{\partial^2 l}{\partial \alpha_1^2} \right\} = \sum_{i=0}^1 \sum_{j=1}^{n_i} \alpha_1^{-1} H_{ij} T_{ij} - \frac{\varphi (H_{ij}^2 - (\Lambda_i(T_{ij}, \theta))^2) + 2 (H_{ij} - \Lambda_i(T_{ij}, \theta))}{\alpha_1^2 (1 + \varphi \Lambda_i(T_{ij}, \theta))}$$

Um den Eintrag $\mathcal{I}(\nu)_{11}^{-1}$ zu berechnen, kann man folgendes Ergebnis nutzen, das aus einer einfachen Formel für die Berechnung der Inversen einer 3x3 Matrix stammt:

$$\mathcal{I}(\nu)^{-1} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}^{-1} = \frac{1}{\det(\mathcal{I}(\nu))} \begin{bmatrix} (ei - fh) & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \end{bmatrix}$$

$$\det(\mathcal{I}(\nu)) = aei + bfg + cdh - gec - afh - idb$$

Da $\mathcal{I}(\nu)$ symmetrisch ist und $a = c$, ergibt sich

$$\det(\mathcal{I}(\nu)) = aei + 2abf - ea^2 - af^2 - ib^2$$

daher ist,

$$\begin{aligned} \mathcal{I}(\nu)_{11}^{-1} &= \frac{(ei - f^2)}{aei + 2abf - ea^2 - af^2 - ib^2} = \frac{i(ei - f^2)}{aei^2 + 2abfi - ea^2i - af^2i - i^2b^2} \\ &= \frac{i(ei - f^2)}{aei^2 + 2abfi - ea^2i - af^2i - i^2b^2 + a^2f^2 - a^2f^2} = \frac{i(ei - f^2)}{(ai - a^2)(ei - f^2) - (ib - af)^2} \end{aligned}$$

$$= \frac{E \left\{ -\frac{\partial^2 l}{\partial \alpha_0^2} \right\} \left[E \left\{ -\frac{\partial^2 l}{\partial \alpha_1 \alpha_1} \right\} E \left\{ -\frac{\partial^2 l}{\partial \alpha_0^2} \right\} - E \left\{ -\frac{\partial^2 l}{\partial \alpha_1 \alpha_0} \right\}^2 \right]}{\left[E \left\{ -\frac{\partial^2 l}{\partial \beta_1^2} \right\} E \left\{ -\frac{\partial^2 l}{\partial \alpha_0^2} \right\} - E \left\{ -\frac{\partial^2 l}{\partial \beta_1 \alpha_0} \right\}^2 \right] \left[E \left\{ -\frac{\partial^2 l}{\partial \alpha_1^2} \right\} E \left\{ -\frac{\partial^2 l}{\partial \alpha_0^2} \right\} - E \left\{ -\frac{\partial^2 l}{\partial \alpha_1 \alpha_0} \right\}^2 \right] - \left[E \left\{ -\frac{\partial^2 l}{\partial \beta_1 \alpha_1} \right\} E \left\{ -\frac{\partial^2 l}{\partial \alpha_0^2} \right\} - E \left\{ -\frac{\partial^2 l}{\partial \beta_1 \alpha_0} \right\} E \left\{ -\frac{\partial^2 l}{\partial \alpha_1 \alpha_0} \right\} \right]^2}$$

$$\begin{aligned}
[(\beta_1\beta_1)(\alpha_0\alpha_0) - (\beta_1\alpha_0)^2] &= \sum_{j=1}^{n_1} \frac{\Lambda_1(T_{1j})}{(1 + \varphi\Lambda_1(T_{1j}, \theta))} \left(\sum_{j=1}^{n_0} \frac{\Lambda_0(T_{0j}, \theta)}{(1 + \varphi\Lambda_0(T_{0j}, \theta))} + \sum_{j=1}^{n_1} \frac{\Lambda_1(T_{1j}, \theta)}{(1 + \varphi\Lambda_1(T_{1j}, \theta))} \right) - \left(\sum_{j=1}^{n_1} \frac{\Lambda_1(T_{1j})}{(1 + \varphi\Lambda_1(T_{1j}, \theta))} \right)^2 \\
&= \sum_{j=1}^{n_1} \frac{\Lambda_1(T_{1j})}{(1 + \varphi\Lambda_1(T_{1j}, \theta))} \sum_{j=1}^{n_0} \frac{\Lambda_0(T_{0j}, \theta)}{(1 + \varphi\Lambda_0(T_{0j}, \theta))}
\end{aligned}$$

$$\begin{aligned}
[(\beta_1\alpha_1)(\alpha_0\alpha_0) - (\beta_1\alpha_0)(\alpha_1\alpha_0)] &= \sum_{j=1}^{n_1} \frac{H_{1j} - \Lambda_1(T_{1j}, \theta)}{\alpha_1(1 + \varphi\Lambda_1(T_{1j}, \theta))} \left(\sum_{j=1}^{n_0} \frac{\Lambda_0(T_{0j}, \theta)}{(1 + \varphi\Lambda_0(T_{0j}, \theta))} + \sum_{j=1}^{n_1} \frac{\Lambda_1(T_{1j}, \theta)}{(1 + \varphi\Lambda_1(T_{1j}, \theta))} \right) \\
&\quad - \sum_{j=1}^{n_1} \frac{\Lambda_1(T_{1j})}{(1 + \varphi\Lambda_1(T_{1j}, \theta))} \left(\sum_{j=1}^{n_0} \frac{H_{0j} - \Lambda_0(T_{0j}, \theta)}{\alpha_1(1 + \varphi\Lambda_0(T_{0j}, \theta))} + \sum_{j=1}^{n_1} \frac{H_{1j} - \Lambda_1(T_{1j}, \theta)}{\alpha_1(1 + \varphi\Lambda_1(T_{1j}, \theta))} \right) \\
&= \sum_{j=1}^{n_0} \frac{\Lambda_0(T_{0j})}{(1 + \varphi\Lambda_0(T_{0j}, \theta))} \sum_{j=1}^{n_1} \frac{H_{1j} - \Lambda_1(T_{1j}, \theta)}{\alpha_1(1 + \varphi\Lambda_1(T_{1j}, \theta))} - \sum_{j=1}^{n_1} \frac{\Lambda_1(T_{1j})}{(1 + \varphi\Lambda_1(T_{1j}, \theta))} \sum_{j=1}^{n_0} \frac{H_{0j} - \Lambda_0(T_{0j}, \theta)}{\alpha_1(1 + \varphi\Lambda_0(T_{0j}, \theta))} \\
&= \sum_{j=1}^{n_0} \frac{\Lambda_0(T_{0j})}{(1 + \varphi\Lambda_0(T_{0j}, \theta))} \sum_{j=1}^{n_1} \frac{H_{1j}}{\alpha_1(1 + \varphi\Lambda_1(T_{1j}, \theta))} - \sum_{j=1}^{n_1} \frac{\Lambda_1(T_{1j})}{(1 + \varphi\Lambda_1(T_{1j}, \theta))} \sum_{j=1}^{n_0} \frac{H_{0j}}{\alpha_1(1 + \varphi\Lambda_0(T_{0j}, \theta))}
\end{aligned}$$

