

# **Flexible Designs in klinischen Prüfungen mit binärer und ordinaler Zielvariable**

**Dissertation**

zur Erlangung des Grades  
eines Doktors der Naturwissenschaften  
der Technischen Universität Dortmund

Der Fakultät Statistik  
der Technischen Universität Dortmund

vorgelegt von

**Wibke Stansen**

Ingelheim am Rhein 2010

1. Gutachter: Prof. Dr. Joachim Hartung

2. Gutachter: Prof. Dr. Roland Fried

Tag der mündlichen Prüfung: 28. Mai 2010

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>5</b>
<b>2</b>	<b>Flexible Designs</b>	<b>10</b>
2.1	Kombinationsmethoden . . . . .	10
2.2	Adaptiv gruppensequentielle klinische Studien . . . . .	14
2.2.1	Verallgemeinertes Pocock-Design . . . . .	15
2.2.2	Verallgemeinertes O'Brien/Fleming-Design . . . . .	17
2.2.3	Andere Designs . . . . .	19
2.2.4	Vergleich der adaptiv gruppensequentiellen Designs . . . . .	23
2.3	Self-Designing klinischer Studien . . . . .	28
2.4	Adaptive Wahl der Fallzahl und der Gewichtung . . . . .	30
2.4.1	Wahl der Fallzahl . . . . .	30
2.4.2	Wahl der Gewichtung im Self-Designing . . . . .	34
2.5	Schätzen in flexiblen Designs . . . . .	35
2.5.1	Intervallschätzung . . . . .	36
2.5.2	Punktschätzung . . . . .	44
<b>3</b>	<b>Flexible Designs in klinischen Studien mit qualitativer Zielvariable</b>	<b>48</b>
3.1	Binäre Zielvariable . . . . .	48
3.1.1	Die Risikodifferenz . . . . .	49
3.1.2	Fallzahlplanung . . . . .	52
3.1.3	Beispiel . . . . .	56
3.1.4	Simulationsstudien . . . . .	61
3.2	Ordinale Zielvariable . . . . .	77
3.2.1	Die verallgemeinerte Risikodifferenz . . . . .	78
3.2.2	Fallzahlplanung . . . . .	82
3.2.3	Beispiel . . . . .	85
3.2.4	Simulationsstudien . . . . .	87
<b>4</b>	<b>Zusammenfassung</b>	<b>95</b>
	<b>Literaturverzeichnis</b>	<b>99</b>

<b>Anhang</b>	<b>103</b>
<b>A Kritische Werte für adaptiv gruppensequentielle Designs</b>	<b>103</b>
<b>B Verteilungen der finalen Teststatistiken im Self-Designing</b>	<b>111</b>
B.1 Verteilung der finalen Teststatistik $Z_k$ . . . . .	112
B.2 Verteilung der finalen Teststatistik $S_k$ . . . . .	114
<b>C Ergebnisse der Simulationsstudien</b>	<b>117</b>
C.1 Binäre Zielvariable . . . . .	117
C.2 Ordinale Zielvariable . . . . .	149

# 1 Einleitung

Die Zulassung von neuen, erfolgversprechenden medizinischen Behandlungen erfordert den Nachweis der Wirksamkeit und Verträglichkeit. Hierzu werden klinische Studien durchgeführt, in welchen die neu entwickelten Therapien systematisch erprobt und überprüft werden. Die Studiengröße wird häufig basierend auf Vorkenntnissen bereits in der Planungsphase festgelegt. Erst nach Beobachtung aller vorgesehenen Patienten wird die statistische Auswertung durchgeführt. Da die Gesamtlaufzeiten einiger Studien jedoch mehrere Jahre betragen, wäre es unangemessen, nicht in deren Verlauf zu überprüfen, ob bereits zu einem früheren Zeitpunkt signifikante Therapieunterschiede bestehen. Eine solche Vorgehensweise hat wichtige ethische und ökonomische Vorteile von hohem Stellenwert. In ethischer Hinsicht ist es erstrebenswert, die Überlegenheit einer neuen Behandlung so früh wie möglich zu erkennen, da dies dazu führen kann, die Studie eher zu beenden und allen Patienten die bessere Therapie schnellstmöglich zukommen zu lassen. Frühzeitig sollte auch festgestellt werden, wenn eine Behandlung gesundheitsschädigend ist, um keinen Patienten einem unvermeidbaren Risiko auszusetzen. Außerdem können durch Auswertungen während des Studienverlaufs Nebenwirkungen früher aufgedeckt und somit besser überwacht werden. Vom ökonomischen Standpunkt ist das frühere Beenden einer Studie interessant, da dies zu großen Kosteneinsparungen führen kann.

Problematisch ist bei der mehrfachen Überprüfung im Verlauf der Studie, ob schon zu einem früheren Zeitpunkt ein Therapieunterschied nachweisbar ist, dass zu jedem Analysezeitpunkt die Möglichkeit besteht, den Fehler I. oder II. Art zu begehen. Berücksichtigung finden die Auswirkungen von Zwischenauswertungen auf die Irrtumswahrscheinlichkeiten bei der Anwendung sequentieller Verfahren. Diese halten trotz mehrfacher Durchführung von Tests im Studienverlauf das Gesamtniveau ein. Die von Wald (1947) entwickelten streng sequentiellen Pläne sind aus organisatorischen Gründen in der Praxis nicht geeignet, da diese eine Analyse nach jeder neuen Beobachtung vorsehen. Eine sinnvolle Erweiterung dieses Vorgehens stellen die erstmals von Armitage, McPherson und Rowe (1969) vorgeschlagenen gruppensequentiellen Verfahren dar, welche eine Zwischenauswertung nach einer bestimmten Anzahl von Patienten vorschreiben. Nach jeder Sequenz wird eine Entscheidung darüber getroffen, ob die Studie beendet oder weitergeführt werden soll. Hierbei kann das vorzeitige Beenden einer Studie aufgrund der Überlegenheit einer Therapie erfolgen oder wegen der Vermutung, dass kein Behandlungsunterschied existiert.

Eine breite Akzeptanz für die medizinische Anwendung haben die für zweiseitige Testprobleme entwickelten gruppensequentiellen Pläne von Pocock (1977) sowie von O'Brien und Fleming (1979) gefunden. Aus diesem Grund sind diese als klassische Verfahren zu verstehen. Wichtige methodische Voraussetzung für die Anwendung dieser Designs sind die Festlegung des maximalen Stichprobenumfangs sowie der Anzahl der durchzuführenden Zwischenauswertungen vor Studienbeginn. Außerdem wird verlangt, dass die einzelnen Sequenzen gleich groß sind. Besonders die letzte Einschränkung ist aus organisatorischen Gründen nicht immer sinnvoll. Deshalb wurden Erweiterungen gruppensequentieller Pläne entwickelt. Ein bekanntes Beispiel ist der so genannte „use function“-Ansatz von Lan und DeMets (1983), bei welchem in der Planungsphase nur die maximale Fallzahl bestimmt werden muss. Zusätzlich ist die Angabe einer use function oder  $\alpha$  spending function notwendig, mittels welcher die Wahrscheinlichkeit für den Fehler I. Art auf die einzelnen Stufen aufgeteilt wird. Dieser Ansatz ist immer dann nützlich, wenn im Voraus nicht bekannt ist, wie viele Patienten in den einzelnen Sequenzen rekrutiert werden sollen.

Ein großer Nachteil bleibt jedoch bei allen gruppensequentiellen Verfahren bestehen. Der in den Zwischenauswertungen beobachtete Effekt darf nicht für die Bestimmung der Umfänge der folgenden Sequenzen verwendet werden und darf somit auch nicht in die Fallzahlkalkulation der gesamten Studie einfließen. Ungünstig ist dies deshalb, da in klinischen Studien häufig der für die Planung der Stichprobengröße zugrundegelegte nachzuweisende Behandlungseffekt fehlspezifiziert wird. Wird dieser überschätzt, können sich beispielsweise Fallzahlen ergeben, welche zu klein sind, um klinisch relevante Unterschiede zwischen den Therapien aufzudecken. Das Hinzufügen von weiteren Patienten zur Erreichung statistischer Signifikanz verändert jedoch die Wahrscheinlichkeit für das Auftreten des Fehlers I. Art und stellt somit die Gültigkeit der gesamten Studie in Frage. In ökonomischer Hinsicht ist es jedoch wünschenswert den Umfang der folgenden Stufe kleiner zu planen, falls sich bei einer Zwischenauswertung herausstellt, dass das Ergebnis schon nahe der Signifikanz ist.

Das Anpassen des Stichprobenumfangs basierend auf den vorliegenden Daten wird durch die Verwendung adaptiver Pläne, welche eine Erweiterung der gruppensequentiellen Verfahren darstellen, ermöglicht. Hervorzuheben sind zu diesem Thema die von Bauer (1989), Bauer und Köhne (1994), Bauer und Röhmel (1995), Bauer und Kieser (1999) sowie Proschan und Hunsberger (1995) beschriebenen Vorgehensweisen. Hier dürfen die in den

Zwischenauswertungen gewonnenen Informationen in die Stichprobenkalkulation der restlichen Studie einfließen. Ermöglicht wird die Adaption der Fallzahl durch eine Transformation der Teststatistik. Beispielsweise wird in Bauer (1989) eine Prüfgröße vorgestellt, welche durch Kombination der p-Werte der einzelnen Sequenzen mittels der von Fisher (1932) vorgeschlagenen Produktregel entsteht.

Eine weitere Form adaptiver Pläne ist das so genannte Self-Designing klinischer Studien, welches erstmals von Fisher (1998) sowie von Shen und Fisher (1999) für normalverteilte Zielgrößen vorgeschlagen wurde. Bei diesen müssen die Anzahl durchzuführender Sequenzen, die Umfänge der einzelnen Stufen sowie der maximale Stichprobenumfang vor Studienbeginn nicht festgelegt werden. Das Verfahren ermittelt die Anzahl der zu rekrutierenden Patienten und die Gewichtung der Daten der folgenden Sequenzen schrittweise, basierend auf den vor der aktuellen Stufe gesammelten Informationen. Durch Kombination der Ergebnisse mit der inversen Normal- sowie der inversen  $\chi^2$ -Methode, wird das Self-Designing für beliebige Zielvariablen in Hartung (2001, 2006) sowie Hartung (2000) und Hartung und Knapp (2003) erweitert.

Neben der datengesteuerten Anpassung der Fallzahl können auch andere Adaptionen am Studiendesign vorgenommen werden. Beispielsweise in Fisher (1998) werden verschiedene Möglichkeiten erwähnt. So können in einer Studie zum Vergleich mehrerer Therapien solche Behandlungsarme entfernt werden, welche frühzeitig signifikante Effekte zeigen oder nicht wirksam sind. Auch möglich ist das nachträgliche Hinzufügen von Therapiegruppen. Stellt sich im Verlauf der klinischen Studie heraus, dass einige der teilnehmenden Studienzentren einen Behandlungseffekt nachweisen können und andere nicht, ist es sinnvoll, mehr Patienten in den erfolgreicherer Einheiten zu rekrutieren. Auch die Prüfgröße sowie das Testproblem können im Studienverlauf geändert werden. Finden zu viele Änderungen am Studiendesign statt, kann dies zu einer Ineffizienz des Verfahrens im klinischen Sinne führen (vgl. Tsiatis und Mehta (2003)).

Die in der Literatur vorgestellten flexiblen Methoden basieren in der Regel auf der Annahme normalverteilter Beobachtungen. Jedoch sind in der klinischen Forschung häufig Kategorien als Ausprägungen einer Zielgröße von Interesse. Die Zielsetzung dieser Arbeit ist die intensive Untersuchung von flexiblen Designs in klinischen Prüfungen mit binärer und ordinaler Zielvariable, wobei sowohl adaptiv gruppensequentielle Verfahren als auch

das Self-Designing klinischer Studien betrachtet und miteinander verglichen werden.

In Kapitel 2 werden die gewichtete inverse Normalmethode sowie die verallgemeinerte inverse  $\chi^2$ -Methode vorgestellt, welche in dieser Arbeit zur Kombination der p-Werte der einzelnen Sequenzen verwendet werden. Finden die so entstehenden Prüfgrößen in gruppensequentiellen Designs mit vorab festgelegter Gewichtung Verwendung, besteht die Möglichkeit Änderungen am Studiendesign im Verlauf der Untersuchung vorzunehmen. Der adaptive Charakter der hier vorgestellten Methoden geht aus der datengesteuerten Wahl der Fallzahl hervor. Zudem kann in jeder Sequenz überprüft werden, ob die Nullhypothese abgelehnt und die Untersuchung somit vorzeitig beendet werden kann. Die durch das mehrfache Testen im Studienverlauf erforderliche Anpassung der kritischen Werte wird für die von Pocock (1977), O'Brien und Fleming (1979) sowie Wang und Tsia-tis (1987) vorgeschlagenen Vorgehensweisen vorgestellt. Neben diesen Strategien wird das sehr einfache Prinzip des sequentiellen Testens von Haybittle (1971) und Peto et al. (1976) beschrieben. Die Designs vom Wang/Tsatis- sowie vom Haybittle/Peto-Typ werden in dieser Arbeit für die inverse  $\chi^2$ -Methode erweitert. Im Self-Designing klinischer Studien erfolgt neben der adaptiven Ermittlung der Fallzahl eine auf den vorliegenden Beobachtungen basierende Wahl der Gewichte. Die Gesamtanzahl der Sequenzen ist hierbei nicht fest vorgegeben. Neben der Überprüfung von Testproblemen ist die Abschätzung des Ausmaßes eines Unterschieds von Interesse, wofür Methoden zur Intervall- und Punktschätzung vorgestellt werden. Da Konfidenzintervalle für nicht-sequentielle Designs die gewünschte Überdeckungswahrscheinlichkeit nicht einhalten, sollten so genannte Repeated Konfidenzintervalle (vgl. Jennison und Turnbull (1989), Hartung und Knapp (2006)) verwendet werden. Im Fall keines vorzeitigen Studienabbruchs hält die Sequenz dieser Intervalle das festgelegte Konfidenzniveau ein, andernfalls ist die resultierende Sequenz der Intervalle konservativ. Für die Punktschätzung des unbekanntem Parameters in den einzelnen Sequenzen basierend auf allen vorliegenden Beobachtungen werden in der Meta-Analyse verwendete Methoden vorgestellt. Zudem wird ein Punktschätzer hergeleitet, der unverzerrt ist für den unbekanntem Parameter, falls in den Zwischenauswertungen lediglich eine Anpassung der Fallzahl und kein Testen der Nullhypothese erfolgt.

Die vorgestellten flexiblen Designs werden in Kapitel 3 bei Vorliegen von binären und ordinalen Zielvariablen untersucht. Zunächst wird der binäre Fall betrachtet, in welchem die Zielvariable zwei Ausprägungen besitzt. Um zu überprüfen, ob ein Unterschied zwischen

den beiden Behandlungsgruppen vorhanden ist, wird die Risikodifferenz verwendet. Basierend auf dieser werden Prüfgrößen mit und ohne Stetigkeitskorrektur vorgestellt (vgl. Schouten et al. (1980)). Können die Realisationen in mehr als zwei Kategorien auftreten, handelt es sich um eine ordinale Zielgröße und die Darstellung des Unterschieds erfolgt mittels der verallgemeinerten Risikodifferenz (vgl. Edwardes und Baltzan (2000)). Basierend auf dieser wird die Ermittlung der notwendigen Fallzahl vorgestellt. Sowohl für den binären als auch für den ordinalen Fall wird die Ermittlung des notwendigen Umfangs auf den sequentiellen Fall übertragen. In Simulationsstudien werden die adaptiv gruppen-sequentiellen Designs und das Self-Designing bei Vorliegen einer qualitativen Zielvariablen untersucht, wobei insbesondere von Interesse ist, ob Prüfgrößen, welche bereits in nicht-sequentiellen Designs das vorgegebene Testniveau nur asymptotisch einhalten, in flexiblen Designs Anwendung finden können.

In Kapitel 4 erfolgt abschließend eine zusammenfassende Darstellung und Diskussion der Ergebnisse der vorliegenden Arbeit.

## 2 Flexible Designs

In diesem Kapitel werden flexible Designs vorgestellt, welche die gewichtete inverse Normalmethode und die verallgemeinerte inverse  $\chi^2$ -Methode zur Kombination der Ergebnisse der einzelnen Sequenzen verwenden. Die adaptiv gruppensequentiellen Testverfahren in Kapitel 2.2 ermöglichen für eine fest vorgegebene Stufenanzahl eine Anpassung der Fallzahl im Verlauf der Untersuchung. Außerdem kann in den Zwischenauswertungen überprüft werden, ob die Studie mit einer vorzeitigen Ablehnung der Nullhypothese beendet werden kann. Eine Erweiterung dieser Testverfahren stellt das Self-Designing klinischer Studien dar, welche in Kapitel 2.3 vorgestellt werden. Die Gesamtanzahl der Sequenzen ist hier nicht durch eine fest vorgegebene Zahl beschränkt, sondern kann ebenso wie die Fallzahl im Verlauf der Studie basierend auf den vorliegenden Beobachtungen ermittelt werden. Bei Kombination der p-Werte mittels der verallgemeinerten inversen  $\chi^2$ -Methode kann die Überprüfung des Testproblems in jeder Stufe erfolgen, bei Verwendung der gewichteten inversen Normalmethode ist dies nur am Studienende unter Einhaltung des globalen Niveaus möglich (vgl. Hartung (2001)). Die Möglichkeit eines vorzeitigen Beendens der Untersuchung mit Beibehalten der Nullhypothese findet in dieser Arbeit keine Berücksichtigung.

### 2.1 Kombinationsmethoden

Um zwei Therapien hinsichtlich ihrer Effektivität miteinander zu vergleichen, wird eine klinische Studie mit Parallelgruppendesign geplant. In Gruppe 1 wird das Novum und in Gruppe 2 die Standardtherapie oder ein Placebo verabreicht. Die Zuweisung der Patienten zu diesen beiden unabhängigen Gruppen erfolgt mittels Randomisation. Der Unterschied zwischen den beiden Behandlungen wird mit eines in Abhängigkeit der vorliegenden Situation geeigneten Parameters  $\theta$  gemessen. Hierbei kann  $\theta$  beispielsweise die Differenz von zwei Erwartungswerten normalverteilter Zielgrößen oder von zwei Erfolgswahrscheinlichkeiten im Fall binärer Zielvariablen darstellen. Um zu überprüfen, ob die neue Therapie der Standardbehandlung überlegen ist, wird zunächst das einseitige Testproblem

$$H_0 : \theta = 0 \quad \text{versus} \quad H_1 : \theta > 0 \quad (2.1)$$

in  $K$  unabhängigen Sequenzen, bezeichnet mit Stufe(1), ..., Stufe( $K$ ), formuliert. Der Ausdruck Stufe(0) steht für a priori Informationen. Unter dieser Nullhypothese gelten die

im Folgenden beschriebenen Verteilungen exakt. Dies ist im Fall von Nullhypothesen der Form  $H_0 : \theta \leq 0$  oder  $H_0 : \theta \geq 0$  im Allgemeinen nicht erfüllt. Solche Hypothesen werden in dieser Arbeit nicht betrachtet, hierfür sei auf Hartung und Knapp (2006) verwiesen. In Kapitel 2.5.1 wird die Überprüfung zweiseitiger Testprobleme im Zusammenhang mit der Konstruktion von zweiseitigen Konfidenzintervallen für flexible Designs vorgestellt.

In Stufe  $i$ ,  $i = 1, \dots, K$ , erfolgt das Testen von  $H_0$  gegen  $H_1$  mittels der Teststatistik  $T_i$ , wobei große Werte von  $T_i$  für eine Ablehnung der Nullhypothese sprechen. Es wird die Annahme getroffen, dass  $T_i$  unter  $H_0$  eine stetige Verteilungsfunktion  $F_{i,0}$  besitzt. Der zugehörige p-Wert ist durch

$$p_i = 1 - F_{i,0}(T_i) \quad (2.2)$$

gegeben. Unter  $H_0$  gilt

$$p_i \sim U(0, 1) \quad \text{und} \quad 1 - p_i = F_{i,0}(T_i) \sim U(0, 1), \quad (2.3)$$

wobei  $U(0, 1)$  die Gleichverteilung auf dem Intervall  $(0, 1)$  bezeichnet.

Sei  $\lambda_G$  ein globales Gewicht, welches in Teilen  $\lambda_i$  über die  $K$  Sequenzen verteilt ist,

$$\lambda_{\Sigma}(K) = \sum_{i=1}^K \lambda_i = \lambda_G, \quad \lambda_i > 0, \quad i = 1, \dots, K. \quad (2.4)$$

Es wird hier die Annahme getroffen, dass die Gewichtungen den Stufen vor Studienbeginn zugewiesen werden. Bei den adaptiv gruppensequentiellen Designs wird in dieser Arbeit nur der Fall der Gleichgewichtung aller Sequenzen betrachtet, wobei  $\lambda_i = 1$  oder  $\lambda_i = 2$  für  $i = 1, \dots, K$  gewählt werden. Bei der inversen Normalmethode mit identischer Gewichtung der einzelnen Stufen spielt der Wert des Gewichts keine Rolle. Dies wird bei der Vorstellung der einzelnen Designs näher erläutert. Erfolgt eine unterschiedliche Gewichtung der einzelnen Sequenzen, kann beispielsweise durch Zuweisung eines großen Gewichts zu früheren Stufen die Wahrscheinlichkeit für einen vorzeitigen Studienabbruch erhöht werden (vgl. Wassmer (2001)).

Im Folgenden werden die gewichtete inverse Normalmethode sowie die verallgemeinerte inverse  $\chi^2$ -Methode zur Kombination der Ergebnisse der einzelnen Stufen vorgestellt (siehe Hartung (2006)).

Gewichtete inverse Normalmethode

Erfolgt die Transformation der  $p$ -Werte mittels der gewichteten inversen Normalmethode, resultiert

$$z_i(\lambda_i) = F_{N(0, \lambda_i)}^{-1}(1 - p_i) \stackrel{H_0}{\sim} N(0, \lambda_i), \quad i = 1, \dots, K. \quad (2.5)$$

Unter der Annahme, dass die einzelnen Sequenzen unabhängig voneinander sind, ist die Kombinationsstatistik bis zur Stufe  $j$  gegeben durch

$$Z_j = \sum_{i=1}^j z_i(\lambda_i) \stackrel{H_0}{\sim} N(0, \lambda_{\Sigma}(j)), \quad \lambda_{\Sigma}(j) = \sum_{i=1}^j \lambda_i, \quad j = 1, \dots, K. \quad (2.6)$$

Bezeichne  $\Phi^{-1}$  die Inverse der Verteilungsfunktion der Standardnormalverteilung  $\Phi$ , dann gilt  $F_{N(0, \lambda_i)}^{-1} = \sqrt{\lambda_i} \Phi^{-1}$ , und  $Z_j$  kann dargestellt werden als

$$Z_j = \sum_{i=1}^j \sqrt{\lambda_i} \Phi^{-1}(1 - p_i) \stackrel{H_0}{\sim} \sqrt{\lambda_{\Sigma}(j)} N(0, 1), \quad j = 1, \dots, K. \quad (2.7)$$

Die gemeinsame Verteilung von  $Z_1, \dots, Z_K$  ist die multivariate Normalverteilung mit Erwartungswertvektor  $\vec{\mu} = (\mu_1, \dots, \mu_K)$ , mit

$$\mu_j = E(Z_j) = \sum_{i=1}^j \sqrt{\lambda_i} E[\Phi^{-1}(1 - p_i)], \quad j = 1, \dots, K, \quad (2.8)$$

und Kovarianzmatrix  $\Sigma$ , d.h.  $(Z_1, \dots, Z_K) \sim N_K(\vec{\mu}, \Sigma)$ . Die einzelnen Kombinationsstatistiken sind voneinander abhängig, wobei die Kovarianz von  $Z_j$  und  $Z_k$ ,  $j, k \in \{1, \dots, K\}$  und  $j < k$ , gegeben ist durch

$$\text{Cov}(Z_j, Z_k) = \sum_{i=1}^j \lambda_i. \quad (2.9)$$

Die Kovarianz der unter der Nullhypothese auf die  $N(0, 1)$ -Verteilung standardisierten Statistiken  $Z_j/\sqrt{\lambda_{\Sigma}(j)}$  und  $Z_k/\sqrt{\lambda_{\Sigma}(k)}$ ,  $j, k \in \{1, \dots, K\}$  und  $j < k$ , ist

$$\begin{aligned} \text{Cov} \left( \frac{Z_j}{\sqrt{\sum_{i=1}^j \lambda_i}}, \frac{Z_k}{\sqrt{\sum_{i=1}^k \lambda_i}} \right) &= \frac{\sum_{i=1}^j \lambda_i \text{Var}(\Phi^{-1}(1 - p_i))}{\sqrt{\sum_{i=1}^j \lambda_i} \sqrt{\sum_{i=1}^k \lambda_i}} \\ &= \frac{\sum_{i=1}^j \lambda_i}{\sqrt{\sum_{i=1}^j \lambda_i} \sqrt{\sum_{i=1}^k \lambda_i}} = \frac{\sqrt{\sum_{i=1}^j \lambda_i}}{\sqrt{\sum_{i=1}^k \lambda_i}}. \end{aligned} \quad (2.10)$$

### Verallgemeinerte inverse $\chi^2$ -Methode

Werden die p-Werte mittels der verallgemeinerten inversen  $\chi^2$ -Methode transformiert, resultiert

$$q_i(\lambda_i) = F_{\chi^2(\lambda_i)}^{-1}(1 - p_i) \stackrel{H_0}{\sim} \chi^2(\lambda_i), \quad i = 1, \dots, K, \quad (2.11)$$

wobei die Beobachtungen in Form von Freiheitsgraden gewichtet werden. Unter der Annahme, dass die einzelnen Sequenzen unabhängig voneinander sind, ist die Kombinationsstatistik bis zur Stufe  $j$  gegeben durch

$$S_j = \sum_{i=1}^j q_i(\lambda_i) \stackrel{H_0}{\sim} \chi^2(\lambda_{\Sigma}(j)), \quad \lambda_{\Sigma}(j) = \sum_{i=1}^j \lambda_i, \quad j = 1, \dots, K. \quad (2.12)$$

### Spezialfälle

Werden in (2.7) die Gewichte  $\lambda_i = 1$ ,  $i = 1, \dots, j$ , gewählt, resultiert die Kombinationsstatistik der ungewichteten inversen Normalmethode, welche durch

$$Z_j = \sum_{i=1}^j z_i(1) = \sum_{i=1}^j \Phi^{-1}(1 - p_i) \sim \sqrt{j} N(0, 1) \quad (2.13)$$

gegeben ist. Diese von Stouffer et al. (1949) angewandte Methode wird im Weiteren als Stouffer-Methode bezeichnet.

Erfolgt bei Verwendung der inversen  $\chi^2$ -Methode die Gewichtung in jeder Sequenz mit  $\lambda_i = 2$ ,  $i = 1, \dots, K$ , resultiert die Kombinationsstatistik von Fisher (1932)

$$S_j = \sum_{i=1}^j q_i(2) = \sum_{i=1}^j -2 \log(p_i) \stackrel{H_0}{\sim} \chi^2(2j). \quad (2.14)$$

Diese findet in den zwei- und dreistufigen adaptiven Designs von Bauer und Köhne (1994) sowie Bauer und Röhmel (1995) Verwendung.

## Verteilung der p–Werte

Obwohl in den verschiedenen Studienabschnitten voneinander unabhängige Beobachtungen realisiert werden, ist die Unabhängigkeit der p–Werte bei datenabhängigen Designänderungen nicht unbedingt gewährleistet. Erfolgt die Überprüfung des Testproblems in beiden Sequenzen eines zweistufigen Designs beispielsweise mit Rangtests, kann die diskrete Verteilung von  $p_2$  unter der Nullhypothese von  $p_1$  abhängen. Dies ist bei einer auf dem Wert von  $p_1$  basierenden Wahl des Umfangs  $n_2$  der zweiten Sequenz der Fall. Um Änderungen am Studiendesign in den Zwischenauswertungen vornehmen zu können ohne dass die Wahrscheinlichkeit für einen Fehler I. Art erhöht wird, sollten die p–Werte „pclud“ sein (vgl. Brannath et al. (2002)). Die p–Werte  $p_1, \dots, p_j$  der Stufen  $1, \dots, j$  sind „pclud“, wenn die Verteilung von  $p_1$  und die bedingte Verteilung der p–Werte  $p_j$  der Sequenzen  $j \geq 2$  gegeben  $p_1, \dots, p_{j-1}$  unter der Nullhypothese stochastisch größer oder gleich der Gleichverteilung auf dem Intervall  $(0, 1)$  sind, d.h.

$$P_{H_0}(p_1 \leq \alpha) \leq \alpha \quad \text{und} \quad P_{H_0}(p_j \leq \alpha \mid p_1, \dots, p_{j-1}) \leq \alpha, \quad \text{für } j \geq 2. \quad (2.15)$$

Sind die Realisationen der einzelnen Sequenzen unabhängig voneinander und erfolgt die Berechnung der p–Werte mittels konservativen Tests, wird das Niveau eingehalten. Liegen unabhängige und  $U(0, 1)$ –verteilte p–Werte vor, wird das Niveau  $\alpha$  voll ausgeschöpft.

## 2.2 Adaptiv gruppensequentielle klinische Studien

Finden die beiden Kombinationsmethoden aus dem vorhergehenden Kapitel 2.1 in gruppensequentiellen Designs Verwendung, können unter der Annahme, dass die Verteilung der p–Werte der einzelnen Sequenzen „pclud“ ist, Änderungen am Studiendesign vorgenommen werden. In dieser Arbeit geht der adaptive Charakter der Methoden aus der möglichen datengesteuerten Wahl der Stichprobenumfänge hervor. In Stufe  $i$ ,  $i = 1, \dots, K$ , kann die Fallzahl  $n_i$  basierend auf den Ergebnissen der Sequenzen  $1, \dots, i - 1$  ermittelt werden, d.h.

$$n_i = \hat{n}(i - 1) = \hat{n}(\text{Stufe}(0), \dots, \text{Stufe}(i - 1)), \quad i = 1, \dots, K. \quad (2.16)$$

Die maximale Anzahl der Sequenzen  $K$  und die Gewichte  $\lambda_i > 0$ ,  $i = 1, \dots, K$ , mit  $\lambda_\Sigma(K) = \sum_{i=1}^K \lambda_i = \lambda_G$ , werden vor Studienbeginn festgelegt.

Da die irrtümliche Ablehnung der Nullhypothese nicht nur einmal, sondern in jeder Stufe erfolgen kann, wird das Niveau  $\alpha$  bei Verwendung der kritischen Werte für nichtsequentielle Designs nicht eingehalten. Dies ist der so genannte „Multiple Looks“–Effekt.

Die bekanntesten Verfahren zur somit notwendigen Adjustierung der Entscheidungsbe-  
reiche wurden von Pocock (1977) sowie O'Brien und Fleming (1979) vorgeschlagen, wel-  
che sehr unterschiedliche Strategien verfolgen. Im Pocock-Design werden durch Verwen-  
dung konstanter kritischer Schranken schon zu Beginn der Untersuchung große Teile  
des globalen Niveaus vergeben. Im O'Brien/Fleming-Design führen die monoton fal-  
lenden kritischen Werte dazu, dass in der letzten Stufe der Großteil des globalen Ni-  
veaus verbraucht wird. Bei einer allgemeineren Klasse von kritischen Schranken gemäß  
Wang und Tsiatis (1987) können durch Einführung eines zusätzlichen Parameters kri-  
tische Werte unterschiedlicher Form ermittelt werden, wobei die Designs vom Pocock-  
sowie O'Brien/Fleming-Typ als Spezialfälle in dieser enthalten sind. Eine sehr einfache  
Strategie wird von Haybittle (1971) und Peto et. al (1976) verfolgt. Bei Verwendung ei-  
ner konstanten kritischen Schranke in den einzelnen Sequenzen, kann in der letzten Stufe  
approximativ ein Test zum nicht-adjustierten kritischen Wert durchgeführt werden.

Findet die inverse Normalmethode Anwendung, können die kritischen Werte der klas-  
sischen gruppensequentiellen Designs verwendet werden (vgl. Lehmaner und Wassmer  
(1999), Wassmer (2001)). Die Designs vom Pocock- und O'Brien/Fleming-Typ bei Kom-  
bination der Ergebnisse mit der verallgemeinerten inversen  $\chi^2$ -Methode werden in Har-  
tung (2006) vorgestellt. Die Erweiterung des Wang/Tsiatis- sowie Haybittle/Peto-Designs  
für diese Kombinationsmethode erfolgt in Kapitel 2.2.3.

### 2.2.1 Verallgemeinertes Pocock-Design

Soll im Verlauf der Untersuchung mehrfach unter Einhaltung des globalen Niveaus  $\alpha$   
überprüft werden, ob  $H_0$  abgelehnt werden kann, müssen die kritischen Werte vor Stu-  
dienbeginn geeignet gewählt werden. Diese werden bei Verwendung der Statistiken  $Z_j$   
aus (2.7) bzw.  $S_j$  aus (2.12) im Weiteren mit  $cv_{Z,PO}(\alpha, K, \{\lambda_i\})$  bzw.  $cv_{S,PO}(\alpha, K, \{\lambda_i\})$   
bezeichnet. Seien

$$X_i \sim N(0, 1) \quad \text{und} \quad Y_i \sim \chi^2(\lambda_i), \quad i = 1, \dots, K, \quad (2.17)$$

voneinander unabhängige Zufallsvariablen, welche  $z_i(1) = \Phi^{-1}(1 - p_i)$  und  $q_i(\lambda_i)$  unter der  
Nullhypothese repräsentieren. Dann sind die kritischen Werte zum Niveau  $\alpha$ ,  $0 < \alpha < 1$ ,

durch die folgenden Bedingungen definiert (vgl. Hartung (2006))

$$P \left( \sum_{i=1}^j \sqrt{\lambda_i} X_i / \sqrt{\sum_{i=1}^j \lambda_i} \leq cv_{Z,PO}(\alpha, K, \{\lambda_i\}) \quad \forall j = 1, \dots, K \right) = 1 - \alpha, \quad (2.18)$$

$$P \left( F_{\chi^2(\lambda_G)}^{-1} \left( F_{\chi^2(\lambda_\Sigma(j))} \left( \sum_{i=1}^j Y_i \right) \right) \leq cv_{S,PO}(\alpha, K, \{\lambda_i\}) \quad \forall j = 1, \dots, K \right) = 1 - \alpha. \quad (2.19)$$

Die Idee des Pocock-Designs ist die Standardisierung der Teststatistik auf eine bestimmte Verteilung, so dass der gleiche kritische Wert in allen  $K$  Stufen verwendet werden kann. In (2.18) wird mit

$$F_{N(0, \lambda_G)}^{-1} \left( F_{N(0, \lambda_\Sigma(j))} \left( \sum_{i=1}^j \sqrt{\lambda_i} X_i \right) \right) / \sqrt{\lambda_G}$$

für  $j = 1, \dots, K$  eine  $N(0, 1)$ -verteilte Variable erzeugt, da

$$F_{N(0, \lambda_\Sigma(j))} \left( \sum_{i=1}^j \sqrt{\lambda_i} X_i \right) \sim U(0, 1)$$

und

$$F_{N(0, \lambda_G)}^{-1} \left( F_{N(0, \lambda_\Sigma(j))} \left( \sum_{i=1}^j \sqrt{\lambda_i} X_i \right) \right) \sim N(0, \lambda_G).$$

Die Statistik aus (2.19) stammt aus einer  $\chi^2(\lambda_G)$ -Verteilung für alle  $j = 1, \dots, K$ , da gilt

$$F_{\chi^2(\lambda_\Sigma(j))} \left( \sum_{i=1}^j Y_i \right) \sim U(0, 1)$$

und somit

$$F_{\chi^2(\lambda_G)}^{-1} \left( F_{\chi^2(\lambda_\Sigma(j))} \left( \sum_{i=1}^j Y_i \right) \right) \sim \chi^2(\lambda_G).$$

Die kritischen Werte können aus Tabellen, bspw. aus Pocock (1977), durch numerische Integration oder mittels Simulationen erhalten werden. In Anhang A finden sich in Tabelle A.1 mit der statistischen Software R (R Development Core Team (2007)), Version 2.6.2, ermittelte kritische Werte mit vier Nachkommastellen. Werden alle Sequenzen gleich gewichtet, d.h.  $\lambda_i = \lambda$  für  $i = 1, \dots, K$ , resultieren bei der inversen Normalmethode für verschiedene Werte für  $\lambda$  die gleichen kritischen Werte aufgrund der Standardisierung der

Kombinationsstatistik auf die  $N(0, 1)$ -Verteilung.

Die Testentscheidungen in den Sequenzen  $j, j = 1, \dots, K$ , sind im Folgenden für die gewichtete inverse Normalmethode und die verallgemeinerte inverse  $\chi^2$ -Methode beschrieben.

Gewichtete inverse Normalmethode:

Unter Verwendung der Statistik  $Z_j$  aus (2.7) lauten die Testentscheidungen:

- Ablehnen von  $H_0$  in Stufe  $j$ , falls

$$Z_j / \sqrt{\lambda_{\Sigma}(j)} > cv_{Z,PO}(\alpha, K, \{\lambda_i\}) \quad (2.20)$$

und  $Z_i / \sqrt{\lambda_{\Sigma}(i)} \leq cv_{Z,PO}(\alpha, K, \{\lambda_i\})$  für  $i < j, i, j \in \{1, \dots, K\}$ .

- Beibehalten von  $H_0$ , falls die letzte Sequenz  $K$  erreicht ist und

$$Z_K / \sqrt{\lambda_{\Sigma}(K)} \leq cv_{Z,PO}(\alpha, K, \{\lambda_i\}). \quad (2.21)$$

Verallgemeinerte inverse  $\chi^2$ -Methode:

Mit  $S_j$  aus (2.12) lauten die Testentscheidungen:

- Ablehnen von  $H_0$  in Stufe  $j$ , falls

$$F_{\chi^2(\lambda_G)}^{-1} \left( F_{\chi^2(\lambda_{\Sigma}(j))} (S_j) \right) > cv_{S,PO}(\alpha, K, \{\lambda_j\}) \quad (2.22)$$

und  $F_{\chi^2(\lambda_G)}^{-1} (F_{\chi^2(\lambda_{\Sigma}(i))} (S_i)) \leq cv_{S,PO}(\alpha, K, \{\lambda_i\})$  für  $i < j, i, j \in \{1, \dots, K\}$ .