$$\begin{aligned}
[(\alpha_1 \alpha_1)(\alpha_0 \alpha_0) - (\alpha_1 \alpha_0)^2] &= \sum_{i=0}^1 \sum_{j=1}^{n_i} \frac{\Lambda_i(T_{ij}, \theta)}{(1 + \varphi \Lambda_i(T_{ij}, \theta))} \left(\sum_{i=0}^1 \sum_{j=1}^{n_i} \alpha_1^{-1} H_{ij} T_{ij} - \frac{\varphi (H_{ij}^2 - (\Lambda_i(T_{ij}, \theta))^2) + 2(H_{ij} - \Lambda_i(T_{ij}, \theta))}{\alpha_1^2 (1 + \varphi \Lambda_i(T_{ij}, \theta))} \right) - \\
&\quad \left(\sum_{i=0}^1 \sum_{j=1}^{n_i} \frac{H_{ij} - \Lambda_i(T_{ij}, \theta)}{\alpha_1 (1 + \varphi \Lambda_i(T_{ij}, \theta))} \right)^2 \\
&= \sum_{i=0}^1 \sum_{j=1}^{n_i} \frac{\Lambda_i(T_{ij}, \theta)}{(1 + \varphi \Lambda_i(T_{ij}, \theta))} \left(\sum_{i=0}^1 \sum_{j=1}^{n_i} \alpha_1^{-1} H_{ij} T_{ij} - \frac{\varphi (H_{ij}^2 - (\Lambda_i(T_{ij}, \theta))^2)}{\alpha_1^2 (1 + \varphi \Lambda_i(T_{ij}, \theta))} \right) - \\
&\quad \sum_{i=0}^1 \sum_{j=1}^{n_i} \frac{\Lambda_i(T_{ij}, \theta)}{(1 + \varphi \Lambda_i(T_{ij}, \theta))} \left(2 \frac{1}{\alpha_1} \sum_{i=0}^1 \sum_{j=1}^{n_i} \frac{H_{ij}}{\alpha_1 (1 + \varphi \Lambda_i(T_{ij}, \theta))} - 2 \frac{1}{\alpha_1} \sum_{j=1}^{n_i} \frac{\Lambda_i(T_{ij}, \theta)}{\alpha_1 (1 + \varphi \Lambda_i(T_{ij}, \theta))} \right) - \\
&\quad \left(\sum_{i=0}^1 \sum_{j=1}^{n_i} \frac{H_{ij}}{\alpha_1 (1 + \varphi \Lambda_i(T_{ij}, \theta))} - \sum_{i=0}^1 \sum_{j=1}^{n_i} \frac{\Lambda_i(T_{ij}, \theta)}{\alpha_1 (1 + \varphi \Lambda_i(T_{ij}, \theta))} \right)^2 \\
&= \sum_{i=0}^1 \sum_{j=1}^{n_i} \frac{\Lambda_i(T_{ij}, \theta)}{(1 + \varphi \Lambda_i(T_{ij}, \theta))} \left(\sum_{i=0}^1 \sum_{j=1}^{n_i} \alpha_1^{-1} H_{ij} T_{ij} - \frac{\varphi (H_{ij}^2 - (\Lambda_i(T_{ij}, \theta))^2)}{\alpha_1^2 (1 + \varphi \Lambda_i(T_{ij}, \theta))} \right) - \\
&\quad 2 \sum_{i=0}^1 \sum_{j=1}^{n_i} \frac{H_{ij}}{\alpha_1 (1 + \varphi \Lambda_i(T_{ij}, \theta))} \sum_{i=0}^1 \sum_{j=1}^{n_i} \frac{\Lambda_i(T_{ij}, \theta)}{\alpha_1 (1 + \varphi \Lambda_i(T_{ij}, \theta))} + 2 \left(\sum_{j=1}^{n_i} \frac{\Lambda_i(T_{ij}, \theta)}{\alpha_1 (1 + \varphi \Lambda_i(T_{ij}, \theta))} \right)^2 - \\
&\quad \left(\sum_{i=0}^1 \sum_{j=1}^{n_i} \frac{H_{ij}}{\alpha_1 (1 + \varphi \Lambda_i(T_{ij}, \theta))} \right)^2 + 2 \sum_{i=0}^1 \sum_{j=1}^{n_i} \frac{\Lambda_i(T_{ij}, \theta)}{\alpha_1 (1 + \varphi \Lambda_i(T_{ij}, \theta))} \sum_{i=0}^1 \sum_{j=1}^{n_i} \frac{H_{ij}}{\alpha_1 (1 + \varphi \Lambda_i(T_{ij}, \theta))} - \\
&\quad \left(\sum_{i=0}^1 \sum_{j=1}^{n_i} \frac{\Lambda_i(T_{ij}, \theta)}{\alpha_1 (1 + \varphi \Lambda_i(T_{ij}, \theta))} \right)^2 \\
&= \sum_{i=0}^1 \sum_{j=1}^{n_i} \frac{\Lambda_i(T_{ij}, \theta)}{(1 + \varphi \Lambda_i(T_{ij}, \theta))} \left(\sum_{i=0}^1 \sum_{j=1}^{n_i} \alpha_1^{-1} H_{ij} T_{ij} - \frac{\varphi (H_{ij}^2 - (\Lambda_i(T_{ij}, \theta))^2)}{\alpha_1^2 (1 + \varphi \Lambda_i(T_{ij}, \theta))} \right) + \\
&\quad \left(\sum_{j=1}^{n_i} \frac{\Lambda_i(T_{ij}, \theta)}{\alpha_1 (1 + \varphi \Lambda_i(T_{ij}, \theta))} \right)^2 - \left(\sum_{i=0}^1 \sum_{j=1}^{n_i} \frac{H_{ij}}{\alpha_1 (1 + \varphi \Lambda_i(T_{ij}, \theta))} \right)^2
\end{aligned}$$

$$I_0 = \sum_{j=1}^{n_0} \frac{\Lambda_0(T_{0j}, \theta)}{(1 + \varphi \Lambda_0(T_{0j}, \theta))}$$

$$I_1 = \sum_{j=1}^{n_1} \frac{\Lambda_1(T_{1j}, \theta)}{(1 + \varphi \Lambda_1(T_{1j}, \theta))}$$

$$H_{ij} = \exp(\alpha_0) \exp(x_i \beta_1) \exp(\alpha_1 T_{ij}) T_{ij}$$

$$H_{1j} = \exp(\alpha_0) \exp(\beta_1) \exp(\alpha_1 T_{1j}) T_{1j}$$

$$H_{0j} = \exp(\alpha_0) \exp(\alpha_1 T_{0j}) T_{0j}$$

$$\mathcal{I}(\nu)^{-1} = \frac{E \left\{ -\frac{\partial^2 l}{\partial \alpha_1^2} \right\} \left[E \left\{ -\frac{\partial^2 l}{\partial \alpha_1 \alpha_1} \right\} E \left\{ -\frac{\partial^2 l}{\partial \alpha_0^2} \right\} - E \left\{ -\frac{\partial^2 l}{\partial \alpha_1 \alpha_0} \right\} \right]^2}{\left[E \left\{ -\frac{\partial^2 l}{\partial \beta_1^2} \right\} E \left\{ -\frac{\partial^2 l}{\partial \alpha_0^2} \right\} - E \left\{ -\frac{\partial^2 l}{\partial \beta_1 \alpha_0} \right\} \right]^2 \left[E \left\{ -\frac{\partial^2 l}{\partial \alpha_1^2} \right\} E \left\{ -\frac{\partial^2 l}{\partial \alpha_0^2} \right\} - E \left\{ -\frac{\partial^2 l}{\partial \alpha_1 \alpha_0} \right\} \right]^2 - \left[E \left\{ -\frac{\partial^2 l}{\partial \beta_1 \alpha_1} \right\} E \left\{ -\frac{\partial^2 l}{\partial \alpha_0^2} \right\} - E \left\{ -\frac{\partial^2 l}{\partial \beta_1 \alpha_0} \right\} \right]^2}$$

$$= \frac{(I_0 + I_1) \left(\sum_{i=0}^1 \sum_{j=1}^{n_i} \frac{\Lambda_i(T_{ij}, \theta)}{(1 + \varphi \Lambda_i(T_{ij}, \theta))} \left(\sum_{i=0}^1 \sum_{j=1}^{n_i} \alpha_1^{-1} H_{ij} T_{ij} - \frac{\varphi (H_{ij}^2 - (\Lambda_i(T_{ij}, \theta))^2)}{\alpha_1^2 (1 + \varphi \Lambda_i(T_{ij}, \theta))} \right) \right) + \left(\sum_{j=1}^{n_1} \frac{H_{1j}}{\alpha_1 (1 + \varphi \Lambda_1(T_{1j}, \theta))} \right)^2 - \left(\sum_{j=1}^{n_1} \frac{\Lambda_1(T_{1j}, \theta)}{\alpha_1 (1 + \varphi \Lambda_1(T_{1j}, \theta))} \right)^2}{(I_0 + I_1) \left(\sum_{i=0}^1 \sum_{j=1}^{n_i} \frac{\Lambda_i(T_{ij}, \theta)}{(1 + \varphi \Lambda_i(T_{ij}, \theta))} \left(\sum_{i=0}^1 \sum_{j=1}^{n_i} \alpha_1^{-1} H_{ij} T_{ij} - \frac{\varphi (H_{ij}^2 - (\Lambda_i(T_{ij}, \theta))^2)}{\alpha_1^2 (1 + \varphi \Lambda_i(T_{ij}, \theta))} \right) \right) + \left(\sum_{j=1}^{n_1} \frac{H_{1j}}{\alpha_1 (1 + \varphi \Lambda_1(T_{1j}, \theta))} \right)^2 - \left(\sum_{j=1}^{n_1} \frac{\Lambda_1(T_{1j}, \theta)}{\alpha_1 (1 + \varphi \Lambda_1(T_{1j}, \theta))} \right)^2} - \left(I_0 \sum_{j=1}^{n_0} \frac{H_{0j}}{\alpha_1 (1 + \varphi \Lambda_0(T_{0j}, \theta))} - I_1 \sum_{j=1}^{n_0} \frac{H_{0j}}{\alpha_1 (1 + \varphi \Lambda_0(T_{0j}, \theta))} \right)^2 \quad (29)$$

$$= \frac{(I_0 + I_1) \left(\sum_{i=0}^1 \sum_{j=1}^{n_i} \alpha_1^{-1} H_{ij} T_{ij} - \frac{\varphi (H_{ij}^2 - (\Lambda_i(T_{ij}, \theta))^2)}{\alpha_1^2 (1 + \varphi \Lambda_i(T_{ij}, \theta))} \right) + \alpha_1^{-2} (I_0 + I_1)^2 - \left(\sum_{i=0}^1 \sum_{j=1}^{n_i} \frac{H_{ij}}{\alpha_1 (1 + \varphi \Lambda_i(T_{ij}, \theta))} \right)^2}{(I_0 + I_1) \left(\sum_{i=0}^1 \sum_{j=1}^{n_i} \alpha_1^{-1} H_{ij} T_{ij} - \frac{\varphi (H_{ij}^2 - (\Lambda_i(T_{ij}, \theta))^2)}{\alpha_1^2 (1 + \varphi \Lambda_i(T_{ij}, \theta))} \right) + \alpha_1^{-2} (I_0 + I_1)^2 - \left(\sum_{i=0}^1 \sum_{j=1}^{n_i} \frac{H_{ij}}{\alpha_1 (1 + \varphi \Lambda_i(T_{ij}, \theta))} \right)^2} - \left(I_0 \sum_{j=1}^{n_0} \frac{H_{0j}}{\alpha_1 (1 + \varphi \Lambda_0(T_{0j}, \theta))} - I_1 \sum_{j=1}^{n_0} \frac{H_{0j}}{\alpha_1 (1 + \varphi \Lambda_0(T_{0j}, \theta))} \right)^2 \quad (30)$$

$$\mathcal{I}(\nu)_{11}^{-1} = \frac{(I_0 + I_1) [W]}{I_0 I_1 [W] - \left(I_0 \sum_{j=1}^{n_1} \frac{H_{1j}}{\alpha_1 (1 + \varphi \Lambda_1(T_{1j}, \theta))} - I_1 \sum_{j=1}^{n_0} \frac{H_{0j}}{\alpha_1 (1 + \varphi \Lambda_0(T_{0j}, \theta))} \right)^2} \quad (31)$$

mit

$$[W] = \left[E \left\{ -\frac{\partial^2 l}{\partial \alpha_1 \alpha_1} \right\} E \left\{ -\frac{\partial^2 l}{\partial \alpha_0^2} \right\} - E \left\{ -\frac{\partial^2 l}{\partial \alpha_1 \alpha_0} \right\} \right]^2 = \left((I_0 + I_1) \left(\sum_{i=0}^1 \sum_{j=1}^{n_i} \alpha_1^{-1} H_{ij} T_{ij} - \frac{\varphi (H_{ij}^2 - (\Lambda_i(T_{ij}, \theta))^2)}{\alpha_1^2 (1 + \varphi \Lambda_i(T_{ij}, \theta))} \right) + \alpha_1^{-2} (I_0 + I_1)^2 - \left(\sum_{i=0}^1 \sum_{j=1}^{n_i} \frac{H_{ij}}{\alpha_1 (1 + \varphi \Lambda_i(T_{ij}, \theta))} \right)^2 \right)^2$$

Anhang B (R-Code)

Simulationsfunktion

```

1 # Simulation of Poisson process with time trend
2
3 simPP <- function(n0, theta, effect, phi = NULL, k = 1,
4                 recrYRS, flwYRS, interimYRS = NULL){
5
6   # overall number of participants
7   n <- n0*(1 + k)
8
9   # participant ID
10  id <- 1:n
11
12  # recruitment times for linear recruitment
13  rec.time <- c(seq(0, recrYRS, length.out = n0), seq(0, recrYRS, length.out = k*n0))
14
15  # observation times
16  if(is.null(interimYRS)) {obs.time <- rep(flwYRS, n)} else {
17
18      obs.time <- interimYRS - rec.time
19      obs.time[obs.time < 0] <- 0
20      obs.time[obs.time > flwYRS] <- flwYRS
21
22  }
23
24  # label treatment group
25  group <- c(rep(0,n0), rep(1,n0*k))
26
27  # between-patient heterogeneity (gamma distributed rates)
28  # gamma-distributed with expectation 1 and variance phi
29  if(is.null(phi) || phi == 0){u <- rep(1, n)} else {
30      u <- rgamma(n, shape = 1/phi, scale = phi)}
31
32  # lambda at time t = 0
33  lambda <- u*exp(theta[1])*exp(log(effect)*group)
34
35  # function to simulate event times from stationary poisson process
36  # (for e.count = 0 the e.times = numeric(0) )
37  events <- function(lmbd, years){
38      rate <- lmbd*years
39      e.count <- rpois(1, rate)
40      e.tms <- runif(e.count, min = 0, max = years)
41      e.tms <- sort(e.tms)
42      return(e.tms)
43  }
44
45  pp <- list()
46
47  for(i in 1:n){pp[[i]] <- list(u = u[i],
48                          lambda = lambda[i],
49                          e.times.stat = events(lambda[i], obs.time[i]),
50                          group = group[i])
51  }
52
53  # thinning
54  for(i in 1:n){
55      # generate random numbers from uniform distribution between 0 and 1
56      pp[[i]]$z <- runif(length(pp[[i]]$e.times.stat))
57
58      # calculate (non-stationary) lambda for each stationary event time
59      pp[[i]]$lambda_Xk <- pp[[i]]$u*exp(theta[1] + theta[2]*pp[[i]]$e.times.stat)*exp(log(effect)*group[i])
60
61      # keep event time if z <= (lambda_Xk/lambda), i.e. delete event time if z > (lambda_Xk/lambda)
62      pp[[i]]$e.times <- pp[[i]]$e.times.stat[pp[[i]]$z <= (pp[[i]]$lambda_Xk/pp[[i]]$lambda)]
63
64      # number of events N(T)
65      pp[[i]]$n.events <- length(pp[[i]]$e.times)
66  }
67
68  n.events = unlist(lapply(pp, '[', 'n.events'))
69  e.times = lapply(pp, '[', 'e.times')
70
71  # make start / stop columns

```

```
72 for(i in 1:n){e.times[[i]] <- c(0, rep(e.times[[i]], each = 2), NA)}
73 e.times.mat <- matrix(unlist(e.times), ncol = 2, byrow = TRUE)
74 colnames(e.times.mat) <- c("start", "stop")
75
76 # store variables in data-frame
77 d <- data.frame( id = id,
78                 group = group,
79                 n.events = n.events,
80                 obs.time = obs.time,
81                 rec.time = rec.time,
82                 u = u,
83                 lambda = lambda)
84
85 # add row for each event
86 d <- d[rep.int(seq_len(nrow(d)), (n.events+1)), ]
87 d <- cbind(d, e.times.mat)
88
89 # censoring variable
90 d$event <- 1
91 d$event[is.na(d$stop)] <- 0
92
93 d$stop[is.na(d$stop)] <- obs.time
94
95 # interarrival times
96 d$gaptime <- d$stop - d$start
97
98 # exclude cases with:
99 # - "no censoring" (event occurs at very last time point during observation time)
100 # - observation time = 0
101 d <- subset(d, d$start != d$stop)
102
103 return(d)
104 }
```

Log-Likelihood

```
1 # log-likelihood
2 LL <- function(param, d){
3   alpha0 <- param[1]
4   alpha1 <- param[2]
5   phi <- param[3]
6   # beta = 0 in blinded estimation
7   beta <- if(is.na(param[4])) {beta <- 0} else {beta <- param[4]}
8
9   if(phi < 1e-6) phi <- 1e-6
10  #if(phi < 1e-6) return(-Inf)
11
12  d_nr <- subset(d, d$start == 0)
13
14  y_ij <- d_nr$mon.n.events
15  T_ij <- d_nr$mon.obs.time
16  t_ijk <- d$stop[d$mon.event == 1]
17
18  # no information about treatment allocation in blinded estimation
19  if(is.na(param[4])) {x_i <- 0} else {x_i <- d_nr$group}
20
21  x <- sum(alpha0 + alpha1*t_ijk) +
22      sum(y_ij*(log(phi) + x_i*beta)) +
23      sum(lgamma(y_ij + 1/phi) - lgamma(1/phi)) -
24      sum((y_ij + 1/phi)*log(1 + (phi/alpha1)*exp(alpha0)*exp(x_i*beta)*(exp(alpha1*T_ij) - 1)))
25
26  return(x)
27 }
```