- Beibehalten von  $H_0$ , falls die letzte Sequenz  $K$  erreicht ist und

$$F_{\chi^2(\lambda_G)}^{-1} \left( F_{\chi^2(\lambda_{\Sigma}(K))} (S_K) \right) \leq cv_{S,PO}(\alpha, K, \{\lambda_i\}). \quad (2.23)$$

### 2.2.2 Verallgemeinertes O'Brien/Fleming-Design

Im O'Brien und Fleming (1979)-Design sind die kritischen Werte bei Betrachtung der auf die  $N(0, 1)$ - bzw.  $\chi^2(\lambda_G)$ -Verteilung standardisierten Kombinationsstatistiken monoton

fallend. Erfolgt hingegen keine Standardisierung der Prüfgrößen auf eine bestimmte Verteilung, sind die mit  $cv_{Z, OBF}(\alpha, K, \{\lambda_i\})$  und  $cv_{S, OBF}(\alpha, K, \{\lambda_i\})$  bezeichneten Grenzen in allen Sequenzen identisch. Diese lassen sich zum Niveau  $\alpha$ ,  $0 < \alpha < 1$ , mit  $X_i$  und  $Y_i$  aus (2.17) aus den folgenden beiden Gleichungen bestimmen (vgl. Hartung (2006))

$$P\left(\sum_{i=1}^j \sqrt{\lambda_i} X_i \leq cv_{Z, OBF}(\alpha, K, \{\lambda_i\}) \quad \forall j = 1, \dots, K\right) = 1 - \alpha, \quad (2.24)$$

$$P\left(\sum_{i=1}^j Y_i \leq cv_{S, OBF}(\alpha, K, \{\lambda_i\}) \quad \forall j = 1, \dots, K\right) = 1 - \alpha. \quad (2.25)$$

Bei der letzten Bedingung ist zu beachten, dass  $Y_i$  als  $\chi^2$ -Zufallsvariable nicht-negativ ist und falls für eine Konstante  $c$  gilt:  $\sum_{i=1}^K Y_i \leq c$ , dann ist auch  $\sum_{i=1}^j Y_i \leq c$  für alle  $j = 1, \dots, K$ , da  $\sum_{i=1}^j Y_i \leq \sum_{i=1}^K Y_i$ . Somit kann der kritische Wert direkt bestimmt werden durch

$$cv_{S, OBF}(\alpha, K, \{\lambda_i\}) = \chi^2(\lambda_G)_{1-\alpha}, \quad \lambda_G = \sum_{i=1}^K \lambda_i. \quad (2.26)$$

In Anhang A in Tabelle A.1 befinden sich die kritischen Werte für das O'Brien/Fleming-Design mit inverser Normalmethode.

Die Testentscheidungen in den Sequenzen  $j$ ,  $j = 1, \dots, K$ , sind im Folgenden für die gewichtete inverse Normalmethode und die verallgemeinerte inverse  $\chi^2$ -Methode beschrieben.

#### Gewichtete inverse Normalmethode:

Mit  $Z_j$  aus (2.7) lauten die Testentscheidungen:

- Ablehnen von  $H_0$  in Stufe  $j$ , falls

$$Z_j > cv_{Z, OBF}(\alpha, K, \{\lambda_i\}) \quad (2.27)$$

und  $Z_i \leq cv_{Z, OBF}(\alpha, K, \{\lambda_i\})$  für  $i < j$ ,  $i, j \in \{1, \dots, K\}$ .

- Beibehalten von  $H_0$ , falls die letzte Stufe  $K$  erreicht ist und

$$Z_K \leq cv_{Z, OBF}(\alpha, K, \{\lambda_i\}). \quad (2.28)$$

Verallgemeinerte inverse  $\chi^2$ -Methode:

Bei Verwendung von  $S_j$  aus (2.12) lauten die Testentscheidungen:

- Ablehnen von  $H_0$  in Stufe  $j$ , falls

$$S_j > cv_{S, OBF}(\alpha, K, \{\lambda_i\}) \quad (2.29)$$

und  $S_i \leq cv_{S, OBF}(\alpha, K, \{\lambda_i\})$  für  $i < j$ ,  $i, j \in \{1, \dots, K\}$ .

- Beibehalten von  $H_0$ , falls die letzte Stufe  $K$  erreicht ist und

$$S_K \leq cv_{S, OBF}(\alpha, K, \{\lambda_i\}). \quad (2.30)$$

Bei der inversen Normalmethode können die kritischen Werte bei Gleichgewichtung aller Stufen, d.h.  $\lambda_i = \lambda$  für  $i = 1, \dots, K$ , mittels  $cv_{Z, OBF}(\alpha, K, \{\lambda\}) = \sqrt{\lambda} cv_{Z, OBF}(\alpha, K, \{1\})$  berechnet werden. Da die Überprüfung des Testproblems in den einzelnen Sequenzen mit  $Z_j = \sum_{i=1}^j \sqrt{\lambda_i} \Phi^{-1}(1 - p_i) = \sqrt{\lambda} \sum_{i=1}^j \Phi^{-1}(1 - p_i)$  erfolgt, spielt der Wert von  $\lambda$  bei Gleichgewichtung der Stufen keine Rolle.

### 2.2.3 Andere Designs

#### Verallgemeinertes Wang/Tsiatis-Design

Eine allgemeinere von Wang- und Tsiatis (1987) vorgeschlagene Klasse von kritischen Werten hängt zusätzlich von dem Parameter  $\Delta$  ab. Die mit  $cv_{Z, WT}(\alpha, K, \{\lambda_i\}, \Delta)$  sowie  $cv_{S, WT}(\alpha, K, \{\lambda_i\}, \Delta)$  bezeichneten Grenzen lassen sich zum Niveau  $\alpha$ ,  $0 < \alpha < 1$ , mit  $X_i$  und  $Y_i$  aus (2.17), aus den folgenden beiden Gleichungen bestimmen

$$P \left( \sum_{i=1}^j \sqrt{\lambda_i} X_i / \left( \sum_{i=1}^j \lambda_i \right)^\Delta \leq cv_{Z, WT}(\alpha, K, \{\lambda_i\}, \Delta) \quad \forall j = 1, \dots, K \right) = 1 - \alpha, \quad (2.31)$$

$$P \left( F_{\chi^2(\lambda_\Sigma(j)^{1-2\Delta} \lambda_G^{2\Delta})}^{-1} \left( F_{\chi^2(\lambda_\Sigma(j))} \left( \sum_{i=1}^j Y_i \right) \right) \leq cv_{S, WT}(\alpha, K, \{\lambda_i\}, \Delta) \quad \forall j = 1, \dots, K \right) = 1 - \alpha. \quad (2.32)$$

Ist  $\Delta = 0$  resultiert das O'Brien/Fleming-Design, im Fall von  $\Delta = 0.5$  ergeben sich die kritischen Werte von Pocock. Im Wang/Tsiatis-Design ist die Prüfgröße in Stufe  $j$ ,

$j = 1, \dots, K$ , gemäß  $N(0, (\lambda_{\Sigma}(j))^{1-2\Delta})$ -verteilt bei Verwendung der inversen Normalmethode und bei Kombination der Ergebnisse mit der inversen  $\chi^2$ -Methode stammt diese aus einer  $\chi^2(\lambda_{\Sigma}(j)^{1-2\Delta} \lambda_G^2 \Delta)$ -Verteilung. In Anhang A sind in Tabelle A.2 kritische Werte für  $\Delta = 0.0, 0.1, \dots, 0.7$  verzeichnet.

Die Testentscheidungen in den Sequenzen  $j, j = 1, \dots, K$ , sind im Folgenden für die gewichtete inverse Normalmethode und die verallgemeinerte inverse  $\chi^2$ -Methode beschrieben.

#### Gewichtete inverse Normalmethode:

Unter Verwendung der Statistik  $Z_j$  aus (2.7) lauten die Testentscheidungen:

- Ablehnen von  $H_0$  in Stufe  $j$ , falls

$$Z_j / (\lambda_{\Sigma}(j))^{\Delta} > cv_{Z,WT}(\alpha, K, \{\lambda_i\}, \Delta) \quad (2.33)$$

und  $Z_i / (\lambda_{\Sigma}(i))^{\Delta} \leq cv_{Z,WT}(\alpha, K, \{\lambda_i\}, \Delta)$  für  $i < j, i, j \in \{1, \dots, K\}$ .

- Beibehalten von  $H_0$ , falls die letzte Sequenz  $K$  erreicht ist und

$$Z_K / (\lambda_{\Sigma}(K))^{\Delta} \leq cv_{Z,WT}(\alpha, K, \{\lambda_i\}, \Delta). \quad (2.34)$$

#### Verallgemeinerte inverse $\chi^2$ -Methode:

Mit  $S_j$  aus (2.12) lauten die Testentscheidungen:

- Ablehnen von  $H_0$  in Stufe  $j$ , falls

$$F_{\chi^2(\lambda_{\Sigma}(j)^{1-2\Delta} \lambda_G^2 \Delta)}^{-1} \left( F_{\chi^2(\lambda_{\Sigma}(j))} (S_j) \right) > cv_{S,WT}(\alpha, K, \{\lambda_j\}, \Delta) \quad (2.35)$$

und  $F_{\chi^2(\lambda_{\Sigma}(i)^{1-2\Delta} \lambda_G^2 \Delta)}^{-1} (F_{\chi^2(\lambda_{\Sigma}(i))} (S_i)) \leq cv_{S,WT}(\alpha, K, \{\lambda_i\}, \Delta)$  für  $i < j, i, j \in \{1, \dots, K\}$ .

- Beibehalten von  $H_0$ , falls die letzte Sequenz  $K$  erreicht ist und

$$F_{\chi^2(\lambda_{\Sigma}(K)^{1-2\Delta} \lambda_G^2 \Delta)}^{-1} \left( F_{\chi^2(\lambda_{\Sigma}(K))} (S_K) \right) \leq cv_{S,WT}(\alpha, K, \{\lambda_i\}, \Delta). \quad (2.36)$$

Bei der inversen Normalmethode können die kritischen Werte bei gleichen Gewichten in allen Stufen, d.h.  $\lambda_i = \lambda$  für  $i = 1, \dots, K$ , berechnet werden mittels  $cv_{Z,WT}(\alpha, K, \{\lambda\}, \Delta) =$

$cv_{Z,WT}(\alpha, K, \{1\}, \Delta) \lambda^{0.5-\Delta}$ . Da die Überprüfung des Testproblems in den einzelnen Sequenzen mit  $Z_j = \sum_{i=1}^j \sqrt{\lambda_i} \Phi^{-1}(1-p_i) / (\lambda_{\Sigma}(j))^{\Delta} = \lambda^{0.5-\Delta} \sum_{i=1}^j \Phi^{-1}(1-p_i) / j^{\Delta}$  erfolgt, spielt der Wert von  $\lambda$  bei Gleichgewichtung der Stufen keine Rolle.

### Haybittle/Peto-Design

Eine sehr einfache Form des sequentiellen Testens schlagen Haybittle (1971) und Peto et al. (1976) vor. Bei Verwendung der ungewichteten inversen Normalmethode als Kombinationsmethode erfolgt eine Ablehnung der Nullhypothese in den Stufen  $j < K$ , falls für die standardisierte Stouffer-Statistik gilt

$$\begin{aligned} \frac{Z_j}{\sqrt{j}} &= \frac{\sum_{i=1}^j \Phi^{-1}(1-p_i)}{\sqrt{j}} > 3 \approx \Phi^{-1}(0.9987) \\ \Leftrightarrow \sum_{i=1}^j \Phi^{-1}(1-p_i) &> \sqrt{j} \Phi^{-1}(0.9987) = F_{N(0,j)}^{-1}(0.9987) \end{aligned} \quad (2.37)$$

Ist die Anzahl der gewählten Sequenzen klein, ist eine vorzeitige Ablehnung der Nullhypothese sehr unwahrscheinlich. Somit kann in der letzten Sequenz approximativ ein Test zum Niveau  $\alpha$  mit dem kritischen Wert eines nichtsequentiellen einstufigen Designs durchgeführt werden. Soll das globale Niveau  $\alpha$  hingegen exakt eingehalten werden, kann eine Anpassung des kritischen Werts der Stufe  $K$  erfolgen (vgl. Jennison und Turnbull (2000)). Mit  $X_i$  aus (2.17) wird dieser gewählt als

$$\begin{aligned} P \left( \sum_{i=1}^j X_i / \sqrt{j} \leq 3 \text{ für } j = 1, \dots, K-1 \quad \wedge \quad \sum_{i=1}^K X_i / \sqrt{K} \leq cv_{Z,HP}(\alpha, K, \{1\}) \right) \\ = 1 - \alpha. \end{aligned} \quad (2.38)$$

Für Stufenanzahlen  $K \geq 7$  können keine kritischen Werte zum globalen Niveau  $\alpha = 0.005$  ermittelt werden. Dies liegt daran, dass bei Verwendung des konstanten Werts 3 das bis zur sechsten Stufe verbrauchte Niveau

$$1 - P(Z_1/\sqrt{1} \leq 3, \dots, Z_6/\sqrt{6} \leq 3) = 0.0053$$

bereits größer als der Wert 0.005 ist.

Dieses einfache Prinzip der Ermittlung von kritischen Werten kann auf die inverse  $\chi^2$ -Methode übertragen werden. Für die Überprüfung des Testproblems wird in den Sequen-

zen  $j < k$  die ungewichtete Prüfgröße der inversen  $\chi^2$ -Methode  $S_j = \sum_{i=1}^j q_i(1)$  betrachtet. Als kritische Schranke wird  $F_{\chi^2(j)}^{-1}(0.9987)$ , das 0.9987-Quantil der  $\chi^2$ -Verteilung mit  $j$  Freiheitsgraden, verwendet. Um insgesamt einen Test zum globalen Niveau  $\alpha$  zu erhalten, wird in Stufe  $K$  der adjustierte kritische Wert mit  $Y_i$  aus (2.17) und  $\lambda_i = 1$  gewählt als

$$P \left( \sum_{i=1}^j Y_i \leq F_{\chi^2(j)}^{-1}(0.9987) \text{ für } j = 1, \dots, K-1 \quad \wedge \quad \sum_{i=1}^K Y_i \leq cv_{S,HP}(\alpha, K, \{1\}) \right) = 1 - \alpha. \quad (2.39)$$

Approximativ kann bei dieser Vorgehensweise in der letzten Sequenz das  $(1 - \alpha)$ -Quantil der  $\chi^2(K)$ -Verteilung als kritische Grenze verwendet werden. Für Designs mit  $K \geq 10$  Stufen können keine kritischen Werte zum globalen Niveau  $\alpha = 0.005$  bestimmt werden, da bereits in der neunten Sequenz gilt

$$1 - P \left( S_1 \leq F_{\chi^2(1)}^{-1}(0.9987), \dots, S_9 \leq F_{\chi^2(9)}^{-1}(0.9987) \right) = 0.0051.$$

Die kritischen Werte für das Haybittle/Peto-Design mit inverser Normal- sowie  $\chi^2$ -Methode befinden sich in Anhang A in Tabelle A.3.

Bei Verwendung des Wertes  $\sqrt{j} \cdot 3$  ( $\approx \sqrt{j} \cdot \Phi^{-1}(0.9987)$ ) bezogen auf die Stouffer-Statistik und des  $F_{\chi^2(j)}^{-1}(0.9987)$ -Quantils bezogen auf die Prüfgröße der ungewichteten  $\chi^2$ -Methode zur Überprüfung des Testproblems in den Sequenzen  $j = 1, \dots, K - 1$  ist die Wahrscheinlichkeit für eine vorzeitige Ablehnung der Nullhypothese sehr gering. Im Folgenden wird das Haybittle/Peto-Design leicht modifiziert. Hierfür wird als kritische Schranke in den Stufen  $j = 1, \dots, K - 1$  das  $F_{N(0,j)}^{-1}(0.995)$ - bzw. das  $F_{\chi^2(j)}^{-1}(0.995)$ -Quantil verwendet, welches in Standardtabellen zu finden ist (vgl. Hartung et al. (1999)). Für die Normalverteilung entspricht dieses Quantil ungefähr dem Wert  $\sqrt{j} \cdot 2.5758$ . Für die beiden Kombinationsmethoden erfolgt die Bestimmung des kritischen Wertes in Stufe  $K$  mit  $X_i$  und  $Y_i$  aus (2.17) und  $\lambda_i = 1$  durch

$$P \left( \sum_{i=1}^j X_i / \sqrt{j} \leq 2.5758 \text{ für } j = 1, \dots, K-1 \quad \wedge \quad \sum_{i=1}^K X_i / \sqrt{K} \leq cv_{Z,HP}^*(\alpha, K, \{1\}) \right) = 1 - \alpha, \quad (2.40)$$

$$P \left( \sum_{i=1}^j Y_i \leq F_{\chi^2(j)}^{-1}(0.995) \text{ für } j = 1, \dots, K-1 \quad \wedge \quad \sum_{i=1}^K Y_i \leq cv_{S,HP}^*(\alpha, K, \{1\}) \right) = 1 - \alpha. \quad (2.41)$$

Bei beiden Kombinationsmethoden können zu dem globalen Niveau  $\alpha = 0.005$  keine kritischen Werte ermittelt werden, da das in der ersten Stufe verbrauchte Niveau bereits ungefähr dem Wert 0.005 entspricht. Die kritischen Werte  $cv_{Z,HP}^*(\alpha, K, \{1\})$  und  $cv_{S,HP}^*(\alpha, K, \{1\})$  für  $\alpha = 0.05$  und  $\alpha = 0.025$  finden sich in Anhang A in Tabelle A.3. Hierbei ist zu beachten, dass bei Verwendung der inversen Normalmethode für  $\alpha = 0.025$  keine kritischen Werte für die Stufen 15 und 20 ermittelt werden können, da das nach 12 Sequenzen vergebene Niveau diesen Wert bereits übertrifft. Für die inverse  $\chi^2$ -Methode ist dies erst nach 20 Stufen der Fall.

#### 2.2.4 Vergleich der adaptiv gruppensequentiellen Designs

Im Folgenden werden die vorgestellten adaptiv gruppensequentiellen Designs hinsichtlich der Vergabe des globalen Fehlers I. Art in den einzelnen Sequenzen miteinander verglichen, d.h. es werden das bis zu einer Stufe vergebene Niveau  $\alpha(j)$  sowie das pro Stufe verbrauchte Niveau  $\alpha_j$ ,  $j = 1, \dots, K$ , betrachtet. Die Angabe der kritischen Werte erfolgt jeweils bezogen auf die standardisierte Prüfgröße

$$Z_j = \frac{\sum_{i=1}^j \sqrt{\lambda_i} \Phi^{-1}(1 - p_i)}{\sqrt{\lambda_\Sigma(j)}} \stackrel{H_0}{\sim} N(0, 1), \quad j = 1, \dots, K,$$

für die inverse Normalmethode, sowie

$$S_j^* = F_{\chi^2(\lambda_G)}^{-1}(F_{\chi^2(\lambda_\Sigma(j))}(S_j)) \stackrel{H_0}{\sim} \chi^2(\lambda_G), \quad j = 1, \dots, K,$$

für die inverse  $\chi^2$ -Methode. Im Wang/Tsiatis-Design (WT) mit inverser Normalmethode ist die gemeinsame Verteilung der Statistiken  $Z_1/(\lambda_\Sigma(1))^\Delta, \dots, Z_K/(\lambda_\Sigma(K))^\Delta$  die multivariate Normalverteilung mit Erwartungswertvektor  $\vec{\mu}_{WT}$  und Kovarianzmatrix  $\Sigma_{WT}$ . Die Kovarianz von  $Z_j/(\lambda_\Sigma(j))^\Delta$  und  $Z_k/(\lambda_\Sigma(k))^\Delta$ ,  $j, k \in \{1, \dots, K\}$  und  $j < k$ , ist

$$Cov(Z_j/(\lambda_\Sigma(j))^\Delta, Z_k/(\lambda_\Sigma(k))^\Delta) = (\lambda_\Sigma(j))^{1-\Delta}/(\lambda_\Sigma(k))^\Delta. \quad (2.42)$$

Somit kann das bis zur Stufe  $j$ ,  $j = 1, \dots, K$ , vergebene Niveau ermittelt werden aus

$$\begin{aligned} \alpha(j) &= P_{H_0}(Z_i/(\lambda_\Sigma(i))^\Delta > cv_{Z,WT}(\alpha, K, \{\lambda_i\}, \Delta) \text{ für mind. ein } i \in \{1, \dots, j\}) \\ &= 1 - F_{N_j(\vec{0}, \Sigma_{WT})}(cv_{Z,WT}(\alpha, K, \{\lambda_i\}, \Delta), \dots, cv_{Z,WT}(\alpha, K, \{\lambda_i\}, \Delta)). \end{aligned} \quad (2.43)$$

Für  $\Delta = 0$  resultiert das im O'Brien/Fleming-Design und für  $\Delta = 0.5$  das im Pocock-Design bis zur Stufe  $j$  verbrauchte Niveau. Im Haybittle/Peto-Design wird dieses berechnet durch

$$\alpha(j) = \begin{cases} 1 - F_{N_j(\vec{0}, \Sigma)}(3, \dots, 3), & j < K, \\ 1 - F_{N_j(\vec{0}, \Sigma)}(3, \dots, 3, cv_{Z,HP}(\alpha, K, \{1\})), & j = K. \end{cases} \quad (2.44)$$

Hierbei ist  $\Sigma$  die Kovarianzmatrix mit den Kovarianzen aus (2.10). Für das modifizierte Haybittle/Peto-Design erfolgt die Berechnung analog.

Bei Kombination der Ergebnisse der einzelnen Sequenzen mittels der inversen  $\chi^2$ -Methode ist das bis zur Stufe  $j$ ,  $j = 1, \dots, K$ , verbrauchte Niveau im Wang/Tsiatis-Design

$$\alpha(j) = P_{H_0} \left( F_{\chi^2(\lambda_{\Sigma}^{(i)})^{1-2\Delta}}^{-1} (F_{\chi^2(\lambda_{\Sigma}^{(i)})}(S_i)) > cv_{S,WT}(\alpha, K, \{\lambda_i\}, \Delta) \right. \\ \left. \text{für mind. ein } i \in \{1, \dots, j\} \right). \quad (2.45)$$

Bei Verwendung des Haybittle/Peto-Designs wird dieses berechnet durch

$$\alpha(j) = \begin{cases} 1 - P_{H_0} \left( S_1 \leq F_{\chi^2(1)}^{-1}(0.9987), \dots, S_j \leq F_{\chi^2(j)}^{-1}(0.9987) \right), & j < K, \\ 1 - P_{H_0} \left( S_1 \leq F_{\chi^2(1)}^{-1}(0.9987), \dots, S_{K-1} \leq F_{\chi^2(K-1)}^{-1}(0.9987), \right. \\ \left. S_K \leq cv_{S,HP}(\alpha, K, \{1\}) \right), & j = K. \end{cases} \quad (2.46)$$

Die Berechnung von  $\alpha(j)$  für das modifizierte Design erfolgt analog.

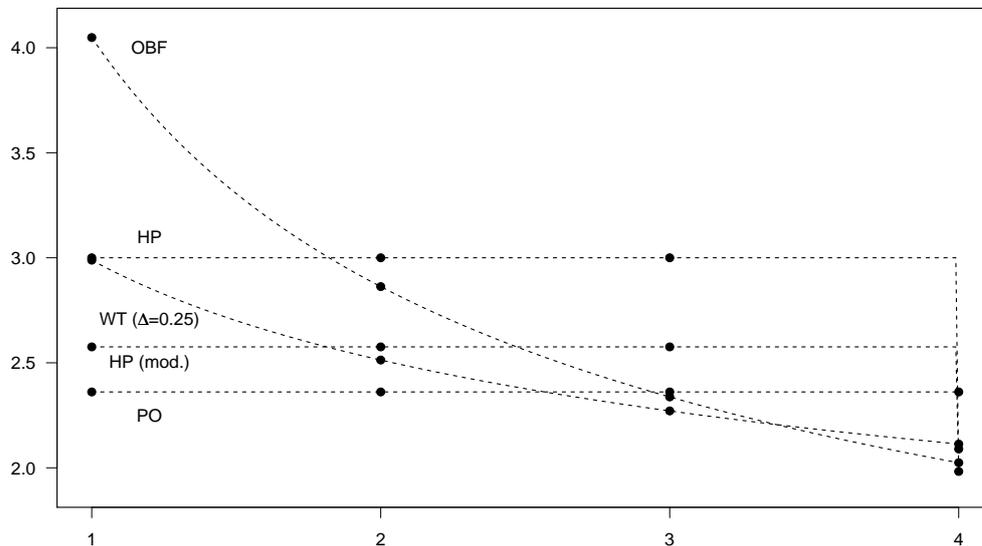
Das pro Stufe  $j = 1, \dots, K$  verbrauchte Niveau ist

$$\alpha_j = \alpha(j) - \alpha(j-1). \quad (2.47)$$

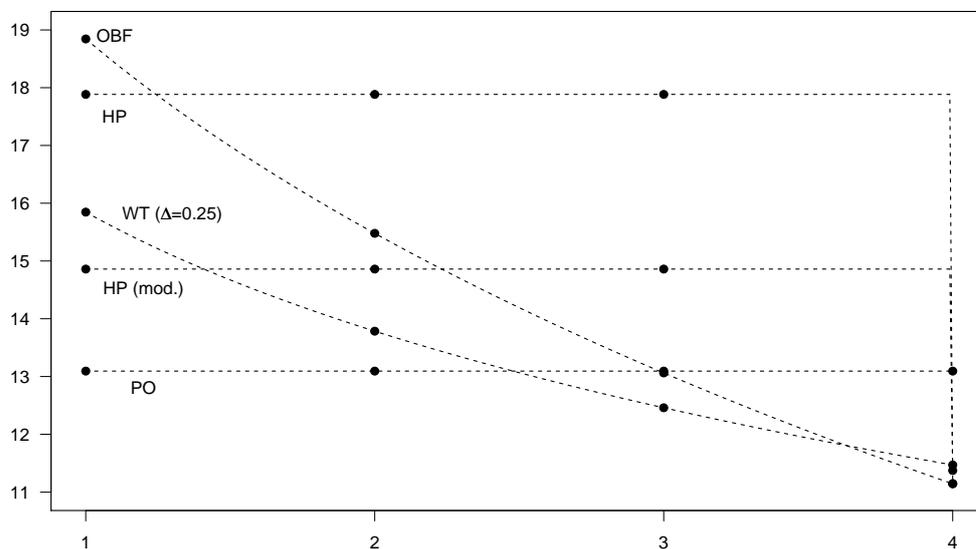
Exemplarisch wird die Vergabe des globalen Niveaus anhand einer klinischen Prüfung mit flexiblem Studiendesign mit  $K = 4$  Sequenzen und Gewichten  $\lambda_j = 1$ ,  $j = 1, \dots, 4$ , für  $\alpha = 0.025$  veranschaulicht. Die kritischen Werte der Stufen  $j = 1, \dots, 4$  bei Verwendung der inversen Normal- sowie  $\chi^2$ -Methode sind in Abbildung 2.1 zu sehen.

**Abbildung 2.1:** Einseitige kritische Werte in einer Studie mit flexiblem Design mit  $K = 4$  Sequenzen, globalem Niveau  $\alpha = 0.025$  und Gewichten  $\lambda_j = 1, j = 1, \dots, 4$

(a) Bezogen auf die standardisierte Stouffer-Statistik  $Z_j/\sqrt{\lambda_{\Sigma}(j)}$



(b) Bezogen auf die standardisierte Statistik  $F_{\chi^2(\lambda_G)}^{-1}(F_{\lambda_{\Sigma}(j)}(S_j))$  der inversen  $\chi^2$ -Methode



Die Aufteilung des globalen Niveaus in den verschiedenen Designs ist in Tabelle 2.1 zusammengefasst. In der zweiten Spalte findet sich der kritische Wert bezogen auf die standardisierte Prüfgröße  $Z_j/\sqrt{\lambda_{\Sigma}(j)}$  bzw.  $F_{\chi^2(\lambda_G)}^{-1}(F_{\lambda_{\Sigma}(j)}(S_j))$ , in der dritten und vierten Spalte sind das bis Stufe  $j$  sowie das pro Stufe  $j, j = 1, \dots, 4$ , vergebene Niveau angegeben.

**Tabelle 2.1:** Vergabe des globalen Niveaus  $\alpha = 0.025$  in einer Studie mit flexiblem Design mit  $K = 4$  Stufen und Gewichten  $\lambda_j = 1$ ,  $j = 1, \dots, 4$ , mit  $\alpha(j)$  dem bis Stufe  $j$  vergebenem Niveau und  $\alpha_j$  dem pro Stufe  $j$  verbrauchten Niveau

<b>Inverse Normalmethode</b>				
Pocock	$j$	$cv_{Z,PO}$	$\alpha(j)$	$\alpha_j$
	1	2.3613	0.0091	0.0091
	2	2.3613	0.0158	0.0067
	3	2.3613	0.0209	0.0051
	4	2.3613	0.0250	0.0041
O'Brien/Fleming	$j$	$cv_{Z,OF}/\sqrt{\lambda_{\Sigma}(j)}$	$\alpha(j)$	$\alpha_j$
	1	4.0486	<0.0001	<0.0001
	2	2.8628	0.0021	0.0021
	3	2.3375	0.0105	0.0083
	4	2.0243	0.0250	0.0145
Wang/Tsiatis	$j$	$cv_{Z,WT} \frac{\lambda_{\Sigma}(j)^{\Delta}}{\sqrt{\lambda_{\Sigma}(j)}}$	$\alpha(j)$	$\alpha_j$
$\Delta = 0.25$	1	2.9887	0.0014	0.0014
	2	2.5132	0.0068	0.0054
	3	2.2709	0.0152	0.0083
	4	2.1133	0.0250	0.0098
$\Delta = 0.4$	1	2.5651	0.0052	0.0052
	2	2.3933	0.0119	0.0067
	3	2.2982	0.0186	0.0067
	4	2.2330	0.0250	0.0064
Haybittle/Peto	$j$	$cv_{Z,HP}$	$\alpha(j)$	$\alpha_j$
	1	3.0000	0.0013	0.0013
	2	3.0000	0.0025	0.0011
	3	3.0000	0.0034	0.0010
	4	1.9828	0.0250	0.0216
Haybittle/Peto (mod.)	$j$	$cv_{Z,HP}^*$	$\alpha(j)$	$\alpha_j$
	1	2.5758	0.0050	0.0050
	2	2.5758	0.0088	0.0038
	3	2.5758	0.0118	0.0030
	4	2.0897	0.0250	0.0132
<b>Inverse <math>\chi^2</math>-Methode</b>				
Pocock	$j$	$cv_{S,PO}$	$\alpha(j)$	$\alpha_j$
	1	13.0939	0.0108	0.0108
	2	13.0939	0.0170	0.0062
	3	13.0939	0.0214	0.0044
	4	13.0939	0.0250	0.0036

O'Brien/Fleming	$j$	$F_{\chi^2(\lambda_G)}^{-1}(F_{\chi^2(\lambda_\Sigma(j))}(cv_{S, OBF}))$	$\alpha(j)$	$\alpha_j$
	1	18.8441	0.0009	0.0009
	2	15.4790	0.0038	0.0029
	3	13.0622	0.0111	0.0073
	4	11.1433	0.0251	0.0140
Wang/Tsiatis	$j$	$F_{\chi^2(\lambda_G)}^{-1}(F_{\chi^2(\lambda_\Sigma(j)^{1-2\Delta} \lambda_G^{2\Delta})}(cv_{S, WT}))$	$\alpha(j)$	$\alpha_j$
$\Delta = 0.25$	1	15.9163	0.0031	0.0031
	2	13.8516	0.0084	0.0053
	3	12.5238	0.0158	0.0074
	4	11.5312	0.0250	0.0092
$\Delta = 0.4$	1	14.1449	0.0067	0.0067
	2	13.2068	0.0129	0.0062
	3	12.6387	0.0191	0.0062
	4	12.2274	0.0250	0.0059
Haybittle/Peto	$j$	$F_{\chi^2(\lambda_G)}^{-1}(F_{\chi^2(\lambda_\Sigma(j))}(cv_{S, HP}))$	$\alpha(j)$	$\alpha_j$
	1	17.8844	0.0013	0.0013
	2	17.8844	0.0021	0.0008
	3	17.8844	0.0027	0.0006
	4	11.1584	0.0249	0.0222
Haybittle/Peto (mod.)	$j$	$F_{\chi^2(\lambda_G)}^{-1}(F_{\chi^2(\lambda_\Sigma(j))}(cv_{S, HP}^*))$	$\alpha(j)$	$\alpha_j$
	1	14.8603	0.0050	0.0050
	2	14.8603	0.0079	0.0029
	3	14.8603	0.0101	0.0022
	4	11.3729	0.0249	0.0148

Die verschiedenen Designs verhalten sich für die inverse Normal- und  $\chi^2$ -Methode sehr ähnlich, so dass die folgenden Ausführungen für beide Kombinationsmethoden zutreffen. Im Pocock-Design erfolgt durch die konstanten kritischen Schranken bereits in den ersten Sequenzen eine hohe Vergabe des globalen Niveaus, hingegen wird im O'Brien/Fleming-Design der Großteil von diesem erst in der letzten Stufe verbraucht und somit ein Beenden der Studie in einer frühen Stufe erschwert. Bei Verwendung der kritischen Schranken gemäß Wang und Tsiatis erfolgt die Vergabe des globalen Niveaus in Abhängigkeit von der Wahl des Parameters  $\Delta$ . Ist dieser nahe null, so ähnelt das Verhalten eher dem O'Brien/Fleming-Design, für große Werte von  $\Delta$  wird bereits in den früheren Stufen ein Großteil des globalen Niveaus verbraucht. Im Haybittle/Peto-Design wird fast das gesamte globale Niveau  $\alpha$  in der letzten Stufe vergeben. In dem modifizierten Haybittle/Peto-

Design wird durch Verwendung des festen Wertes 2.5758 (anstelle des Wertes 3) bei der inversen Normalmethode bzw. des  $F_{\chi^2(\lambda_\Sigma(j))}^{-1}(0.995)$ -Quantils (anstelle des 0.9987-Quantils) bei der inversen  $\chi^2$ -Methode ein wesentlich größerer Anteil des globalen Niveaus bereits in den ersten drei Sequenzen verbraucht.

### 2.3 Self-Designing klinischer Studien

Das Testproblem (2.1) soll erneut in  $j$ ,  $j = 1, \dots, k$ , voneinander unabhängigen Stufen, bezeichnet mit Stufe(1), ..., Stufe( $k$ ), untersucht werden. Bei dem Self-Designing klinischer Studien ist im Gegensatz zu den adaptiv gruppensequentiellen Designs aus Kapitel 2.2 die Anzahl der Sequenzen  $k$  nicht notwendigerweise durch eine vorgegebene Zahl  $K$  beschränkt. Hier können sowohl die Stichprobenumfänge  $n_j$  als auch die Gewichte  $\lambda_j$  basierend auf den bis vor dieser Stufe  $j$  vorliegenden Daten gewählt werden. Bezeichnet Stufe(0) wieder a priori Information, bedeutet dies

$$n_j = \hat{n}(j-1) = \hat{n}(\text{Stufe}(0), \dots, \text{Stufe}(j-1)) \quad (2.48)$$

und

$$\lambda_j = \hat{\lambda}(j-1) = \hat{\lambda}(\text{Stufe}(0), \dots, \text{Stufe}(j-1)),$$

wobei

$$\lambda_j \leq \lambda_G - \lambda_\Sigma(j-1), \quad \lambda_\Sigma(j) = \sum_{i=1}^j \lambda_i, \quad \lambda_\Sigma(0) = 0, \quad \lambda_i > 0, \quad j = 1, \dots, k. \quad (2.49)$$

Es ist zu beachten, dass die Gesamtanzahl durchgeführter Stufen  $k$  zufällig ist und sich im Verlauf der sequentiellen Studie realisiert. Unter der Voraussetzung, dass für eine Sequenz nichtnegativer Gewichte  $\lambda_1, \dots, \lambda_k, \dots$  unter  $H_0$  mit Wahrscheinlichkeit Eins ein endliches (zufälliges)  $k$  existiert, ist

$$\lambda_\Sigma(k) = \sum_{j=1}^k \lambda_j = \sum_{j=1}^{\infty} \lambda_j = \lambda_G \quad (f.s.). \quad (2.50)$$

Dies wird beispielsweise erfüllt, wenn ein minimales Stufengewicht  $\lambda_{\min} > 0$  eingeführt wird, da in diesem Fall gilt

$$k \leq \lambda_G / \lambda_{\min}. \quad (2.51)$$

Die Testentscheidungen in den Sequenzen  $j, j = 1, \dots, K$ , sind im Folgenden für die gewichtete inverse Normalmethode bzw. die verallgemeinerte inverse  $\chi^2$ -Methode beschrieben.

Gewichtete inverse Normalmethode:

Das Testen von  $H_0$  kann bei Verwendung der Kombinationsstatistik  $Z_j$  aus (2.7) nur am Studienende erfolgen (vgl. Hartung (2001)). Ist das globale Gewicht  $\lambda_G$  komplett auf die einzelnen Sequenzen aufgeteilt, d.h. in Stufe  $k$  gilt

$$\lambda_{\Sigma}(k) = \sum_{j=1}^k \lambda_j = \lambda_G, \quad (2.52)$$

endet die Studie in dieser Sequenz. Anschließend erfolgt die Berechnung der finalen Teststatistik  $Z_k$ . Unter der Bedingung  $\lambda_{\Sigma}(k) = \lambda_G$  ist  $Z_k$  gemäß  $N(0, \lambda_G)$ -verteilt. Der Nachweis befindet sich in Anhang B.1. Die Testentscheidung in der letzten Stufe  $k$  lautet:

- Ablehnen von  $H_0$ , falls  $\lambda_{\Sigma}(k) = \lambda_G$  und

$$Z_k > \sqrt{\lambda_G} \Phi^{-1}(1 - \alpha). \quad (2.53)$$

- Beibehalten von  $H_0$ , falls  $\lambda_{\Sigma}(k) = \lambda_G$  und

$$Z_k \leq \sqrt{\lambda_G} \Phi^{-1}(1 - \alpha). \quad (2.54)$$

Durch das einmalige Testen am Studienende wird das globale Niveau  $\alpha$  ohne Adjustierung des kritischen Werts kontrolliert.

Verallgemeinerte inverse  $\chi^2$ -Methode:

Erfolgt die Kombination der Ergebnisse der einzelnen Sequenzen mit der Statistik  $S_j$  aus (2.12), kann in jeder Sequenz  $j, j = 1, \dots, k$ , das einseitige Testproblem (2.1) überprüft werden (vgl. Hartung (2000), Hartung und Knapp (2003)). Die möglichen Testentscheidungen sind:

- Ablehnen von  $H_0$  in Stufe  $j$ , falls

$$S_j > \chi^2(\lambda_G)_{1-\alpha} \quad (2.55)$$

und  $S_i \leq \chi^2(\lambda_G)_{1-\alpha}$  für  $i < j$  und  $i, j \in \{1, \dots, k\}$ .

- Beibehalten von  $H_0$  in Stufe  $k$ , falls  $\lambda_\Sigma(k) = \lambda_G$  und

$$S_k \leq \chi^2(\lambda_G)_{1-\alpha}. \quad (2.56)$$

Das globale Niveau  $\alpha$  wird kontrolliert, da wegen  $q_j(\lambda_j) \geq 0 \forall j$  gilt:  $S_j \geq S_{j-1}$  für  $j \in \{1, \dots, k\}$ . Der Nachweis der  $\chi^2(\lambda_G)$ -Verteilung der Teststatistik  $S_k$ , falls  $\lambda_\Sigma(k) = \lambda_G$  ist, befindet sich in Anhang B.2.

## 2.4 Adaptive Wahl der Fallzahl und der Gewichtung

Im Folgenden wird die Bestimmung der Fallzahl sowie der Gewichtung in den einzelnen Sequenzen der flexiblen Designs basierend auf den vorliegenden Daten erläutert. Da bei den adaptiv gruppensequentiellen Verfahren die Gewichte der einzelnen Stufen vor Studienbeginn festgelegt werden, trifft die Beschreibung der datenbasierten Wahl der Gewichtung nur für das Self-Designing zu.

### 2.4.1 Wahl der Fallzahl

Für gegebene Fehlerwahrscheinlichkeiten I. und II. Art,  $\alpha$  und  $\beta$ , sei  $f_j(\alpha, \beta)$  eine Stichprobenfunktion, welche für die in Stufe  $j$ ,  $j = 1, \dots, k$ , als  $T_j$  vorliegende Teststatistik  $T$  geeignet ist. Der Index  $j$  bei  $f_j(\alpha, \beta)$  bedeutet, dass die einbezogenen Parameter durch Schätzungen aktualisiert werden, in welche die Informationen der Sequenzen  $1, \dots, j$  einfließen. Für die Stufe  $j = 0$  wird Vorwissen benötigt. Die adaptive Berechnung der Fallzahl wird basierend auf der Betrachtung der bedingten Power vorgestellt (vgl. Hartung (2001)). Diese gibt die Wahrscheinlichkeit an, dass in der aktuellen Sequenz  $j$  die Nullhypothese abgelehnt wird, gegeben die Beobachtungen der vorhergehenden Stufen  $1, \dots, j - 1$ ,  $j = 1, \dots, k$ . Als bedingte Fehlerwahrscheinlichkeit I. Art wird die Wahrscheinlichkeit bezeichnet, dass die Nullhypothese verworfen wird, obwohl diese zutrifft und die Informationen der Sequenzen  $1, \dots, j - 1$  vorliegen. Für die Schätzung von dieser werden im Folgenden zwei Strategien vorgestellt (vgl. Hartung (2006)). Bei den adaptiv gruppensequentiellen Designs werden das Wang/Tsiatis-Design, welches das Pocock- sowie O'Brien/Fleming-Design als Spezialfälle enthält, sowie das Haybittle/Peto-Design

betrachtet. Strategie I kann nicht im Self-Designing angewendet werden, da hierfür die Gewichtungen der Sequenzen vor Studienbeginn bekannt sein müssen.

### Gewichtete inverse Normalmethode

◇ Strategie I: Planung mit dem vorab festgelegten Gewicht  $\lambda_j$

Liegen die Beobachtungen der Sequenzen  $1, \dots, j-1$  vor und soll in der nächsten Stufe  $j$  mittels der Statistik

$$\hat{Z}_j = Z_{j-1} + \sqrt{\lambda_j} \Phi^{-1}(1 - \hat{p}_j) \quad (2.57)$$

der kritische Wert  $(\lambda_{\Sigma}(j))^{\Delta} cv_{Z,WT}(\alpha, K, \{\lambda_i\}, \Delta)$  im Wang/Tsiatis-Design (WT) überschritten werden, wobei  $Z_0 = 0$  und  $\lambda_{\Sigma}(0) = 0$  gilt, kann die bedingte Fehlerwahrscheinlichkeit I. Art geschätzt werden durch

$$\begin{aligned} \hat{p}_j &= P_{H_0} \left( \frac{\hat{Z}_j}{(\lambda_{\Sigma}(j))^{\Delta}} > cv_{Z,WT}(\alpha, K, \{\lambda_i\}, \Delta) \mid Z_{j-1} \right) \\ &= 1 - \Phi \left( \frac{(\lambda_{\Sigma}(j))^{\Delta} cv_{Z,WT}(\alpha, K, \{\lambda_i\}, \Delta) - Z_{j-1}}{\sqrt{\lambda_j}} \right). \end{aligned} \quad (2.58)$$

Im Haybittle/Peto-Design kann die Schätzung dieser Wahrscheinlichkeit mittels

$$\hat{p}_j = \begin{cases} 1 - \Phi \left( \sqrt{j} \left( 3 - \sum_{i=1}^{j-1} \Phi^{-1}(1 - p_i) \right) \right), & j < K, \\ 1 - \Phi \left( \sqrt{j} \left( cv_{Z,HP}(\alpha, K, \{1\}) - \sum_{i=1}^{j-1} \Phi^{-1}(1 - p_i) \right) \right), & j = K, \end{cases} \quad (2.59)$$

erfolgen.

◇ Strategie II: Planung mit dem Restgewicht  $\lambda_G - \lambda_{\Sigma}(j-1)$

Liegen die Beobachtungen der Sequenzen  $1, \dots, j-1$  vor und soll in der nächsten Stufe  $j$  mittels der Statistik

$$\hat{Z}_j = Z_{j-1} + \sqrt{\lambda_G - \lambda_{\Sigma}(j-1)} \Phi^{-1}(1 - \hat{p}_j) \quad (2.60)$$

der kritische Wert  $(\lambda_G)^{\Delta} cv_{Z,WT}(\alpha, K, \{\lambda_i\}, \Delta)$  im Wang/Tsiatis-Design (WT) bzw.  $\sqrt{\lambda_G} \Phi^{-1}(1 - \alpha)$  im Self-Designing (SD) überschritten werden, wobei  $Z_0 = 0$  und

$\lambda_\Sigma(0) = 0$  gilt, kann die bedingte Fehlerwahrscheinlichkeit I. Art geschätzt werden durch

$$\begin{aligned} \text{WT: } \hat{p}_j &= 1 - \Phi \left( \frac{(\lambda_G)^\Delta \text{cv}_{Z,WT}(\alpha, K, \{\lambda_i\}, \Delta) - Z_{j-1}}{\sqrt{\lambda_G - \lambda_\Sigma(j-1)}} \right), \\ \text{bzw.} & \\ \text{SD: } \hat{p}_j &= 1 - \Phi \left( \frac{\sqrt{\lambda_G} \Phi^{-1}(1 - \alpha) - Z_{j-1}}{\sqrt{\lambda_G - \lambda_\Sigma(j-1)}} \right). \end{aligned} \quad (2.61)$$

Im Haybittle/Peto-Design kann die Schätzung dieser Wahrscheinlichkeit mittels

$$\hat{p}_j = \begin{cases} 1 - \Phi \left( \frac{\sqrt{K} \left( 3 - \sum_{i=1}^{j-1} \Phi^{-1}(1 - p_i) \right)}{\sqrt{K - (j-1)}} \right), & j < K, \\ 1 - \Phi \left( \frac{\sqrt{K} \text{cv}_{Z,HP}(\alpha, K, \{1\}) - \sum_{i=1}^{j-1} \Phi^{-1}(1 - p_i)}{\sqrt{K - (j-1)}} \right), & j = K, \end{cases} \quad (2.62)$$

erfolgen.

### Verallgemeinerte inverse $\chi^2$ -Methode

◇ Strategie I: Planung mit dem vorab festgelegten Gewicht  $\lambda_j$

Liegen die Beobachtungen der Sequenzen  $1, \dots, j-1$  vor und soll in der nächsten Stufe  $j$  mittels der Statistik

$$\hat{S}_j = S_{j-1} + F_{\chi^2(\lambda_j)}^{-1}(1 - \hat{p}_j) \quad (2.63)$$

der kritische Wert  $\text{cv}_{S,WT}^*(\alpha, K, \{\lambda_i\}, \Delta)$  im Wang/Tsiatis-Design (WT) überschritten werden, mit  $S_0 = 0$  und  $\lambda_\Sigma(0) = 0$ , kann die bedingte Fehlerwahrscheinlichkeit I. Art geschätzt werden durch

$$\begin{aligned} \hat{p}_j &= P_{H_0} \left( \hat{S}_j > \text{cv}_{S,WT}^*(\alpha, K, \{\lambda_i\}, \Delta) \mid S_{j-1} \right) \\ &= 1 - F_{\chi^2(\lambda_j)} \left( \text{cv}_{S,WT}^*(\alpha, K, \{\lambda_i\}, \Delta) - S_{j-1} \right), \end{aligned} \quad (2.64)$$

mit  $\text{cv}_{S,WT}^*(\alpha, K, \{\lambda_i\}, \Delta) = F_{\chi^2(\lambda_\Sigma(j))}^{-1} \left( F_{\chi^2(\lambda_\Sigma(j))^{1-2\Delta} \lambda_G^{2\Delta}}(\text{cv}_{S,WT}(\alpha, K, \{\lambda_i\}, \Delta)) \right)$ .

Im Haybittle/Peto-Design kann die Schätzung dieser Wahrscheinlichkeit mittels

$$\hat{p}_j = \begin{cases} 1 - F_{\chi^2(1)} \left( F_{\chi^2(j)}^{-1}(0.9987) - S_{j-1} \right), & j < K, \\ 1 - F_{\chi^2(1)} (cv_{S,HP}(\alpha, K, \{1\}) - S_{j-1}), & j = K, \end{cases} \quad (2.65)$$

erfolgen.

◇ Strategie II: Planung mit dem Restgewicht  $\lambda_G - \lambda_{\Sigma}(j-1)$

Liegen die Beobachtungen der Sequenzen  $1, \dots, j-1$  vor und soll in der nächsten Stufe  $j$  mittels der Statistik

$$\hat{S}_j = S_{j-1} + F_{\chi^2(\lambda_G - \lambda_{\Sigma}(j-1))}^{-1}(1 - \hat{p}_j) \quad (2.66)$$

der kritische Wert  $cv_{S,WT}^*(\alpha, K, \{\lambda_i\}, \Delta)$  im Wang/Tsiatis-Design (WT) überschritten werden, mit  $S_0 = 0$  und  $\lambda_{\Sigma}(0) = 0$ , kann die bedingte Fehlerwahrscheinlichkeit I. Art geschätzt werden durch

$$\hat{p}_j = 1 - F_{\chi^2(\lambda_G - \lambda_{\Sigma}(j-1))} (cv_{S,WT}(\alpha, K, \{\lambda_i\}, \Delta) - S_{j-1}). \quad (2.67)$$

Für  $\Delta = 0$  resultiert

$$\hat{p}_j = 1 - F_{\chi^2(\lambda_G - \lambda_{\Sigma}(j-1))} (cv_S - S_{j-1}),$$

wobei  $cv_S = \chi^2(\lambda_G)_{1-\alpha}$  ist. Dies ist der Schätzer für die bedingte Fehlerwahrscheinlichkeit I. Art im O'Brien/Fleming-Design und im Self-Designing.

Im Haybittle/Peto-Design kann diese Wahrscheinlichkeit mittels

$$\hat{p}_j = \begin{cases} 1 - F_{\chi^2(K-(j-1))} \left( F_{\chi^2(j)}^{-1}(0.9987) - S_{j-1} \right), & j < K, \\ 1 - F_{\chi^2(K-(j-1))} (cv_{S,HP}(\alpha, K, \{1\}) - S_{j-1}), & j = K, \end{cases} \quad (2.68)$$

geschätzt werden.

Erfolgt die Berechnung des Umfangs der Stufe  $j$ ,  $j = 1, \dots, k$ , als

$$M_j(\beta) = f_{j-1}(\hat{p}_j, \beta), \quad j \geq 1, \quad (2.69)$$

wird der entsprechende kritische Wert mit der (bedingten) Power  $1 - \beta$  überschritten. Die bedingte Power gibt die Wahrscheinlichkeit an, dass in Sequenz  $j$  die Nullhypothese

abgelehnt wird, gegeben die Beobachtungen der vorhergehenden Stufen  $1, \dots, j - 1$ . Da sich in  $f_{j-1}$  einbezogene Parameterschätzungen eventuell noch nicht stabilisiert haben, sollte im Allgemeinen nur ein Teil von  $M_j(\beta)$  als Fallzahl  $n_j$  für die Stufe  $j$  verwendet werden, d.h.

$$n_j = \varepsilon_j M_j(\beta), \quad \text{mit } 0 < \varepsilon_j \leq 1. \quad (2.70)$$

Der Umfang  $n_j$  wird als natürliche Zahl oberhalb des berechneten Werts gewählt, so dass diese Anzahl in das Randomisationsschema der klinischen Studie passt. Die Definition der in (2.70) verwendeten Sequenz  $\{\varepsilon_j\}$  sollte vor Studienbeginn erfolgen. In den adaptiv gruppensequentiellen Designs werden die Gewichte der  $j = 1, \dots, K$  Stufen vor Studienbeginn festgelegt. Somit kann die Sequenz  $\{\varepsilon_j\}$  gewählt werden als

$$\varepsilon_j = \frac{\lambda_j}{\lambda_G - \lambda_\Sigma(j-1)}, \quad j = 1, \dots, K, \quad (2.71)$$

wobei dies einer zu dem in Stufe  $j$  verbliebenen Gewicht proportionalen Aufteilung der Fallzahl  $n_j$  entspricht. Eine weitere Möglichkeit für die Wahl der Sequenz  $\{\varepsilon_j\}$  wird in Kapitel 2.4.2 im Zusammenhang mit der adaptiven Wahl der Gewichte im Self-Designing vorgestellt.

Um sehr kleine oder sehr große Fallzahlen zu vermeiden, ist die Einführung eines minimalen als auch eines maximalen Umfangs,  $n_{j,\min}$  und  $n_{j,\max}$ , in Sequenz  $j$  sinnvoll mit  $n_{j,\min} \leq n_j \leq n_{j,\max}$ . Insgesamt ist zu beachten, dass größere Umfänge in den früheren Sequenzen zu einer höheren Power in diesen Stufen führen, was die Wahrscheinlichkeit für ein schnelleres Ende des Verfahrens erhöht (vgl. Hartung (2001)).

### 2.4.2 Wahl der Gewichtung im Self-Designing

Bei dem Self-Designing klinischer Studien muss nur das Gewicht der ersten Sequenz vor Beginn der Untersuchung festgelegt werden. Alle weiteren Gewichte können basierend auf den vorliegenden Beobachtungen ermittelt werden. Einer proportionalen Aufteilung der in Stufe  $j$  verbliebenen Varianz der finalen  $N(0, \lambda_G)$ - bzw.  $\chi^2(\lambda_G)$ -Testverteilung entspricht die Wahl des Gewichts durch

$$\lambda_j = (\lambda_G - \lambda_\Sigma(j-1)) \varepsilon_j, \quad \varepsilon_j = \frac{n_j}{M_j(\beta)}, \quad j \geq 1. \quad (2.72)$$

Die Sequenz  $\{\varepsilon_j\}$  kann hier wie folgt ermittelt werden (siehe Hartung (2001)). Für eine Fehlerrate II. Art  $\beta_g$ , mit  $\beta \leq \beta_g < 1$ , kann in Analogie zu (2.69), wobei  $\beta$  durch  $\beta_g$  ersetzt wird, die Fallzahl durch

$$m_j(\beta_g) = f_{j-1}(\hat{p}_j, \beta_g), \quad j \geq 1, \quad (2.73)$$

berechnet werden. Somit kann die Sequenz  $\{\varepsilon_j\}$  bestimmt werden als

$$\varepsilon_j = \frac{m_j(\beta_g)}{M_j(\beta)}, \quad j \geq 1. \quad (2.74)$$

Dieses Vorgehen ist verwandt mit der Power-Spending-Prozedur aus Bauer (1992). Es ist zu beachten, dass  $\beta_g$  nur dann eine untere Grenze für den Fehler II. Art in Stufe  $j$  ist, falls  $\lambda_j < \lambda_G - \lambda_\Sigma(j-1)$  ist. Große Werte für  $\beta_g$  erhöhen die Anzahl der durchzuführenden Sequenzen. Dies wird durch die Verwendung eines minimalen Stufenumfangs  $n_{j,\min}$  adjustiert.

Nehmen  $m_j(\beta_g)$  und  $M_j(\beta)$  in Stufe  $j$ ,  $j = 1, \dots, k$ , den gleichen Wert an oder gilt für das ermittelte Gewicht  $\lambda_j$

$$\lambda_G - \lambda_j - \lambda_\Sigma(j-1) < \lambda_{\min}, \quad (2.75)$$

erfolgt die Gewichtung in Sequenz  $j$  mit  $\lambda_j = \lambda_G - \lambda_\Sigma(j-1)$ . Somit ist das globale Gewicht  $\lambda_G$  verbraucht und die Studie wird nach dieser Stufe beendet. Anstelle der Einschränkung der sequentiellen Umfänge  $n_j$  durch  $n_{j,\min}$  und  $n_{j,\max}$  können direkt  $m_j(\beta_g)$  sowie  $M_j(\beta)$  restringiert werden

$$m_{j,\min} \leq m_j(\beta_g) \leq m_{j,\max} \quad \text{sowie} \quad M_{j,\min} \leq M_j(\beta) \leq M_{j,\max}.$$

## 2.5 Schätzen in flexiblen Designs

Neben der Überprüfung von Testproblemen ist die Abschätzung des Ausmaßes eines vorhandenen Wirkungsunterschieds von Interesse. Geeignete statistische Methoden hierfür sind die Punktschätzung und Intervallschätzung (Konfidenzschätzung) des unbekanntem Parameters, wobei entweder ein einzelner Schätzwert oder ein Intervall angegeben wird.

### 2.5.1 Intervallschätzung

Soll in den einzelnen Sequenzen eines mehrstufig durchgeführten Designs ein Konfidenzintervall für den unbekannt Parameter konstruiert werden, ist von der Verwendung von Konfidenzintervallen für nicht-sequentielle Designs abzuraten. Diese halten aufgrund des bereits bei dem mehrfachen Testen der Nullhypothese beschriebenen „Multiple Looks“-Effekts das vorgegebene Konfidenzniveau nicht ein. Deshalb schlagen Jennison und Turnbull (1984, 1989) die Verwendung von so genannten Repeated Konfidenzintervallen vor. Diese werden so konstruiert, dass die Sequenz aller geplanten Intervallen das gewünschte Konfidenzniveau besitzt. Erfolgt ein vorzeitiger Abbruch der Studie ist die Sequenz der beobachteten Intervalle konservativ. Diese Eigenschaft gilt nicht für das Self-Designing mit inverser Normalmethode. Hier kann nur in der letzten Stufe des Verfahrens nach Vergabe des gesamten globalen Gewichts ein Konfidenzintervall konstruiert werden, welches das vorgegebene Konfidenzniveau kontrolliert.

#### Gewichtete inverse Normalmethode

Sei  $\hat{\theta}_i$  ein Schätzer für  $\theta$  in Stufe  $i$ ,  $i = 1, \dots, k$ . Die für die Überprüfung des Testproblems geeignete Teststatistik  $T_i = f_i(\hat{\theta}_i)$  sei als Funktion von  $\hat{\theta}_i$  (streng) monoton steigend in  $\hat{\theta}_i$ . Somit ist  $T_i(\theta) = g_i(\hat{\theta}_i - \theta)$  als Funktion des Abstands zwischen der  $i$ -ten Schätzung  $\hat{\theta}_i$  und  $\theta$  (streng) monoton fallend in  $\theta$ . Die Kombinationsstatistik, welche als Funktion von  $\theta$  bis zur Sequenz  $j$  durch

$$Z_j(\theta) = \sum_{i=1}^j \sqrt{\lambda_i} \Phi^{-1} [F_{i,0}(T_i(\theta))], \quad j = 1, \dots, k, \quad (2.76)$$

gegeben ist, ist (streng) monoton fallend in  $\theta$ .

#### Adaptiv gruppensequentielle Designs

Um in den einzelnen Sequenzen eines adaptiv gruppensequentiellen Designs zweiseitige Repeated Konfidenzintervalle ermitteln zu können, werden die kritischen Werte eines mehrstufig durchgeführten zweiseitigen Testverfahrens benötigt. Mittels diesen kann ein Testproblem der Form

$$H_0 : \theta = 0 \quad \text{vs.} \quad H_1 : \theta \neq 0 \quad (2.77)$$

überprüft werden. Um directionale Konflikte zu vermeiden, falls die Beobachtungen der einzelnen Sequenzen in verschiedene Richtungen zeigen, werden nur einseitige p-Werte zur Kombination der Ergebnisse betrachtet (vgl. Wassmer (2001)).

In Kapitel 2.2 wird die Herleitung der kritischen Werte für adaptiv gruppensequentielle Designs mit der gewichteten inversen Normalmethode im einseitigen Fall beschrieben. In Analogie erfolgt die Bestimmung der zweiseitigen kritischen Werte. Zur Unterscheidung werden die folgenden Bezeichnungen für diese gewählt

$$\begin{aligned} \text{einseitig:} \quad & cv_{Z, OBF}(\alpha, K, \{\lambda_i\}), cv_{Z, PO}(\alpha, K, \{\lambda_i\}), cv_{Z, WT}(\alpha, K, \{\lambda_i\}, \Delta), \\ & cv_{Z, HP}(\alpha, K, \{1\}), \\ \text{zweiseitig:} \quad & cv_{Z, OBF}(\alpha, K, \{\lambda_i\})_2, cv_{Z, PO}(\alpha, K, \{\lambda_i\})_2, cv_{Z, WT}(\alpha, K, \{\lambda_i\}, \Delta)_2, \\ & cv_{Z, HP}(\alpha, K, \{1\})_2. \end{aligned}$$

Die Herleitung der zweiseitigen kritischen Werte wird im Folgenden für das Wang/Tsiatis-Design beschrieben. Seien  $X_i \sim N(0, 1)$ ,  $i = 1, \dots, K$ , voneinander unabhängige Zufallsvariablen, welche  $z_i(1) = \Phi^{-1}(1 - p_i)$  unter der Nullhypothese repräsentieren. Die zweiseitigen kritischen Werte zum Niveau  $2\alpha$  sind durch die Bedingung

$$P \left( \frac{\left| \sum_{i=1}^j \sqrt{\lambda_i} X_i \right|}{\left( \sum_{i=1}^j \lambda_i \right)^\Delta} \leq cv_{Z, WT}(2\alpha, K, \{\lambda_i\}, \Delta)_2 \quad \forall j = 1, \dots, K \right) = 1 - 2\alpha \quad (2.78)$$

definiert. In einem nicht-sequentiellem Design können für die Überprüfung eines zweiseitigen Testproblems zum Niveau  $2\alpha$  die einseitigen kritischen Werte zum Niveau  $\alpha$  verwendet werden. Im Mehrdimensionalen ist dies im Allgemeinen nicht möglich (vgl. Wassmer (2001)). Beispielhaft wird dies anhand des zweistufigen O'Brien/Fleming-Designs ( $\Delta = 0$ ) illustriert.

Die gemeinsame Verteilung der Statistiken  $Z_1, \dots, Z_K$  ist die multivariate Normalverteilung mit Erwartungswertvektor  $\vec{\mu}_{OBF}$  und Kovarianzmatrix  $\Sigma_{OBF}$ . Für ein Design mit  $K = 2$  Sequenzen und Gewichten  $\lambda_1$  und  $\lambda_2$  ist diese gegeben durch

$$\Sigma_{OBF} = \begin{pmatrix} \lambda_1 & \lambda_1 \\ \lambda_1 & \lambda_1 + \lambda_2 \end{pmatrix}$$

und die Berechnung der ein- und zweiseitigen kritischen Werte  $cv_{Z, OBF}(\alpha, 2, \{\lambda_i\}) = cv_{Z, OBF}$  und  $cv_{Z, OBF}(2\alpha, 2, \{\lambda_i\})_2 = cv_{Z, OBF, 2}$ , erfolgt mittels

$$\begin{aligned} \text{einseitig:} \quad & F_{N(\vec{0}, \Sigma_{OBF})}(cv_{Z, OBF}, cv_{Z, OBF}) \stackrel{!}{=} 1 - \alpha, \\ \text{zweiseitig:} \quad & F_{N(\vec{0}, \Sigma_{OBF})}(cv_{Z, OBF, 2}, cv_{Z, OBF, 2}) + F_{N(\vec{0}, \Sigma_{OBF})}(-cv_{Z, OBF, 2}, -cv_{Z, OBF, 2}) \\ & - F_{N(\vec{0}, \Sigma_{OBF})}(-cv_{Z, OBF, 2}, cv_{Z, OBF, 2}) - F_{N(\vec{0}, \Sigma_{OBF})}(cv_{Z, OBF, 2}, -cv_{Z, OBF, 2}) \stackrel{!}{=} 1 - 2\alpha. \end{aligned}$$

In Tabelle 2.2 werden für verschiedene Werte für  $\alpha$  sowie für unterschiedliche Gewichte

ein- und zweiseitige kritische Werte miteinander verglichen.

**Tabelle 2.2:** Vergleich ein- und zweiseitiger kritischer Werte im O'Brien/Fleming-Design mit gewichteter inverser Normalmethode mit Gewichten  $\lambda_1 = \lambda_2 = 1$  sowie  $\lambda_1 = 2$  und  $\lambda_2 = 8$  für  $\alpha = 0.005, \dots, 0.4$

$\alpha$	0.005	0.025	0.05	0.1	0.2	0.3	0.4
Gewichte $\lambda_1 = \lambda_2 = 1$							
$cv_{Z, OBF}(\alpha, 2, \{\lambda_i\})$	3.6481	2.7965	2.3730	1.8993	1.3488	0.9668	0.6491
$cv_{Z, OBF}(2\alpha, 2, \{\lambda_i\})_2$	3.6481	2.7965	2.3730	1.8993	1.3482	0.9604	0.6110
Gewichte $\lambda_1 = 2$ und $\lambda_2 = 8$							
$cv_{Z, OBF}(\alpha, 2, \{\lambda_i\})$	8.1455	6.1981	5.2031	4.0687	2.7819	1.9726	1.3629
$cv_{Z, OBF}(2\alpha, 2, \{\lambda_i\})_2$	8.1455	6.1981	5.2031	4.0686	2.7778	1.9390	1.2240

Die Ergebnisse in Tabelle 2.2 weisen darauf hin, dass sich die ein- und zweiseitigen kritischen Werte im Fall gleicher Gewichte in den einzelnen Sequenzen erst für sehr große Werte für das Signifikanzniveau  $\alpha$  voneinander unterscheiden. Sind die Gewichte  $\lambda_1$  und  $\lambda_2$  ungleich Eins und nicht identisch in den beiden Stufen, resultieren bereits für kleinere Werte für  $\alpha$  verschiedene ein- und zweiseitige kritische Werte. Bei der Planung eines adaptiv gruppensequentiellen Designs sollte deshalb genau überprüft werden, welche kritischen Werte im gewählten Design Anwendung finden.

Die zweiseitigen Repeated Konfidenzintervalle für den unbekannt Parameter  $\theta$  sind definiert als eine Sequenz von Intervallen  $KI_j$ ,  $j = 1, \dots, K$ , für welche die gemeinsame Überdeckungswahrscheinlichkeit zu einem festgelegten Konfidenzniveau  $(1 - 2\alpha)$  eingehalten wird

$$P(\theta \in KI_j \quad \forall j = 1, \dots, K) = 1 - 2\alpha. \quad (2.79)$$

Dies wird durch die Inversion eines zweiseitigen gruppensequentiellen Tests zum Niveau  $2\alpha$  ermöglicht (vgl. Jennison und Turnbull (1984, 1989)). Die Berechnung der einzelnen Konfidenzintervalle  $KI_j$  basiert auf den bis zur Stufe  $j$  vorliegenden Informationen.

In Stufe  $j$ ,  $j = 1, \dots, K$ , führt der Konfidenzbereich

$$KI_j : \left\{ \theta \left| \frac{|Z_j(\theta)|}{\left(\sum_{i=1}^j \lambda_i\right)^\Delta} \leq cv_{Z, WT}(2\alpha, K, \{\lambda_i\}, \Delta)_2 \right. \right\} \quad (2.80)$$

zu dem zweiseitigen Konfidenzintervall

$$KI_j : [\theta_{L,j}, \theta_{U,j}], \quad (2.81)$$

mit

$$\begin{aligned} \theta_{L,j} \text{ löst: } Z_j(\theta_{L,j}) &= + \left( \sum_{i=1}^j \lambda_i \right)^\Delta cv_{Z,WT}(2\alpha, K, \{\lambda_i\}, \Delta)_2, \\ \theta_{U,j} \text{ löst: } Z_j(\theta_{U,j}) &= - \left( \sum_{i=1}^j \lambda_i \right)^\Delta cv_{Z,WT}(2\alpha, K, \{\lambda_i\}, \Delta)_2. \end{aligned} \quad (2.82)$$

Im Fall des Haybittle/Peto-Designs können die Repeated Konfidenzintervalle ermittelt werden aus

$$KI_j : \begin{cases} \left\{ \theta \left| \frac{|\sum_{i=1}^j \Phi^{-1}(1-p_i)|}{\sqrt{j}} \leq 3 \right. \right\}, & \text{falls } j < K, \\ \left\{ \theta \left| \frac{|\sum_{i=1}^j \Phi^{-1}(1-p_i)|}{\sqrt{j}} \leq cv_{Z,HP}(\alpha, K, \{1\})_2 \right. \right\}, & \text{falls } j = K. \end{cases} \quad (2.83)$$

Die Lösung der Konfidenzgrenzen erfolgt analog zu dem Vorhergehenden.

Bei Beenden der Untersuchung bevor alle  $K$  geplanten Sequenzen beobachtet wurden, ist die resultierende Sequenz der Repeated Konfidenzintervalle konservativ. Da die gemeinsame Überdeckungswahrscheinlichkeit der Sequenz aller geplanten Intervalle gemäß (2.79) gerade  $(1 - 2\alpha)$  ist, ist der Konfidenzoeffizient der einzelnen Konfidenzintervalle  $KI_j$ ,  $j = 1, \dots, K$ , mindestens  $(1 - 2\alpha)$ , d.h.

$$P(\theta \in KI_j) \geq 1 - 2\alpha \quad \forall j = 1, \dots, K \quad (2.84)$$

(vgl. Jennison und Turnbull (2000)). Aufgrund der Definition der Repeated Konfidenzintervalle besteht eine Dualität zum Testen, da in dem Intervall der Stufe  $j$ ,  $j = 1, \dots, K$ , genau die Werte für den unbekannt Parameter  $\theta$  enthalten sind, für die der zugehörige gruppensequentielle Test nicht abgelehnt werden kann.

### Self-Designing

Aufgrund der unbekannt Verteilung der Kombinationsstatistik  $Z_j$  in den Stufen  $j < k$  mit  $\lambda_\Sigma(j) < \lambda_G$ , kann in Analogie zum einmaligen Testen am Studienende nur in Sequenz  $k$  mit  $\lambda_\Sigma(k) = \lambda_G$  ein Konfidenzintervall konstruiert werden, welches das vorgegebene Konfidenzniveau kontrolliert. Im Folgenden wird eine Möglichkeit vorgestellt, in allen

Stufen einen Vertrauensbereich anzugeben. Dieses Vorgehen ist analog zu der Ermittlung von Repeated Konfidenzintervallen für das Self-Designing mit inverser  $\chi^2$ -Methode (vgl. Hartung und Knapp (2006)).

Ist in Stufe  $j$  das globale Gewicht noch nicht verbraucht, d.h.  $\lambda_\Sigma(j) < \lambda_G$ , kann mit beliebigen Umfängen  $n_{1,j+1}$  und  $n_{2,j+1}$  formal die Statistik  $T_{j+1}$  für eine weitere Stufe ( $j+1$ ) definiert werden. Dieser wird das gesamte verbliebene Gewicht  $\lambda_{j+1} = \lambda_G - \lambda_\Sigma(j)$  zugewiesen. Somit gilt

$$Z_{j+1}(\theta) \sim N(0, \lambda_G), \quad \text{falls } \sum_{i=1}^j \lambda_i < \lambda_G. \quad (2.85)$$

Unter diesen Annahmen kann in Stufe  $j$ ,  $j = 1, \dots, k$ , mit dem Konfidenzbereich

$$KI_j : \left\{ \theta \mid |Z_j(\theta)| \leq \sqrt{\lambda_G} \Phi^{-1}(1 - \alpha) \right\} \quad (2.86)$$

ein zweiseitiges  $(1 - 2\alpha)$ -Konfidenzintervall ermittelt werden

$$KI_j : [\theta_{L,j}, \theta_{U,j}], \quad (2.87)$$

mit

$$\begin{aligned} \theta_{L,j} \text{ löst: } & Z_j(\theta_{L,j}) = \sqrt{\lambda_G} \Phi^{-1}(1 - \alpha), \\ \theta_{U,j} \text{ löst: } & Z_j(\theta_{U,j}) = \sqrt{\lambda_G} \Phi^{-1}(\alpha). \end{aligned} \quad (2.88)$$

Es ist beachten, dass das vorgegebene Konfidenzniveau sowohl für die Intervalle der Stufen  $j$  mit  $\lambda_\Sigma(j) < \lambda_G$  als auch für die Sequenz der Intervalle nicht kontrolliert wird. Die empirische Überdeckungswahrscheinlichkeit der Sequenz der Konfidenzintervalle wird in den Kapiteln 3.1.4 und 3.2.4 in Simulationen untersucht. Das finale Intervall  $KI_k$  mit  $\lambda_\Sigma(k) = \lambda_G$  besitzt aufgrund der  $N(0, \lambda_G)$ -Verteilung der finalen Kombinationsstatistik die Überdeckungswahrscheinlichkeit  $(1 - 2\alpha)$ , da gilt

$$P(\sqrt{\lambda_G} \Phi^{-1}(\alpha) \leq Z_k(\theta) \leq \sqrt{\lambda_G} \Phi^{-1}(1 - \alpha)) = 1 - 2\alpha.$$

Einseitige Konfidenzintervalle können in den adaptiv gruppensequentiellen Designs sowie im Self-Designing durch Betrachtung der entsprechenden einseitigen Konfidenzbereiche in Analogie zu dem Vorhergehenden bestimmt werden.

### Verallgemeinerte inverse $\chi^2$ -Methode

Sei  $\hat{\theta}_i$  ein Schätzer für  $\theta$  in Stufe  $i$ ,  $i = 1, \dots, k$ . Die für die Überprüfung des Testproblems geeignete Teststatistik  $T_i = f_i(\hat{\theta}_i)$  sei als Funktion von  $\hat{\theta}_i$  (streng) monoton steigend in  $\hat{\theta}_i$ . Somit ist  $T_i(\theta) = g_i(\hat{\theta}_i - \theta)$  als Funktion des Abstands zwischen der  $i$ -ten Schätzung  $\hat{\theta}_i$  und  $\theta$  (streng) monoton fallend in  $\theta$ . Die Kombinationsstatistik bis zur Stufe  $j$  ist als Funktion von  $\theta$  gegeben durch

$$S_j(\theta) = \sum_{i=1}^j F_{\chi^2(\lambda_i)}^{-1}(F_{0,i}(T_i(\theta))), \quad j = 1, \dots, k, \quad (2.89)$$

und (streng) monoton fallend in  $\theta$ . Eine weitere Statistik, welche alle Ergebnisse bis zur Sequenz  $j$  kombiniert, ist

$$R_j(\theta) = \sum_{i=1}^j F_{\chi^2(\lambda_i)}^{-1}(1 - F_{0,i}(T_i(\theta))), \quad j = 1, \dots, k, \quad (2.90)$$

(vgl. Hartung und Knapp (2006)). Diese kann durch Betrachtung des zu dem einseitigen Testproblem (2.1) entgegengesetzten Paares von Hypothesen motiviert werden und ist (streng) monoton steigend in  $\theta$ . Bei den adaptiv gruppensequentiellen Designs werden vor Studienbeginn die Gewichtung der einzelnen Stufen festgelegt und  $S_j(\theta)$  und  $R_j(\theta)$  sind in Stufe  $j$  gemäß  $\chi^2(\lambda_\Sigma(j))$ -verteilt. Bei dem Self-Designing erfolgt die Vergabe der Gewichte im Verlauf der Untersuchung. Ist in Sequenz  $k$  das gesamte globale Gewicht vergeben, gilt

$$S_k(\theta) \sim \chi^2(\lambda_G) \quad \text{bzw.} \quad R_k(\theta) \sim \chi^2(\lambda_G), \quad \text{falls} \quad \sum_{i=1}^k \lambda_i = \lambda_G. \quad (2.91)$$

Endet das Verfahren in Stufe  $k$  und es ist  $\lambda_\Sigma(k) < \lambda_G$ , kann mit beliebigen Umfängen  $n_{1,k+1}$  und  $n_{2,k+1}$  formal die Statistik  $T_{k+1}$  für eine weitere Stufe ( $k+1$ ) definiert werden, welcher das gesamte verbliebene Gewicht  $\lambda_{k+1} = \lambda_G - \lambda_\Sigma(k)$  zugewiesen wird. Somit gilt

$$S_{k+1}(\theta) \sim \chi^2(\lambda_G) \quad \text{bzw.} \quad R_{k+1}(\theta) \sim \chi^2(\lambda_G), \quad \text{falls} \quad \sum_{i=1}^k \lambda_i < \lambda_G. \quad (2.92)$$

Es ist zu beachten, dass die Stufe ( $k+1$ ) in der Praxis nicht durchgeführt werden muss, da das Testen der Nullhypothese in jeder Sequenz möglich ist und somit die Möglichkeit des vorzeitigen Beendens der Studie besteht. Die Konstruktion der ein- und zweiseitigen Repeated Konfidenzintervalle wird im Folgenden vorgestellt.