Startwerte

```
1 startVals <- function(x, x_nr, mon.time, blnd = TRUE){
2
3   # how many events in first half year's (or less) observations of each patient?
4   sumE_halftime <- sum(x$mon.event[x$stop <= 0.5])
5
6   # how long is the observation time (sum time of all patients)?
7   sumT_halftime <- sum(sapply(x_nr$mon.obs.time, function(x) min(x, 0.5)))
8
9   # alpha 0
10  a0.strt <- log(sumE_halftime / sumT_halftime)
11
12  if(sumE_halftime == 0){a0.strt <- log(1e-6)}
13
14  # alpha 1
15  a1.strt <- -0.1
16
17  # phi
18  phi.strt <- (var(x_nr$mon.n.events) - mean(x_nr$mon.n.events)) / mean(x_nr$mon.n.events)^2
19
20  if(phi.strt < 1e-6) phi.strt <- 1e-6
21
22  # start values: blinded
23  sv <- c(a0.strt = a0.strt, a1.strt = a1.strt, phi.strt = phi.strt)
24
25  # start values: unblinded
26  if(!blnd) {
27    b1.strt <- log( mean(x_nr$mon.n.events[x_nr$group == 1]) / mean(x_nr$mon.n.events[x_nr$group == 0]) )
28
29    if(sum(x_nr$mon.n.events[x_nr$group == 1]) == 0) stop("group 1 has zero events")
30    if(sum(x_nr$mon.n.events[x_nr$group == 0]) == 0) stop("group 0 has zero events")
31
32    sv <- c(sv, b1.strt = b1.strt)
33  }
34
35  return(sv)
36 }
```

BCM-TREND

```

1 bcmtrend = function(n0, Tbl, Ta0, Ta1, Tphi, recrYRS, flwYRS, interimYRS,
2                     n.sim, FIcrit, firstlook, tstep){
3
4   ### save place for data
5
6   # estimates
7   Eb1 = rep(NA, n.sim) # estimate effect
8   Eb1_Se = rep(NA, n.sim) # estimate standard error of effect
9   Ea0 = rep(NA, n.sim)
10  Ea1 = rep(NA, n.sim)
11  Ephi = rep(NA, n.sim)
12
13  # likelihood ratio test
14  LR = rep(NA, n.sim)
15
16  # number of patients
17  N = rep(NA, n.sim)
18
19  # convergence
20  conv.monloop = rep(NA, n.sim) # convergence in monitoring loop
21  conv.monloop.ct = rep(NA, n.sim) # convergence in monitoring loop
22  conv = rep(NA, n.sim) # convergence full model
23  conv0 = rep(NA, n.sim) # convergence null model
24
25  # count step 2 for calculating alternative start values
26  strtvls.step2.monloop = rep(NA, n.sim)
27  strtvls.step2.monloop.ct = rep(NA, n.sim)
28  strtvls.step2.x1 = rep(NA, n.sim)
29  strtvls.step2.x0 = rep(NA, n.sim)
30
31  # warnings
32  warns.monloop = rep(NA, n.sim)
33  warns.monloop.ct = rep(NA, n.sim)
34  warns.m1 = rep(NA, n.sim)
35  warns.m0 = rep(NA, n.sim)
36
37  warns.min.theta.monloop = rep(NA, n.sim)
38  warns.theta.m1 = rep(NA, n.sim)
39  warns.theta.m0 = rep(NA, n.sim)
40
41  # a1 was zero
42  no.trend.monloop = rep(NA, n.sim)
43  no.trend.m1 = rep(NA, n.sim)
44
45  # phi was < 1e-6
46  phi.too.small.monloop = rep(NA, n.sim)
47  phi.too.small.m1 = rep(NA, n.sim)
48
49  # average follow-up time and study length
50  avgFU = rep(NA, n.sim)
51  SL = rep(NA, n.sim)
52
53  # observed information (blinded)
54  obsFImon = rep(0, n.sim)
55
56  simnr = rep(NA, n.sim)
57
58  ### maximal observation time
59  max.obs.time = flwYRS
60
61  ### n.sim loop
62  for(j in 1:n.sim){
63
64    my.objects <- c("a0.hat",
65                  "a0.hat.0",
66                  "b1b1",
67                  "a0a0",
68                  "a1a1",
69                  "a1a0",
70                  "a1b1",
71                  "a0b1",
72                  "b1_SE",
73                  "d_bcm",
74                  "d_bcm_nr",

```



```

75     "D0",
76     "D1",
77     "FisherInfo",
78     "fisher.info.inv.11.bld",
79     "fisher.info.inv.11",
80     "H0",
81     "H1",
82     "index",
83     "L0",
84     "L1",
85     "mstrt.vls",
86     "m0",
87     "m",
88     "mx",
89     "mm",
90     "ma0.hat",
91     "mphi.hat",
92     "mprm",
93     "mDO",
94     "mD1",
95     "mD",
96     "mHO",
97     "mH1",
98     "mTj",
99     "mb1b1",
100    "ma0a0",
101    "mala1",
102    "mala0",
103    "ma1b1",
104    "ma0b1",
105    "mLO",
106    "mL1",
107    "mFisherInfo",
108    "mon.time",
109    "phi.hat",
110    "phi.hat.0",
111    "prm",
112    "rnms",
113    "strt.vls",
114    "simdat",
115    "Tj",
116    "T1j",
117    "T0j",
118    "unbl.FisherInfo",
119    "x",
120    "x0")
121
122    rm(list = ls()[ls() %in% my.objects])
123
124 # Count warnings/convergence
125
126     # count non-convergence in monitoring loop
127     ncount <- 0
128     nyesno <- 0
129
130     # count warnigns for glm.nb() start vals in monitoring loop
131     wncount <- 0
132     wnyesno <- 0
133
134     # count step 2 and step 3 for calculating alternative start values
135     step2 <- 0
136     step2.count <- 0
137     step2.x1 <- 0
138     step2.x0 <- 0
139
140     # warnings for step2 in final analysis
141     wnyesno.ml <- 0
142     wnyesno.m0 <- 0
143
144     # save theta when iteration limit is reached
145     wnthetamin <- Inf
146     wntheta.ml <- NA
147     wntheta.m0 <- NA
148
149     # count a1 was zero
150     a1.was.zero <- 0

```

```

151         a1.was.zero.ml <- 0
152
153         # count phi < 1e-6
154         phi.blw.lemin6 <- 0
155         phi.blw.lemin6.ml <- 0
156
157
158
159 # Initialize variables
160         # start point for monitoring
161         k <- 0
162         # start value for Fisher Information
163         mFisherInfo <- 0
164
165 ### simulate data until "endpoint at planning stage"
166         simdat <- simPP(n0 = n0,
167             theta = c(Ta0, Ta1),
168             effect = exp(Tb1),
169             phi = Tphi,
170             recrYRS = recrYRS,
171             fllwYRS = fllwYRS,
172             interimYRS = interimYRS)
173
174         if(sum(is.na(simdat))>0) stop("there is a missing value in the data set")
175
176         # absolute event times (= on investigator's time scale)
177         simdat$abs.start <- simdat$start + simdat$rec.time
178         simdat$abs.stop <- simdat$stop + simdat$rec.time
179
180         #####
181         ### repeat monitoring until Information reaches critical value or study ends ###
182
183         while(mFisherInfo < FIcrit){
184
185             # monitoring time
186             mon.time <- firstlook + k*step
187
188             # subset data for blinded continuous monitoring
189             # (= throw out events with event time > monitoring time)
190
191             # all events which start before mon.time and censored events
192             d_bcm <- simdat[simdat$abs.start <= mon.time, ]
193
194             # categorize events (event = 1, censored = 0)
195             d_bcm$mon.event <- ifelse(d_bcm$abs.stop <= mon.time & d_bcm$stop < max.obs.time, 1, 0)
196
197             # observation times at current monitoring time point
198             d_bcm$mon.obs.time <- mon.time - d_bcm$rec.time
199
200             # exclude patients with zero observation time (i.e. rec.time == mon.time)
201             d_bcm <- d_bcm[d_bcm$mon.obs.time != 0, ]
202
203             # upper limit for observation times (i.e. maximal follow-up time)
204             d_bcm$mon.obs.time[d_bcm$mon.obs.time > max.obs.time] <- max.obs.time
205
206             # reduce data to n rows (= one row per patient)
207             rnms <- as.numeric(row.names(d_bcm))
208             index <- rnms %% 1 == 0
209             d_bcm_nr <- d_bcm[index, c("id", "group", "n.events", "mon.obs.time")]
210
211             # number of events at current monitoring time point
212             d_bcm_nr$mon.n.events <- sapply(split(d_bcm_nr$mon.event, d_bcm_nr$id), sum)
213
214             # go one step ahead if no events occurred until monitoring time point
215             if(sum(d_bcm_nr$mon.n.events) == 0) {k <- k + 1;
216                 if(mon.time - (recrYRS + fllwYRS) >= 0) break;
217                 next}
218
219             # export number of events at current monitoring time point to d_bcm
220             d_bcm$mon.n.events <- NA
221             split(d_bcm$mon.n.events, d_bcm$id) <- split(d_bcm_nr$mon.n.events, d_bcm_nr$id)
222
223 ## bcm trend
224         # start values
225         mstrt.vls <- startVals(d_bcm, d_bcm_nr, mon.time, blnd = TRUE)
226