Adaptiv gruppensequentielle Designs

In Stufe  $j$ ,  $j = 1, \dots, K$ , führen die Konfidenzbereiche

$$KI_{L,j} : \{\theta \mid S_j(\theta) \leq cv_{S,WT}^*(\alpha, K, \{\lambda_i\}, \Delta)\} \quad (2.93)$$

und

$$KI_{U,j} : \{\theta \mid R_j(\theta) \leq cv_{S,WT}^*(\alpha, K, \{\lambda_i\}, \Delta)\}, \quad (2.94)$$

mit  $cv_{S,WT}^*(\alpha, K, \{\lambda_i\}, \Delta) = F_{\chi^2(\lambda_\Sigma(j))}^{-1}(F_{\chi^2(\lambda_\Sigma(j)^{1-2\Delta} \lambda_G^2 \Delta)}(cv_{S,WT}(\alpha, K, \{\lambda_i\}, \Delta)))$ , jeweils zu einseitigen Konfidenzintervallen, welche gegeben sind durch

$$KI_{L,j} : [\theta_{L,j}, \infty), \text{ mit } \theta_{L,j} \text{ löst } S_j(\theta_{L,j}) = cv_{S,WT}^*(\alpha, K, \{\lambda_i\}, \Delta) \quad (2.95)$$

und

$$KI_{U,j} : (-\infty, \theta_{U,j}], \text{ mit } \theta_{U,j} \text{ löst } R_j(\theta_{U,j}) = cv_{S,WT}^*(\alpha, K, \{\lambda_i\}, \Delta). \quad (2.96)$$

Aus der Schnittmenge der beiden einseitigen Konfidenzintervalle (2.95) und (2.96) resultiert das zweiseitige Konfidenzintervall für  $\theta$

$$KI_j : KI_{L,j} \cap KI_{U,j} = [\theta_{L,j}, \theta_{U,j}]. \quad (2.97)$$

Im Haybittle/Peto-Design erfolgt die Herleitung der zweiseitigen Konfidenzintervalle analog basierend auf der Betrachtung der Konfidenzbereiche

$$KI_{L,j} : \begin{cases} \left\{ \theta \mid \sum_{i=1}^j F_{\chi^2(i)}^{-1}(F_{0,i}(T_i(\theta))) \leq F_{\chi^2(j)}^{-1}(0.9987) \right\}, & j < K, \\ \left\{ \theta \mid \sum_{i=1}^j F_{\chi^2(i)}^{-1}(F_{0,i}(T_i(\theta))) \leq cv_{S,HP}(\alpha, K, \{1\}) \right\}, & j = K, \end{cases} \quad (2.98)$$

und

$$KI_{U,j} : \begin{cases} \left\{ \theta \mid \sum_{i=1}^j F_{\chi^2(i)}^{-1}(1 - F_{0,i}(T_i(\theta))) \leq F_{\chi^2(j)}^{-1}(0.9987) \right\}, & j < K, \\ \left\{ \theta \mid \sum_{i=1}^j F_{\chi^2(i)}^{-1}(1 - F_{0,i}(T_i(\theta))) \leq cv_{S,HP}(\alpha, K, \{1\}) \right\}, & j = K. \end{cases} \quad (2.99)$$

In Analogie zu den Ausführungen bei der inversen Normalmethode ist die resultierende Sequenz der ein- sowie zweiseitigen Repeated Konfidenzintervalle konservativ, falls die Studie vor Beobachtung aller geplanten Sequenzen  $K$  beendet wird. Der Konfidenzoeffizient der einzelnen Konfidenzintervalle  $KI_j$ ,  $j = 1, \dots, K$ , ist mindestens  $(1 - 2\alpha)$ .

Self-Designing

In Stufe  $j$ ,  $j = 1, \dots, k$ , führen die Konfidenzbereiche

$$KI_{L,j} : \{\theta \mid S_j(\theta) \leq \chi^2(\lambda_G)_{1-\alpha}\} \quad (2.100)$$

und

$$KI_{U,j} : \{\theta \mid R_j(\theta) \leq \chi^2(\lambda_G)_{1-\alpha}\} \quad (2.101)$$

jeweils zu einseitigen Konfidenzintervallen, welche gegeben sind durch

$$KI_{L,j} : [\theta_{L,j}, \infty), \text{ mit } \theta_{L,j} \text{ löst } S_j(\theta_{L,j}) = \chi^2(\lambda_G)_{1-\alpha} \quad (2.102)$$

und

$$KI_{U,j} : (-\infty, \theta_{U,j}], \text{ mit } \theta_{U,j} \text{ löst } R_j(\theta_{U,j}) = \chi^2(\lambda_G)_{1-\alpha}. \quad (2.103)$$

Aus der Schnittmenge der beiden einseitigen Intervalle (2.102) und (2.103) resultiert das zweiseitige Konfidenzintervall für  $\theta$

$$KI_j : KI_{L,j} \cap KI_{U,j} = [\theta_{L,j}, \theta_{U,j}]. \quad (2.104)$$

Der Konfidenzkoeffizient der einseitigen Konfidenzintervalle  $KI_{L,j}$  und  $KI_{U,j}$  ist mindestens  $(1 - \alpha)$  für  $j = 1, \dots, k$  und exakt  $(1 - \alpha)$  für  $j = k$ , falls  $\lambda_{\Sigma}(k) = \lambda_G$  gilt. Im Fall der zweiseitigen Intervalle  $KI_{L,j}$ ,  $j = 1, \dots, k$ , ist der Konfidenzkoeffizient größer oder gleich  $(1 - 2\alpha)$  (vgl. Hartung und Knapp (2006)).

Eine besondere Eigenschaft weisen das O'Brien/Fleming-Design und das Self-Designing bei Verwendung der inversen  $\chi^2$ -Methode auf. Hier sind sowohl die ein- als auch die zweiseitigen Repeated Konfidenzintervalle ineinander verschachtelt, wobei jeweils das Intervall der Stufe  $j + 1$  in demjenigen der vorhergehenden Sequenz  $j$  enthalten ist (vgl. Hartung und Knapp (2006)). Im Fall der zweiseitigen Repeated Konfidenzintervalle bedeutet dies

$$KI_j \supset KI_{j+1}, \quad j = 1, \dots, k - 1. \quad (2.105)$$

Diese Eigenschaft resultiert aus dem Monotonieverhalten der Kombinationsstatistiken der verallgemeinerten inversen  $\chi^2$ -Methode, welche als Funktion von  $\theta$  durch  $S_j(\theta)$  und  $R_j(\theta)$  gegeben sind. Beide Statistiken sind monoton steigend in dem Index  $j$ , d.h. es gilt

$$S_j(\theta) < S_{j+1}(\theta) \quad \text{sowie} \quad R_j(\theta) < R_{j+1}(\theta).$$

Sowohl  $S_j(\theta)$  als auch  $S_{j+1}(\theta)$  werden für die Ermittlung der unteren Konfidenzgrenze mit dem gleichen kritischen Wert in Beziehung gesetzt. Dies und die Tatsache, dass  $S_{j+1}(\theta)$  monoton fallend in  $\theta$  ist, führen zu der Konsequenz, dass  $\theta_{L,j+1} > \theta_{L,j}$  ist. Ebenso werden  $R_j(\theta)$  als auch  $R_{j+1}(\theta)$  für die Ermittlung der oberen Konfidenzgrenze mit dem gleichen kritischen Wert in Beziehung gesetzt. Zusammen mit der Tatsache, dass  $R_{j+1}(\theta)$  monoton steigend in  $\theta$  ist, resultiert dies in  $\theta_{U,j+1} < \theta_{U,j}$ .

### 2.5.2 Punktschätzung

Mittels der Punktschätzung kann ein einzelner Schätzwert für den unbekannt Parameter  $\theta$  angegeben werden. In flexiblen Designs erfolgt eine Unterteilung der Studie in einzelne Sequenzen. Um eine für die gesamte Untersuchung gemeinsame Schätzung von  $\theta$  zu ermitteln, eignen sich in der Meta-Analyse verwendete Methoden. In der Meta-Analyse werden mehrere Studien miteinander verbunden, um die Genauigkeit der statistischen Aussagen zu erhöhen. Die Schätzer der einzelnen Untersuchungen können durch Berechnung eines gewichteten arithmetischen Mittels zu einem gemeinsamen Schätzer kombiniert werden. Die Gewichtung kann auf verschiedene Weise erfolgen, wobei häufig die inverse Varianz der studienspezifischen Schätzer oder die Fallzahlen der einzelnen Studien verwendet werden. Im Folgenden werden das homogene Modell mit festen Effekten sowie Möglichkeiten zur Punktschätzung des unbekannt Parameters vorgestellt (siehe Hartung und Knapp (2005)).

#### Das homogene Modell mit festen Effekten

Seien die Schätzer  $\hat{\theta}_j$ ,  $j = 1, \dots, k$ , der  $k$  Studien voneinander unabhängige normalverteilte Zufallsvariablen mit einem gemeinsamen Erwartungswert  $\theta$  und Varianzen  $Var(\hat{\theta}_j)$ . Es wird die Annahme getroffen, dass ein für  $Var(\hat{\theta}_j)$  unverzerrter Schätzer  $\widehat{Var}(\hat{\theta}_j)$  vorliegt, welcher skaliert  $\chi^2$ -verteilt ist mit  $\lambda_j$  Freiheitsgraden. Außerdem sei  $\widehat{Var}(\hat{\theta}_j)$  stochastisch unabhängig von  $\hat{\theta}_j$ . Somit ist das Modell mit festen Effekten definiert durch

$$\hat{\theta}_j \sim N(\theta, Var(\hat{\theta}_j)), \quad \lambda_j \widehat{Var}(\hat{\theta}_j)/Var(\hat{\theta}_j) \sim \chi^2(\lambda_j), \quad j = 1, \dots, k. \quad (2.106)$$

Der Gauss-Markov Schätzer für  $\theta$  in diesem Modell ist durch

$$\hat{\theta}_{(GM),k} = \sum_{j=1}^k \frac{\omega_j}{\omega_{\Sigma}(k)} \hat{\theta}_j, \quad \omega_j = \frac{1}{\text{Var}(\hat{\theta}_j)}, \quad \omega_{\Sigma}(k) = \sum_{j=1}^k \omega_j, \quad (2.107)$$

gegeben. Dieser Schätzer hat unter allen für  $\theta$  unverzerrten Schätzern minimale Varianz. Die Eigenschaft der Unverzerrtheit gilt auch, wenn die unbekanntenen Gewichte durch ihre Schätzer ersetzt werden (vgl. Graybill und Deal (1959)). Der Graybill–Deal Schätzer für  $\theta$  ist

$$\hat{\theta}_{(1),k} = \sum_{j=1}^k w_{(1)j} \hat{\theta}_j, \quad w_{(1)j} = \frac{\hat{\omega}_j}{\hat{\omega}_{\Sigma}(k)}, \quad \hat{\omega}_j = \frac{1}{\widehat{\text{Var}}(\hat{\theta}_j)}, \quad \hat{\omega}_{\Sigma}(k) = \sum_{j=1}^k \hat{\omega}_j, \quad (2.108)$$

wobei  $\sum_{j=1}^k w_{(1)j} = 1$  und  $E(\hat{\theta}_{(1),k}) = \sum_{j=1}^k w_{(1)j} E(\hat{\theta}_j) = \theta$  gilt.

Alternativ kann die Gewichtung der Schätzer  $\hat{\theta}_j$  der einzelnen Untersuchungen mittels

$$w_{(2)j} = \frac{m_j}{\sum_{j=1}^k m_j}$$

erfolgen, wobei  $m_j = 2/(1/n_{1j} + 1/n_{2j})$  das harmonische Mittel der Gruppenumfänge  $n_{1j}$  und  $n_{2j}$  der Studie  $j$ ,  $j = 1, \dots, k$ , ist (vgl. Hartung und Knapp (2005)). Der den  $k$  klinischen Prüfungen gemeinsame Gesamteffekt der Behandlung kann geschätzt werden durch

$$\hat{\theta}_{(2),k} = \sum_{j=1}^k w_{(2)j} \hat{\theta}_j = \sum_{j=1}^k \frac{m_j}{m_{\Sigma}(k)} \hat{\theta}_j, \quad (2.109)$$

mit  $\sum_{j=1}^k w_{(2)j} = 1$  und  $E(\hat{\theta}_{(2),k}) = \theta$ .

## Übertragung auf klinische Prüfungen mit flexiblem Studiendesign

In klinischen Studien mit flexiblem Design wird für die Kombination der einzelnen Sequenzen implizit ein Modell mit festen Effekten zugrunde gelegt (siehe Hartung (2006)). Anstelle von verschiedenen Studien werden hier die Ergebnisse von mehreren unabhängigen Studienteilen miteinander kombiniert. Somit finden die vorgestellten Schätzer Verwendung, wobei hier in der Regel die Unverzerrtheit für  $\theta$  nicht erfüllt ist. In klinischen Prüfungen mit flexiblem Design gibt es verschiedene Verzerrungsquellen. Soll am Ende

der Untersuchung eine Schätzung für  $\theta$  angegeben werden und findet ein vorzeitiger Abbruch Berücksichtigung, erfolgt eine Selektion der Daten. Hierdurch kann es passieren, dass die Schätzung des interessierenden Parameters auf vergleichsweise extremen Daten basiert (siehe Coburger und Wassmer (2001, 2003)). Erfolgt in den Zwischenauswertungen lediglich eine Anpassung der Fallzahl, besteht eine positive Korrelation zwischen den Gewichtsfunktionen  $w_{(i)j}$ ,  $i = 1, 2$ , und den Stufenumfängen.

Die durch eine Anpassung der Fallzahl im Studienverlauf entstehende Verzerrung der Schätzung wird beispielhaft anhand des Schätzers  $\hat{\theta}_{(2),k}$  illustriert. Hierfür wird eine in  $k$  Stufen durchgeführte klinische Studie betrachtet, wobei ein vorzeitiges Beenden der Untersuchung nicht ermöglicht wird. In einem balancierten Design, d.h.  $m_j = n_j/2$ , gilt für den Erwartungswert von  $\hat{\theta}_{(2),k}$  in der letzten Stufe  $k$

$$\begin{aligned}
E(\hat{\theta}_{(2),k}) &= E \left[ E \left( \frac{\sum_{j=1}^k m_j \hat{\theta}_j}{\sum_{j=1}^k m_j} \mid \hat{\theta}_1, \dots, \hat{\theta}_{k-1}; \widehat{Var}(\hat{\theta}_1), \dots, \widehat{Var}(\hat{\theta}_{j-1}) \right) \right] \\
&= E(w_{(2)1}\hat{\theta}_1) + E(w_{(2)2}\hat{\theta}_2) + \dots \\
&\quad + E \left[ \underbrace{E \left( \left( 1 - \sum_{j=1}^{k-1} w_{(2)j} \right) \hat{\theta}_k \mid \hat{\theta}_1, \dots, \hat{\theta}_{k-1}; \widehat{Var}(\hat{\theta}_1), \dots, \widehat{Var}(\hat{\theta}_{j-1}) \right)} \right] \\
&= E \left[ \left( 1 - \sum_{j=1}^{k-1} w_{(2)j} \right) E \left( \hat{\theta}_k \mid \hat{\theta}_1, \dots, \hat{\theta}_{k-1}; \widehat{Var}(\hat{\theta}_1), \dots, \widehat{Var}(\hat{\theta}_{j-1}) \right) \right] \\
&= E \left[ \left( 1 - \sum_{j=1}^{k-1} w_{(2)j} \right) \theta \right] \\
&= E \left[ \theta + w_{(2)1}(\hat{\theta}_1 - \theta) + \dots + w_{(2)k-1}(\hat{\theta}_{k-1} - \theta) \right] \\
&= \theta + Cov(w_{(2)1}, \hat{\theta}_1) + \dots + Cov(w_{(2)k-1}, \hat{\theta}_{k-1}).
\end{aligned}$$

Der Umfang  $n_j$  der Stufe  $j$ ,  $j = 2, \dots, k$ , ist eine in  $\hat{\theta}_{j-1}$  fallende Funktion, so dass das Gewicht  $w_{(2)j-1}$  eine in  $\hat{\theta}_{j-1}$  steigende Funktion und  $Cov(w_{(2)j-1}, \hat{\theta}_{j-1})$  positiv ist. Dies bedeutet, dass der Schätzer  $\hat{\theta}_{(2),k}$  den wahren Wert  $\theta$  überschätzt.

Soll im Anschluss an ein  $k$ -stufiges Design ohne Möglichkeit des vorzeitigen Beendens der Studie eine unverzerrte Schätzung für  $\theta$  angegeben werden, wird eine Gewichtsfunktion  $w_j \in [0, 1]$  benötigt, die unkorreliert mit  $\hat{\theta}_j$  ist für  $j = 1, \dots, k-1$ . Da beim Self-Designing auch die Gewichte  $\lambda_j$ ,  $j = 1, \dots, k$ , der einzelnen Sequenzen in Abhängigkeit der in den Zwischenauswertungen vorliegenden Daten ermittelt werden, darf  $w_j$  weder von  $n_j$  noch von  $\lambda_j$  abhängen. Ein Vorschlag für eine Gewichtsfunktion ist

$$w_{(3)j} = \frac{m_j}{m_1 + \dots + m_j + m_0 + m_{j+2} + \dots + m_k} = \frac{m_j}{\sum_{i=1}^k m_i - m_{j+1} + m_0}, \quad (2.110)$$

$j = 1, \dots, k-1$ , und  $w_{(3)k} = 1 - \sum_{j=1}^{k-1} w_{(3)j}$ . Mit  $m_0 = 2/(1/n_1 + 1/n_2)$  wird das harmonische Mittel der Gruppenumfänge einer klinischen Untersuchung mit einem festem Studiendesign bezeichnet. Der resultierende Schätzer ist

$$\hat{\theta}_{(3),k} = \sum_{j=1}^k w_{(3)j} \hat{\theta}_j. \quad (2.111)$$

Basierend auf der vorhergehenden Betrachtung kann gezeigt werden, dass  $\hat{\theta}_{(3),k}$  unverzerrt für  $\theta$  ist, falls in den Zwischenauswertungen lediglich eine Anpassung der Fallzahl und kein Testen der Nullhypothese erfolgt. In diesem Fall gilt  $E(\hat{\theta}_{(3),k}) = \theta$ , da  $Cov(w_{(3)j}, \hat{\theta}_j) = 0$  für  $j = 1, \dots, k-1$  ist.

### 3 Flexible Designs in klinischen Studien mit qualitativer Zielvariable

In der Literatur erfolgt die Vorstellung von flexiblen Designs in der Regel für normalverteilte Beobachtungen. Jedoch sind in der klinischen Forschung häufig Kategorien als Ausprägungen einer Zielgröße von Interesse. In Kapitel 3.1 wird das Vorliegen einer binären Variable untersucht, welche nur zwei Werte annehmen kann. Beispielsweise soll überprüft werden, ob ein Medikament wirkt oder nicht wirkt. Eine Erweiterung auf mehrere Kategorien wird in Kapitel 3.2 betrachtet. Dies ist zum Beispiel bei der Einstufung der Verträglichkeit einer Behandlung in die Kategorien sehr gut, gut, mittel, schlecht und sehr schlecht der Fall. Ob flexible Designs in klinischen Prüfungen mit binärer und ordinaler Zielgröße Anwendung finden können, wird in diesem Kapitel intensiv untersucht. Hierbei ist insbesondere die Überprüfung von Teststatistiken von Interesse, welche bereits in nicht-sequentiellen Designs das vorgegebene Testniveau nur asymptotisch einhalten.

#### 3.1 Binäre Zielvariable

Zwei Behandlungen sollen hinsichtlich des Auftretens bzw. des Nichtauftretens eines festgelegten Ereignisses miteinander verglichen werden. In einer sequentiell durchgeführten klinischen Studie mit  $n_j = n_{1j} + n_{2j}$  Patienten in Stufe  $j$ ,  $j = 1, \dots, k$ , sei  $Y_{Ji,j}$  die binäre Zielgröße des Patienten  $i$ ,  $i = 1, \dots, n_{Jj}$ , der Behandlungsgruppe  $J$ ,  $J = 1, 2$ . Diese Variable kann nur die Werte 1 und 0 annehmen, d.h.

$$Y_{Ji,j} = \begin{cases} 1, & \text{falls das interessierende Ereignis für Patient } i, i = 1, \dots, n_{Jj}, \\ & \text{der Gruppe } J, J = 1, 2, \text{ und Stufe } j, j = 1, \dots, k, \text{ eintritt,} \\ 0, & \text{sonst.} \end{cases} \quad (3.1)$$

Die Wahrscheinlichkeit, dass das interessierende Ereignis für Patient  $i$  der Gruppe  $J$  und Stufe  $j$  eintritt sei  $p_J = P(Y_{Ji,j} = 1)$ . Um zu überprüfen, bei wie vielen Patienten in Stufe  $j$  die Ausprägung 1 aufgetreten ist, wird die Summe der Variablen  $Y_{Ji,j}$  betrachtet. Für diese gilt

$$Y_{Jj} = \sum_{i=1}^{n_{Jj}} Y_{Ji,j} \sim B(n_{Jj}, p_J), \quad J = 1, 2, \quad j = 1, \dots, k. \quad (3.2)$$

Hierbei bezeichnet  $B(n_{Jj}, p_J)$  die Binomialverteilung mit den Parametern  $n_{Jj}$  und  $p_J$ . Der Erwartungswert dieser binomialverteilten Zufallsvariablen ist  $E(Y_{Jj}) = n_{Jj}p_J$  und

die Varianz ist  $Var(Y_{Jj}) = n_{Jj}p_J(1 - p_J)$ ,  $J = 1, 2$  (vgl. Hartung et al. (1999)).

In Tabelle 3.1 ist zu sehen, wie die absoluten Häufigkeiten für das Auftreten der Ausprägungen 1 und 0 in Stufe  $j$ ,  $j = 1, \dots, k$ , in einer  $(2 \times 2)$ -Kontingenztafel, auch Vierfeldertafel genannt, festgehalten werden können (vgl. Hartung et al. (1999)).

**Tabelle 3.1:** Vierfeldertafel der beobachteten absoluten Häufigkeiten pro Gruppe  $J$ ,  $J = 1, 2$ , in Stufe  $j$ ,  $j = 1, \dots, k$ , einer klinischen Studie mit binärer Zielvariable

Gruppe	1	0	Umfang
1	$n_{11,j}$	$n_{12,j}$	$n_{11,j} + n_{12,j} = n_{1j}$
2	$n_{21,j}$	$n_{22,j}$	$n_{21,j} + n_{22,j} = n_{2j}$
	$n_{11,j} + n_{21,j}$	$n_{12,j} + n_{22,j}$	$n_j$

Die Randsummen  $n_{1j}$  und  $n_{2j}$  entsprechen den Stichprobenumfängen der beiden Behandlungsgruppen in Sequenz  $j$ , welche in einem festen Verhältnis  $r = n_{2j}/n_{1j}$  zueinander stehen. Somit ist  $n_{1j} + n_{2j} = (1 + r)n_{1j} = n_j$  und bei gleicher Aufteilung der Patienten auf beide Gruppen gilt  $n_{1j} = n_{2j} = n_j/2$ . Die unbekannte Ereigniswahrscheinlichkeit  $p_J$  der Behandlung  $J$ ,  $J = 1, 2$ , kann in der Sequenz  $j$ ,  $j = 1, \dots, k$ , mittels dem Schätzer

$$\hat{p}_{Jj} = Y_{Jj}/n_{Jj} \quad (3.3)$$

erwartungstreu geschätzt werden, da gilt  $E(\hat{p}_{Jj}) = n_{Jj}p_J/n_{Jj} = p_J$ . Die Varianz von  $\hat{p}_{Jj}$ ,  $J = 1, 2$ ,  $j = 1, \dots, k$ , ist

$$Var(\hat{p}_{Jj}) = Var(Y_{Jj})/n_{Jj}^2 = p_J(1 - p_J)/n_{Jj}. \quad (3.4)$$

Im Folgenden wird die Risikodifferenz  $D$  vorgestellt, welche für die Messung des Behandlungsunterschieds bei binären Beobachtungen verwendet werden kann. Der in Kapitel 2 mit  $\theta$  bezeichnete Parameter wird in diesem Zusammenhang mit  $D$  benannt.

### 3.1.1 Die Risikodifferenz

Die Risikodifferenz  $D = p_1 - p_2$  ist die Differenz der Wahrscheinlichkeiten für das interessierende Ereignis unter den beiden Therapien mit  $-1 \leq D \leq 1$  (vgl. Schumacher und Schulgen (2002)). Diese wird in Stufe  $j = 1, \dots, k$  durch

$$\hat{D}_j = \hat{p}_{1j} - \hat{p}_{2j} = \frac{Y_{1j}}{n_{1j}} - \frac{Y_{2j}}{n_{2j}} \quad (3.5)$$

erwartungstreu geschätzt, da  $E(\hat{D}_j) = n_{1j}p_1/n_{1j} - n_{2j}p_2/n_{2j} = D$  ist. Die Varianz von  $\hat{D}_j$  ist

$$\text{Var}(\hat{D}_j) = \text{Var}(\hat{p}_{1j}) + \text{Var}(\hat{p}_{2j}) = \frac{p_1(1-p_1)}{n_{1j}} + \frac{p_2(1-p_2)}{n_{2j}} \quad (3.6)$$

und für  $n_{1j} \geq 2$  und  $n_{2j} \geq 2$  ist ein unverzerrter Schätzer für  $\text{Var}(\hat{D}_j)$  gegeben durch

$$\widehat{\text{Var}}(\hat{D}_j) = \frac{\hat{p}_{1j}(1-\hat{p}_{1j})}{n_{1j}-1} + \frac{\hat{p}_{2j}(1-\hat{p}_{2j})}{n_{2j}-1} = \sum_{J=1}^2 \frac{1}{n_{Jj}-1} \frac{Y_{Jj}}{n_{Jj}} \frac{(n_{Jj}-Y_{Jj})}{n_{Jj}} \quad (3.7)$$

(vgl. Hartung und Knapp (2001)). Ist in beiden Behandlungsgruppen die Anzahl der Ereignisse  $Y_{Jj}$  entweder Null oder gleich dem Gruppenumfang  $n_{Jj}$ , ist die Varianzschätzung von  $\hat{D}_j$  gerade Null. Um dieses Problem zu umgehen, wird im Folgenden der modifizierte Schätzer

$$\widehat{\text{Var}}^*(\hat{D}_j) = \sum_{J=1}^2 \widehat{\text{Var}}^*(\hat{p}_{Jj}) = \sum_{J=1}^2 \frac{1}{n_{Jj}-1} \frac{Y_{Jj} + \varepsilon}{n_{Jj} + 2\varepsilon} \frac{(n_{Jj} - Y_{Jj}) + \varepsilon}{n_{Jj} + 2\varepsilon}, \quad \varepsilon > 0, \quad (3.8)$$

verwendet, welcher die Nicht-Negativität des Varianzschätzers gewährleistet und die Symmetrie der Varianz um den Punkt  $p_J = 0.5$ ,  $J = 1, 2$ , erhält (vgl. Hartung und Knapp (2004)). Für den Parameter  $\varepsilon$  wird in dieser Arbeit der Wert  $1/16$  gewählt.

Um zu überprüfen, ob die Erfolgswahrscheinlichkeit in Gruppe 1 größer als in Gruppe 2 ist, wird für die Risikodifferenz das einseitige Testproblem

$$H_0 : D = p_1 - p_2 = 0 \quad \text{versus} \quad H_1 : D > 0 \quad (3.9)$$

formuliert. Im Folgenden werden Prüfgrößen vorgestellt, welche asymptotisch äquivalent zu der  $\chi^2$ -Statistik in  $(2 \times 2)$ -Kontingenztafeln sind. Eine solche Teststatistik ist durch

$$T_{(1),j} = \frac{\hat{D}_j}{\sqrt{\widehat{\text{Var}}^*(\hat{D}_j)}} = \frac{\left( \frac{Y_{1j}}{n_{1j}} - \frac{Y_{2j}}{n_{2j}} \right)}{\sqrt{\sum_{J=1}^2 \frac{1}{n_{Jj}-1} \frac{Y_{Jj} + \varepsilon}{n_{Jj} + 2\varepsilon} \frac{(n_{Jj} - Y_{Jj}) + \varepsilon}{n_{Jj} + 2\varepsilon}}}. \quad (3.10)$$

gegeben. Mittels der Ungleichung von Jensen (siehe Mood et al. (1974))

$$E \left( \sqrt{\widehat{\text{Var}}^*(\hat{D}_j)} \right) \leq \sqrt{\text{Var}(\hat{D}_j)} \quad (3.11)$$

zeigt sich, dass der Standardfehler von  $\hat{D}_j$  systematisch unterschätzt wird. Aufgrund dessen ist die Erhöhung des Nenners der Prüfgröße von Interesse. Wird  $\widehat{\text{Var}}^*(\hat{p}_{Jj})$  mit dem

Faktor  $(n_{Jj} - 1)/(n_{Jj} - 2)$ ,  $J = 1, 2$ , multipliziert, kann die Varianz von  $\hat{D}_j = \hat{p}_{1j} - \hat{p}_{2j}$  alternativ geschätzt werden durch (vgl. Hartung und Knapp (2004))

$$\widehat{Var}^{**}(\hat{D}_j) = \sum_{J=1}^2 \frac{1}{n_{Jj} - 2} \frac{Y_{Jj} + \varepsilon}{n_{Jj} + 2\varepsilon} \frac{(n_{Jj} - Y_{Jj}) + \varepsilon}{n_{Jj} + 2\varepsilon}. \quad (3.12)$$

Somit ist eine weitere Teststatistik gegeben durch

$$T_{(2),j} = \frac{\hat{D}_j}{\sqrt{\widehat{Var}^{**}(\hat{D}_j)}} = \frac{\left(\frac{Y_{1j}}{n_{1j}} - \frac{Y_{2j}}{n_{2j}}\right)}{\sqrt{\sum_{J=1}^2 \frac{1}{n_{Jj} - 2} \frac{Y_{Jj} + \varepsilon}{n_{Jj} + 2\varepsilon} \frac{(n_{Jj} - Y_{Jj}) + \varepsilon}{n_{Jj} + 2\varepsilon}}}. \quad (3.13)$$

Da die vorhergehenden Teststatistiken das Niveau nicht immer einhalten, kann insbesondere für kleine Stichprobenumfänge eine Stetigkeitskorrektur durchgeführt werden. Bei Anwendung der Yates-Korrektur,  $n_j/(2 n_{1j}n_{2j})$ , resultiert die Prüfgröße (vgl. Hartung et al. (1999))

$$T_{(3),j} = \frac{\hat{D}_j \pm \frac{n_j}{2 n_{1j}n_{2j}}}{\sqrt{\widehat{Var}^*(\hat{D}_j)}} = \frac{\left(\frac{Y_{1j}}{n_{1j}} - \frac{Y_{2j}}{n_{2j}}\right) \pm \frac{n_j}{2 n_{1j}n_{2j}}}{\sqrt{\sum_{J=1}^2 \frac{1}{n_{Jj} - 1} \frac{Y_{Jj} + \varepsilon}{n_{Jj} + 2\varepsilon} \frac{(n_{Jj} - Y_{Jj}) + \varepsilon}{n_{Jj} + 2\varepsilon}}}. \quad (3.14)$$

Falls der von Schouten et al. (1980) vorgeschlagene Korrekturterm,  $\min(n_{1j}, n_{2j})/(2 n_{1j}n_{2j})$ , Berücksichtigung findet, ist die zugehörige Statistik durch

$$T_{(4),j} = \frac{\hat{D}_j \pm \frac{\min(n_{1j}, n_{2j})}{2 n_{1j}n_{2j}}}{\sqrt{\widehat{Var}^*(\hat{D}_j)}} = \frac{\left(\frac{Y_{1j}}{n_{1j}} - \frac{Y_{2j}}{n_{2j}}\right) \pm \frac{\min(n_{1j}, n_{2j})}{2 n_{1j}n_{2j}}}{\sqrt{\sum_{J=1}^2 \frac{1}{n_{Jj} - 1} \frac{Y_{Jj} + \varepsilon}{n_{Jj} + 2\varepsilon} \frac{(n_{Jj} - Y_{Jj}) + \varepsilon}{n_{Jj} + 2\varepsilon}}}. \quad (3.15)$$

gegeben. Diese besitzt ein weniger konservatives Testverhalten als die Yates-korrigierte Prüfgröße. Für die Überprüfung des obigen einseitigen Testproblems wird der Korrekturterm subtrahiert. Soll ein zweiseitiges  $(1 - 2\alpha)$ -Konfidenzintervall für den unbekannt Parameter bestimmt werden, erfolgt bei der Berechnung der unteren bzw. oberen Konfidenzgrenze die Subtraktion bzw. Addition des Korrekturterms.

Die Statistiken  $T_{(i),j}$ ,  $i = 1, \dots, 4$ , sind unter  $H_0$  bei großen Stichprobenumfängen  $n_{1j}$  und  $n_{2j}$  approximativ standardnormalverteilt und (streng) monoton steigend in  $\hat{D}_j$ ,  $j = 1, \dots, k$ . Als Funktion von  $D$  ist  $T_{(i),j}(D)$ ,  $i = 1, \dots, 4$ , (streng) monoton fallend in  $D$ . Alle

vier vorgestellten Teststatistiken werden in Kapitel 3.1.4 in Simulationsstudien von klinischen Prüfungen mit flexiblem Design bei Betrachtung einer binären Zielgröße untersucht.

### 3.1.2 Fallzahlplanung

Die Herleitung der Fallzahl für eine binäre Zielvariable wird zunächst anhand eines einstufigen nicht-sequentiellen Designs illustriert. Anschließend werden die vorgestellten Stichprobenformeln auf eine in mehreren Sequenzen durchgeführte klinische Studie übertragen.

#### Stichprobenfunktionen in einem nicht-sequentiellen Design

Für die Herleitung einer approximativen Formel für den Stichprobenumfang wird eine normalverteilte Zufallsvariable  $Z$  betrachtet, welche unter der Nullhypothese den Erwartungswert  $E_{H_0}(Z) = \mu_0^*$  und die Varianz  $Var_{H_0}(Z) = \sigma_0^{*2}$  und unter der Alternativhypothese den Erwartungswert  $E_{H_1}(Z) = \mu_1^*$  und die Varianz  $Var_{H_1}(Z) = \sigma_1^{*2}$  besitzt (siehe Bock (1998)). Somit gilt

$$U^* = \frac{Z - \mu_0^*}{\sigma_0} \stackrel{H_0}{\sim} N(0, 1) \quad (3.16)$$

und

$$V^* = \frac{Z - \mu_1^*}{\sigma_1} \stackrel{H_1}{\sim} N(0, 1). \quad (3.17)$$

Es besteht die Beziehung

$$U^* = V^* \frac{\sigma_1^*}{\sigma_0^*} + \frac{\mu_1^* - \mu_0^*}{\sigma_0^*}. \quad (3.18)$$

Die Nullhypothese  $H_0 : \mu_0^* = \mu_1^*$  wird zugunsten der Alternative  $H_1 : \mu_1^* > \mu_0^*$  zum Niveau  $\alpha$  verworfen, falls  $U^* > z_{1-\alpha}$  ist, wobei  $z_{1-\alpha}$  das  $(1 - \alpha)$ -Quantil der Standardnormalverteilung bezeichnet. Die Wahrscheinlichkeit der Ablehnung der Nullhypothese, falls die Alternativhypothese zutrifft, wird als Power bezeichnet. Für einen Test mit Power  $1 - \beta$  bedeutet dies

$$P_{H_1}(U^* > z_{1-\alpha}) = 1 - \beta \quad (3.19)$$

$$\Leftrightarrow P_{H_1} \left( \frac{V^* \sigma_1^* + \mu_1^* - \mu_0^*}{\sigma_0^*} > z_{1-\alpha} \right) = 1 - \beta \Leftrightarrow 1 - P_{H_1} \left( V^* \leq z_{1-\alpha} \frac{\sigma_0^*}{\sigma_1^*} - \frac{\mu_1^* - \mu_0^*}{\sigma_1^*} \right) = 1 - \beta$$

$$\Leftrightarrow \Phi \left( z_{1-\alpha} \frac{\sigma_0^*}{\sigma_1^*} - \frac{\mu_1^* - \mu_0^*}{\sigma_1^*} \right) = \beta \Leftrightarrow z_{1-\alpha} \frac{\sigma_0^*}{\sigma_1^*} - \frac{\mu_1^* - \mu_0^*}{\sigma_1^*} = -z_{1-\beta},$$

mit  $-z_{1-\beta} = z_\beta$  dem  $\beta$ -Quantil der Standardnormalverteilung. Die hieraus zur Bestimmung des erforderlichen Stichprobenumfangs resultierende Grundgleichung ist

$$z_{1-\alpha} \sigma_0^* + z_{1-\beta} \sigma_1^* = \mu_1^* - \mu_0^*. \quad (3.20)$$

Hierbei muss bekannt sein, auf welche Art  $\mu_1^* - \mu_0^*$ ,  $\sigma_0^*$  und  $\sigma_1^*$  von der Fallzahl  $n$  abhängen. Durch verschiedene Spezifikationen der Standardabweichungen ergeben sich unterschiedliche Stichprobenformeln.

Im Folgenden wird die Herleitung einer Fallzahlformel für die Risikodifferenz  $D$  vorgestellt basierend auf der Annahme unterschiedlicher Varianzen von  $D$  unter der Null- und Alternativhypothese. Seien  $Z = \hat{p}_1 - \hat{p}_2$ ,  $\mu_0^* = 0$  und  $\mu_1^* = p_1 - p_2 = D$ . Die Varianz von  $Z$  wird unter der Alternativhypothese festgelegt als

$$\sigma_1^{*2} = \frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}$$

(vgl. Formel (3.6)). Unter der Nullhypothese wird die Varianz gewählt als

$$\sigma_0^{*2} = \bar{p}(1-\bar{p}) \left( \frac{1}{n_1} + \frac{1}{n_2} \right),$$

mit  $\bar{p} = (n_1 p_1 + n_2 p_2) / (n_1 + n_2)$  (vgl. Bock (1998)). Für ein festes Verhältnis  $r = n_2 / n_1$  der Umfänge der beiden Gruppen gilt

$$\bar{p} = \frac{p_1 + r p_2}{r + 1}.$$

Werden die oben festgelegten Größen in die Grundgleichung (3.20) eingesetzt und nach dem Gesamtumfang  $n = n_1 + n_2 = (r + 1)n_1$  aufgelöst

$$\begin{aligned} & z_{1-\alpha} \sqrt{\bar{p}(1-\bar{p}) \left( \frac{1}{n_1} + \frac{1}{n_2} \right)} + z_{1-\beta} \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}} = D \\ \Leftrightarrow & z_{1-\alpha} \sqrt{\bar{p}(1-\bar{p})(r+1)} + z_{1-\beta} \sqrt{r p_1(1-p_1) + p_2(1-p_2)} = D \sqrt{r n_1} \\ \Leftrightarrow & n_1 = \frac{1}{r} \frac{\left( z_{1-\alpha} \sqrt{\bar{p}(1-\bar{p})(r+1)} + z_{1-\beta} \sqrt{r p_1(1-p_1) + p_2(1-p_2)} \right)^2}{D^2} \\ \Rightarrow & n = \frac{r+1}{r} \frac{\left( z_{1-\alpha} \sqrt{\bar{p}(1-\bar{p})(r+1)} + z_{1-\beta} \sqrt{r p_1(1-p_1) + p_2(1-p_2)} \right)^2}{D^2}, \end{aligned}$$

resultiert die Stichprobenfunktion

$$f_D(\alpha, \beta)_{bin(1)} = n = \frac{r+1}{r} \frac{\max\left\{0, z_{1-\alpha} \sqrt{\bar{p}(1-\bar{p})(r+1)} + z_{1-\beta} \sqrt{r p_1(1-p_1) + p_2(1-p_2)}\right\}^2}{D^2}. \quad (3.21)$$

Bei Betrachtung gleicher Varianzen von  $Z$  sowohl unter der Null- als auch unter der Alternativhypothese werden die Größen der Grundgleichung (3.20) zur Fallzahlplanung gewählt als  $Z = \hat{p}_1 - \hat{p}_2$ ,  $\mu_0^* = 0$ ,  $\mu_1^* = p_1 - p_2 = D$  und die Varianz von  $Z$  sei gemäß (3.6) unter  $H_0$  und  $H_1$

$$\sigma_0^{*2} = \sigma_1^{*2} = \frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}.$$

Mit  $n_2 = r n_1$  ist die Stichprobenfunktion gegeben durch

$$f_D(\alpha, \beta)_{bin(2)} = n = \frac{r+1}{r} \frac{\max\{0, z_{1-\alpha} + z_{1-\beta}\}^2 (r p_1(1-p_1) + p_2(1-p_2))}{D^2}. \quad (3.22)$$

In Tabelle 3.2 sind mit beiden vorgestellten Stichprobenfunktionen berechnete Fallzahlen angegeben, zum einen für eine feste Risikodifferenz  $D = 0.2$  und zum anderen für eine feste Ereigniswahrscheinlichkeit  $p_1 = 0.7$ .