```

```

227 # maximisation of the log-likelihood with the pooled data
228 mx <- optim(par = c(mstrt.vls["a0.strt"], mstrt.vls["a1.strt"], mstrt.vls["phi.strt"]), fn = LL,
229 d = d_bcm, control = list(fnscale = -1))
230
231 # step2: choose new starting values if algorithm does not converge
232 # *****
233 if(mx$convergence > 0){
234   step2 <- 1
235   step2.count <- step2.count + 1
236
237   mm <- MASS::glm.nb(mon.n.events ~ 1 + offset(log(mon.obs.time)),
238 data = d_bcm_nr)
239   ma0.hat <- coef(mm)[1]
240
241   # do poisson model if warning
242   if(!is.null(mm$th.warn) && mm$th.warn == "iteration limit reached"){
243     # count cases with warnigns
244     wnyesno <- 1
245     wncount <- wncount + 1
246     wnthetamin <- min(wnthetamin, mm$theta)
247
248     # poisson model
249     mm <- glm(mon.n.events ~ 1 + offset(log(mon.obs.time)),
250 family = poisson(link = "log"), data = d_bcm_nr)
251
252     ma0.hat <- coef(mm)[1]
253   }
254
255   mphi.hat <- ifelse(1/mm$theta > 1e-6, 1/mm$theta, 1e-6)
256
257   mx <- optim(par = c(ma0.hat, mstrt.vls["a1.strt"], mphi.hat), fn = LL, d = d_bcm,
258 control = list(fnscale = -1))
259 }
260
261
262 if(mx$convergence != 0) {nyesno <- 1;
263 ncount <- ncount + 1}
264
265 #Blinded Estimation of parameters
266 #Estimates
267 mprm <- mx$par
268 names(mprm) <- c("a0", "a1", "phi")
269
270 # Set lower bound for phi
271 if(mprm["phi"] < 1e-6) {mprm["phi"] <- 1e-6
272 phi.blw.1emin6 <- 1}
273
274 # Set Dm to exp(a0)*t when a1 is close to zero
275 if(mprm["a1"] > -1e-6 & mprm["a1"] < 1e-6) {
276   # count
277   a1.was.zero <- a1.was.zero + 1
278
279   mD <- exp(mprm["a0"])*d_bcm_nr$mon.obs.time
280
281 } else {
282
283   mD <- (1/mprm["a1"])*exp(mprm["a0"])*(exp(mprm["a1"]*d_bcm_nr$mon.obs.time) - 1)
284 }
285
286 #Blinded estimation of cumulative rates Delta0 and Delta1
287 mD0 <- mD*(1 + expct.RR)
288 mD1 <- mD0*expct.RR
289
290 mTj <- d_bcm_nr$mon.obs.time
291
292 # rename cumulative rates
293 mL0 <- mD0
294 mL1 <- mD1
295
296 # term
297 mH0 <- exp(mprm["a0"])*exp(mprm["a1"]*mTj)*mTj
298 mH1 <- mH0*expct.RR
299
300 mb1b1 <- 0.5*(sum(mL1/(1+mprm["phi"]*mL1)))
301
302 ma1a1 <- 0.5*(

```

```

303      sum(1/mprm["a1"]*mH0*mTj - (mprm["phi"]*(mH0^2-mL0^2) +
304      2*(mH0 - mL0))/((mprm["a1"]^2)*(1+mprm["phi"]*mL0))) +
305      sum(1/mprm["a1"]*mH1*mTj - (mprm["phi"]*(mH1^2-mL1^2) + 2*(mH1 - mL1))/((mprm["a1"]^2)*(1+mprm["phi"]*mL1))))
306
307      mala0 <- 0.5*(sum((mH0-mL0)/(mprm["a1"]*(1+mprm["phi"]*mL0))) + sum((mH1-mL1)/(mprm["a1"]*(1+mprm["phi"]*mL1))))
308      malb1 <- 0.5*(sum((mH1-mL1)/(mprm["a1"]*(1+mprm["phi"]*mL1))))
309      ma0b1 <- 0.5*(sum(mL1/(1+mprm["phi"]*mL1)))
310      ma0a0 <- 0.5*(sum(mL1/(1+mprm["phi"]*mL1)) + sum(mL0/(1+mprm["phi"]*mL0)))
311
312      fisher.info.inv.ll.bld <- ma0a0*(mala1*ma0a0 - mala0^2)/
313      ((mb1b1*ma0a0 - ma0b1^2)*(mala1*ma0a0 - mala0^2)-(malb1*ma0a0 - ma0b1*mala0)^2)
314      #####
315
316      #Blinded estimation of Fisher Info
317      mFisherInfo <- 1/fisher.info.inv.ll.bld
318      **
319
320      ### go one month ahead
321      k <- k + 1
322
323      ### stop if all patients reach max. follow-up time
324      ### (absolute monitoring time - (recruitment period + follow-up))
325      if(mon.time - (recrYRS + flwYRS) >= 0 ) break
326    }
327
328      ### monitoring stop #####
329      #####
330
331      oFlm <- mFisherInfo
332
333      ### Wald-Test
334
335      # start values
336      strt.vls <- startVals(d_bcm, d_bcm_nr, mon.time, blnd = FALSE)
337
338      # maximisation of the log-likelihood
339      x <- optim(par = c(strt.vls["a0.strt"], strt.vls["a1.strt"], strt.vls["phi.strt"], strt.vls["b1.strt"]), fn = LL,
340      d = d_bcm,
341      control = list(fnscale = -1))
342
343      # Choose new starting values if algorithm does not converge
344      # step2#####
345      if(x$convergence > 0){
346
347        step2.x1 <- 1
348
349        d_bcm_nr$groupf <- as.factor(d_bcm_nr$group)
350
351        m <- MASS::glm.nb(mon.n.events ~ groupf + offset(log(mon.obs.time)), data = d_bcm_nr)
352        a0.hat <- coef(m)[1]
353        b.hat <- coef(m)[2]
354        phi.hat <- ifelse(1/m$theta > 1e-6, 1/m$theta, 1e-6)
355
356        if(!is.null(m$th.warn) && m$th.warn == "iteration limit reached"){
357          # count cases with warnings
358          wnyesno.ml <- 1
359          wntheta.ml <- m$theta
360
361          d_bcm_nr$groupf <- factor(d_bcm_nr$group)
362
363          # poisson model
364          m <- glm(mon.n.events ~ groupf + offset(log(mon.obs.time)),
365          family = poisson(link = "log"),
366          data = d_bcm_nr)
367
368          a0.hat <- coef(m)[1]
369          b.hat <- coef(m)[2]
370          phi.hat <- 1e-6
371        }
372
373
374
375      x <- optim(par = c(a0.hat, strt.vls["a1.strt"], phi.hat, b.hat), fn = LL, d = d_bcm,
376      control = list(fnscale = -1))
377
378