**Tabelle 3.2:** Bestimmung der Fallzahl einer klinischen Studie mit binärer Zielvariable mit  $f_D(\alpha, \beta)_{bin(1)}$  und  $f_D(\alpha, \beta)_{bin(2)}$  für eine feste Risikodifferenz  $D = 0.2$  bzw. eine feste Ereigniswahrscheinlichkeit  $p_1 = 0.7$  (Fehlerraten I. und II. Art  $\alpha = 0.025$  und  $\beta = 0.1$ )

$D$	$p_1$	$p_2$	$f_D(\alpha, \beta)_{bin(1)}$	$f_D(\alpha, \beta)_{bin(2)}$	$p_1$	$p_2$	$D$	$f_D(\alpha, \beta)_{bin(1)}$	$f_D(\alpha, \beta)_{bin(2)}$
0.2	0.3	0.1	164	158	0.7	0.1	0.6	24	18
	0.4	0.2	218	212		0.2	0.5	38	32
	0.5	0.3	248	242		0.3	0.4	62	56
	0.6	0.4	260	254		0.4	0.3	112	106
	0.7	0.5	248	242		0.5	0.2	248	242
0.8	0.6	218	212	0.6	0.1	954	946		

Bei Verwendung der Fallzahlformel  $f_D(\alpha, \beta)_{bin(2)}$  resultieren geringere Umfänge als bei Berechnung des Umfangs mit  $f_D(\alpha, \beta)_{bin(1)}$ .

## Übertragung der Stichprobenfunktionen auf sequentielle Designs

In einer klinischen Studie mit adaptivem Design können die Umfänge der einzelnen Sequenzen basierend auf den vorliegenden Informationen ermittelt werden. Unter Einbeziehung von Schätzungen der unbekanntenen Ereigniswahrscheinlichkeiten  $p_1$  und  $p_2$  in die Fallzahlkalkulation, können die Annahmen für diese Größen überprüft und zu kleine oder zu große Stichprobenumfänge der gesamten Studie vermieden werden.

In Stufe  $j$ ,  $j = 1, \dots, k$ , kann die unbekanntene Ereigniswahrscheinlichkeit  $p_J$  der Gruppe  $J = 1, 2$  geschätzt werden durch ein gewichtetes Mittel der Schätzungen  $\hat{p}_{Jj}$  der einzelnen Sequenzen. Eine Möglichkeit ist die Verwendung des Schätzers (2.109) aus Kapitel 2.5.2, welcher für die Ereigniswahrscheinlichkeit  $p_J$ ,  $J = 1, 2$ , gegeben ist durch

$$\hat{p}_{Jj}^* = \sum_{i=1}^j m_i \hat{p}_{Ji} / \sum_{i=1}^j m_i, \quad (3.23)$$

mit  $m_i = 2/(1/n_{1i} + 1/n_{2i})$ ,  $i = 1, \dots, j$ . Hiermit kann die Risikodifferenz basierend auf allen vorhandenen Informationen mittels

$$\hat{D}_j^* = \hat{p}_{1j}^* - \hat{p}_{2j}^* \quad (3.24)$$

geschätzt werden. Mit diesen Schätzern resultiert die Stichprobenfunktion

$$f_{D,j}(\alpha, \beta)_{bin(1)} = n_{j+1} = \frac{r+1}{r} \max \left\{ 0, z_{1-\alpha} \sqrt{\hat{p}_j^*(1-\hat{p}_j^*)(r+1)} \right. \\ \left. + z_{1-\beta} \sqrt{r \hat{p}_{1j}^*(1-\hat{p}_{1j}^*) + \hat{p}_{2j}^*(1-\hat{p}_{2j}^*)} \right\}^2 / \hat{D}_j^{*2}, \quad (3.25)$$

$j = 0, \dots, k-1$ , mit

$$\hat{p}_j^* = \frac{\hat{p}_{1j}^* + r \hat{p}_{2j}^*}{r+1}. \quad (3.26)$$

Für  $j = 0$  entspricht  $f_{D,0}(\alpha, \beta)_{bin(1)}$  gerade der Stichprobenfunktion  $f_D(\alpha, \beta)_{bin(1)}$  eines einstufigen nicht-sequentiellen Studiendesigns, wobei  $D_0 > 0$  einen auf Vorwissen basierenden Wert für die unbekanntene Risikodifferenz  $D$  bezeichnet. Um den Wertebereich des Schätzers  $\hat{D}_j^*$  auf die Alternativhypothese zu restringieren, wird  $\hat{D}_j^* = \rho$ ,  $\rho > 0$ , gewählt, falls Werte kleiner als  $\rho$  auftreten. In dieser Arbeit wird  $\rho$  auf den Wert 0.001 gesetzt. Die Übertragung der Fallzahl  $f_D(\alpha, \beta)_{bin(2)}$  auf den sequentiellen Fall kann analog erfolgen.

Mit der sequentiellen Variante der ausgewählten Stichprobenfunktion kann der Umfang der Stufe  $(j + 1)$ ,  $j = 1, \dots, k - 1$ , mittels der in Kapitel 2.4.1 beschriebenen Vorgehensweise bestimmt werden. Ziel der auf der bedingten Power basierenden Fallzahlplanung ist, dass unter Berücksichtigung der Informationen der Sequenzen  $1, \dots, j$  die Ablehnung der Nullhypothese in der folgenden Stufe  $(j + 1)$  mit einer bestimmten Wahrscheinlichkeit ermöglicht werden soll.

### 3.1.3 Beispiel

Das folgende Beispiel soll klinische Prüfungen mit flexiblem Studiendesign veranschaulichen. Das adaptiv gruppensequentielle Design vom O'Brien/Fleming-Typ sowie das Self-Designing werden betrachtet, wobei zur Kombination der Ergebnisse der einzelnen Sequenzen sowohl die inverse Normal- als auch die inverse  $\chi^2$ -Methode verwendet werden. Insbesondere die Konstruktion der Repeated Konfidenzintervalle wird veranschaulicht und Unterschiede zwischen den beiden Kombinationsmethoden sowie den verschiedenen Designs werden aufgezeigt. Bei dem O'Brien/Fleming-Design und bei dem Self-Designing mit inverser  $\chi^2$ -Methode sind die Repeated Konfidenzintervalle ineinander verschachtelt, d.h.  $K_{j+1} \subset K_j$ ,  $j = 1, \dots, k - 1$  (vgl. Kapitel 2.5.1). Dies ist in den anderen vorgestellten Designs nicht unbedingt der Fall.

In einer klinischen Studie werden ein neuartiges Medikament (1) und eine Standardtherapie (2) mit gleicher Patientenzahl in beiden Gruppen untersucht. Primär von Interesse ist, ob die Prüfmedikation „wirkt“ oder „nicht wirkt“, d.h. die primäre Zielvariable ist binär. Die Messung des Behandlungsunterschieds erfolgt mittels der Risikodifferenz  $D = p_1 - p_2$ , wobei  $p_1$  und  $p_2$  die Erfolgswahrscheinlichkeiten in den beiden Gruppen sind. Für die Standardtherapie ist bekannt, dass mit einer Wahrscheinlichkeit von 50% ( $p_2 = 0.5$ ) ein Behandlungserfolg eintritt und von der neuen Therapie wird erwartet, dass die Erfolgswahrscheinlichkeit 70% ( $p_1 = 0.7$ ) beträgt. Folglich wird angenommen, dass der Unterschied zwischen den beiden Therapien  $D = 0.7 - 0.5 = 0.2$  ist. Das einseitige Testproblem

$$H_0 : D = 0 \quad \text{vs.} \quad H_1 : D > 0$$

wird in den Sequenzen  $j$ ,  $j = 1, \dots, k$ , mittels der Teststatistik  $T_{(2),j}$  überprüft.

Als Studiendesigns werden das O'Brien/Fleming-Design mit  $K = 5$  Sequenzen, Gewichten  $\lambda_1 = \dots = \lambda_5 = 1$  und kritischen Werten  $cv_{Z, OBF}(0.025, 5, \{1\}) = 4.5617$  bzw.  $cv_{S, OBF}(0.025, 5, \{1\}) = \chi^2(5)_{0.975} = 12.8325$  sowie das Self-Designing mit globalem Gewicht  $\lambda_G = 10$ , Gewicht der ersten Stufen  $\lambda_1 = 2$ , minimalem Stufengewicht  $\lambda_{\min} = 1$  und kritischen Werten  $cv_Z = 1.9600$  bzw.  $cv_S = \chi^2(10)_{0.975} = 20.4832$  gewählt. In Stufe  $j = 1, \dots, 5$  wird die Fallzahl ermittelt als  $n_j = \varepsilon_j M_j(\beta)$  mit  $\varepsilon_j = m_j(\beta_g)/M_j(\beta)$  und  $\beta_g = 0.3$ . In einem nicht-sequentiellen Design resultiert mittels der Stichprobenfunktion  $f_D(\alpha, \beta)_{bin(1)}$  mit  $\alpha = 0.025$  und  $\beta = 0.9$  der Umfang  $n = 248$ . Für Gruppe 1 und 2 wurden jeweils insgesamt  $n_J = 200$ ,  $J = 1, 2$ , Beobachtungen aus  $B(n_J, p_J)$ -Verteilungen erzeugt. Von diesen werden in Stufe  $j$ ,  $j = 1, \dots, k$ , der verschiedenen Designs  $n_{Jj}$  Werte pro Behandlungsgruppe sequentiell ausgewählt.

Im Folgenden werden die Abläufe bei Verwendung des O'Brien/Fleming-Designs mit inverser Normalmethode beschrieben. Die Ergebnisse der anderen Designs finden sich in Tabelle 3.3.

In der ersten Stufe werden  $n_1 = 248/K \approx 50$  Patienten behandelt. Der Anteil der Responder, d.h. der Patienten für die ein Behandlungserfolg eingetreten ist, ist  $\hat{p}_{11} = 17/25 = 0.680$  in Gruppe 1 und  $\hat{p}_{21} = 12/25 = 0.480$  in Gruppe 2. Mit dem Schätzer für die Risikodifferenz  $\hat{D}_{(2),j} = (\sum_{i=1}^j n_i \hat{D}_i) / (\sum_{i=1}^j n_i)$  aus (2.109) aus Kapitel 2.5.2 resultiert in der ersten Sequenz  $\hat{D}_{(2),1} = 0.200$ . Der beobachtete Wert der Teststatistik ist  $\hat{T}_{(2),1} = 1.403$  mit zugehörigem p-Wert  $p_1 = 1 - \Phi(\hat{T}_{(2),1}) = 0.080$ . Da  $\hat{Z}_1 = 1.403 < cv_{Z, OBF}(0.025, 5, \{1\})$  ist, wird eine weitere Sequenz durchgeführt.

Die bedingte Fehlerwahrscheinlichkeit der Stufe 2 wird geschätzt als  $\hat{p}_2 = 0.057$  und mit  $m_2(0.3) = 108$  und  $M_2(0.1) = 196$  resultiert  $n_2 = \varepsilon_1 M_2(0.1) = 108$ . Die relativen Häufigkeiten in Stufe 2 sind  $\hat{p}_{12} = 17/54 = 0.315$  und  $\hat{p}_{22} = 15/54 = 0.278$ . Somit sind  $\hat{T}_{(2),2} = 0.414$ ,  $p_2 = 0.340$  und  $\hat{Z}_2 = 1.817$ . Basierend auf den Informationen von beiden Sequenzen resultiert als Schätzung für die Risikodifferenz  $\hat{D}_{(2),2} = 0.089$ . Da der kritische Wert in Stufe 2 nicht überschritten wird, erfolgt die Planung einer weiteren Sequenz.

Mit dem bedingten p-Wert  $\hat{p}_3 = 0.056$  der Stufe 3 ergeben sich  $m_3(0.3) = 124$  und  $M_3(0.1) = 248$ . Mit somit  $n_3 = 124$  Patienten resultieren  $\hat{p}_{13} = 52/62 = 0.839$  und  $\hat{p}_{23} = 28/62 = 0.452$ . Der Schätzwert für die Risikodifferenz ist  $\hat{D}_{(2),3} = 0.220$ . Es sind  $\hat{T}_{(2),3} = 4.842$  und  $p_3 < 0.001$ . Der Wert der Kombinationsstatistik  $\hat{Z}_3 = 6.659$  ist größer als der kritische Wert und somit kann  $H_0$  abgelehnt und die Studie vorzeitig beendet werden.

**Tabelle 3.3:** Ergebnisse des fiktiven Beispiels einer klinischen Prüfung mit binärer Zielvariable unter Verwendung des O'Brien/Fleming-Designs und des Self-Designings mit inverser Normal- sowie  $\chi^2$ -Methode

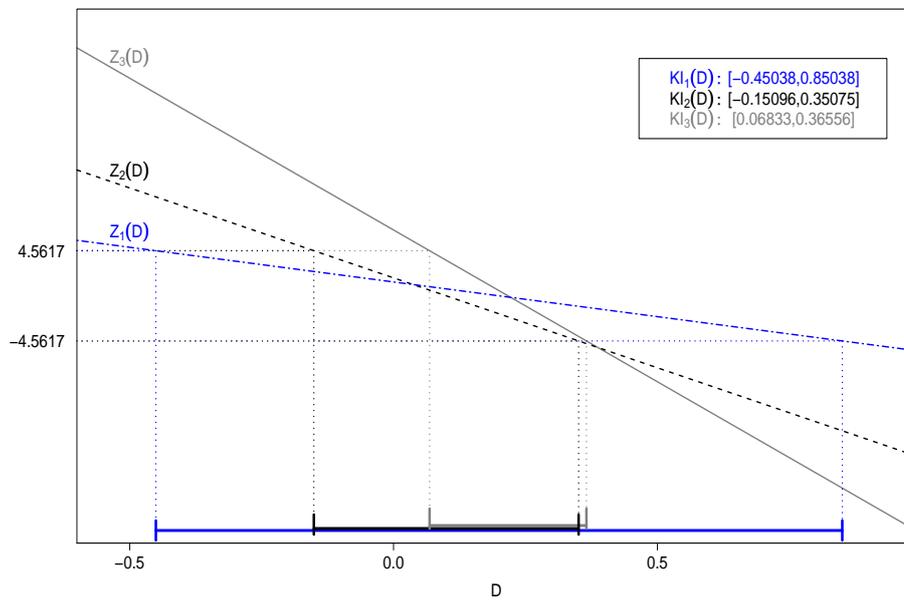
Inverse Normalmethode				Inverse $\chi^2$ -Methode				
O'Brien/Fleming-Design								
$j$	1	2	3	$j$	1	2	3	
$n_j$	50	108	124	$n_j$	50	120	124	
$n_{11,j}$	17	17	52	$n_{11,j}$	17	23	46	
$n_{21,j}$	12	15	28	$n_{21,j}$	12	21	22	
$\hat{p}_{1j}$	0.680	0.315	0.839	$\hat{p}_{1j}$	0.680	0.383	0.742	
$\hat{p}_{2j}$	0.480	0.278	0.452	$\hat{p}_{2j}$	0.480	0.350	0.355	
$T_{(2),j}$	1.403	0.414	4.842	$T_{(2),j}$	1.403	0.373	4.623	
$p_j$	0.080	0.340	<0.001	$p_j$	0.080	0.355	<0.001	
$z_j$	1.403	0.414	4.842	$q_j$	3.058	0.857	22.700	
$Z_j$	1.403	1.817	6.659	$S_j$	3.058	3.915	26.615	
$\hat{p}_j$		0.057	0.056	$\hat{p}_j$		0.044	0.030	
$m_j(\beta_g)$		108	124	$m_j(\beta_g)$		120	124	
$M_j(\beta)$		196	248	$M_j(\beta)$		214	248	
$\hat{D}_{(2),j}$	0.200	0.089	0.220	$\hat{D}_{(2),j}$	0.200	0.082	0.211	
$\hat{D}_{L,j}$	-0.450	-0.151	0.068	$\hat{D}_{L,j}$	-0.284	-0.152	0.118	
$\hat{D}_{U,j}$	0.850	0.351	0.366	$\hat{D}_{U,j}$	0.684	0.316	0.315	
Self-Designing								
$j$	1	2	3	4	$j$	1	2	3
$n_j$	50	98	124	6	$n_j$	50	114	124
$n_{11,j}$	17	12	57	0	$n_{11,j}$	17	20	49
$n_{21,j}$	12	10	33	0	$n_{21,j}$	12	18	25
$\hat{p}_{1j}$	0.680	0.245	0.919	0.000	$\hat{p}_{1j}$	0.680	0.351	0.790
$\hat{p}_{2j}$	0.480	0.204	0.532	0.000	$\hat{p}_{2j}$	0.480	0.316	0.403
$\lambda_j$	2.000	4.261	1.870	1.870	$\lambda_j$	2.000	4.471	1.765
$T_{(2),j}$	1.403	0.474	5.269	0.000	$T_{(2),j}$	1.403	0.390	4.702
$p_j$	0.080	0.318	<0.001	0.500	$p_j$	0.080	0.348	<0.001
$z_j$	1.403	0.474	5.269	0.000	$q_j$	5.043	4.991	26.338
$Z_j/\sqrt{\lambda_{\Sigma}(j)}$	1.403	1.184	3.566	3.215	$S_j$	5.043	10.034	36.371
$\hat{p}_j$		0.068	0.047	0.998	$\hat{p}_j$		0.051	0.024
$m_j(\beta_g)$		98	124	6	$m_j(\beta_g)$		112.344	124
$M_j(\beta)$		184	248	6	$M_j(\beta)$		203.311	248
$\hat{D}_{(2),j}$	0.200	0.095	0.228	0.223	$\hat{D}_{(2),j}$	0.200	0.085	0.215
$\hat{D}_{L,j}$	-0.425	-0.095	0.076	0.067	$\hat{D}_{L,j}$	-0.366	-0.131	0.104
$\hat{D}_{U,j}$	0.825	0.270	0.312	0.275	$\hat{D}_{U,j}$	0.766	0.293	0.291

Das adaptiv gruppensequentielle Design vom O'Brien/Fleming-Typ mit inverser Normal- und  $\chi^2$ -Methode sowie das Self-Designing mit inverser  $\chi^2$ -Methode enden nach Durchführung von drei Sequenzen mit der Ablehnung der Nullhypothese. Die resultierenden Gesamtumfänge sind  $n = \sum_{j=1}^3 n_j = 282$ ,  $n = 294$  und  $n = 288$ . Bei dem Self-Designing mit inverser Normalmethode kann die Überprüfung des Testproblems unter Einhaltung des globalen Niveaus erst erfolgen, wenn das gesamte globale Gewicht  $\lambda_G$  verbraucht ist, d.h.  $\lambda_{\Sigma}(j) = \lambda_G$ . Da dies in der dritten Stufe nicht erfüllt ist, muss eine weitere Sequenz durchgeführt werden. Nach Beobachtung von sechs weiteren Patienten in der vierten Sequenz ist mit  $\lambda_{\Sigma}(4) = \lambda_G$  das Studienende erreicht und die Nullhypothese kann abgelehnt werden, da  $Z_4/\sqrt{\lambda_{\Sigma}(4)} > \Phi^{-1}(1 - \alpha)$  ist. Der Gesamtumfang ist mit  $n = 278$  im Vergleich zu den anderen Designs am geringsten. Die Schätzung von  $D$  mittels dem Schätzer  $\widehat{D}_{(2),j}$  fällt in der zweiten Stufe wesentlich geringer aus als der vor Studienbeginn vermutete Wert  $D = 0.2$ . Somit resultiert für die dritte Sequenz bei allen Designs ein hoher Stichprobenumfang, welcher durch den maximalen Umfang beschränkt wird.

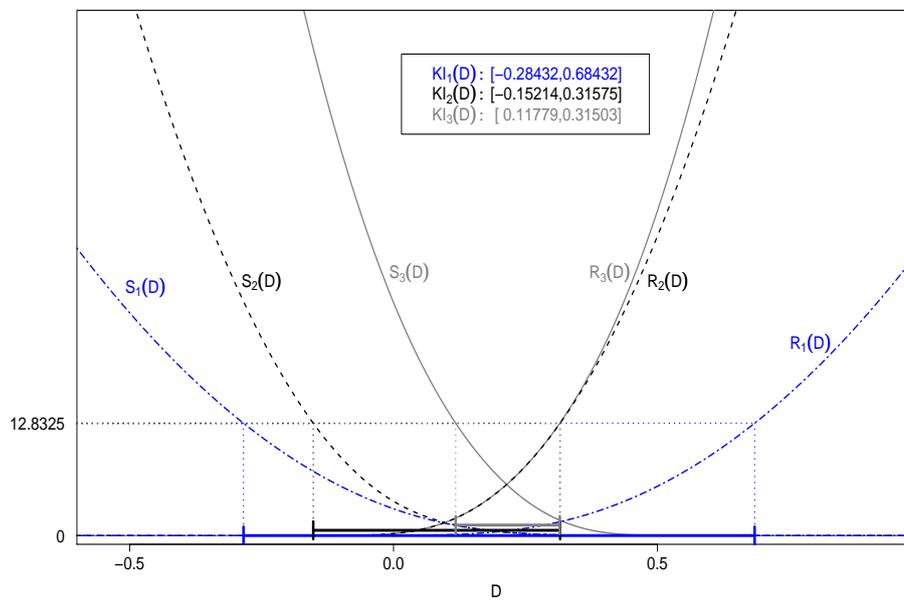
Mit diesem Beispiel wird eine extreme Situation veranschaulicht, da in einer Sequenz stark abweichende Ergebnisse resultieren. Dies hat Auswirkungen auf die Gesamtanzahl zu beobachtender Patienten, da für die Planung der weiteren Stufen diese Information einbezogen wird. Das mittlere Verhalten der in Kapitel 2 vorgestellten flexiblen Designs bei Betrachtung einer binären Zielvariable wird in Kapitel 3.1.4 in Simulationen untersucht.

Anhand des O'Brien/Fleming-Designs mit inverser Normal- sowie  $\chi^2$ -Methode wird die Konstruktion der Repeated 95%-Konfidenzintervalle für  $D$  in den drei beobachteten Sequenzen veranschaulicht. In Übereinstimmung mit der Testentscheidung enthalten die Intervalle der dritten Stufe nicht die Null. Bei dem O'Brien/Fleming-Design mit inverser  $\chi^2$ -Methode sind die Konfidenzintervalle ineinander verschachtelt, was in Abbildung 3.2 zu sehen ist. Für flexible Designs mit der inversen Normalmethode ist dies nicht unbedingt erfüllt. In diesem Beispiel ist das Intervall der dritten Stufe nicht komplett in dem der vorherigen Sequenz enthalten (vgl. Abbildung 3.1).

**Abbildung 3.1:** Repeated 95%–Konfidenzintervalle für die Risikodifferenz im O’Brien/Fleming–Design mit inverser Normalmethode



**Abbildung 3.2:** Repeated 95%–Konfidenzintervalle für die Risikodifferenz im O’Brien/Fleming–Design mit inverser  $\chi^2$ –Methode



### 3.1.4 Simulationsstudien

Ziel der folgenden Simulationen ist die Untersuchung der adaptiv gruppensequentiellen Designs vom Pocock-, O'Brien/Fleming-, Wang/Tsiatis- und Haybittle/Peto-Typ sowie des Self-Designing klinischer Studien bei Vorliegen einer binären Zielvariablen. Sowohl die inverse Normal- als auch die inverse  $\chi^2$ -Methode werden für die Kombination der Ergebnisse der einzelnen Stufen betrachtet. Die Überprüfung des einseitigen Testproblems  $H_0 : D = 0$  vs.  $H_1 : D > 0$  erfolgt in den Sequenzen  $j$ ,  $j = 1, \dots, k$ , mit den in Kapitel 3.1.1 vorgestellten Prüfgrößen  $T_{(1),j}$ ,  $T_{(2),j}$ ,  $T_{(3),j}$  und  $T_{(4),j}$ . Die verschiedenen Designs werden sowohl unter der Null- als auch unter der Alternativhypothese untersucht. Von Interesse sind die empirische Ablehnwahrscheinlichkeit, der mittlere sowie mediane Stufenumfang, die durchschnittlich benötigte Stufenanzahl und die empirische Überdeckungswahrscheinlichkeit der Sequenz der realisierten Konfidenzintervalle sowie des Intervalls der letzten Stufe des flexiblen Verfahrens.

Für die Planung der Studie werden zwei Szenarien betrachtet. In Szenario I wird angenommen, dass die Risikodifferenz  $D = 0.2$  ist, wobei die Ereigniswahrscheinlichkeiten in den beiden Gruppen als  $p_1 = 0.7$  und  $p_2 = 0.5$  vermutet werden. In einem nicht-sequentiellen Design würde mit den Fehlerwahrscheinlichkeiten I. und II. Art  $\alpha = 0.025$  und  $\beta = 0.1$  unter Verwendung der Fallzahlformel (3.1.2) der Umfang  $n_{fest} = f_D(0.025, 0.1)_{bin(1)} = 248$  resultieren. In Szenario II wird die Risikodifferenz als  $D = 0.1$  mit  $p_1 = 0.9$  und  $p_2 = 0.8$  vermutet. Dies würde zu einem Umfang  $n_{fest} = f_D(0.025, 0.1)_{bin(1)} = 532$  führen. Die für  $p_1$  und  $p_2$  angenommenen Werte für die Planung der Studie sind hypothetisch und können sich als falsch herausstellen. Als tatsächlich unter der Null- und Alternativhypothese vorliegende Ereignisraten  $p_1$  und  $p_2$  werden betrachtet

Unter  $H_0$  : 1a)  $p_1 = 0.2, p_2 = 0.2, D = 0.0$ ,

1b)  $p_1 = 0.5, p_2 = 0.5, D = 0.0$ ,

1c)  $p_1 = 0.8, p_2 = 0.8, D = 0.0$ ,

Unter  $H_1$  : 2a)  $p_1 = 0.3, p_2 = 0.2, D = 0.1$  ( $n_{fest} = 784$ ),

2b)  $p_1 = 0.9, p_2 = 0.8, D = 0.1$  ( $n_{fest} = 532$ ),

2c)  $p_1 = 0.7, p_2 = 0.5, D = 0.2$  ( $n_{fest} = 248$ ),

2d)  $p_1 = 0.3, p_2 = 0.1, D = 0.2$  ( $n_{fest} = 164$ ),

2e)  $p_1 = 0.6, p_2 = 0.3, D = 0.3$  ( $n_{fest} = 112$ ),

2f)  $p_1 = 0.9, p_2 = 0.6, D = 0.3$  ( $n_{fest} = 84$ ).

Unter der Alternativhypothese werden für Szenario I eine Unter- sowie Überschätzung der Risikodifferenz bei gleichzeitiger Fehleinschätzung der Größe der Ereigniswahrscheinlichkeiten der beiden Gruppen untersucht (Fälle 2a und 2b sowie 2e und 2f). Für Szenario II wird in den Fällen 2c, 2d, 2e und 2f betrachtet, dass die wahre Risikodifferenz größer als angenommen ist bei meist falschen Annahmen für  $p_1$  und  $p_2$ . In Fall 2d bei Szenario I und in Fall 2a bei Szenario II ist der für die Risikodifferenz vermutete Wert richtig, jedoch sind die für  $p_1$  und  $p_2$  angenommenen Werte falsch. Zudem wird der Fall betrachtet, dass die vor Studienbeginn getroffenen Annahmen zutreffen (Fall 2c in Szenario I und Fall 2b in Szenario II). Die in einem nicht-sequentiellen Design notwendigen Fallzahlen sind in Klammern angegeben.

Als maximale Stufenanzahl wird bei den adaptiv gruppensequentiellen Designs  $K = 3$  sowie  $K = 5$  gewählt. Bei dem Self-Designing wird das globale Gewicht festgelegt als  $\lambda_G = 5$  und  $\lambda_G = 10$ . Die für die Berechnung des Umfangs  $m_j(\beta_g)$ ,  $j = 1, \dots, k$ , notwendige Fehlerwahrscheinlichkeit  $\beta_g$  wird auf die Werte 0.3 und 0.8 gesetzt. Der minimale Stufenumfang wird als  $n_{j,\min} = 6$  gewählt, da hiermit alle vier Teststatistiken berechenbar sind.

In Szenario I wird für die erste Stufe der Umfang  $n_1 = n_{fest}/K = 248/3 \approx 84$  beziehungsweise  $n_1 = 248/5 \approx 50$  betrachtet. Als maximale Werte für die Fallzahlen  $m_j(\beta_g)$  und  $M_j(\beta)$  werden im Fall von drei Sequenzen  $m_{j,\max} = 2/5 n_{fest} \approx 100$  oder  $m_{j,\max} = 2/3 n_{fest} \approx 166$  und  $M_{j,\max} = n_{fest}/2 \approx 124$  oder  $M_{j,\max} = n_{fest} = 248$  und im Fall von fünf Sequenzen  $m_{j,\max} = n_{fest}/4 \approx 62$  oder  $m_{j,\max} = n_{fest}/2 \approx 124$  und  $M_{j,\max} = n_{fest}/3 \approx 84$  oder  $M_{j,\max} = n_{fest} = 248$  verwendet. In Szenario II wird für die erste Stufe der Umfang  $n_1 = n_{fest}/K = 532/3 \approx 178$  beziehungsweise  $n_1 = 532/5 \approx 108$  betrachtet. Als maximale Werte für die Fallzahlen  $m_j(\beta_g)$  und  $M_j(\beta)$  werden im Fall von drei Sequenzen  $m_{j,\max} = 2/5 n_{fest} \approx 214$  oder  $m_{j,\max} = 2/3 n_{fest} \approx 356$  und  $M_{j,\max} = n_{fest}/2 \approx 266$  oder  $M_{j,\max} = n_{fest} = 532$  und im Fall von fünf Sequenzen  $m_{j,\max} = n_{fest}/4 \approx 134$  oder  $m_{j,\max} = n_{fest}/2 \approx 266$  und  $M_{j,\max} = n_{fest}/3 \approx 178$  oder  $M_{j,\max} = n_{fest} = 532$  verwendet. Insgesamt werden die in Tabelle 3.4 angegebenen Konstellationen der Parameter für Szenario I und II in Simulationen mit 10000 Wiederholungen betrachtet.

**Tabelle 3.4:** Einstellungen der Parameter in den Simulationsstudien flexibler Designs in klinischen Prüfungen mit binärer Zielvariable (Fehlerraten I. und II. Art  $\alpha = 0.025$  und  $\beta = 0.1$ , minimaler Stufenumfang  $n_{j,\min} = 6$ )

Konstellation	$\beta_g$	$K$	$\lambda_i$	$\lambda_G$	$\lambda_1$	$\lambda_{\min}$	Szenario I			Szenario II		
							$n_1$	$m_{j,\max}$	$M_{j,\max}$	$n_1$	$m_{j,\max}$	$M_{j,\max}$
I	0.3	5	1	10	2	1	50	62	84	108	134	178
II	0.3	5	1	10	2	1	50	124	248	108	266	532
III	0.8	5	1	10	2	1	50	62	84	108	134	178
IV	0.8	5	1	10	2	1	50	124	248	108	266	532
V	0.3	3	1	5	1	1	84	100	124	178	214	266
VI	0.3	3	1	5	1	1	84	166	248	178	356	532
VII	0.8	3	1	5	1	1	84	100	124	178	214	266
VIII	0.8	3	1	5	1	1	84	166	248	178	356	532

Die Ergebnisse von Konstellation I unter der Null- sowie unter der Alternativhypothese sind für Szenario I in den Tabellen 3.5 sowie 3.7 und für Szenario II in den Tabellen 3.6 sowie 3.8 zu sehen. Die Resultate der Konstellationen II-VIII befinden sich für Szenario I in Anhang C.1, für Szenario II werden diese in dieser Arbeit nicht berichtet. In der ersten Spalte der Tabellen ist die verwendete Teststatistik (TS) verzeichnet, in der zweiten Spalte die empirische Ablehnwahrscheinlichkeit eines nicht-sequentiellen Designs (NSD), in der dritten Spalte das verwendete flexible Design (PO=Pocock, OBF=O'Brien/Fleming, WT=Wang/Tsiatis, HP=Haybittle/Peto, SD=Self-Designing) und in den folgenden Spalten die empirische Ablehnwahrscheinlichkeit, der mittlere sowie mediane Stufenumfang, die durchschnittlich benötigte Stufenanzahl und die empirische Überdeckungswahrscheinlichkeit der Sequenz der realisierten Konfidenzintervalle sowie des Intervalls der letzten Stufe des flexiblen Verfahrens für die inverse Normal- sowie für die inverse  $\chi^2$ -Methode.

**Tabelle 3.5:** Empirisches Niveau  $\hat{\alpha}\%$  im nicht-sequentiellen Design (NSD) sowie  $\hat{\alpha}\%$ , mittlerer und medianer Umfang (ASN = „Average Sample Number“,  $N_{med}$ ), mittlere Stufenanzahl  $\bar{k}$  und empirische Überdeckungswahrscheinlichkeiten  $(1 - 2\hat{\alpha})\%$  der Sequenz der beobachteten Intervalle  $\{KI_j\}$  sowie des finalen Intervalls  $KI_k$  in flexiblen Designs mit binärer Zielvariable unter der Nullhypothese (Szenario I, Konstellation I)

TS	NSD	Design	Inverse Normalmethode					Inverse $\chi^2$ -Methode						
			$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$ $\{KI_j\}$ $KI_k$	$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$ $\{KI_j\}$ $KI_k$		
1a) $p_1 = p_2 = 0.2, D = 0.0$														
$T_{(1)}$	2.54	PO	2.64	315.08	320	4.93	94.39	96.47	2.89	314.72	320	4.93	93.95	95.89
		OBF	2.41	316.35	320	4.98	94.88	95.41	2.70	315.66	320	4.98	94.58	94.58
		WT	2.60	316.24	320	4.97	94.75	95.75	2.80	315.75	320	4.96	94.45	94.91
		HP	2.61	318.34	320	4.99	94.67	94.92	2.60	315.07	320	4.99	94.70	94.73
		SD	2.57	195.07	196	3.05	93.77	94.78	2.74	194.75	196	3.03	94.49	94.49
$T_{(2)}$	2.40	PO	2.36	315.61	320	4.94	94.96	96.83	2.47	315.56	320	4.94	94.84	96.41
		OBF	2.18	316.46	320	4.99	95.45	95.93	2.46	316.02	320	4.98	95.18	95.18
		WT	2.21	316.48	320	4.97	95.40	96.38	2.41	316.26	320	4.97	95.12	95.49
		HP	2.54	318.68	320	4.99	95.13	95.33	2.35	315.38	320	4.99	95.32	95.33
		SD	2.58	195.20	196	3.05	94.20	95.05	2.53	195.00	196	3.03	94.91	94.91
$T_{(3)}$	1.58	PO	0.61	318.50	320	4.98	98.61	99.25	0.78	318.28	320	4.98	98.47	99.04
		OBF	0.28	318.42	320	5.00	99.31	99.43	0.53	318.18	320	4.99	98.92	98.92
		WT	0.45	318.41	320	4.99	99.09	99.30	0.65	318.40	320	4.99	98.68	98.79
		HP	0.38	319.48	320	5.00	99.21	99.25	0.57	317.40	320	5.00	98.84	98.85
		SD	0.51	196.23	196	3.03	98.24	98.78	0.65	195.54	196	3.02	98.46	98.46
$T_{(4)}$	2.00	PO	1.38	317.27	320	4.96	96.94	98.26	1.41	317.31	320	4.96	96.84	97.99
		OBF	0.95	317.69	320	4.99	97.63	97.85	1.30	317.47	320	4.99	96.71	96.71
		WT	1.13	317.76	320	4.98	97.42	98.03	1.36	317.54	320	4.98	96.69	97.09
		HP	1.06	319.26	320	5.00	97.37	97.52	1.23	316.43	320	4.99	96.78	96.80
		SD	1.13	195.93	196	3.04	95.97	96.78	1.43	195.40	196	3.02	96.32	96.32
1b) $p_1 = p_2 = 0.5, D = 0.0$														
$T_{(1)}$	2.39	PO	2.47	315.63	320	4.94	94.64	96.66	3.04	314.55	320	4.92	93.79	95.60
		OBF	2.58	316.02	320	4.98	94.69	95.17	3.09	315.21	320	4.97	93.84	93.84
		WT	2.49	316.16	320	4.97	94.70	95.78	3.05	315.22	320	4.96	93.74	94.26
		HP	2.83	318.09	320	4.98	94.32	94.71	3.04	314.75	320	4.98	93.94	94.02
		SD	2.65	195.21	196	3.05	93.67	94.68	2.80	194.75	196	3.03	94.20	94.20
$T_{(2)}$	2.39	PO	2.27	315.92	320	4.94	95.24	96.98	2.66	315.16	320	4.93	94.45	96.08
		OBF	2.31	316.34	320	4.98	95.17	95.74	2.68	315.58	320	4.98	94.50	94.50
		WT	2.36	316.39	320	4.97	95.11	96.06	2.75	315.74	320	4.96	94.24	94.76
		HP	2.41	318.54	320	4.99	95.01	95.27	2.62	315.37	320	4.99	94.74	94.76
		SD	2.59	195.36	196	3.05	94.13	94.96	2.64	195.03	196	3.03	94.45	94.45
$T_{(3)}$	1.64	PO	0.91	317.90	320	4.97	97.99	98.92	1.37	317.03	320	4.96	97.29	98.26
		OBF	0.50	318.12	320	5.00	98.81	99.00	0.97	317.72	320	4.99	97.89	97.89
		WT	0.79	318.03	320	4.99	98.45	98.82	1.18	317.68	320	4.98	97.63	97.93
		HP	0.65	319.26	320	4.99	98.57	98.84	1.03	316.86	320	4.99	97.88	97.90
		SD	0.70	195.87	196	3.03	97.62	98.32	0.95	195.27	196	3.02	97.95	97.95
$T_{(4)}$	2.39	PO	1.71	316.51	320	4.95	96.48	97.83	1.97	316.25	320	4.95	95.94	97.17
		OBF	1.26	317.20	320	4.99	96.81	97.21	1.72	316.50	320	4.98	95.86	95.86
		WT	1.43	317.18	320	4.98	96.65	97.41	1.79	316.83	320	4.98	95.86	96.30
		HP	1.48	318.85	320	4.99	96.63	96.88	1.85	316.00	320	4.99	95.88	95.88
		SD	1.48	195.96	196	3.05	95.67	96.50	1.87	195.57	196	3.03	95.66	95.66

TS	NSD	Design	Inverse Normalmethode						Inverse $\chi^2$ -Methode					
			$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$ { $KI_j$ } $KI_k$		$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$ { $KI_j$ } $KI_k$	
1c) $p_1 = p_2 = 0.8, D = 0.0$														
$T_{(1)}$	2.60	PO	2.77	314.91	320	4.93	94.73	96.39	3.02	314.28	320	4.92	94.01	95.89
		OBF	2.60	316.24	320	4.98	94.85	95.14	3.08	315.51	320	4.98	94.24	94.24
		WT	2.67	315.99	320	4.97	94.57	95.44	3.01	315.52	320	4.96	94.16	94.67
		HP	2.85	318.15	320	4.98	94.35	94.65	2.71	314.84	320	4.99	94.53	94.55
		SD	2.58	195.04	196	3.05	94.04	95.11	2.78	194.82	196	3.03	94.63	94.63
$T_{(2)}$	2.48	PO	2.47	315.42	320	4.94	95.13	96.73	2.59	315.02	320	4.93	94.89	96.47
		OBF	2.60	316.24	320	4.98	94.98	95.31	2.58	316.04	320	4.98	95.05	95.05
		WT	2.44	316.51	320	4.97	95.13	95.96	2.75	315.92	320	4.96	94.86	95.29
		HP	2.50	318.46	320	4.99	95.14	95.28	2.27	315.23	320	4.99	95.25	95.25
		SD	2.50	195.19	196	3.05	94.31	95.25	2.51	195.09	196	3.03	95.11	95.11
$T_{(3)}$	1.47	PO	0.71	318.33	320	4.98	98.63	99.20	0.75	318.28	320	4.98	98.42	99.06
		OBF	0.33	318.47	320	5.00	99.27	99.42	0.47	318.17	320	5.00	98.82	98.82
		WT	0.50	318.48	320	4.99	99.00	99.32	0.59	318.34	320	4.99	98.65	98.83
		HP	0.28	319.73	320	5.00	99.24	99.35	0.49	317.53	320	5.00	98.88	98.89
		SD	0.49	196.17	196	3.03	98.35	98.85	0.60	195.54	196	3.02	98.60	98.60
$T_{(4)}$	2.01	PO	1.66	316.67	320	4.96	96.73	98.04	1.57	316.74	320	4.96	96.54	97.74
		OBF	1.11	317.66	320	4.99	97.26	97.53	1.16	317.26	320	4.99	96.64	96.64
		WT	1.24	317.70	320	4.99	97.32	97.91	1.31	317.32	320	4.98	96.72	97.12
		HP	0.91	319.25	320	4.99	97.54	97.67	1.28	316.31	320	4.99	96.78	96.78
		SD	1.19	195.94	196	3.04	96.22	97.00	1.31	195.47	196	3.02	96.51	96.51

**Tabelle 3.6:** Empirisches Niveau  $\hat{\alpha}\%$  im nicht-sequentiellen Design (NSD) sowie  $\hat{\alpha}\%$ , mittlerer und medianer Umfang (ASN = „Average Sample Number“,  $N_{med}$ ), mittlere Stufenanzahl  $\bar{k}$  und empirische Überdeckungswahrscheinlichkeiten  $(1 - 2\hat{\alpha})\%$  der Sequenz der beobachteten Intervalle  $\{KI_j\}$  sowie des finalen Intervalls  $KI_k$  in flexiblen Designs mit binärer Zielvariable unter der Nullhypothese (Szenario II, Konstellation I)

TS	NSD	Design	Inverse Normalmethode						Inverse $\chi^2$ -Methode					
			$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$ { $KI_j$ } $KI_k$		$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$ { $KI_j$ } $KI_k$	
1a) $p_1 = p_2 = 0.2, D = 0.0$														
$T_{(1)}$	2.54	PO	2.46	678.56	688	4.94	95.01	96.79	2.68	677.36	688	4.93	94.60	96.15
		OBF	2.40	679.70	688	4.98	95.02	95.53	2.48	679.08	688	4.98	94.99	94.99
		WT	2.26	680.46	688	4.97	95.30	96.15	2.54	679.62	688	4.97	94.81	95.31
		HP	2.51	684.69	688	4.99	94.98	95.14	2.55	677.01	688	4.99	94.88	94.92
		SD	2.50	417.62	420	3.04	93.95	94.92	2.70	416.74	420	3.02	94.60	94.60
$T_{(2)}$	2.40	PO	2.25	679.66	688	4.95	95.38	97.07	2.37	678.92	688	4.94	95.09	96.55
		OBF	2.43	680.28	688	4.98	95.20	95.67	2.49	678.96	688	4.98	95.14	95.14
		WT	2.12	680.72	688	4.97	95.50	96.38	2.50	679.51	688	4.97	94.93	95.44
		HP	2.41	684.91	688	4.99	95.12	95.29	2.35	677.52	688	4.99	95.29	95.33
		SD	2.44	417.82	420	3.04	94.17	95.06	2.53	417.29	420	3.03	94.89	94.89
$T_{(3)}$	1.58	PO	0.95	683.37	688	4.97	98.09	98.90	1.07	683.30	688	4.97	97.80	98.56
		OBF	0.61	683.68	688	5.00	98.62	98.83	0.87	683.47	688	4.99	98.11	98.11
		WT	0.72	683.99	688	4.99	98.45	98.84	0.98	683.44	688	4.99	97.93	98.10
		HP	0.82	686.39	688	4.99	98.38	98.45	0.87	681.72	688	4.99	98.14	98.15
		SD	0.79	418.62	420	3.02	97.50	98.20	0.99	418.34	420	3.02	97.97	97.97

TS	NSD	Design	Inverse Normalmethode						Inverse $\chi^2$ -Methode					
			$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$		$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$	
	$\hat{\alpha}\%$						$\{KI_j\}$	$KI_k$					$\{KI_j\}$	$KI_k$
$T_{(4)}$	2.00	PO	1.57	681.53	688	4.96	96.76	97.97	1.71	680.96	688	4.96	96.41	97.56
		OBF	1.32	682.24	688	4.99	96.97	97.36	1.48	681.49	688	4.99	96.30	96.30
		WT	1.41	682.18	688	4.98	96.96	97.60	1.58	682.06	688	4.98	96.52	96.92
		HP	1.43	686.07	688	4.99	96.82	96.93	1.57	679.33	688	4.99	96.27	96.33
		SD	1.38	418.45	420	3.03	95.75	96.58	1.54	417.63	420	3.02	96.33	96.33
1b) $p_1 = p_2 = 0.5, D = 0.0$														
$T_{(1)}$	2.39	PO	2.84	677.08	688	4.93	94.84	96.37	3.07	675.77	688	4.92	93.86	95.77
		OBF	2.76	679.39	688	4.98	94.52	94.83	2.67	678.02	688	4.98	94.54	94.54
		WT	2.65	679.80	688	4.97	94.69	95.56	2.71	678.24	688	4.96	94.53	94.85
		HP	2.84	684.32	688	4.99	94.62	94.84	2.83	676.37	688	4.98	94.41	94.43
		SD	2.64	417.53	420	3.04	93.85	94.74	2.59	416.94	420	3.02	94.49	94.49
$T_{(2)}$	2.39	PO	2.70	677.68	688	4.93	95.03	96.54	2.59	677.89	688	4.94	94.96	96.34
		OBF	2.52	679.89	688	4.98	94.82	95.15	2.49	678.40	688	4.98	94.84	94.84
		WT	2.48	680.28	688	4.97	95.00	95.75	2.52	678.93	688	4.96	94.97	95.28
		HP	2.66	684.49	688	4.99	94.86	95.06	2.67	676.76	688	4.99	94.73	94.74
		SD	2.50	417.50	420	3.04	94.08	94.92	2.55	417.07	420	3.02	94.74	94.74
$T_{(3)}$	1.64	PO	1.21	682.68	688	4.97	97.59	98.45	1.64	680.59	688	4.95	96.71	97.85
		OBF	1.05	683.02	688	4.99	97.89	98.08	1.15	682.18	688	4.99	97.54	97.54
		WT	1.14	683.23	688	4.99	97.81	98.23	1.23	682.55	688	4.98	97.29	97.57
		HP	1.16	686.06	688	4.99	97.86	97.96	1.26	680.53	688	4.99	97.37	97.39
		SD	1.05	418.68	420	3.03	97.03	97.62	1.29	417.99	420	3.02	97.37	97.37
$T_{(4)}$	2.39	PO	2.09	679.59	688	4.95	96.04	97.36	1.97	679.57	688	4.95	95.70	97.10
		OBF	1.76	681.55	688	4.99	96.29	96.65	1.71	680.30	688	4.99	95.78	95.78
		WT	1.70	681.91	688	4.98	96.41	97.06	1.70	680.91	688	4.98	96.03	96.30
		HP	1.91	685.12	688	4.99	96.08	96.21	1.90	678.62	688	4.99	95.64	95.65
		SD	1.65	418.32	420	3.03	95.41	96.10	1.89	417.93	420	3.02	95.79	95.79
1c) $p_1 = p_2 = 0.8, D = 0.0$														
$T_{(1)}$	2.60	PO	2.60	677.81	688	4.93	94.78	96.63	2.67	677.26	688	4.93	94.69	96.35
		OBF	2.59	680.04	688	4.98	94.92	95.40	2.68	678.42	688	4.98	94.74	94.74
		WT	2.45	680.26	688	4.97	95.13	96.02	2.82	678.66	688	4.96	94.48	94.94
		HP	2.52	684.37	688	4.99	95.02	95.21	2.45	677.22	688	4.99	94.88	94.88
		SD	2.61	417.34	420	3.03	94.02	94.90	2.81	417.09	420	3.02	94.68	94.68
$T_{(2)}$	2.48	PO	2.44	678.58	688	4.94	95.20	96.85	2.47	678.23	688	4.94	94.95	96.57
		OBF	2.38	680.12	688	4.98	95.16	95.63	2.62	678.93	688	4.98	94.93	94.93
		WT	2.31	680.59	688	4.97	95.42	96.31	2.63	678.91	688	4.96	94.90	95.40
		HP	2.34	684.55	688	4.99	95.20	95.40	2.15	677.55	688	4.99	95.46	95.47
		SD	2.51	417.45	420	3.03	94.23	95.18	2.74	417.24	420	3.02	94.88	94.88
$T_{(3)}$	1.47	PO	1.08	683.18	688	4.97	97.89	98.77	1.17	682.60	688	4.97	97.58	98.49
		OBF	0.63	683.67	688	5.00	98.63	98.78	0.87	683.13	688	4.99	98.08	98.08
		WT	0.79	683.98	688	4.99	98.36	98.77	1.00	683.24	688	4.99	97.94	98.15
		HP	0.70	686.78	688	5.00	98.60	98.66	0.88	681.26	688	5.00	98.24	98.25
		SD	0.76	418.70	420	3.02	97.63	98.25	1.00	418.33	420	3.01	97.89	97.89
$T_{(4)}$	2.01	PO	1.70	680.81	688	4.96	96.68	97.93	1.82	680.16	688	4.95	96.29	97.48
		OBF	1.42	681.77	688	4.99	96.77	97.16	1.58	681.09	688	4.99	96.16	96.16
		WT	1.45	682.17	688	4.98	96.77	97.45	1.74	680.99	688	4.98	96.16	96.65
		HP	1.43	685.72	688	4.99	96.84	96.95	1.58	679.65	688	5.00	96.48	96.51
		SD	1.56	418.21	420	3.03	95.61	96.55	1.74	418.04	420	3.02	96.11	96.11

Für die einzelnen Designs resultieren für die inverse Normal- sowie für die inverse  $\chi^2$ -Methode vergleichbare Ergebnisse, wobei das empirische Niveau bei Verwendung der inversen Normalmethode tendenziell geringer als für die inverse  $\chi^2$ -Methode ausfällt.

Erfolgt die Überprüfung des Testproblems in dem nicht-sequentiellen Design mit  $T_{(1)}$ , wird das vorgegebene Niveau  $\alpha = 2.5\%$  leicht überschritten, bei Verwendung der Prüfgröße  $T_{(2)}$  ist das empirische Niveau kleiner als  $\alpha$ . Das Testverhalten der Prüfgrößen mit Stetigkeitskorrektur,  $T_{(3)}$  und  $T_{(4)}$ , ist sehr konservativ, das resultierende empirische Niveau liegt insbesondere für  $T_{(3)}$  weit unterhalb des für  $\alpha$  festgelegten Werts von  $2.5\%$ .

In den adaptiv gruppensequentiellen Designs sowie im Self-Designing zeigt sich, dass sowohl ein liberales als auch ein konservatives Testverhalten im Vergleich zu nicht-sequentiellen Designs verstärkt wird. Dies wird insbesondere im Fall der beiden konservativen Prüfgrößen  $T_{(3)}$  und  $T_{(4)}$  deutlich. Bei Verwendung von  $T_{(3)}$  ist das empirische Niveau in den meisten Fällen kleiner als  $1\%$  und für  $T_{(4)}$  geringer als  $1.5\%$  in Szenario I. In Szenario II ist eine Erhöhung dieser Werte aufgrund des größeren mittleren Stichprobenumfangs zu erkennen. Im Fall von  $T_{(1)}$  wird das vorgegebene Niveau häufig überschritten, wobei dies bei der inversen  $\chi^2$ -Methode stärker ausfällt als bei der inversen Normalmethode. Das für die Prüfgröße  $T_{(2)}$  resultierende empirische Niveau liegt für beide Szenarien meistens in der Nähe des Wertes  $2.5\%$ .

Die adaptiv gruppensequentiellen Designs enden im Mittel nach Durchführung der vorgegebenen maximalen Stufenanzahl  $K = 5$ . Für das Self-Designing werden durchschnittlich nur drei Sequenzen benötigt. Der mittlere und der mediane Gesamtumfang (ASN,  $N_{med}$ ) der Studie sind unter der Nullhypothese ungefähr gleich groß. Bei den adaptiv gruppensequentiellen Designs sind ASN und  $N_{med}$  im Vergleich zu dem Self-Designing wesentlich größer, da bei ersteren im Mittel zwei Sequenzen mehr beobachtet werden. Im Vergleich zu dem nicht-sequentiellen Design fällt der Gesamtumfang in den flexiblen Designs größer aus, da der vor Studienbeginn angenommene Unterschied  $D = 0.2$  nicht vorliegt. Aufgrund der Wahl höherer minimaler und maximaler Umfänge in den einzelnen Sequenzen bei Szenario II, sind die mittleren sowie medianen Fallzahlen größer als bei Szenario I. Die empirische Überdeckungswahrscheinlichkeit der Sequenz der Konfidenzintervalle ist unter  $H_0$  im Fall der Teststatistiken  $T_{(1)}$  und  $T_{(2)}$  in der Nähe des für das Konfidenzniveau vorgegebenen Wertes von  $95\%$ , bei Verwendung der Prüfgrößen  $T_{(3)}$  und  $T_{(4)}$  fällt diese

aufgrund des konservativen Testverhaltens dieser Statistiken wesentlich größer aus.

In den Tabellen 3.7 und 3.8 sind die Simulationsergebnisse der Parameterkonstellation I für Szenario I und II unter der Alternativhypothese zu sehen.

**Tabelle 3.7:** Empirische Power  $(1 - \hat{\beta})\%$  im nicht-sequentiellen Design (NSD) sowie  $(1 - \hat{\beta})\%$ , mittlerer und medianer Umfang (ASN = „Average Sample Number“,  $N_{med}$ ), mittlere Stufenanzahl  $\bar{k}$  und empirische Überdeckungswahrscheinlichkeiten  $(1 - 2\hat{\alpha})\%$  der Sequenz der beobachteten Intervalle  $\{KI_j\}$  sowie des finalen Intervalls  $KI_k$  in flexiblen Designs mit binärer Zielvariable unter der Alternativhypothese (Szenario I, Konstellation I)

TS	NSD	Design	Inverse Normalmethode						Inverse $\chi^2$ -Methode					
			$(1 - \hat{\beta})\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$	$\{KI_j\}$	$KI_k$	$(1 - \hat{\beta})\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$
2a) $p_1 = 0.3, p_2 = 0.2, D = 0.1$														
$T_{(1)}$	44.87	PO	43.72	264.21	320	4.23	95.99	97.91	37.62	269.68	320	4.30	95.28	96.77
		OBF	49.56	270.32	320	4.62	97.06	97.52	44.77	271.00	320	4.64	96.20	96.20
		WT	48.92	266.96	320	4.41	96.86	97.75	44.14	270.03	320	4.46	95.97	96.38
		HP	52.89	285.52	320	4.69	96.73	96.91	44.71	274.93	320	4.82	95.95	95.99
		SD	32.26	181.64	196	3.27	93.85	94.94	32.56	180.61	196	3.06	96.16	96.16
$T_{(2)}$	44.61	PO	42.08	267.00	320	4.27	96.48	98.24	35.78	273.00	320	4.35	95.57	97.07
		OBF	48.06	272.30	320	4.64	97.34	97.77	42.68	274.13	320	4.68	96.60	96.60
		WT	47.57	269.39	320	4.44	97.01	97.80	41.74	273.55	320	4.51	96.34	96.63
		HP	51.57	287.41	320	4.72	97.06	97.18	42.09	277.59	320	4.85	96.49	96.54
		SD	31.64	182.82	196	3.27	95.11	95.89	31.15	182.15	196	3.08	96.47	96.47
$T_{(3)}$	38.32	PO	21.54	292.21	320	4.64	98.81	99.41	18.60	295.15	320	4.67	98.08	98.88
		OBF	22.30	295.42	320	4.89	99.52	99.62	21.42	295.14	320	4.88	98.73	98.73
		WT	22.98	294.40	320	4.79	99.36	99.66	21.84	295.03	320	4.78	98.78	98.95
		HP	25.88	303.64	320	4.89	99.30	99.42	21.74	296.08	320	4.94	98.83	98.85
		SD	12.50	191.13	196	3.20	98.73	99.08	15.28	189.12	196	3.09	98.87	98.87
$T_{(4)}$	41.87	PO	31.89	279.22	320	4.45	97.75	98.87	27.35	283.65	320	4.50	96.74	97.97
		OBF	35.08	283.62	320	4.77	98.40	98.64	31.57	284.17	320	4.78	97.52	97.52
		WT	35.50	282.56	320	4.63	98.16	98.80	31.83	283.98	320	4.64	97.45	97.77
		HP	38.08	295.91	320	4.81	98.28	98.43	32.46	286.85	320	4.89	97.40	97.40
		SD	21.46	187.88	196	3.25	96.48	97.21	22.88	186.40	196	3.09	97.55	97.55
2b) $p_1 = 0.9, p_2 = 0.8, D = 0.1$														
$T_{(1)}$	61.24	PO	60.04	241.02	320	3.91	95.73	98.19	52.57	248.43	320	4.00	95.54	97.30
		OBF	66.34	249.33	282	4.44	96.65	97.17	59.92	252.63	302	4.53	95.96	95.96
		WT	66.17	244.97	278	4.16	96.31	97.58	59.45	250.95	312	4.27	95.62	96.14
		HP	69.60	268.60	318	4.51	96.38	96.84	60.27	257.72	292	4.78	95.67	95.74
		SD	46.29	175.16	196	3.37	94.07	94.87	45.45	174.42	196	3.09	95.83	95.83
$T_{(2)}$	59.87	PO	57.96	246.03	320	3.99	95.98	98.36	49.88	254.36	320	4.09	95.74	97.46
		OBF	64.86	251.97	284	4.49	96.84	97.30	57.01	256.48	314	4.60	96.47	96.47
		WT	64.32	248.37	288	4.21	96.68	97.78	56.86	254.72	318	4.32	96.27	96.81
		HP	68.96	272.09	320	4.57	96.54	96.96	57.87	262.35	298	4.85	96.18	96.27
		SD	45.36	176.17	196	3.36	95.09	95.72	43.45	176.53	196	3.12	96.24	96.24

TS	NSD (1 - β)%	Design	Inverse Normalmethode						Inverse χ <sup>2</sup> -Methode					
			(1 - β̂)%	ASN	N <sub>med</sub>	k̄	(1 - 2 α̂)%	{KI <sub>j</sub> }	KI <sub>k</sub>	(1 - β̂)%	ASN	N <sub>med</sub>	k̄	(1 - 2 α̂)%
T <sub>(3)</sub>	53.20	PO	28.09	285.80	320	4.57	99.05	99.74	23.56	289.87	320	4.61	98.50	99.43
		OBF	28.45	289.23	320	4.88	99.60	99.71	26.63	288.95	320	4.88	99.07	99.07
		WT	30.74	287.47	320	4.74	99.49	99.80	28.11	289.02	320	4.75	98.95	99.16
		HP	33.25	298.03	320	4.87	99.59	99.68	28.31	290.08	320	4.95	99.02	99.05
		SD	15.56	191.25	196	3.27	98.99	99.26	19.40	189.29	196	3.15	99.01	99.01
T <sub>(4)</sub>	57.41	PO	43.79	266.80	320	4.28	97.66	99.10	37.19	272.52	320	4.36	97.22	98.36
		OBF	47.27	272.70	320	4.71	98.46	98.81	41.82	273.97	320	4.75	97.83	97.83
		WT	48.34	269.48	320	4.51	98.47	99.07	42.22	273.25	320	4.58	97.69	98.02
		HP	50.85	284.91	320	4.73	98.27	98.43	42.83	276.71	320	4.89	97.82	97.85
		SD	27.17	185.51	196	3.32	97.05	97.56	30.75	183.91	196	3.14	97.63	97.63
2c) p <sub>1</sub> = 0.7, p <sub>2</sub> = 0.5, D = 0.2														
T <sub>(1)</sub>	89.91	PO	91.53	170.17	174	2.90	96.03	97.77	87.39	176.34	174	2.97	95.69	97.11
		OBF	92.22	183.69	174	3.79	96.38	96.87	90.18	183.18	174	3.78	95.33	95.33
		WT	93.49	177.10	174	3.29	97.00	97.94	90.41	181.58	174	3.38	96.24	96.57
		HP	94.61	205.16	236	3.62	96.76	96.95	91.22	190.69	190	4.12	94.42	94.50
		SD	74.35	149.65	148	3.46	91.40	92.27	76.72	145.86	148	2.92	94.90	94.90
T <sub>(2)</sub>	89.63	PO	90.83	173.80	174	2.96	96.43	98.08	86.29	180.57	174	3.03	96.00	97.18
		OBF	91.75	186.37	174	3.85	96.67	97.16	88.98	188.48	174	3.91	95.74	95.74
		WT	92.80	180.14	174	3.35	97.47	98.30	89.58	186.36	174	3.47	96.71	97.00
		HP	94.28	209.60	236	3.70	97.57	97.74	89.67	196.87	196	4.28	94.59	94.63
		SD	74.60	151.06	152	3.49	93.73	94.31	75.71	148.54	160	2.97	95.77	95.77
T <sub>(3)</sub>	87.47	PO	78.19	207.29	218	3.48	98.75	99.29	72.94	211.84	236	3.51	98.03	98.63
		OBF	75.02	221.72	225	4.37	98.90	99.06	73.69	220.60	234	4.32	97.20	97.20
		WT	79.80	213.32	216	3.89	99.23	99.55	76.07	218.44	234	3.97	98.25	98.48
		HP	82.29	234.69	242	4.14	99.23	99.31	76.58	225.18	236	4.53	97.06	97.08
		SD	50.23	167.42	196	3.47	97.49	97.97	58.24	164.08	196	3.10	97.90	97.90
T <sub>(4)</sub>	88.99	PO	85.99	187.70	174	3.17	97.42	98.52	80.98	195.73	174	3.27	96.63	97.71
		OBF	85.59	201.90	192	4.08	97.65	97.93	82.91	202.02	196	4.07	96.41	96.41
		WT	88.18	194.14	174	3.57	98.21	98.69	83.85	199.93	190	3.69	97.26	97.50
		HP	89.21	221.49	236	3.90	98.19	98.31	84.53	209.34	216	4.39	95.22	95.26
		SD	63.56	159.22	185	3.48	94.83	95.48	68.40	155.52	194	3.01	96.57	96.57
2d) p <sub>1</sub> = 0.3, p <sub>2</sub> = 0.1, D = 0.2														
T <sub>(1)</sub>	98.45	PO	98.82	129.37	112	2.30	95.79	98.20	97.63	134.26	112	2.36	95.45	97.30
		OBF	97.76	144.37	134	3.43	96.25	96.87	97.38	143.35	124	3.45	95.25	95.25
		WT	99.09	137.60	116	2.78	96.66	97.95	98.40	140.46	118	2.87	96.15	96.66
		HP	99.28	163.73	174	2.94	96.70	96.91	98.16	152.37	144	3.92	94.20	94.41
		SD	88.68	129.08	118	3.58	92.81	93.69	91.02	124.36	112	2.84	94.84	94.84
T <sub>(2)</sub>	98.31	PO	98.65	131.64	112	2.34	96.25	98.41	97.25	139.30	112	2.45	96.29	97.86
		OBF	97.24	147.14	136	3.54	96.68	97.24	96.80	148.74	132	3.61	95.56	95.56
		WT	98.96	140.11	130	2.83	96.98	98.17	97.94	144.32	126	2.96	96.79	97.18
		HP	99.41	169.57	174	3.06	97.08	97.32	97.86	158.55	150	4.09	95.16	95.33
		SD	89.23	130.72	118	3.62	94.70	95.39	90.56	128.26	112	2.96	95.48	95.48
T <sub>(3)</sub>	97.56	PO	93.14	167.20	164	2.93	98.83	99.54	89.57	177.85	174	3.07	98.46	99.19
		OBF	81.33	195.19	190	4.29	99.25	99.42	83.17	188.88	176	4.23	98.17	98.17
		WT	90.91	180.25	174	3.59	99.36	99.68	88.63	184.45	174	3.65	98.94	99.11
		HP	89.72	202.47	222	3.72	99.43	99.53	88.37	195.51	192	4.47	98.01	98.05
		SD	57.63	157.99	166	3.61	98.59	99.03	71.15	151.72	168	3.16	98.64	98.64

TS	NSD (1 - β̂)%	Design	Inverse Normalmethode						Inverse χ <sup>2</sup> -Methode					
			(1 - β̂)%	ASN	N <sub>med</sub>	$\bar{k}$	(1 - 2 α̂)%		(1 - β̂)%	ASN	N <sub>med</sub>	$\bar{k}$	(1 - 2 α̂)%	
							{KI <sub>j</sub> }	KI <sub>k</sub>					{KI <sub>j</sub> }	KI <sub>k</sub>
T <sub>(4)</sub>	97.91	PO	96.99	146.03	112	2.57	97.48	99.02	94.71	154.88	112	2.69	97.27	98.40
		OBF	93.03	168.96	162	3.87	97.83	98.15	91.70	166.27	158	3.86	96.77	96.77
		WT	96.88	157.46	148	3.19	98.19	98.84	94.74	161.84	154	3.31	97.54	97.85
		HP	95.99	182.84	174	3.32	98.45	98.67	94.76	173.57	170	4.23	96.25	96.34
		SD	72.96	145.35	136	3.63	96.20	96.78	80.58	139.04	126	3.03	97.01	97.01
2e) p <sub>1</sub> = 0.6, p <sub>2</sub> = 0.3, D = 0.3														
T <sub>(1)</sub>	99.85	PO	99.91	100.18	106	1.85	96.63	98.01	99.73	100.14	102	1.82	96.58	97.67
		OBF	99.31	112.75	110	3.00	95.25	95.66	99.52	107.81	102	2.79	93.30	93.30
		WT	99.90	106.39	98	2.30	97.83	98.46	99.78	105.40	98	2.29	97.00	97.17
		HP	99.94	124.60	112	2.26	97.95	98.08	99.82	112.41	106	2.99	92.42	92.48
		SD	93.38	107.97	106	3.43	89.08	89.83	96.67	101.13	94	2.53	93.59	93.59
T <sub>(2)</sub>	99.85	PO	99.90	101.43	108	1.87	96.82	98.05	99.70	103.83	112	1.90	97.00	97.69
		OBF	99.17	114.81	112	3.08	96.56	97.04	99.09	111.91	110	2.97	93.68	93.68
		WT	99.85	108.64	104	2.35	98.18	98.80	99.67	108.24	106	2.37	97.35	97.52
		HP	99.96	128.72	112	2.35	98.15	98.29	99.72	118.23	112	3.31	93.30	93.37
		SD	94.10	109.87	110	3.49	91.61	92.18	96.86	103.40	102	2.62	95.15	95.15
T <sub>(3)</sub>	99.69	PO	99.45	117.80	112	2.18	98.77	99.19	98.93	118.81	112	2.14	98.33	98.85
		OBF	93.02	141.53	132	3.69	98.97	99.07	94.77	133.42	118	3.44	98.77	98.77
		WT	98.49	127.39	112	2.79	99.00	99.21	97.83	129.74	112	2.88	98.59	98.68
		HP	98.94	146.62	150	2.72	99.34	99.39	97.68	137.88	132	3.59	98.76	98.82
		SD	78.87	126.75	118	3.56	96.76	97.43	88.27	118.53	112	2.85	97.32	97.32
T <sub>(4)</sub>	99.85	PO	99.79	107.03	112	1.96	97.36	98.30	99.57	111.15	112	2.02	97.39	97.89
		OBF	97.50	124.75	113	3.32	96.28	96.60	98.00	120.40	112	3.13	94.00	94.00
		WT	99.62	114.83	112	2.47	98.57	99.03	99.02	116.90	112	2.60	97.27	97.40
		HP	99.65	136.49	112	2.51	98.58	98.69	99.21	125.74	112	3.43	92.68	92.77
		SD	87.40	116.73	118	3.49	92.79	93.63	93.84	109.50	112	2.69	95.16	95.16
2f) p <sub>1</sub> = 0.9, p <sub>2</sub> = 0.6, D = 0.3														
T <sub>(1)</sub>	99.99	PO	100.00	79.25	50	1.50	97.14	98.56	100.00	79.54	50	1.51	96.73	97.69
		OBF	99.73	91.91	82	2.74	95.57	95.99	99.84	86.06	70	2.53	93.90	93.90
		WT	99.98	86.11	76	1.98	97.77	98.48	99.99	83.08	66	1.94	96.57	96.94
		HP	100.00	99.82	108	1.85	99.05	99.43	99.94	89.45	78	2.76	92.51	92.73
		SD	95.41	92.12	84	3.38	88.94	89.98	98.64	83.46	70	2.38	93.52	93.52
T <sub>(2)</sub>	99.99	PO	100.00	80.04	50	1.52	97.34	98.66	100.00	81.23	50	1.54	96.95	97.80
		OBF	99.62	93.84	84	2.89	97.15	97.50	99.60	89.61	78	2.71	94.67	94.67
		WT	99.96	87.11	78	2.00	98.20	98.87	99.91	85.49	72	2.05	96.91	97.13
		HP	100.00	102.80	112	1.92	99.11	99.41	99.91	94.42	88	2.97	94.64	94.86
		SD	95.93	93.08	84	3.45	92.40	93.20	98.45	85.79	74	2.52	95.21	95.21
T <sub>(3)</sub>	99.97	PO	99.96	92.12	88	1.77	98.67	99.46	99.91	95.07	86	1.81	98.43	99.04
		OBF	91.40	124.08	116	3.70	99.46	99.51	94.50	110.91	106	3.33	99.24	99.24
		WT	98.82	105.05	102	2.52	99.34	99.68	98.47	104.91	104	2.54	98.83	98.96
		HP	99.57	119.47	112	2.28	99.50	99.64	97.99	115.11	112	3.48	99.18	99.22
		SD	74.81	114.94	114	3.61	97.35	98.08	89.33	102.60	100	2.81	97.64	97.64
T <sub>(4)</sub>	99.99	PO	100.00	84.92	50	1.62	98.02	99.15	99.97	86.43	50	1.64	97.59	98.28
		OBF	98.37	105.24	98	3.17	97.17	97.47	98.26	97.67	90	2.90	95.38	95.38
		WT	99.80	93.79	86	2.22	98.60	99.10	99.60	93.16	84	2.27	97.53	97.69
		HP	99.91	109.11	112	2.05	98.98	99.17	99.43	101.91	94	3.15	94.39	94.49
		SD	85.05	102.85	96	3.51	94.03	94.76	92.90	92.65	84	2.62	95.17	95.17

**Tabelle 3.8:** Empirische Power  $(1 - \hat{\beta})\%$  im nicht-sequentiellen Design (NSD) sowie  $(1 - \hat{\beta})\%$ , mittlerer und medianer Umfang (ASN = „Average Sample Number“,  $N_{med}$ ), mittlere Stufenanzahl  $\bar{k}$  und empirische Überdeckungswahrscheinlichkeiten  $(1 - 2\hat{\alpha})\%$  der Sequenz der beobachteten Intervalle  $\{KI_j\}$  sowie des finalen Intervalls  $KI_k$  in flexiblen Designs mit binärer Zielvariable unter der Alternativhypothese (Szenario II, Konstellation I)

TS	NSD	Design	Inverse Normalmethode						Inverse $\chi^2$ -Methode					
			$(1 - \hat{\beta})\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$		$(1 - \hat{\beta})\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$	
			$\{KI_j\}$						$KI_k$					
2a) $p_1 = 0.3, p_2 = 0.2, D = 0.1$														
$T_{(1)}$	44.87	PO	76.71	455.64	510	3.49	96.39	98.29	69.67	474.48	510	3.61	96.11	97.65
		OBF	81.08	475.21	492	4.17	97.36	97.71	75.93	480.44	510	4.26	96.77	96.77
		WT	81.36	464.83	482	3.80	97.10	97.90	75.56	479.83	510	3.93	96.72	97.01
		HP	83.82	523.44	520	4.19	97.13	97.34	76.91	492.37	516	4.57	96.42	96.46
		SD	58.11	354.57	420	3.37	94.18	94.95	59.05	350.96	420	3.05	96.13	96.13
$T_{(2)}$	44.61	PO	76.05	461.37	510	3.54	96.55	98.36	68.69	479.77	510	3.65	96.29	97.71
		OBF	80.33	478.92	500	4.20	97.56	97.89	74.85	486.23	510	4.31	96.87	96.87
		WT	80.92	468.55	490	3.83	97.18	98.03	74.48	483.84	510	3.96	96.87	97.16
		HP	83.64	526.74	522	4.21	97.24	97.45	76.22	497.28	516	4.61	96.69	96.72
		SD	57.95	354.95	420	3.38	94.99	95.62	58.34	352.62	420	3.08	96.63	96.63
$T_{(3)}$	38.32	PO	60.76	517.89	664	3.93	98.39	99.24	53.90	533.08	688	4.01	97.98	98.76
		OBF	62.01	537.77	579	4.56	99.23	99.39	57.15	538.62	616	4.59	98.42	98.42
		WT	64.76	525.33	584	4.24	99.15	99.51	59.22	538.38	640	4.32	98.53	98.79
		HP	66.91	567.91	624	4.50	99.04	99.13	59.74	546.68	610	4.77	98.22	98.24
		SD	37.34	378.64	420	3.33	98.11	98.61	41.50	374.79	420	3.13	98.48	98.48
$T_{(4)}$	41.87	PO	69.17	488.62	510	3.73	97.64	98.93	62.10	504.83	660	3.82	96.96	98.17
		OBF	72.52	507.14	510	4.37	98.26	98.58	66.94	512.27	538	4.45	97.52	97.52
		WT	73.77	495.31	510	4.02	98.13	98.67	67.77	510.03	536	4.13	97.54	97.88
		HP	75.28	547.40	560	4.36	98.11	98.21	69.09	520.80	534	4.68	97.19	97.21
		SD	46.87	367.58	420	3.36	96.11	96.72	50.49	363.72	420	3.10	97.33	97.33
2b) $p_1 = 0.9, p_2 = 0.8, D = 0.1$														
$T_{(1)}$	61.24	PO	91.84	367.66	376	2.90	96.12	98.33	86.99	387.85	376	3.03	96.26	97.90
		OBF	92.63	396.83	376	3.86	97.05	97.48	89.86	399.12	376	3.93	96.84	96.84
		WT	93.71	382.78	376	3.33	96.90	97.90	90.71	394.97	376	3.45	96.82	97.32
		HP	94.80	448.57	510	3.69	96.74	97.00	90.68	416.18	412	4.37	96.57	96.70
		SD	74.87	322.18	328	3.52	94.72	95.47	76.45	318.76	356	3.02	96.27	96.27
$T_{(2)}$	59.87	PO	91.52	373.88	376	2.96	96.32	98.35	86.56	391.75	376	3.06	96.27	97.96
		OBF	92.18	399.74	376	3.89	97.16	97.61	89.10	405.66	376	4.02	97.07	97.07
		WT	93.39	386.78	376	3.37	97.15	98.04	90.18	400.05	376	3.50	97.06	97.45
		HP	94.91	451.15	510	3.72	96.90	97.19	90.05	424.34	422	4.43	96.93	97.06
		SD	75.82	323.74	334	3.54	95.51	96.21	75.92	321.40	366	3.08	96.74	96.74
$T_{(3)}$	53.20	PO	79.90	445.94	468	3.48	98.67	99.47	72.95	465.40	510	3.58	98.26	99.15
		OBF	73.52	482.49	488	4.43	99.37	99.53	71.35	481.92	506	4.47	98.98	98.98
		WT	80.30	461.33	468	3.94	99.22	99.57	76.06	476.10	510	4.04	98.90	99.09
		HP	79.65	511.86	516	4.21	99.31	99.48	76.11	492.81	516	4.67	98.65	98.69
		SD	47.66	362.86	412	3.44	98.58	98.92	55.79	354.59	420	3.17	98.56	98.56
$T_{(4)}$	57.41	PO	86.69	406.44	376	3.20	97.59	98.95	81.24	426.10	376	3.31	97.41	98.62
		OBF	85.10	438.89	422	4.14	98.16	98.51	81.01	443.43	436	4.25	97.97	97.97
		WT	88.64	419.86	376	3.62	98.26	98.82	84.70	434.28	416	3.75	97.99	98.35
		HP	87.75	479.67	510	3.95	98.29	98.52	84.45	455.46	470	4.54	97.59	97.64