```

```

379   }
380
381 # Estimates
382   prm <- x$par
383
384   if(length(x$par) < 4) stop("less than 4 parameters in final analysis")
385
386   names(prm) <- c("a0", "a1", "phi", "beta1")
387
388
389   # Set lower bound for phi
390   if(prm["phi"] < 1e-6){ prm["phi"] <- 1e-6
391     phi.blw.lemin6.ml <- 1 }
392
393   # Set D0 to exp(a0) when a1 is close to zero
394   if(prm["a1"] > -1e-6 && prm["a1"] < 1e-6){
395     a1.was.zero.ml <- 1
396
397     D0 <- exp(prm["a0"])*d_bcm_nr$mon.obs.time[d_bcm_nr$group == 0]
398
399     } else{
400
401     # Cumulative rate Delta0
402     D0 <- (1/prm["a1"])*exp(prm["a0"])*(exp(prm["a1"]*d_bcm_nr$mon.obs.time[d_bcm_nr$group == 0]) - 1)
403   }
404
405   # Cumulative rate Delta1
406   D1 <- D0*exp(prm["beta1"])
407
408   # Fisher Info
409
410   T0j <- d_bcm_nr$mon.obs.time[d_bcm_nr$group == 0]
411   T1j <- d_bcm_nr$mon.obs.time[d_bcm_nr$group == 1]
412
413   # rename cumulative rates
414   L0 <- D0
415   L1 <- D1
416
417   # term
418   H0 <- exp(prm["a0"])*exp(prm["a1"]*T0j)*T0j
419   H1 <- H0*exp(prm["beta1"])
420
421   b1b1 <- (sum(L1/(1+prm["phi"]*L1)))
422   a1a1 <- (sum(1/prm["a1"]*H0*T0j - (prm["phi"]*(H0^2-L0^2) + 2*(H0 - L0))/ ((prm["a1"]^2)*(1+prm["phi"]*L0))) +
423     sum(1/prm["a1"]*H1*T1j - (prm["phi"]*(H1^2-L1^2) + 2*(H1 - L1))/((prm["a1"]^2)*(1+prm["phi"]*L1))))
424   a1a0 <- (sum((H0-L0)/(prm["a1"]*(1+prm["phi"]*L0))) + sum((H1-L1)/(prm["a1"]*(1+prm["phi"]*L1))))
425   a1b1 <- (sum((H1-L1)/(prm["a1"]*(1+prm["phi"]*L1))))
426   a0b1 <- (sum(L1/(1+prm["phi"]*L1)))
427   a0a0 <- (sum(L1/(1+prm["phi"]*L1)) + sum(L0/(1+prm["phi"]*L0)))
428
429   fisher.info.inv.ll <- a0a0*(a1a1*a0a0 - a1a0^2)/
430     ((b1b1*a0a0 - a0b1^2)*(a1a1*a0a0 - a1a0^2)-(a1b1*a0a0 - a0b1*a1a0)^2)
431
432
433   # standard error of b1
434   b1_SE <- sqrt(fisher.info.inv.ll)
435
436   EFlm <- 1/fisher.info.inv.ll
437   # TFI
438
439
440 ### LR-Test
441   x0 <- optim(par = c(strt.vls["a0.strt"], strt.vls["a1.strt"], strt.vls["phi.strt"]), fn = LL, d = d_bcm,
442     control = list(fnscale = -1))
443
444   if(x0$convergence > 0){
445
446     step2.x0 <- 1
447
448     m0 <- MASS::glm.nb(mon.n.events ~ 1 + offset(log(mon.obs.time)), data = d_bcm_nr)
449     a0.hat.0 <- coef(m0)[1]
450
451     if(!is.null(m0$th.warn) && m0$th.warn == "iteration limit reached"){
452       wnyesno.m0 <- 1
453       wntheta.m0 <- m0$theta
454

```

```

455     # poisson model
456     m0 <- glm(mon.n.events ~ 1 + offset(log(mon.obs.time)),
457             family = poisson(link = "log"),
458             data = d_bcm_nr)
459
460     a0.hat.0 <- coef(m0)[1]
461   }
462
463   phi.hat.0 <- ifelse(1/m0$theta > 1e-6, 1/m0$theta, 1e-6)
464
465   x0 <- optim(par = c(a0.hat.0, strt.vls["a1.strt"], phi.hat.0), fn = LL, d = d_bcm,
466             control = list(fnscale = -1))
467
468   }
469
470   ### save estimated parameters
471   Ea0[j] = x$par[1]
472   Ea1[j] = x$par[2]
473   Ephi[j] = x$par[3]
474   Eb1[j] = x$par[4]
475
476   warns.monloop[j] = wnyesno
477   warns.monloop.ct[j] = wncount
478
479   warns.ml[j] = wnyesno.ml
480   warns.m0[j] = wnyesno.m0
481
482   conv.monloop[j] = nyesno
483   conv.monloop.ct[j] = ncount
484
485   conv[j] = x$convergence
486   conv0[j] = x0$convergence
487
488   strtvls.step2.monloop[j] = step2
489   strtvls.step2.monloop.ct[j] = step2.count
490   strtvls.step2.x1[j] = step2.x1
491   strtvls.step2.x0[j] = step2.x0
492
493   no.trend.monloop[j] = a1.was.zero
494   no.trend.ml[j] = a1.was.zero.ml
495
496   phi.too.small.monloop[j] = phi.blw.1emin6
497   phi.too.small.ml[j] = phi.blw.1emin6.ml
498
499   warns.min.theta.monloop[j] = wnthetamin
500   warns.theta.ml[j] = wntheta.ml
501   warns.theta.m0[j] = wntheta.m0
502
503
504   Eb1_Se[j] = b1_SE
505   LR[j] = 2*(x$value - x0$value)
506
507   N[j] = length(d_bcm_nr$id)
508
509   avgFU[j] = mean(d_bcm_nr$mon.obs.time)
510   SL[j] = mon.time
511   obsFImon[j] = oFImon
512   simnr[j] = j
513
514   }
515
516   Res <- data.frame(
517
518     Tb1,
519     Eb1,
520     Eb1_Se,
521     LR,
522
523     Ta0,
524     Ea0,
525     Ta1,
526     Ea1,
527     Tphi,
528     Ephi,
529
530     conv.monloop,

```

```

531     conv.monloop.ct ,
532     conv ,
533     conv0 ,
534
535         warns.monloop ,
536     warns.monloop.ct ,
537     warns.ml ,
538     warns.m0 ,
539
540         strtvls.step2.monloop.ct ,
541     strtvls.step2.x1 ,
542     strtvls.step2.x0 ,
543
544     no.trend.monloop ,
545     no.trend.ml ,
546
547     phi.too.small.monloop ,
548     phi.too.small.ml ,
549
550     warns.min.theta.monloop ,
551     warns.theta.ml ,
552     warns.theta.m0 ,
553
554     N ,
555     avgFU ,
556     SL ,
557     obsFImon ,
558     SimNr = simnr)
559
560 #### make results table
561 ## effective number of simulations
562 n.sim.res <- max(Res$SimNr, na.rm = TRUE)
563
564 ## Wald Test
565 #Betas and standard errors
566 betas <- Ebl
567 betas_se <- Ebl_Se
568 betas_true <- Tbl
569
570 #z-values
571 z <- betas/betas_se
572 #How many times H0 was rejected?
573 pstarType2wald <- mean(z < qnorm(0.025))
574 #Monte Carlo error
575 mcerrType2wald <- sqrt((pstarType2wald*(1-pstarType2wald))/n.sim)
576
577 ## Likelihood-Ratio Test
578 lr_sig <- (LR > qchisq(0.95, 1) & betas < 0)
579 #How many times H0 was rejected?
580 pstarType2lira <- mean(lr_sig)
581 #Monte Carlo error
582 mcerrType2lira <- sqrt((pstarType2lira*(1-pstarType2lira))/n.sim)
583
584 ## coverage probability for confidence intervals
585 CIlower <- Ebl - qnorm(1 - 0.025)*Ebl_Se
586 CIupper <- Ebl + qnorm(1 - 0.025)*Ebl_Se
587 CIconvertsTbl <- (Tbl >= CIlower & Tbl <= CIupper)
588 CI_CoverProb <- mean(CIconvertsTbl)*100
589
590 ## Bias
591 Biasa0 <- mean(Ea0) - Ta0
592 Biasa1 <- mean(Ea1) - Ta1
593 Biasphi <- mean(Ephi) - Tphi
594 Biasb1 <- mean(Eb1) - Tbl
595
596 ## follow-up times
597 FU <- c(m = mean(avgFU*12),
598       sd = sd(avgFU*12),
599       quantiles5pcrt = quantile(avgFU*12, probs = c(0,0.05,0.25,0.5,0.75,0.95,1)),
600       prctmax = mean(avgFU == flwYRS)*100)
601
602 ## study length
603 ST <- c(m = mean(SL*12),
604       sd = sd(SL*12),
605       quantiles5pcrt = quantile(SL*12, probs = c(0,0.05,0.25,0.5,0.75,0.95,1)),
606       prctmax = mean(SL == (flwYRS+recrYRS))*100)