TS	NSD (1 - β)%	Design	Inverse Normalmethode						Inverse χ <sup>2</sup> -Methode					
			(1 - β̂)%	ASN	N <sub>med</sub>	k̄	(1 - 2 α̂)%	(1 - β̂)%	ASN	N <sub>med</sub>	k̄	(1 - 2 α̂)%		
							{KI <sub>j</sub> }	KI <sub>k</sub>					{KI <sub>j</sub> }	KI <sub>k</sub>
		SD	60.14	344.96	384	3.47	96.93	97.42	65.49	338.39	404	3.13	97.60	97.60
2c) p <sub>1</sub> = 0.7, p <sub>2</sub> = 0.5, D = 0.2														
T <sub>(1)</sub>	89.91	PO	99.84	221.49	226	1.89	96.90	98.48	99.63	222.33	218	1.87	96.53	97.52
		OBF	98.79	250.52	242	3.12	97.33	97.66	98.78	240.26	236	2.99	95.82	95.82
		WT	99.86	235.75	230	2.36	98.37	98.94	99.67	236.10	236	2.43	97.91	98.11
		HP	99.82	280.87	242	2.37	97.88	98.05	99.66	251.70	240	3.29	94.81	94.95
		SD	92.84	237.47	232	3.70	90.37	91.03	96.00	224.35	220	2.75	95.09	95.09
T <sub>(2)</sub>	89.63	PO	99.80	223.30	228	1.90	96.86	98.39	99.62	226.70	230	1.91	96.84	97.77
		OBF	98.83	252.54	242	3.16	97.66	97.98	98.61	243.58	240	3.08	96.20	96.20
		WT	99.79	238.21	236	2.40	98.40	98.98	99.59	238.19	240	2.46	98.15	98.34
		HP	99.84	285.30	242	2.41	97.99	98.16	99.63	256.68	242	3.40	95.19	95.29
		SD	93.76	239.07	234	3.75	93.13	93.64	96.53	228.04	228	2.84	95.31	95.31
T <sub>(3)</sub>	87.47	PO	99.49	247.17	242	2.11	98.38	99.10	98.96	252.64	242	2.12	98.18	98.63
		OBF	94.03	295.05	274	3.65	98.85	99.05	94.39	282.16	246	3.54	96.73	96.73
		WT	98.93	267.16	242	2.71	99.07	99.47	98.17	272.32	242	2.82	98.74	98.82
		HP	98.64	312.91	302	2.68	98.92	99.02	98.09	294.70	276	3.71	96.98	97.05
		SD	78.77	272.31	248	3.80	96.87	97.38	88.21	254.49	242	3.02	97.54	97.54
T <sub>(4)</sub>	88.99	PO	99.73	231.94	242	1.97	97.56	98.63	99.43	240.19	242	2.02	97.37	98.25
		OBF	97.02	270.41	250	3.37	97.71	97.99	97.44	259.82	242	3.27	96.38	96.38
		WT	99.59	250.64	242	2.54	98.46	98.94	99.11	252.19	242	2.61	98.29	98.47
		HP	99.29	297.42	242	2.53	98.35	98.47	99.10	272.13	242	3.56	95.15	95.24
		SD	86.19	254.33	248	3.77	93.75	94.31	92.68	239.44	242	2.90	96.36	96.36
2d) p <sub>1</sub> = 0.3, p <sub>2</sub> = 0.1, D = 0.2														
T <sub>(1)</sub>	98.45	PO	100.00	163.10	108	1.44	96.94	98.36	100.00	163.91	108	1.44	96.88	97.94
		OBF	99.34	190.96	168	2.82	97.17	97.66	99.50	176.54	142	2.57	96.77	96.77
		WT	99.99	179.23	156	1.94	98.30	98.98	99.96	172.71	136	1.91	97.74	97.98
		HP	100.00	209.48	214	1.83	99.19	99.58	99.89	183.35	154	2.77	96.40	96.52
		SD	95.38	189.47	170	3.76	92.22	92.92	98.25	172.46	142	2.58	95.80	95.80
T <sub>(2)</sub>	98.31	PO	100.00	166.19	108	1.48	96.97	98.46	100.00	165.86	108	1.46	97.12	98.13
		OBF	99.11	192.65	172	2.91	97.51	97.94	99.34	179.01	148	2.70	97.21	97.21
		WT	100.00	180.69	160	1.97	98.40	99.01	99.91	174.38	138	1.94	98.13	98.28
		HP	100.00	210.35	214	1.84	99.20	99.55	99.90	188.29	164	2.89	97.21	97.39
		SD	96.53	189.86	172	3.85	94.57	95.23	98.61	174.67	148	2.67	96.53	96.53
T <sub>(3)</sub>	97.56	PO	99.99	182.43	108	1.64	98.39	99.37	99.96	183.33	108	1.62	98.42	99.10
		OBF	92.29	244.45	234	3.62	99.26	99.38	93.34	215.51	202	3.34	98.36	98.36
		WT	99.40	204.61	186	2.34	99.34	99.71	98.94	202.83	186	2.38	99.27	99.37
		HP	99.73	234.28	242	2.07	99.49	99.61	98.88	220.98	210	3.29	98.88	98.95
		SD	73.14	239.41	238	3.99	98.13	98.62	87.22	205.88	196	3.02	98.31	98.31
T <sub>(4)</sub>	97.91	PO	100.00	172.99	108	1.54	97.59	98.92	99.99	173.25	108	1.53	97.64	98.56
		OBF	97.38	213.17	196	3.19	97.93	98.20	97.49	194.67	172	2.96	97.54	97.54
		WT	99.89	189.79	168	2.11	98.81	99.30	99.63	186.34	156	2.14	98.54	98.69
		HP	99.92	221.06	230	1.94	99.29	99.52	99.59	201.80	182	3.04	97.46	97.57
		SD	82.11	212.53	200	3.92	95.76	96.35	92.57	188.71	170	2.82	96.98	96.98
2e) p <sub>1</sub> = 0.6, p <sub>2</sub> = 0.3, D = 0.3														
T <sub>(1)</sub>	99.85	PO	100.00	135.90	108	1.24	97.76	98.67	100.00	131.69	108	1.19	97.30	98.07
		OBF	99.77	153.26	134	2.40	96.46	96.67	99.86	139.67	114	1.97	95.07	95.07
		WT	100.00	143.68	108	1.55	99.02	99.27	99.98	138.19	108	1.51	98.26	98.36
		HP	100.00	161.23	108	1.45	99.43	99.66	99.97	140.92	108	2.01	95.34	95.45

TS	NSD	Design	Inverse Normalmethode						Inverse $\chi^2$ -Methode					
			$(1 - \hat{\beta})\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2 \hat{\alpha})\%$	$\{KI_j\}$	$KI_k$	$(1 - \hat{\beta})\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2 \hat{\alpha})\%$
		SD	94.92	156.87	138	3.49	87.12	87.93	99.08	140.45	114	2.14	93.18	93.18
$T_{(2)}$	99.85	PO	100.00	136.32	108	1.24	97.83	98.74	100.00	134.68	108	1.22	97.36	98.06
		OBF	99.67	154.48	136	2.51	98.36	98.49	99.83	141.41	114	2.07	95.25	95.25
		WT	100.00	145.52	108	1.62	99.04	99.27	99.97	138.59	108	1.52	98.76	98.80
		HP	100.00	163.80	108	1.48	99.46	99.68	99.98	144.32	108	2.15	96.43	96.57
		SD	96.30	157.84	140	3.59	90.84	91.49	99.48	141.91	114	2.24	95.45	95.45
$T_{(3)}$	99.69	PO	100.00	143.16	108	1.32	98.71	99.13	100.00	138.87	108	1.25	98.29	98.72
		OBF	96.88	178.45	160	3.05	99.12	99.30	98.28	155.70	126	2.48	99.07	99.07
		WT	99.98	152.76	130	1.73	99.50	99.64	99.82	149.92	108	1.76	99.04	99.07
		HP	100.00	171.26	108	1.56	99.60	99.75	99.79	157.41	108	2.35	99.42	99.46
		SD	83.43	181.44	164	3.79	95.80	96.54	93.93	153.90	120	2.47	97.00	97.00
$T_{(4)}$	99.85	PO	100.00	137.20	108	1.25	98.04	98.84	100.00	137.72	108	1.25	98.04	98.67
		OBF	99.16	162.88	144	2.70	97.48	97.64	99.50	146.81	118	2.23	94.90	94.90
		WT	100.00	148.79	128	1.68	99.18	99.44	99.93	143.40	108	1.66	98.46	98.52
		HP	100.00	166.69	108	1.51	99.48	99.68	99.93	148.96	108	2.19	96.00	96.09
		SD	90.06	167.99	150	3.65	90.71	91.96	97.27	146.61	114	2.30	94.86	94.86
2f) $p_1 = 0.9, p_2 = 0.6, D = 0.3$														
$T_{(1)}$	99.99	PO	100.00	116.77	108	1.07	97.88	98.50	100.00	116.13	108	1.07	97.67	98.21
		OBF	99.88	130.67	120	2.12	96.10	96.64	99.97	120.26	108	1.57	96.75	96.75
		WT	100.00	123.02	108	1.30	99.13	99.42	100.00	118.38	108	1.23	98.78	98.84
		HP	100.00	131.43	108	1.21	99.48	99.74	99.99	120.30	108	1.51	97.16	97.36
		SD	96.17	136.78	126	3.25	87.40	88.73	99.42	121.59	114	1.78	93.31	93.31
$T_{(2)}$	99.99	PO	100.00	117.72	108	1.09	98.02	98.52	100.00	116.61	108	1.07	97.84	98.35
		OBF	99.84	131.55	120	2.24	98.46	98.68	99.96	121.59	108	1.65	97.38	97.38
		WT	100.00	123.33	108	1.30	99.19	99.46	100.00	118.72	108	1.23	98.99	99.10
		HP	100.00	131.61	108	1.21	99.51	99.74	100.00	121.78	108	1.58	98.13	98.25
		SD	96.84	137.11	126	3.34	92.74	93.41	99.56	122.25	114	1.83	97.03	97.03
$T_{(3)}$	99.97	PO	100.00	120.79	108	1.12	98.96	99.40	100.00	120.00	108	1.11	98.75	99.11
		OBF	96.62	158.02	140	2.96	99.60	99.69	98.46	131.48	108	2.06	99.57	99.57
		WT	100.00	129.54	108	1.46	99.64	99.84	99.94	125.58	108	1.41	99.50	99.60
		HP	100.00	137.97	108	1.29	99.78	99.94	99.90	130.38	108	1.81	99.62	99.74
		SD	77.52	164.17	146	3.65	97.58	98.33	92.37	131.68	114	2.12	98.09	98.09
$T_{(4)}$	99.99	PO	100.00	118.69	108	1.10	98.43	98.88	100.00	118.00	108	1.09	98.27	98.70
		OBF	99.47	140.59	126	2.45	98.57	98.76	99.59	124.88	108	1.78	97.15	97.15
		WT	100.00	126.56	108	1.37	99.23	99.53	99.99	121.51	108	1.32	98.99	99.06
		HP	100.00	134.73	108	1.25	99.63	99.82	99.97	125.15	108	1.68	97.88	97.98
		SD	84.11	146.93	132	3.43	93.47	94.69	96.47	125.78	114	1.92	96.71	96.71

Wie bereits unter der Nullhypothese beobachtet werden konnte, resultieren auch unter der Alternativen vergleichbare Ergebnisse für die inverse Normal- sowie für die inverse  $\chi^2$ -Methode bei den einzelnen Designs. Die Beschreibung der Resultate der Simulationen erfolgt zunächst für Szenario I.

Ist die tatsächlich vorliegende Risikodifferenz mit  $D = 0.1$  geringer als vermutet und sind

die Ereigniswahrscheinlichkeiten kleiner (Fall 2a) bzw. größer (Fall 2b) als vor Studienbeginn angenommen, ist der Umfang  $n_{fest} = 248$  nicht ausreichend um diesen Unterschied mit der gewünschten Power von 90% zu entdecken. Im nicht-sequentiellen Design liegt die mit dieser Fallzahl resultierende empirische Power in beiden Fällen unterhalb des vorgegebenen Werts von 90%. Bei den flexiblen Designs gilt dies für Fall 2a sowie einige Konstellationen von Fall 2b, da durch die Beschränkungen der sequentiellen Fallzahlen der mittlere sowie mediane Gesamtumfang der Untersuchung zu gering ausfallen. Durch die Wahl größerer maximaler Umfänge in den einzelnen Sequenzen kann die empirische Power erhöht werden. Diese ist in Konstellation II und Fall 2b bei Verwendung der Prüfgrößen  $T_{(1)}$  und  $T_{(2)}$  ungefähr 90%. Hierbei ist der ASN geringer als der für diesen Fall in einem nicht-sequentiellen Design resultierende Umfang  $n_{fest} = 532$ . Werden die beiden Teststatistiken mit Stetigkeitskorrektur  $T_{(3)}$  und  $T_{(4)}$  verwendet, ist die empirische Power wesentlich kleiner bei größerem mittleren sowie medianen Stichprobenumfang. Für das Haybittle/Peto-Design ist der ASN im Vergleich zu den anderen adaptiv gruppensequentiellen Designs relativ hoch bei nur geringfügig größerer empirischer Power. Insgesamt resultiert für die adaptiv gruppensequentiellen Designs mit inverser Normalmethode in fast allen betrachteten Parameterkonstellationen eine höhere empirische Power als bei der inversen  $\chi^2$ -Methode bei häufig geringerem ASN. In den Szenarien mit  $\beta_g = 0.8$  und bei Verwendung der Prüfgrößen  $T_{(3)}$  und  $T_{(4)}$  ist die empirische Power für die inverse  $\chi^2$ -Methode größer als für die inverse Normalmethode mit Ausnahme des Haybittle/Peto-Designs. Für das Self-Designing resultiert häufig eine kleinere empirische Power bei geringerem ASN als für die adaptiv gruppensequentiellen Designs. Die Ergebnisse für das Self-Designing sind für beide Kombinationsmethoden ähnlich, die mittlere Stufenanzahl ist für die inverse Normalmethode tendenziell etwas größer.

Sind die vor Studienbeginn getroffenen Annahmen für die Ereigniswahrscheinlichkeiten richtig (Fall 2c), ist die empirische Power in den adaptiv gruppensequentiellen Designs größer als 90% bei Verwendung der Teststatistiken  $T_{(1)}$  und  $T_{(2)}$ . Hierbei fällt der durchschnittliche Stichprobenumfang geringer aus als in einem nicht-sequentiellen Design, in welchem die gewünschte Power von 90% annähernd erreicht wird. Die Verwendung der Prüfgrößen  $T_{(3)}$  und  $T_{(4)}$  führt in allen Designs zu einem Powerverlust bei gleichzeitiger Zunahme des mittleren sowie des medianen Umfangs der Studie. Im Self-Designing ist bei beiden Kombinationsmethoden die empirische Power kleiner als in den übrigen flexiblen Designs, wobei dies in den geringeren mittleren Stichprobenumfängen begründet ist.

Insgesamt fallen ASN und  $N_{med}$  kleiner aus als die in einem nicht-sequentiellen Design verwendete Fallzahl  $n_{fest} = 248$ .

Falls die Annahme für die Größe der Risikodifferenz richtig ist, die Werte für die Ereigniswahrscheinlichkeiten mit  $p_1 = 0.3$  und  $p_2 = 0.1$  jedoch geringer ausfallen (Fall 2d), führt die Berechnung der Fallzahl mit  $f_D(\alpha, \beta)_{bin(1)}$  zu dem Umfang  $n_{fest} = 164$ . Bei Verwendung von  $n_{fest} = 248$  in dem nicht-sequentiellen Design ist die empirische Power größer als der vorgegebene Wert von 90%. Dies trifft auch in den flexiblen Designs zu, wobei hier der mittlere und mediane Umfang bei Verwendung der Teststatistiken  $T_{(1)}$  und  $T_{(2)}$  in den meisten Fällen geringer als  $n_{fest} = 164$  sind. Erfolgt die Überprüfung des Testproblems mit den Prüfgrößen  $T_{(3)}$  und  $T_{(4)}$ , kommt es zu einer Erhöhung des medianen und mittleren Umfangs der Studie, welcher insbesondere für  $T_{(3)}$  in der Regel größer als  $n_{fest} = 164$  ist.

Ist die Risikodifferenz mit  $D = 0.3$  größer als vermutet und sind die Ereigniswahrscheinlichkeiten kleiner (Fall 2e) bzw. größer (Fall 2f) als vor Studienbeginn angenommen, resultiert in einem nicht-sequentiellen Design mit  $n_{fest} = 248$  eine empirische Power, welche den vorgegebenen Wert von 90% überschreitet. In den flexiblen Designs ist diese bei Verwendung der Teststatistiken  $T_{(1)}$  und  $T_{(2)}$  in den meisten betrachteten Konstellationen größer als der vorgegebene Wert von 90%. Kommen die Prüfgrößen  $T_{(3)}$  und  $T_{(4)}$  zur Anwendung, verringert sich die empirische Power bei gleichzeitiger Zunahme des mittlerem sowie medianen Stichprobenumfangs. Das Self-Designing unterscheidet sich für die beiden Kombinationsmethoden insbesondere in der mittleren Anzahl durchgeführter Stufen. Da bei Kombination der Ergebnisse mit der inversen Normalmethode eine Überprüfung des Testproblems erst nach Vergabe des globalen Gewichts erfolgen kann, werden meist mehr Sequenzen benötigt als bei Verwendung der inversen  $\chi^2$ -Methode.

Die für Szenario I beschriebenen Beobachtungen werden durch die Ergebnisse von Szenario II bestätigt. In den Fällen, in denen die Risikodifferenz größer ist als vor Studienbeginn angenommen, wird der Unterschied des Self-Designings mit inverser Normalmethode gegenüber den anderen Designs verdeutlicht. Bei diesem kann das Testen der Nullhypothese erst erfolgen, wenn das globale Gewicht verbraucht ist. Somit werden im Mittel mehr Sequenzen als bei den anderen flexiblen Testverfahren benötigt, bei welchen bei Vorliegen eines großen Behandlungsunterschieds die Untersuchung vorzeitig beendet werden kann.

Da in Szenario II der Stichprobenumfang der ersten Stufe sowie die minimale und maximale Fallzahl der einzelnen Sequenzen relativ groß gewählt wurden, fallen der mittlere und mediane Umfang in den flexiblen Designs teilweise recht hoch aus im Vergleich zu der in einem nicht-sequentiellen Design benötigten Fallzahl. Dies ist insbesondere der Fall, wenn die tatsächlich vorliegende Risikodifferenz wesentlich größer ist als vor Studienbeginn vermutet wurde.

Aufgrund der Möglichkeit des vorzeitigen Abbruchs der Studie sind sowohl die Sequenz der Konfidenzintervalle als auch das finale Intervall konservativ (vgl. Kapitel 2.5.1) und in den meisten betrachteten Konstellationen ist die resultierende empirische Überdeckungswahrscheinlichkeit wesentlich größer als der für das Konfidenzniveau festgelegte Wert von 95%. Bei dem Self-Designing mit inverser Normalmethode ist die empirische Überdeckungswahrscheinlichkeit der Sequenz der Konfidenzintervalle teilweise deutlich geringer als 95%, da das Konfidenzniveau in diesem Fall nicht kontrolliert wird.

Allgemein zeigt sich, dass die empirische Power für die Teststatistiken mit Stetigkeitskorrektur  $T_{(3)}$  und  $T_{(4)}$  deutlich kleiner ist als für die Prüfgrößen  $T_{(1)}$  und  $T_{(2)}$ . Unter Berücksichtigung der Ergebnisse unter der Nullhypothese wird von der Verwendung von  $T_{(3)}$  und  $T_{(4)}$  in einem flexiblen Design abgeraten und der Gebrauch von  $T_{(2)}$  empfohlen. Größere Werte des Parameters  $\beta_g$  führen im Allgemeinen zu einer höheren mittleren Stufenanzahl, wobei die empirische Power und der mittlere sowie mediane Umfang geringer ausfallen. Das Haybittle/Peto-Design unterscheidet sich zwischen beiden Kombinationsmethoden. Die empirische Power ist bei Verwendung der inversen Normalmethode in den Fällen 2a und 2b meist deutlich größer als bei den anderen adaptiv gruppensequentiellen Designs. Für die inverse  $\chi^2$ -Methode bewegt sich diese für das Haybittle/Peto-Design in der gleichen Größenordnung wie für die anderen flexiblen Verfahren. Jedoch ist bei diesem bei beiden Kombinationsmethoden der mittlere sowie mediane Umfang der Studie erhöht. Aufgrund dessen wird die Verwendung dieses Designs nicht unbedingt empfohlen. Das mittlere Verhalten der anderen adaptiv gruppensequentiellen Designs liegt sehr nah beieinander. Das Self-Designing bietet den Vorteil, dass unter der Nullhypothese im Mittel weniger Sequenzen benötigt werden als in den adaptiv gruppensequentiellen Designs. Unter der Alternativhypothese resultieren bei einer Überschätzung der Risikodifferenz vor Studienbeginn vergleichsweise kleine Werte für die empirische Power, was in geringeren mittleren und medianen Fallzahlen begründet ist.

Durch Zulassen größerer maximaler Umfänge in den einzelnen Sequenzen kann die empirische Power erhöht werden.

### 3.2 Ordinale Zielvariable

Bei binären Zielvariablen sind nur zwei Ausprägungen möglich, jedoch sind häufig mehr als zwei Werte von Interesse. Beispielsweise soll die Beurteilung der Wirkung eines Medikaments in den Kategorien sehr gut, gut, mittel, schlecht, sehr schlecht erfolgen.

In einer sequentiell durchgeführten klinischen Studie mit  $n_j = n_{1j} + n_{2j}$  Patienten in Stufe  $j$ ,  $j = 1, \dots, k$ , sei  $\tilde{Y}_{Ji,j}$  die ordinale Zielgröße des Patienten  $i$ ,  $i = 1, \dots, n_{Jj}$ , der Behandlungsgruppe  $J$ ,  $J = 1, 2$ . Diese Variable kann Ausprägungen in  $l$ ,  $l = 1, \dots, m$ , Responsekategorien annehmen, d.h.  $\tilde{Y}_{Ji,j} = l$ . Hierbei stellt die erste die beste Kategorie dar. Sei  $p_{Jl} = P(\tilde{Y}_{Ji,j} = l)$  die Wahrscheinlichkeit, dass Patient  $i$ ,  $i = 1, \dots, n_{Jj}$ , der Gruppe  $J$ ,  $J = 1, 2$ , in Stufe  $j$ ,  $j = 1, \dots, k$ , eine Ausprägung in Kategorie  $l$ ,  $l = 1, \dots, m$ , besitzt, welche als identisch über die einzelnen Sequenzen hinweg angenommen wird. Es gilt  $\sum_{l=1}^m p_{Jl} = 1$  für  $J = 1, 2$ . Um zu überprüfen, wie viele der Patienten der Gruppe  $J$  in Stufe  $j$  Ausprägungen in den einzelnen Kategorien besitzen, werden die Zufallsvariablen  $Y_{Jl,j} = \sum_{i=1}^{n_{Jj}} \mathbf{I}_{\{\tilde{Y}_{Ji,j}=l\}}$  betrachtet, wobei  $\mathbf{I}$  die Indikatorfunktion ist. Die Realisationen dieser Variablen sind die Häufigkeiten  $n_{Jl,j}$ , welche gemäß Tabelle 3.9 in einer  $(2 \times m)$ -Kontingenztafel festgehalten werden können (vgl. Hartung et al. (1999)).

**Tabelle 3.9:**  $(2 \times m)$ -Kontingenztafel der beobachteten absoluten Häufigkeiten des Auftretens der Kategorien  $l$ ,  $l = 1, \dots, m$ , pro Gruppe  $J$ ,  $J = 1, 2$ , in Stufe  $j$ ,  $j = 1, \dots, k$ , einer klinischen Studie mit ordinaler Zielvariable

Gruppe	Kategorie			Umfang
	1	...	$m$	
1	$n_{11,j}$	...	$n_{1m,j}$	$n_{1j}$
2	$n_{21,j}$	...	$n_{2m,j}$	$n_{2j}$

Die Verteilung der Zufallsvariablen  $(Y_{J1,j}, \dots, Y_{Jm,j})$  der Gruppe  $J$ ,  $J = 1, 2$ , in Stufe  $j$ ,

$j = 1, \dots, k$ , ist die Multinomialverteilung mit den Parametern  $n_{Jj}, p_{J1}, \dots, p_{Jm}$ , d.h.

$$(Y_{J1,j}, \dots, Y_{Jm,j}) \sim M(n_{Jj}, p_{J1}, \dots, p_{Jm}) \quad (3.27)$$

(vgl. Hartung et al. (1999)). Für die einzelnen Zufallsvariablen gilt  $Y_{Jl,j} \sim B(n_{Jj}, p_{Jl})$ , wobei  $B(n_{Jj}, p_{Jl})$  die Binomialverteilung mit Parametern  $n_{Jj}$  und  $p_{Jl}$  bezeichnet. Es sind  $E(Y_{Jl,j}) = n_{Jj}p_{Jl}$  und  $Var(Y_{Jl,j}) = n_{Jj}p_{Jl}(1 - p_{Jl})$ ,  $J = 1, 2$ ,  $l = 1, \dots, m$ ,  $j = 1, \dots, k$ . Die unbekanntenen Wahrscheinlichkeiten  $p_{Jl}$  können in Gruppe  $J = 1, 2$  durch

$$\hat{p}_{Jl,j} = Y_{Jl,j}/n_{Jj}, \quad l = 1, \dots, m, \quad (3.28)$$

erwartungstreu geschätzt werden, da  $E(Y_{Jl,j}/n_{Jj}) = n_{Jj}p_{Jl}/n_{Jj} = p_{Jl}$  ist. Die Varianz von  $\hat{p}_{Jl,j}$  ist

$$Var(\hat{p}_{Jl,j}) = Var(Y_{Jl,j}/n_{Jj}) = Var(Y_{Jl,j})/n_{Jj}^2 = p_{Jl}(1 - p_{Jl})/n_{Jj}. \quad (3.29)$$

Im Folgenden wird die verallgemeinerte Risikodifferenz vorgestellt. In diesem Zusammenhang wird der Parameter  $\theta$  aus Kapitel 2 mit  $D_G$  bezeichnet.

### 3.2.1 Die verallgemeinerte Risikodifferenz

Die in Kapitel 3.1.1 vorgestellte Risikodifferenz zur Messung des Behandlungsunterschieds bei Betrachtung einer binären Zielgröße kann für den Fall einer  $(2 \times m)$ -Kontingenztafel verallgemeinert werden (vgl. Edwardes und Baltzan (2000)). Die Wahrscheinlichkeit, dass Gruppe 1 häufiger Ausprägungen in besseren Responsekategorie besitzt als Gruppe 2, ist

$$\Pi_1 = \sum_{l=1}^m p_{1l} \sum_{l'>l} p_{2l'}, \quad (3.30)$$

mit  $-1 \leq \Pi_1 \leq 1$ . Hierbei ist  $\sum_{l'>l}$  die Summation über alle Paare mit Werten  $l'$  und  $l$  mit  $l' > l$ . Umgekehrt ist die Wahrscheinlichkeit, dass Gruppe 1 Realisationen in schlechteren Kategorien besitzt als Gruppe 2, durch

$$\Pi_2 = \sum_{l=1}^m p_{1l} \sum_{l'<l} p_{2l'} \quad (3.31)$$

definiert, mit  $-1 \leq \Pi_2 \leq 1$ . Fallen beispielsweise alle Ausprägungen der Gruppe 1 in die erste Kategorie und für Gruppe 2 gibt es nur Realisationen in den Kategorien  $2, \dots, m$ , ist

$\Pi_1 = 1$  und  $\Pi_2 = 0$ . Die Schätzung von  $\Pi_1$  und  $\Pi_2$  kann in Stufe  $j$  mit

$$\hat{\Pi}_{1j} = \sum_{l=1}^m \hat{p}_{1l,j} \sum_{l'>l} \hat{p}_{2l',j} = \frac{\sum_{l=1}^m Y_{1l,j} \sum_{l'>l} Y_{2l',j}}{n_{1j} n_{2j}} \quad (3.32)$$

und

$$\hat{\Pi}_{2j} = \sum_{l=1}^m \hat{p}_{1l,j} \sum_{l'<l} \hat{p}_{2l',j} = \frac{\sum_{l=1}^m Y_{1l,j} \sum_{l'<l} Y_{2l',j}}{n_{1j} n_{2j}} \quad (3.33)$$

erfolgen. Der Erwartungswert von  $\hat{\Pi}_{Jj}$  ist  $E(\hat{\Pi}_{Jj}) = \sum_{l=1}^m E(\hat{p}_{1l,j}) \cdot \sum_{l'>l} E(\hat{p}_{2l',j}) = \Pi_{Jj}$ , da die Schätzer  $\hat{p}_{Jl,j}$ ,  $J = 1, 2$ ,  $l = 1, \dots, m$ , in Stufe  $j$ ,  $j = 1, \dots, k$ , voneinander unabhängig sind. Die Differenz der Wahrscheinlichkeiten  $\Pi_1$  und  $\Pi_2$  ist die von Edwardes und Baltzan (2000) vorgeschlagene verallgemeinerte Risikodifferenz

$$D_G = \Pi_1 - \Pi_2 = \sum_{l=1}^m p_{1l} \sum_{l'>l} p_{2l'} - \sum_{l=1}^m p_{1l} \sum_{l'<l} p_{2l'}, \quad (3.34)$$

mit  $-1 \leq D_G \leq 1$ . Positive Werte von  $D_G$  zeigen an, dass Gruppe 1 häufiger Ausprägungen in besseren Responsekategorien besitzt als Gruppe 2. Diese Differenz entspricht dem Assoziationsmaß von Somers (1962) für ordinale Variablen. In Stufe  $j$  kann  $D_G$  durch

$$\hat{D}_{Gj} = \hat{\Pi}_{1j} - \hat{\Pi}_{2j} = \sum_{l=1}^m \hat{p}_{1l,j} \sum_{l'>l} \hat{p}_{2l',j} - \sum_{l=1}^m \hat{p}_{1l,j} \sum_{l'<l} \hat{p}_{2l',j} \quad (3.35)$$

erwartungstreu geschätzt werden, da gilt  $E(\hat{D}_{Gj}) = E(\hat{\Pi}_{1j}) - E(\hat{\Pi}_{2j}) = \Pi_1 - \Pi_2 = D_G$ .

Die asymptotische Varianz von  $\hat{D}_{Gj}$  wird mittels einer von Goodman und Kruskal (1972) vorgeschlagenen generellen Formel für Assoziationsmaße in  $(2 \times m)$ -Kontingenztafeln ermittelt. Sei  $\xi$  ein Assoziationsmaß, welches als Quotient dargestellt werden kann als

$$\xi = \frac{\nu}{\delta}, \quad (3.36)$$

wobei  $\nu$  und  $\delta$  von den unbekanntenen Zellwahrscheinlichkeiten  $p_{Jl}$  abhängen,  $J = 1, 2$ ,  $l = 1, \dots, m$ . Der Maximum-Likelihood-Schätzer für  $\xi$  sei durch  $\hat{\xi} = \hat{\nu}/\hat{\delta}$  gegeben. Die partiellen Ableitungen von  $\xi$  können mittels der Quotientenregel bestimmt werden

$$\xi'_{Jl} = \partial\xi/\partial p_{Jl} = \frac{1}{\delta^2} (\nu'_{Jl} \delta - \nu \delta'_{Jl}) = \frac{1}{\delta^2} \varphi_{Jl}, \quad (3.37)$$

wobei

$$\nu'_{Jl} = \partial\nu/\partial p_{Jl} \quad \text{und} \quad \delta'_{Jl} = \partial\delta/\partial p_{Jl} \quad (3.38)$$

sind. Die asymptotische Varianz von  $\hat{\xi}$

$$Var(\hat{\xi}) = \frac{1}{\delta^4} \sum_{J=1}^2 \frac{1}{n_J} \left( \sum_{l=1}^m p_{Jl} \varphi_{Jl}^2 - \left( \sum_{l=1}^m p_{Jl} \varphi_{Jl} \right)^2 \right) \quad (3.39)$$

hängt von den unbekanntem Wahrscheinlichkeiten  $p_{Jl}$  sowie von den unbekanntem partiellen Ableitungen von  $\xi$  bzgl.  $p_{Jl}$  ab. Werden diese jeweils durch ihre Maximum-Likelihood-Schätzer ersetzt, ist ein Schätzer für die asymptotische Varianz von  $\hat{\xi}$  gegeben durch

$$\widehat{Var}(\hat{\xi}) = \frac{1}{\hat{\delta}^4} \sum_{J=1}^2 \frac{1}{n_J} \left( \sum_{l=1}^m \hat{p}_{Jl} \hat{\varphi}_{Jl}^2 - \left( \sum_{l=1}^m \hat{p}_{Jl} \hat{\varphi}_{Jl} \right)^2 \right). \quad (3.40)$$

Mittels dieser Vorgehensweise kann die asymptotische Varianz der verallgemeinerten Risikodifferenz wie folgt ermittelt werden. Seien

$$\nu = \sum_{l=1}^m p_{1l} \sum_{l'>l} p_{2l'} - \sum_{l=1}^m p_{1l} \sum_{l'<l} p_{2l'} \quad \text{und} \quad \delta = \sum_{l=1}^m p_{1l} = \sum_{l=1}^m p_{2l}.$$

Die partiellen Ableitungen sind gegeben durch

$$\begin{aligned} \nu'_{1l} &= \partial\nu/\partial p_{1l} = \sum_{l'>l} p_{2l'} - \sum_{l'<l} p_{2l'}, & \nu'_{2l} &= \partial\nu/\partial p_{2l} = \sum_{l'<l} p_{1l'} - \sum_{l'>l} p_{1l'}, \\ \delta'_{1l} &= \partial\delta/\partial p_{1l} = 1, & \delta'_{2l} &= \partial\delta/\partial p_{2l} = 1. \end{aligned}$$

Außerdem sind

$$\varphi_{1l} = \left( \sum_{l'>l} p_{2l'} - \sum_{l'<l} p_{2l'} \right) - \left( \sum_{l=1}^m p_{1l} \sum_{l'>l} p_{2l'} - \sum_{l=1}^m p_{1l} \sum_{l'<l} p_{2l'} \right)$$

und

$$\varphi_{2l} = \left( \sum_{l'<l} p_{1l'} - \sum_{l'>l} p_{1l'} \right) - \left( \sum_{l=1}^m p_{1l} \sum_{l'>l} p_{2l'} - \sum_{l=1}^m p_{1l} \sum_{l'<l} p_{2l'} \right).$$

Es gilt  $\bar{\varphi} = \sum_{J=1}^2 \sum_{l=1}^m p_{Jl} \varphi_{Jl} = 0$ . Somit ist in Stufe  $j$ ,  $j = 1, \dots, k$ , die asymptotische Varianz von  $\hat{D}_{Gj}$  gegeben durch

$$\begin{aligned} Var(\hat{D}_{Gj}) &= \frac{1}{n_{1j}} \sum_{l=1}^m p_{1l} \left\{ \left( \sum_{l'>l} p_{2l'} - \sum_{l'<l} p_{2l'} \right) - D_G \right\}^2 \\ &+ \frac{1}{n_{2j}} \sum_{l=1}^m p_{2l} \left\{ \left( \sum_{l'<l} p_{1l'} - \sum_{l'>l} p_{1l'} \right) - D_G \right\}^2 \end{aligned} \quad (3.41)$$

und kann mittels

$$\widehat{Var}(\widehat{D}_{Gj}) = \frac{1}{(n_{1j} n_{2j})^2} \left\{ \sum_{l=1}^m Y_{1l,j} \left( \left( \sum_{l'>l} Y_{2l',j} - \sum_{l'<l} Y_{2l',j} \right) - n_{2j} \widehat{D}_{Gj} \right)^2 \right. \\ \left. + \sum_{l=1}^m Y_{2l,j} \left( \left( \sum_{l'<l} Y_{1l',j} - \sum_{l'>l} Y_{1l',j} \right) - n_{1j} \widehat{D}_{Gj} \right)^2 \right\} \quad (3.42)$$

geschätzt werden.

Ist von Interesse, ob in Gruppe 1 häufiger Ausprägungen in besseren Responsekategorien auftreten als in Gruppe 2, kann für die verallgemeinerte Risikodifferenz das einseitige Testproblem

$$H_0 : D_G = 0 \quad \text{vs.} \quad H_1 : D_G > 0 \quad (3.43)$$

formuliert werden. Dieses kann in Stufe  $j$ ,  $j = 1, \dots, k$ , mittels der unter der Nullhypothese für große Stichprobenumfänge  $n_{1j}$  und  $n_{2j}$  approximativ standardnormalverteilten Teststatistik

$$T_j = \frac{\widehat{D}_{Gj}}{\sqrt{\widehat{Var}(\widehat{D}_{Gj})}} \stackrel{H_0}{\approx} N(0, 1) \quad (3.44)$$

überprüft werden.

Bei Vorliegen der Beobachtungen in Form einer Vierfeldertafel kann die verallgemeinerte Risikodifferenz  $D_G$  auf die Risikodifferenz  $D$  aus Kapitel 3.1.1 zurückgeführt werden. Mit  $p_{11} = p_1$ ,  $p_{12} = 1 - p_1$ ,  $p_{21} = p_2$  und  $p_{22} = 1 - p_2$  sind

$$D_G = p_{11} p_{22} - p_{12} p_{21} = p_1 (1 - p_2) - p_2 (1 - p_1) = D$$

und

$$\begin{aligned} Var(\widehat{D}_G) &= \frac{p_{11} (p_{22} - D)^2 + p_{12} (-p_{21} - D)^2}{n_1} + \frac{p_{21} (-p_{12} - D)^2 + p_{22} (p_{11} - D)^2}{n_2} \\ &= \frac{p_1 (1 - p_2 - D)^2 + (1 - p_1) (-p_2 - D)^2}{n_1} + \frac{p_2 (-1 - p_1 - D)^2 + (1 - p_2) (p_1 - D)^2}{n_2} \\ &= \frac{1}{n_1} p_1 (1 - p_1) + \frac{1}{n_2} p_2 (1 - p_2) = Var(\widehat{D}). \end{aligned}$$

### 3.2.2 Fallzahlplanung

Die Herleitung einer Stichprobenfunktion bei Vorliegen einer ordinalen Zielgröße wird zunächst anhand eines einstufigen nicht-sequentiellen Designs illustriert. Anschließend wird die vorgestellte Fallzahlformel auf eine in mehreren Sequenzen durchgeführte klinische Studie übertragen.

#### Stichprobenfunktion in einem nicht-sequentiellen Design

Mittels der Grundgleichung (3.20),  $z_{1-\alpha} \sigma_0^* + z_{1-\beta} \sigma_1^* = \mu_1^* - \mu_0^*$ , aus Kapitel 3.1.2 erfolgt die Herleitung einer Stichprobenfunktion basierend auf der Betrachtung der verallgemeinerten Risikodifferenz  $D_G$ .