```

```

607
608 ## standard error of beta1
609 Ose_b1 <- sd(betas) # observed
610 Ese_b1 <- mean(betas_se) # estimated from Fisher Information
611
612 # number of patients
613 N <- c(m = mean(N),
614       sd = sd(N),
615       quantiles5pcrt = quantile(N, probs = c(0,0.05,0.25,0.5,0.75,0.95,1)),
616       prctmax = mean(N == (n0*2))*100)
617
618 ## summary of results
619 opc <- c(Ta0 = Ta0,
620         AEa0 = mean(Ea0),
621         Biasa0 = Biasa0,
622
623         Ta1 = Ta1,
624         AEa1 = mean(Ea1),
625         Biasa1 = Biasa1,
626
627         Tphi = Tphi,
628         AEphi = mean(Ephi),
629         Biasphi = Biasphi,
630
631         Tb1 = Tb1,
632         AEb1 = mean(Eb1),
633         Biasb1 = Biasb1,
634         Obs_SEb1 = Ose_b1,
635         Est_SEb1 = Ese_b1,
636
637         CI_CV_PR = CI_CoverProb,
638         h0rej.wald = pstarType2wald,
639         mcerr.wald = mcerrType2wald,
640         h0rej.lira = pstarType2lira,
641         mcerr.lira = mcerrType2lira,
642
643         studylength = ST,
644         followuptime = FU,
645         N = N,
646
647         sum.nonconv.mon = sum(conv.monloop),
648         sum.nonconv.mon.ct = sum(conv.monloop.ct),
649         sum.nonconv.1 = sum(conv == 1),
650         sum.nonconv0.1 = sum(conv0 == 1),
651         sum.nonconv.10 = sum(conv == 10),
652         sum.nonconv0.10 = sum(conv0 == 10),
653         #
654         sum.warns.mon = sum(warns.monloop),
655         sum.warns.mon.ct = sum(warns.monloop.ct),
656         sum.warns.m1 = sum(warns.m1),
657         sum.warns.m0 = sum(warns.m0),
658         #
659         sum.step2.mon.ct = sum(strtvls.step2.monloop.ct),
660         sum.step2.x1 = sum(strtvls.step2.x1),
661         sum.step2.x0 = sum(strtvls.step2.x0),
662         #
663         sum.no.trend.mon = sum(no.trend.monloop),
664         sum.no.trend.m1 = sum(no.trend.m1),
665         #
666         sum.phi.too.small.mon = sum(phi.too.small.monloop),
667         sum.phi.too.small.m1 = sum(phi.too.small.m1),
668         #
669         sum.na.betas = sum(is.na(betas)),
670         sum.na.betas_se = sum(is.na(betas_se)),
671
672         n.sim.res = n.sim.res,
673         avgobsFImon = mean(obsFImon),
674         obsFImon.rg = range(obsFImon),
675         firstlook = firstlook)
676
677 #return(opc)
678 #return(Res)
679 A <- list(Res, opc)
680 }

```


BCM-CONST

```

1 bcmconst = function(n0, Tbl, Ta0, Ta1, Tphi, recrYRS, flwYRS, interimYRS,
2                       n.sim, FICrit, firstlook, tstep){
3
4   ### save place for data
5
6   # estimates
7   Ebl = rep(NA, n.sim) # estimate effect
8   Ebl_Se = rep(NA, n.sim) # estimate standard error of effect
9   E10 = rep(NA, n.sim) # estimate baseline rate (i.e. Intercept of glm.nb)
10  Ephi = rep(NA, n.sim)
11
12  # likelihood ratio test
13  LR = rep(NA, n.sim) # Chi-Square
14  LR.pval = rep(NA, n.sim) # p-value
15
16  # p-values from glm.nb
17  glm.nb.pval = rep(NA, n.sim)
18
19  # number of patients
20  N = rep(NA, n.sim)
21
22  # convergence
23  non.conv.monloop = rep(NA, n.sim) # convergence in monitoring loop
24  non.conv.monloop.ct = rep(NA, n.sim) # convergence in monitoring loop
25  non.conv.m1 = rep(NA, n.sim) # convergence full model
26  non.conv.m0 = rep(NA, n.sim) # convergence null model
27
28  # warnings
29  warns.monloop = rep(NA, n.sim)
30  warns.monloop.ct = rep(NA, n.sim)
31  warns.m0 = rep(NA, n.sim)
32  warns.m1 = rep(NA, n.sim)
33
34  warns.min.theta.monloop = rep(NA, n.sim)
35  warns.theta.m1 = rep(NA, n.sim)
36  warns.theta.m0 = rep(NA, n.sim)
37
38  # average follow-up time and study length
39  avgFU = rep(NA, n.sim)
40  SL = rep(NA, n.sim)
41
42  # observed Fisher Information (blinded)
43  obsFImon = rep(0, n.sim)
44
45  simnr = rep(NA, n.sim)
46
47  ### maximal observation time
48  max.obs.time = flwYRS
49
50  ### n.sim loop
51  for(j in 1:n.sim){
52
53    my.objects <- c("b.hat",
54                  "d_bcm",
55                  "d_bcm_nr",
56                  "index",
57                  "l0.hat",
58                  "lr.chi.squ",
59                  "m0",
60                  "m1",
61                  "mim0",
62                  "mm",
63                  "mlambda.hat",
64                  "mlambda0.hat",
65                  "mlambda1.hat",
66                  "mphi.hat",
67                  "mI1",
68                  "mI2",
69                  "mon.time",
70                  "mFisherInfo",
71                  "phi.hat",
72                  "pv.chi.squ",
73                  "rnms",
74                  "simdat")

```

```

75
76 rm(list = ls()[ls() %in% my.objects])
77
78
79 # counter for non-convergence in monitoring loop
80 ncount <- 0
81 nyesno <- 0
82
83 # counter for warnings in monitoring loop
84 wncount <- 0
85 wnyesno <- 0
86
87 # warnings in m1
88 wnml <- 0
89
90 # save theta when iteration limit is reached
91 wnthetamin <- Inf
92 wntheta.ml <- NA
93 wntheta.m0 <- NA
94
95 # Initialize variables
96 # start point for monitoring
97 k <- 0
98 # start value for Fisher Information
99 mFisherInfo <- 0
100
101 ### simulate data until "endpoint at planning stage"
102 simdat <- simPP(n0 = n0,
103               theta = c(Ta0, Tal),
104               effect = exp(Tb1),
105               phi = Tphi,
106               recrYRS = recrYRS,
107               fllwYRS = fllwYRS,
108               interimYRS = interimYRS)
109
110 if(sum(is.na(simdat))>0) stop("there is a missing value in the data set")
111
112 # absolute event times (= on investigator's time scale)
113 simdat$abs.start <- simdat$start + simdat$rec.time
114 simdat$abs.stop <- simdat$stop + simdat$rec.time
115
116 #####
117 ### repeat monitoring until Fisher Information reaches critical value or study ends ###
118
119 while(mFisherInfo < FIcrit){
120
121   # monitoring time
122   mon.time <- firstlook + k*step
123
124   # subset data for blinded continuous monitoring
125   # (= throw out events with event time > monitoring time)
126
127   # all events which start before mon.time and censored events
128   d_bcm <- simdat[simdat$abs.start <= mon.time, ]
129
130   # categorize events (event = 1, censored = 0)
131   d_bcm$mon.event <- ifelse(d_bcm$abs.stop <= mon.time & d_bcm$stop < max.obs.time, 1, 0)
132
133   # observation times at current monitoring time point
134   d_bcm$mon.obs.time <- mon.time - d_bcm$rec.time
135
136   # exclude patients with zero observation time (i.e. rec.time == mon.time)
137   d_bcm <- d_bcm[d_bcm$mon.obs.time != 0, ]
138
139   # upper limit for observation times (i.e. maximal follow-up time)
140   d_bcm$mon.obs.time[d_bcm$mon.obs.time > max.obs.time] <- max.obs.time
141
142   # reduce data to n rows (= one row per patient)
143   rnms <- as.numeric(row.names(d_bcm))
144   index <- rnms %% 1 == 0
145   d_bcm_nr <- d_bcm[index, c("id", "group", "n.events", "mon.obs.time")]
146
147   # number of events at current monitoring time point
148   d_bcm_nr$mon.n.events <- sapply(split(d_bcm_nr, d_bcm_nr$id), sum)
149
150   # go one step ahead if no events occurred until monitoring time point

```

```

151   if(sum(d_bcm_nr$mon.n.events) == 0) {k <- k + 1;
152   if(mon.time - (recrYRS + flwYRS) >= 0) break;
153   next}
154
155
156  ** bcm-const
157   # blinded estimates of rates
158   mm <- glm.nb(mon.n.events ~ 1 + offset(log(mon.obs.time)), data = d_bcm_nr)
159   mphi.hat <- 1/mm$theta
160   mlambda.hat <- exp(coef(mm)[1])
161
162   # if model does not converge
163   if(mm$converged == FALSE){
164     nyesno <- 1
165     ncount <- ncount + 1
166   }
167
168   if(!is.null(mm$th.warn) && mm$th.warn == "iteration limit reached"){
169     # count cases with warnigns
170     wnyesno <- 1
171     wncount <- wncount + 1
172     wnthetamin <- min(wnthetamin, mm$theta)
173
174     # poisson model
175     mm <- glm(mon.n.events ~ 1 + offset(log(mon.obs.time)),
176             family = poisson(link = "log"),
177             data = d_bcm_nr)
178
179     mlambda.hat <- exp(coef(mm)[1])
180     mphi.hat <- 0
181   }
182
183   # blinded estimation of annual relapse rates for both groups
184   mlambda0.hat <- 2*mlambda.hat/(1 + expct.RR)
185   mlambda1.hat <- mlambda0.hat*expct.RR
186
187   # Fisher Info
188   mI1 <- 0.5*sum(d_bcm_nr$mon.obs.time*mlambda1.hat/(1 + mphi.hat*d_bcm_nr$mon.obs.time*mlambda1.hat))
189   mI2 <- 0.5*sum(d_bcm_nr$mon.obs.time*mlambda0.hat/(1 + mphi.hat*d_bcm_nr$mon.obs.time*mlambda0.hat))
190
191   mFisherInfo <- 1/(1/mI1 + 1/mI2)
192  **
193
194  #### go one month ahead
195   k <- k + 1
196
197   # stop if all patients reach max. follow-up time
198   # (absolute monitoring time - (recruitment period + follow-up))
199   if(mon.time - (recrYRS + flwYRS) >= 0) break
200 }
201 #####
202
203   oFIm <- mFisherInfo
204
205  #### Wald-Test
206
207   # GLM
208   d_bcm_nr$group = factor(d_bcm_nr$group)
209   m1 <- glm.nb(mon.n.events ~ group + offset(log(mon.obs.time)), data = d_bcm_nr)
210   b.hat <- coef(m1)[2]
211   l0.hat <- coef(m1)[1]
212   phi.hat <- 1/m1$theta
213
214
215   if(!is.null(m1$th.warn) && m1$th.warn == "iteration limit reached"){
216     wnm1 <- 1
217     wntheta.m1 <- m1$theta
218
219     m1 <- glm(mon.n.events ~ group + offset(log(mon.obs.time)),
220             family = poisson(link = "log"),
221             data = d_bcm_nr)
222
223     b.hat <- coef(m1)[2]
224     l0.hat <- coef(m1)[1]
225     phi.hat <- 0
226   }

```

```

227
228   if(!is.null(m0$th.warn) && m0$th.warn == "iteration limit reached"){wntheta.m0 <- m0$theta}
229
230 #### save estimated parameters
231   Eb1[j] = b.hat
232   Eb1_Se[j] = summary(m1)$coefficients[2,2]
233   Ephi[j] = phi.hat
234   E10[j] = 10.hat
235
236   warns.monloop[j] = wnyesno
237   warns.monloop.ct[j] = wncount
238   warns.m0[j] = !is.null(m0$th.warn) && m0$th.warn == "iteration limit reached"
239   warns.m1[j] = wnml
240
241   non.conv.monloop[j] = nyesno
242   non.conv.monloop.ct[j] = ncount
243   non.conv.m1[j] = as.numeric(m1$converged == FALSE)
244   non.conv.m0[j] = as.numeric(m0$converged == FALSE)
245
246   warns.min.theta.monloop[j] = wntheta.min
247   warns.theta.m1[j] = wntheta.m1
248   warns.theta.m0[j] = wntheta.m0
249
250
251   glm.nb.pval[j] = summary(m1)$coefficients[2,4]
252   LR[j] = lr.chi.squ
253   LR.pval[j] = pv.chi.squ
254
255   N[j] = length(d_bcm_nr$id)
256
257   avgFU[j] = mean(d_bcm_nr$mon.obs.time)
258   SL[j] = mon.time
259   obsFImon[j] = oFlm
260   simnr[j] = j
261
262 }
263
264 Res <- data.frame(
265   Tb1,
266   Eb1,
267   Eb1_Se,
268   LR,
269   LR.pval,
270   glm.nb.pval,
271
272   Ta0,
273   E10,
274   Ta1,
275   Tphi,
276   Ephi,
277
278   warns.monloop,
279   warns.monloop.ct,
280
281   warns.m1,
282   warns.m0,
283
284   non.conv.monloop,
285   non.conv.monloop.ct,
286   non.conv.m1,
287   non.conv.m0,
288
289   warns.min.theta.monloop,
290   warns.theta.m1,
291   warns.theta.m0,
292
293   N,
294   avgFU,
295   SL,
296   obsFImon,
297   SimNr = simnr)
298
299
300 #### make results table
301 ## number of simulations
302 n.sim.res <- nrow(Res)

```

```

303
304 ## Wald Test
305 #Betas and standard errors
306 betas <- Eb1
307 betas_se <- Eb1_Se
308 betas_true <- Tbl
309
310 #z-values
311 z <- betas/betas_se
312 #How many times H0 was rejected?
313 pstarType2wald <- mean(z < qnorm(0.025))
314 #Monte Carlo error
315 mcerrType2wald <- sqrt((pstarType2wald*(1-pstarType2wald))/n.sim)
316
317 ## Likelihood-Ratio Test
318 lr_sig <- (LR > qchisq(0.95, 1) & betas < 0)
319 #How many times H0 was rejected?
320 pstarType2lira <- mean(lr_sig)
321 #Monte Carlo error
322 mcerrType2lira <- sqrt((pstarType2lira*(1-pstarType2lira))/n.sim)
323
324 ## coverage probability for confidence intervals
325 CIlower <- Eb1 - qnorm(1 - 0.025)*Eb1_Se
326 CIupper <- Eb1 + qnorm(1 - 0.025)*Eb1_Se
327 CICoversTbl <- (Tbl >= CIlower & Tbl <= CIupper)
328 CI_CoverProb <- mean(CICoversTbl)*100
329
330 ## Bias
331 Biasphi <- mean(Ephi) - Tphi
332 Biasb1 <- mean(Eb1) - Tbl
333
334 ## follow-up times
335 FU <- c(m = mean(avgFU*12),
336         sd = sd(avgFU*12),
337         quantiles5pcrt = quantile(avgFU*12, probs = c(0,0.05,0.25,0.5,0.75,0.95,1)),
338         prctmax = mean(avgFU == flwYRS)*100)
339
340 ## study length
341 ST <- c(m = mean(SL*12),
342         sd = sd(SL*12),
343         quantiles5pcrt = quantile(SL*12, probs = c(0,0.05,0.25,0.5,0.75,0.95,1)),
344         prctmax = mean(SL == (flwYRS+recrYRS))*100)
345
346 ## standard error of betal
347 Ose_b1 <- sd(betas) # observed
348 Ese_b1 <- mean(betas_se) # estimated from glm.nb
349
350 # number of patients
351 N <- c(m = mean(N),
352        sd = sd(N),
353        quantiles5pcrt = quantile(N, probs = c(0,0.05,0.25,0.5,0.75,0.95,1)),
354        prctmax = mean(N == (n0*2))*100)
355
356 ## summary of results
357 opc <- c(Ta0 = Ta0,
358         AEa0 = NA,
359         Biasa0 = NA,
360
361         Ta1 = Ta1,
362         AEa1 = NA,
363         Biasa1 = NA,
364
365         Tphi = Tphi,
366         AEphi = mean(Ephi),
367         Biasphi = Biasphi,
368
369         Tbl = Tbl,
370         AEb1 = mean(Eb1),
371         Biasb1 = Biasb1,
372         Obs_SEb1 = Ose_b1,
373         Est_SEb1 = Ese_b1,
374
375         CI_CV_PR = CI_CoverProb,
376         h0rej.wald = pstarType2wald,
377         mcerr.wald = mcerrType2wald,
378         h0rej.lira = pstarType2lira,

```

```
379         mcerr.lira = mcerrType2lira ,
380
381         studylength = ST,
382         followuptime = FU,
383         N = N,
384
385         # sum.nonconv.mon = sum(non.conv.monloop),
386         # sum.nonconv.mon.ct = sum(non.conv.monloop.ct),
387         # sum.warns.mon = sum(warns.monloop),
388         # sum.warns.mon.ct = sum(warns.monloop.ct),
389         #
390         # sum.nonconv.m1 = sum(non.conv.m1 == 1),
391         # sum.nonconv0.m0 = sum(non.conv.m0 == 1),
392         # sum.warns.m0 = sum(warns.m0),
393         # sum.warns.m1 = sum(warns.m1),
394
395         sum.na.betas = sum(is.na(betas)),
396         sum.na.betas_se = sum(is.na(betas_se)),
397
398         n.sim.res = n.sim.res ,
399         avgobsFImon = mean(obsFImon),
400         obsFImon.rg = range(obsFImon),
401         firstlook = firstlook)
402
403     #return(opc)
404     #return(Res)
405     A <- list(Res, opc)
406 }
```

Eidesstattliche Erklärung

Eidesstattliche Erklärung zur Masterarbeit

Ich versichere, dass ich die Arbeit selbständig und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Alle Stellen, die wörtlich oder sinngemäß aus Veröffentlichungen oder anderen Quellen entnommen sind, sind als solche kenntlich gemacht. Die schriftliche und elektronische Form der Arbeit stimmen überein. Ich stimme der Überprüfung der Arbeit durch eine Plagiatssoftware zu.

Unterschrift :

Ort, Datum :