Seien  $Z = \hat{D}_G = \sum_{l=1}^m \hat{p}_{1l} \sum_{l'>l} \hat{p}_{2l'} - \sum_{l=1}^m \hat{p}_{1l} \sum_{l'<l} \hat{p}_{2l'}$ ,  $\mu_0^* = 0$  und  $\mu_1^* = D_G$ . Die Varianz von  $Z$  wird unter der Null- und Alternativhypothese festgelegt als

$$\begin{aligned} \sigma_0^{*2} = \sigma_1^{*2} &= \frac{1}{n_1} \sum_{l=1}^m p_{1l} \left\{ \left( \sum_{l'>l} p_{2l'} - \sum_{l'<l} p_{2l'} \right) - D_G \right\}^2 \\ &+ \frac{1}{n_2} \sum_{l=1}^m p_{2l} \left\{ \left( \sum_{l'<l} p_{1l'} - \sum_{l'>l} p_{1l'} \right) - D_G \right\}^2. \end{aligned}$$

Werden diese Größen in die Grundgleichung eingesetzt, resultiert mit  $n_2 = r n_1$  die Fallzahlformel

$$f_{D_G}(\alpha, \beta) = n = \frac{r+1}{r} \frac{\max\{0, z_{1-\alpha} + z_{1-\beta}\}^2}{D_G^2} Q, \quad (3.45)$$

mit

$$Q = r \sum_{l=1}^m p_{1l} \left\{ \left( \sum_{l'>l} p_{2l'} - \sum_{l'<l} p_{2l'} \right) - D_G \right\}^2 + \sum_{l=1}^m p_{2l} \left\{ \left( \sum_{l'<l} p_{1l'} - \sum_{l'>l} p_{1l'} \right) - D_G \right\}^2.$$

Diese Fallzahl entspricht für  $m = 2$  der Stichprobenfunktion  $f_D(\alpha, \beta)_{bin(2)}$  aus Kapitel 3.1.2, da  $D_G = p_{11} - p_{21} = p_1 - p_2 = D$  und

$$\begin{aligned} Q &= r p_{11} (p_{22} - D)^2 + r p_{12} (p_{21} + D)^2 + p_{21} (p_{12} + D)^2 + p_{22} (p_{11} - D)^2 \\ &= r p_{11} (1 - p_{21} - p_{11} + p_{21})^2 + r (1 - p_{11}) (p_{21} + p_{11} - p_{21})^2 \\ &\quad + p_{21} (1 - p_{11} + p_{11} - p_{21})^2 + (1 - p_{21}) (p_{11} - p_{11} + p_{21})^2 \\ &= r p_1 (1 - p_1) + p_2 (1 - p_2) \end{aligned}$$

sind, so dass insgesamt gilt

$$f_{D_G}(\alpha, \beta) = \frac{r+1}{r} \frac{\max\{0, z_{1-\alpha} + z_{1-\beta}\}^2 (r p_1 (1-p_1) + p_2 (1-p_2))}{D^2} = f_D(\alpha, \beta)_{bin(2)}.$$

In Tabelle 3.10 sind beispielhaft für  $m = 3, 4$  und  $5$  Kategorien für jeweils drei verschiedene Annahmen für die Ereigniswahrscheinlichkeiten  $p_{Jl}$ ,  $J = 1, 2$ ,  $l = 1, \dots, m$ , mit  $f_{D_G}(\alpha, \beta)$  ermittelte Fallzahlen verzeichnet.

**Tabelle 3.10:** Bestimmung der Fallzahl einer klinischen Studie mit ordinaler Zielvariable bei Betrachtung von  $m = 3, 4$  oder  $5$  Kategorien mittels der Stichprobenfunktion  $f_{D_G}(\alpha, \beta)$  (Fehlerraten I. und II. Art  $\alpha = 0.025$  und  $\beta = 0.1$ )

$J$	$p_{J1}$	$p_{J2}$	$p_{J3}$	$p_{J4}$	$p_{J5}$	$D_G$	$f_{D_G}(\alpha, \beta)$
<i>m = 3 Kategorien</i>							
1	0.50	0.30	0.20	–	–	0.20	296
2	0.33	0.33	0.33	–	–		
1	0.60	0.30	0.10	–	–	0.30	120
2	0.40	0.20	0.40	–	–		
1	0.60	0.20	0.20	–	–	0.44	52
2	0.20	0.30	0.50	–	–		
<i>m = 4 Kategorien</i>							
1	0.18	0.44	0.32	0.06	–	0.21	258
2	0.10	0.35	0.40	0.15	–		
1	0.20	0.40	0.30	0.10	–	0.24	206
2	0.10	0.30	0.40	0.20	–		
1	0.18	0.44	0.28	0.10	–	0.44	52
2	0.02	0.27	0.35	0.36	–		
<i>m = 5 Kategorien</i>							
1	0.30	0.40	0.20	0.09	0.01	0.26	182
2	0.20	0.30	0.20	0.20	0.10		
1	0.25	0.51	0.20	0.03	0.01	0.30	118
2	0.10	0.42	0.37	0.07	0.04		
1	0.18	0.58	0.20	0.03	0.01	0.40	68
2	0.12	0.26	0.34	0.22	0.06		

### Übertragung der Stichprobenfunktion auf sequentielle Designs

In Stufe  $j$ ,  $j = 1, \dots, k$ , können die unbekanntes Wahrscheinlichkeiten der beiden Behandlungsgruppen  $p_{1l}$  und  $p_{2l}$  der Kategorien  $l$ ,  $l = 1, \dots, m$ , geschätzt werden durch ein gewichtetes Mittel der Schätzungen  $\hat{p}_{Jl,j}$  der einzelnen Sequenzen. Eine Möglichkeit ist die Verwendung des Schätzers (2.109) aus Kapitel 2.5.2, welcher für die Ereigniswahrscheinlichkeit  $p_{lJ}$ ,  $J = 1, 2$ ,  $l = 1, \dots, m$ , gegeben ist durch

$$\hat{p}_{Jl,j}^* = \frac{\sum_{i=1}^j m_i \hat{p}_{Jl,i}}{\sum_{i=1}^j m_i}, \quad (3.46)$$

mit  $m_i = 2/(1/n_{1i} + 1/n_{2i})$ ,  $i = 1, \dots, j$ . Hiermit kann die verallgemeinerte Risikodifferenz unter Einbeziehung aller vorhandenen Informationen mittels

$$\hat{D}_{Gj}^* = \sum_{l=1}^m \hat{p}_{1l,j}^* \sum_{l'>l} \hat{p}_{2l',j}^* - \sum_{l=1}^m \hat{p}_{1l,j}^* \sum_{l'<l} \hat{p}_{2l',j}^* \quad (3.47)$$

geschätzt werden. Die resultierende Fallzahl der Sequenz  $(j + 1)$ , welche die Update-Schätzungen der Wahrscheinlichkeiten der einzelnen Kategorien bis zur Stufe  $j$  berücksichtigt, ist durch

$$f_{D_{G,j}}(\alpha, \beta) = n_{j+1} = \frac{r+1}{r} \frac{\max\{0, z_{1-\alpha} + z_{1-\beta}\}^2}{\hat{D}_{Gj}^{*2}} Q_j^* \quad (3.48)$$

gegeben,  $j = 0, 1, \dots, k - 1$ , mit

$$\begin{aligned} Q_j^* = & r \sum_{l=1}^m \hat{p}_{1l,j}^* \left\{ \left( \sum_{l'>l} \hat{p}_{2l',j}^* - \sum_{l'<l} \hat{p}_{2l',j}^* \right) - \hat{D}_{Gj}^* \right\}^2 \\ & + \sum_{l=1}^m \hat{p}_{2l,j}^* \left\{ \left( \sum_{l'<l} \hat{p}_{1l',j}^* - \sum_{l'>l} \hat{p}_{1l',j}^* \right) - \hat{D}_{Gj}^* \right\}^2. \end{aligned}$$

Für  $j = 0$  entspricht  $f_{D_{G,0}}(\alpha, \beta)$  der Stichprobenfunktion  $f_{D_G}(\alpha, \beta)$  eines nicht-sequentiellen Studiendesigns, wobei  $D_{G,0} > 0$  einen auf Vorwissen basierenden Wert für die unbekanntes verallgemeinerte Risikodifferenz  $D_G$  bezeichnet. Um den Wertebereich des Schätzers  $\hat{D}_{G,j}^*$  auf die Alternativhypothese zu restringieren, wird  $\hat{D}_{G,j}^* = \rho$ ,  $\rho > 0$ , gewählt, falls Werte kleiner als  $\rho$  auftreten. In dieser Arbeit wird  $\rho$  auf den Wert 0.001 gesetzt.

Mit der sequentiellen Variante der ausgewählten Stichprobenfunktion kann der Umfang der Stufe  $(j + 1)$ ,  $j = 1, \dots, k - 1$ , mittels der in Kapitel 2.4.1 beschriebenen Vorgehensweise bestimmt werden. Ziel der auf der bedingten Power basierenden Fallzahlplanung ist, dass unter Berücksichtigung der Informationen der Sequenzen  $1, \dots, j$  die Ablehnung der Nullhypothese in der folgenden Stufe  $(j + 1)$  mit einer bestimmten Wahrscheinlichkeit ermöglicht werden soll.

### 3.2.3 Beispiel

Im Rahmen dieses Beispiels werden flexible Designs in klinischen Untersuchungen mit ordinaler Zielgröße veranschaulicht. Das mittlere Verhalten der flexiblen Designs in klinischen Prüfungen mit ordinaler Zielvariable wird in Kapitel 3.2.4 in Simulationen untersucht. Als Studiendesigns werden das Testverfahren vom Wang/Tsiatis-Typ mit  $\Delta = 0.25$ ,  $K = 3$  und  $\lambda_1 = \lambda_2 = \lambda_3 = 1$  sowie das Self-Designing mit  $\lambda_G = 5$ ,  $\lambda_1 = 1$  und  $\lambda_{\min} = 1$  betrachtet. Für die Kombination der Ergebnisse der einzelnen Sequenzen werden sowohl die inverse Normal- als auch die inverse  $\chi^2$ -Methode verwendet.

Von Interesse ist eine ordinale Zielvariable mit  $m = 4$  Kategorien, wobei die erste die beste Kategorie darstellt. Zunächst erfolgt die Herleitung des Umfangs in einem nicht-sequentiellen Studiendesign. Unter Verwendung der Stichprobenfunktion (3.45) aus Kapitel 3.2.2 müssen Annahmen über die Größe der Ereigniswahrscheinlichkeiten  $p_{Jl}$ ,  $J = 1, 2$ ,  $l = 1, \dots, 4$ , getroffen werden. Für die in Gruppe 2 verabreichte Standardtherapie werden aufgrund von vorherigen Untersuchungen die Werte  $p_{21} = 0.15$ ,  $p_{22} = 0.35$ ,  $p_{23} = 0.40$  und  $p_{24} = 0.10$  angenommen. Für die neue Behandlung werden die Ereigniswahrscheinlichkeiten vermutet als  $p_{11} = 0.30$ ,  $p_{12} = 0.40$ ,  $p_{13} = 0.20$  und  $p_{14} = 0.10$ . Die hiermit resultierende verallgemeinerte Risikodifferenz ist  $D_G = 0.23$ . Mit den Fehlerwahrscheinlichkeiten I. und II. Art  $\alpha = 0.025$  und  $\beta = 0.1$  ergibt sich in einem nicht-sequentiellen Design als notwendige Fallzahl  $f_{D_G}(0.025, 0.1) = n_{\text{fest}} = 240$ . Als Umfang der ersten Sequenz wird  $n_1 = n_{\text{fest}}/K = 80$  verwendet mit gleichen Fallzahlen in beiden Gruppen  $n_{1j} = n_{2j} = n_j/2 = 40$ . Von den für Gruppe  $J$ ,  $J = 1, 2$ , erzeugten  $n_J = 200$ , Beobachtungen mit den Ausprägungen 1, 2, 3 und 4 werden in Stufe  $j$ ,  $j = 1, \dots, k$ , der verschiedenen Designs  $n_{Jj}$  Werte pro Behandlungsgruppe sequentiell ausgewählt.

Die Ergebnisse der einzelnen Designs sind in Tabelle 3.11 zu sehen. Hierbei bezeichnet  $S_j^* = F_{\chi^2(\lambda_{\Sigma(j)}^{1-2} \Delta \lambda_G^2 \Delta)}^{-1}(F_{\chi^2(\lambda_{\Sigma(j)})}(S_j))$  die Prüfgröße des Wang/Tsiatis-Designs mit inverser  $\chi^2$ -Methode aus Kapitel 2.2.3.

**Tabelle 3.11:** Ergebnisse des fiktiven Beispiels einer klinischen Prüfung mit ordinaler Zielvariable unter Verwendung des Wang/Tsiatis-Designs und des Self-Designings mit inverser Normal- sowie  $\chi^2$ -Methode

Inverse Normalmethode			Inverse $\chi^2$ -Methode			
Wang/Tsiatis-Design						
$j$	1	2	$j$	1	2	
$n_j$	80	50	$n_j$	80	56	
$n_{11,j}$	10	6	$n_{11,j}$	10	7	
$n_{12,j}$	14	10	$n_{12,j}$	14	12	
$n_{13,j}$	12	7	$n_{13,j}$	12	7	
$n_{14,j}$	4	2	$n_{14,j}$	4	2	
$n_{21,j}$	5	4	$n_{21,j}$	5	4	
$n_{22,j}$	11	6	$n_{22,j}$	11	6	
$n_{23,j}$	16	10	$n_{23,j}$	16	13	
$n_{24,j}$	8	5	$n_{24,j}$	8	5	
$D_{Gj}$	0.245	0.258	$D_{Gj}$	0.245	0.325	
$T_j$	2.062	1.710	$T_j$	2.062	2.338	
$p_j$	0.020	0.044	$p_j$	0.020	0.010	
$Z_j/(\lambda_{\Sigma(j)}^\Delta)$	2.062	3.172	$S_j^*$	7.271	13.230	
$\hat{p}_j$		0.137	$\hat{p}_j$		0.120	
$m_j$		50	$m_j$		56	
$M_j$		108	$M_j$		114	
$\hat{D}_{G(2),j}$	0.245	0.250	$\hat{D}_{G(2),j}$	0.245	0.278	
$\hat{D}_{GL,j}$	-0.081	0.034	$\hat{D}_{GL,j}$	-0.056	0.054	
$\hat{D}_{GU,j}$	0.571	0.467	$\hat{D}_{GU,j}$	0.546	0.505	
Self-Designing						
$j$	1	2	3	$j$	1	2
$n_j$	80	54	12	$n_j$	80	56
$n_{11,j}$	10	7	0	$n_{11,j}$	10	7
$n_{12,j}$	14	11	4	$n_{12,j}$	14	12
$n_{13,j}$	12	7	2	$n_{13,j}$	12	7
$n_{14,j}$	4	2	0	$n_{14,j}$	4	2
$n_{21,j}$	5	4	0	$n_{21,j}$	5	4
$n_{22,j}$	11	6	0	$n_{22,j}$	11	6
$n_{23,j}$	16	12	6	$n_{23,j}$	16	13

Inverse Normalmethode				Inverse $\chi^2$ -Methode		
Self-Designing						
$j$	1	2	3	$j$	1	2
$n_{24,j}$	8	5	0	$n_{24,j}$	8	5
$D_{Gj}$	0.245	0.309	0.667	$D_{Gj}$	0.245	0.325
$T_j$	2.062	2.167	3.464	$T_j$	2.062	2.338
$p_j$	0.020	0.015	0.000	$p_j$	0.020	0.010
$\lambda_j$	1.000	1.895	2.105	$\lambda_j$	1.000	1.931
$Z_j/\sqrt{\lambda_{\Sigma}(j)}$	2.062	2.965	4.504	$S_j$	5.446	14.55
$\hat{p}_j$		0.123	0.676	$\hat{p}_j$		0.117
$m_j$		54	10	$m_j$		56
$M_j$		114	12	$M_j$		116
$\hat{D}_{G(2),j}$	0.245	0.271	0.303	$\hat{D}_{G(2),j}$	0.245	0.278
$\hat{D}_{GL,j}$	-0.276	0.037	-0.276	$\hat{D}_{GL,j}$	-0.159	0.024
$\hat{D}_{GU,j}$	0.766	0.521	0.766	$\hat{D}_{GU,j}$	0.649	0.540

Bei Verwendung des Wang/Tsiatis-Designs sowie des Self-Designings mit inverser  $\chi^2$ -Methode endet die klinische Prüfung nach zwei Sequenzen vorzeitig. Insgesamt drei Stufen werden bei dem Self-Designing mit inverser Normalmethode beobachtet. Hier muss eine weitere Sequenz beobachtet werden, da das globale Gewicht  $\lambda_G = 5$  nach der zweiten Stufe nicht verbraucht ist. In allen vier Designs kann die Nullhypothese abgelehnt werden und für die zweite Sequenz resultieren vergleichbare Ergebnisse.

### 3.2.4 Simulationsstudien

Ziel der folgenden Simulationen ist die Untersuchung der adaptiv gruppensequentiellen Designs vom Pocock-, O'Brien/Fleming-, Wang/Tsiatis- und Haybittle/Peto-Typ sowie des Self-Designings klinischer Studien bei Vorliegen einer ordinalen Zielvariable. Sowohl die inverse Normal- als auch die inverse  $\chi^2$ -Methode finden für die Kombination der Ergebnisse der einzelnen Sequenzen Anwendung. Das einseitige Testproblem  $H_0 : D_G = 0$  vs.  $H_1 : D_G > 0$  wird mit der in Kapitel 3.2.1 vorgestellten Prüfgröße  $T_j$  aus (3.44) überprüft. Die Betrachtung der verschiedenen Designs erfolgt sowohl unter der Null- als auch unter der Alternativhypothese. Von Interesse sind die empirische Ablehnwahrscheinlichkeit, der mittlere sowie mediane Stufenumfang, die durchschnittlich benötigte Stufenanzahl und die empirische Überdeckungswahrscheinlichkeit der Sequenz der realisierten Konfidenzintervalle sowie des Intervalls der letzten Stufe des flexiblen Verfahrens.

Die interessierende ordinale Zielgröße kann Ausprägungen in  $m = 4$  Kategorien annehmen, wobei die erste die beste Kategorie darstellt. Für die Planung der Studie werden für die Wahrscheinlichkeiten  $p_{Jl}$ ,  $J = 1, 2$ ,  $l = 1, \dots, m$ , die folgenden Annahmen getroffen. In Gruppe 1 seien  $p_{11} = 0.18$ ,  $p_{12} = 0.44$ ,  $p_{13} = 0.32$  sowie  $p_{14} = 0.06$  und in Gruppe 2 seien  $p_{21} = 0.10$ ,  $p_{22} = 0.35$ ,  $p_{23} = 0.40$ ,  $p_{24} = 0.15$ . Hiermit ist  $D_G = 0.213$  und mit den Fehlerwahrscheinlichkeiten I. und II. Art  $\alpha = 0.025$  und  $\beta = 0.1$  resultiert unter Verwendung der Fallzahlformel (3.45) aus Kapitel 3.2.2 der Umfang  $n_{fest} = 258$  in einem nicht-sequentiellen Design. Diese Annahmen sind hypothetisch und können sich als falsch herausstellen. Als tatsächlich unter der Null- und Alternativhypothese vorliegende Werte für die Ereigniswahrscheinlichkeiten  $p_{Jl}$ ,  $J = 1, 2$ ,  $l = 1, \dots, m$ , werden die in Tabelle 3.12 verzeichneten Fälle betrachtet.

**Tabelle 3.12:** Tatsächlich unter der Null- und Alternativhypothese vorliegende Ereignisraten  $p_{Jl}$ ,  $J = 1, 2$ ,  $l = 1, \dots, m$ , sowie resultierende verallgemeinerte Risikodifferenz  $D_G$  und Fallzahl  $n_{fest}$  in einem nicht-sequentiellen Design

	$J$	$p_{J1}$	$p_{J2}$	$p_{J3}$	$p_{J4}$	$D_G$	$n_{fest}$
Unter $H_0$							
1a)	1,2	0.25	0.25	0.25	0.25	0	-
1b)	1,2	0.70	0.10	0.10	0.10	0	-
1c)	1,2	0.10	0.10	0.10	0.70	0	-
1d)	1,2	0.05	0.15	0.75	0.05	0	-
1e)	1,2	0.15	0.75	0.05	0.05	0	-
Unter $H_1$							
2a)	1	0.58	0.30	0.10	0.02	0.171	384
	2	0.46	0.28	0.16	0.10		
2b)	1	0.22	0.40	0.34	0.04	0.201	304
	2	0.15	0.35	0.28	0.22		
2c)	1	0.18	0.44	0.32	0.06	0.213	258
	2	0.10	0.35	0.40	0.15		
2d)	1	0.45	0.30	0.20	0.05	0.255	182
	2	0.25	0.34	0.26	0.15		
2e)	1	0.52	0.36	0.08	0.04	0.336	96
	2	0.30	0.30	0.20	0.20		

Unter der Alternativhypothese werden eine Unter- sowie Überschätzung der Risikodifferenz bei gleichzeitiger Fehleinschätzung der Größe der Ereigniswahrscheinlichkeiten der beiden Gruppen untersucht (Fälle 2a und 2b sowie 2d und 2e). Zudem wird der Fall

betrachtet, dass die vor Studienbeginn getroffenen Annahmen zutreffen (Fall 2c). Die in einem nicht-sequentiellen Design notwendigen Fallzahlen sind in Klammern angegeben. Als maximale Stufenanzahl wird bei den adaptiv gruppensequentiellen Designs  $K = 3$  sowie  $K = 5$  gewählt. Bei dem Self-Designing wird das globale Gewicht festgelegt als  $\lambda_G = 5$  und  $\lambda_G = 10$ . Die für die Berechnung des Umfangs  $m_j(\beta_g)$ ,  $j = 1, \dots, k$ , notwendige Fehlerwahrscheinlichkeit  $\beta_g$  wird auf die Werte 0.3 und 0.8 gesetzt.

Für die erste Stufe wird der Umfang  $n_1 = n_{fest}/K = 258/3 \approx 86$  beziehungsweise  $n_1 = 258/5 \approx 52$  betrachtet. Als minimaler Stufenumfang wird  $n_{j,\min} = 6$  gewählt. Als maximale Werte für die Fallzahlen  $m_j(\beta_g)$  und  $M_j(\beta)$  werden im Fall von drei Sequenzen  $m_{j,\max} = 2/5 n_{fest} \approx 104$  oder  $m_{j,\max} = 2/3 n_{fest} \approx 172$  und  $M_{j,\max} = n_{fest}/2 \approx 130$  oder  $M_{j,\max} = n_{fest} = 258$  und im Fall von fünf Sequenzen  $m_{j,\max} = n_{fest}/4 \approx 66$  oder  $m_{j,\max} = n_{fest}/2 \approx 130$  und  $M_{j,\max} = n_{fest}/3 \approx 86$  oder  $M_{j,\max} = n_{fest} = 258$  verwendet. Insgesamt werden die in Tabelle 3.13 angegebenen Konstellationen der Parameter in Simulationen mit 10000 Wiederholungen betrachtet.

**Tabelle 3.13:** Einstellungen der Parameter in den Simulationsstudien flexibler Designs in klinischen Prüfungen mit ordinaler Zielvariable (Fehlerraten I. und II. Art  $\alpha = 0.025$  und  $\beta = 0.1$ , minimaler Stufenumfang  $n_{j,\min} = 6$ )

Konstellation	$\beta_g$	$K$	$\lambda_i$	$\lambda_G$	$\lambda_1$	$\lambda_{\min}$	$n_1$	$m_{j,\max}$	$M_{j,\max}$
I	0.3	5	1	10	2	1	52	66	86
II	0.3	5	1	10	2	1	52	130	258
III	0.8	5	1	10	2	1	52	66	86
IV	0.8	5	1	10	2	1	52	130	258
V	0.3	3	1	5	1	1	86	104	130
VI	0.3	3	1	5	1	1	86	172	258
VII	0.8	3	1	5	1	1	86	104	130
VIII	0.8	3	1	5	1	1	86	172	258

Die Ergebnisse von Konstellation I unter der Null- sowie unter der Alternativhypothese sind in den Tabellen 3.14 und 3.15 zu sehen, die Resultate der Konstellationen II-VIII befinden sich in Anhang C.2. In diesen Tabellen findet sich in der ersten Spalte die empirische Ablehnwahrscheinlichkeit eines nicht-sequentiellen Designs (NSD), in der zweiten Spalte das verwendete flexible Design (PO=Pocock, OBF=O'Brien/Fleming, WT=Wang/Tsiatis, HP=Haybittle/Peto, SD=Self-Designing) und in den folgenden Spal-

ten die empirische Ablehnwahrscheinlichkeit, der mittlere sowie mediane Stufenumfang, die durchschnittlich benötigte Stufenanzahl und die empirische Überdeckungswahrscheinlichkeit der Sequenz der realisierten Konfidenzintervalle sowie des Intervalls der letzten Stufe der verschiedenen flexiblen Verfahren für die inverse Normal– sowie für die inverse  $\chi^2$ –Methode.

**Tabelle 3.14:** Empirisches Niveau  $\hat{\alpha}\%$  im nicht–sequentiellen Design (NSD) sowie  $\hat{\alpha}\%$ , mittlerer und medianer Umfang (ASN = „Average Sample Number“,  $N_{med}$ ), mittlere Stufenanzahl  $\bar{k}$  und empirische Überdeckungswahrscheinlichkeiten  $(1 - 2\hat{\alpha})\%$  der Sequenz der beobachteten Intervalle  $\{KI_j\}$  sowie des finalen Intervalls  $KI_k$  in flexiblen Designs mit ordinaler Zielvariable unter der Nullhypothese (Konstellation I)

NSD	Design	Inverse Normalmethode						Inverse $\chi^2$ –Methode					
		$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$ $\{KI_j\}$	$KI_k$	$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$ $KI_k$	$\{KI_j\}$
1a) $\vec{p}_1 = \vec{p}_2 = (0.25, 0.25, 0.25, 0.25)$ , $D_G = 0.0$													
2.66	PO	3.02	336.27	342	4.92	93.81	95.90	3.49	335.13	342	4.91	93.12	94.89
	OBF	2.94	337.21	342	4.97	93.88	94.42	3.59	336.00	342	4.96	92.88	92.88
	WT	3.07	337.15	342	4.96	93.99	95.08	3.75	336.18	342	4.95	92.47	92.99
	HP	3.24	339.40	342	4.97	93.71	94.10	3.49	335.53	342	4.98	92.68	92.71
	SD	2.76	208.26	210	3.03	93.27	94.31	3.19	207.58	210	3.01	93.61	93.61
1b) $\vec{p}_1 = \vec{p}_2 = (0.70, 0.10, 0.10, 0.10)$ , $D_G = 0.0$													
2.37	PO	3.20	335.66	342	4.91	93.54	95.78	3.17	335.99	342	4.92	93.38	95.47
	OBF	2.96	337.60	342	4.98	94.05	94.56	3.56	336.46	342	4.97	93.09	93.09
	WT	2.73	337.74	342	4.96	94.23	95.36	3.39	336.35	342	4.95	93.32	93.83
	HP	3.10	339.69	342	4.98	93.83	94.09	3.49	335.71	342	4.98	93.19	93.21
	SD	2.95	208.02	210	3.02	92.95	94.00	3.11	207.66	210	3.01	93.43	93.43
1c) $\vec{p}_1 = \vec{p}_2 = (0.10, 0.10, 0.10, 0.70)$ , $D_G = 0.0$													
2.69	PO	2.94	336.21	342	4.92	94.22	96.15	3.36	335.33	342	4.91	93.49	95.34
	OBF	2.69	337.69	342	4.98	94.41	94.84	3.24	336.53	342	4.97	93.17	93.17
	WT	2.89	337.63	342	4.96	94.44	95.37	3.14	336.62	342	4.95	93.60	94.05
	HP	2.83	339.97	342	4.98	94.17	94.56	3.24	335.88	342	4.98	93.52	93.59
	SD	2.67	208.19	210	3.03	93.68	94.72	3.14	207.82	210	3.01	93.80	93.80
1d) $\vec{p}_1 = \vec{p}_2 = (0.05, 0.15, 0.75, 0.05)$ , $D_G = 0.0$													
2.49	PO	3.30	335.93	342	4.92	93.65	95.74	3.53	335.30	342	4.91	93.14	95.04
	OBF	3.00	337.27	342	4.98	94.12	94.75	3.39	336.00	342	4.97	92.94	92.94
	WT	3.34	337.05	342	4.96	93.85	94.84	3.79	335.90	342	4.94	93.13	93.60
	HP	3.42	339.55	342	4.98	93.88	94.18	3.68	334.95	342	4.97	92.80	92.84
	SD	3.10	207.76	210	3.03	92.96	94.12	3.16	207.85	210	3.01	93.36	93.36
1e) $\vec{p}_1 = \vec{p}_2 = (0.15, 0.75, 0.05, 0.05)$ , $D_G = 0.0$													
2.49	PO	3.62	334.73	342	4.90	92.85	95.34	4.09	333.96	342	4.89	92.06	94.27
	OBF	3.26	337.00	342	4.97	93.72	94.24	4.03	335.41	342	4.96	92.17	92.17
	WT	3.57	336.53	342	4.95	93.42	94.56	3.81	335.27	342	4.94	92.46	92.84
	HP	3.33	339.53	342	4.98	93.20	93.65	3.94	334.80	342	4.97	92.17	92.17
	SD	2.92	208.05	210	3.02	92.88	94.09	3.76	207.40	210	3.01	92.86	92.86

Für die einzelnen Designs resultieren für die inverse Normal- sowie für die inverse  $\chi^2$ -Methode vergleichbare Ergebnisse, wobei das empirische Niveau bei Verwendung der inversen Normalmethode tendenziell geringer als für die inverse  $\chi^2$ -Methode ausfällt.

Erfolgt die Überprüfung des Testproblems in dem nicht-sequentiellen Design, besitzt das empirische Niveau ungefähr den für  $\alpha$  vorgegebenen Wert von 2.5%. In den flexiblen Designs ist eine Erhöhung von diesem zu erkennen. Teilweise ist das empirische Niveau wesentlich größer als 2.5%, da die Approximation der Verteilung der Prüfgröße durch die Normalverteilung aufgrund kleiner Umfänge in den einzelnen Sequenzen nicht immer zutreffend zu sein scheint. Hier müssten größere minimale und maximale Umfänge zugelassen werden, was zu höheren mittleren und medianen Umfängen führen kann.

Die adaptiv gruppensequentiellen Designs enden im Mittel nach Durchführung der vorgegebenen maximalen Stufenanzahl  $K = 5$ . Für das Self-Designing werden durchschnittlich nur drei Sequenzen benötigt. Der mittlere sowie der mediane Gesamtumfang (ASN,  $N_{med}$ ) der Studie sind unter der Nullhypothese ungefähr gleich groß. Bei den adaptiv gruppensequentiellen Designs sind ASN und  $N_{med}$  im Vergleich zu dem Self-Designing wesentlich größer, da bei ersteren im Mittel zwei Sequenzen mehr beobachtet werden. Im Vergleich zu dem nicht-sequentiellen Design fällt der Gesamtumfang in den flexiblen Designs größer aus, da der vor Studienbeginn angenommene Unterschied  $D_G = 0.213$  nicht vorliegt. Die empirische Überdeckungswahrscheinlichkeit der Konfidenzintervalle ist unter  $H_0$  ungefähr 95%, da im Mittel fast alle geplanten Sequenzen beobachtet werden.

Die Simulationsergebnisse für Konstellation I unter der Alternativhypothese sind in Tabelle 3.15 zu sehen.

**Tabelle 3.15:** Empirische Power  $(1-\hat{\beta})\%$  im nicht-sequentiellen Design (NSD) sowie  $(1-\hat{\beta})\%$ , mittlerer und medianer Umfang (ASN = „Average Sample Number“,  $N_{med}$ ), mittlere Stufenanzahl  $\bar{k}$  und empirische Überdeckungswahrscheinlichkeiten  $(1-2\hat{\alpha})\%$  der Sequenz der beobachteten Intervalle  $\{KI_j\}$  sowie des finalen Intervalls  $KI_k$  in flexiblen Designs mit ordinaler Zielvariable unter der Alternativhypothese (Konstellation I)

NSD ( $1-\hat{\beta})\%$	Design	Inverse Normalmethode					Inverse $\chi^2$ -Methode						
		$(1-\hat{\beta})\%$	ASN	$N_{med}$	$\bar{k}$	$(1-2\hat{\alpha})\%$ $\{KI_j\}$ $KI_k$	$(1-\hat{\beta})\%$	ASN	$N_{med}$	$\bar{k}$	$(1-2\hat{\alpha})\%$ $KI_k$ $\{KI_j\}$		
2a) $\vec{p}_1 = (0.58, 0.30, 0.10, 0.02)$ , $\vec{p}_2 = (0.46, 0.28, 0.16, 0.10)$ , $D_G = 0.171$													
75.27	PO	79.36	215.46	190	3.30	95.45	97.35	73.14	223.12	256	3.40	94.82	96.24
	OBF	81.59	231.74	236	4.06	95.75	96.15	78.00	231.96	248	4.07	92.39	92.39
	WT	82.34	225.43	228	3.67	96.52	97.41	78.05	228.82	246	3.74	95.11	95.58
	HP	83.86	253.89	262	4.04	95.72	95.96	79.03	235.87	256	4.33	90.05	90.12
	SD	58.47	174.48	210	3.25	88.80	89.78	60.51	171.93	210	2.91	92.96	92.96
2b) $\vec{p}_1 = (0.22, 0.40, 0.34, 0.04)$ , $\vec{p}_2 = (0.15, 0.35, 0.28, 0.22)$ , $D_G = 0.201$													
83.47	PO	87.72	193.48	190	3.00	95.72	97.34	83.02	200.93	190	3.09	94.60	95.97
	OBF	89.02	209.09	192	3.82	94.78	95.15	87.08	206.13	190	3.79	91.80	91.80
	WT	90.52	199.45	190	3.35	96.62	97.64	86.67	205.92	190	3.46	94.69	95.19
	HP	91.35	230.87	256	3.72	95.71	96.00	87.59	212.56	218	4.08	88.73	88.82
	SD	67.85	165.88	182	3.29	87.38	88.43	70.39	161.41	188	2.85	91.72	91.72
2c) $\vec{p}_1 = (0.18, 0.44, 0.32, 0.06)$ , $\vec{p}_2 = (0.10, 0.35, 0.40, 0.15)$ , $D_G = 0.213$													
89.10	PO	92.24	177.89	188	2.78	95.74	97.29	88.17	184.49	190	2.87	94.80	96.09
	OBF	92.73	195.32	190	3.69	94.66	95.10	91.09	191.36	190	3.62	91.83	91.83
	WT	93.91	185.74	184	3.18	96.45	97.40	91.11	189.84	190	3.27	94.89	95.14
	HP	94.67	214.38	248	3.49	96.20	96.38	92.05	196.83	196	3.90	89.54	89.59
	SD	73.90	160.05	164	3.33	86.62	87.75	77.02	154.03	154	2.81	92.07	92.07
2d) $\vec{p}_1 = (0.45, 0.30, 0.20, 0.05)$ , $\vec{p}_2 = (0.25, 0.34, 0.26, 0.15)$ , $D_G = 0.255$													
96.57	PO	98.25	141.69	124	2.27	95.93	97.62	96.85	146.61	124	2.33	95.46	96.74
	OBF	97.49	160.73	146	3.32	93.93	94.37	97.25	155.87	132	3.25	90.15	90.15
	WT	98.50	153.45	132	2.76	96.75	97.66	97.48	153.57	124	2.80	95.24	95.48
	HP	98.75	177.79	190	2.92	96.72	96.90	98.01	161.09	148	3.50	88.21	88.33
	SD	84.63	141.73	130	3.35	83.05	84.26	88.58	135.67	124	2.69	90.78	90.78
2e) $\vec{p}_1 = (0.52, 0.36, 0.08, 0.04)$ , $\vec{p}_2 = (0.30, 0.30, 0.20, 0.20)$ , $D_G = 0.336$													
99.94	PO	100.00	95.41	58	1.59	96.46	97.66	99.97	96.54	58	1.60	96.03	96.91
	OBF	99.39	110.80	100	2.75	91.52	91.94	99.66	102.73	84	2.47	89.16	89.16
	WT	99.98	103.39	92	2.03	97.46	98.03	99.96	100.52	82	2.00	95.62	95.79
	HP	99.97	118.96	124	1.98	97.92	98.09	99.91	105.60	85	2.59	88.45	88.82
	SD	93.49	109.58	100	3.33	76.96	77.89	97.64	98.78	84	2.35	87.00	87.00

Für die einzelnen Designs resultieren für die inverse Normal- sowie für die inverse  $\chi^2$ -Methode unter der Alternativen vergleichbare Ergebnisse. Dies konnte bereits unter der Nullhypothese beobachtet werden.

Ist die tatsächlich vorliegende verallgemeinerte Risikodifferenz geringer als vor Studienbeginn angenommen (Fälle 2a und 2b), ist die Fallzahl  $n_{fest} = 258$  nicht ausreichend

um diesen Unterschied mit der gewünschten Power von 90% zu entdecken. Im nicht-sequentiellen Design liegt die mit diesem Umfang resultierende empirische Power unterhalb des vorgegebenen Werts von 90% in beiden Fällen. Bei den flexiblen Designs gilt dies für Fall 2a sowie einige Parameterkonstellationen von Fall 2b, da durch die Beschränkungen der sequentiellen Fallzahlen der mittlere sowie mediane Gesamtumfang der Untersuchung zu gering ausfallen. Durch die Wahl größerer maximaler Umfänge in den einzelnen Sequenzen kann die empirische Power erhöht werden. In Fall 2b ist diese in einigen Konstellationen ungefähr 90%. Hierbei ist der ASN geringer als der für diesen Fall in einem nicht-sequentiellen Design resultierende Umfang  $n_{fest} = 304$ . Für das Haybittle/Peto-Design ist der ASN im Vergleich zu den anderen adaptiv gruppensequentiellen Designs relativ hoch bei nur geringfügig größerer empirischer Power. Insgesamt resultiert für die adaptiv gruppensequentiellen Designs mit inverser Normalmethode in fast allen betrachteten Parameterkonstellationen eine höhere empirische Power als bei der inversen  $\chi^2$ -Methode bei häufig geringerem ASN. In den Konstellationen mit  $\beta_g = 0.8$  ist die empirische Power für die inverse  $\chi^2$ -Methode oft größer als für die inverse Normalmethode mit Ausnahme des Haybittle/Peto-Designs. Die Ergebnisse für das Self-Designing sind für beide Kombinationsmethoden ähnlich, die mittlere Stufenanzahl ist für die inverse Normalmethode tendenziell etwas größer.

Sind die vor Studienbeginn getroffenen Annahmen für die Ereigniswahrscheinlichkeiten richtig (Fall 2c), ist die empirische Power in den adaptiv gruppensequentiellen Designs größer als 90%. Hierbei fällt der durchschnittliche Stichprobenumfang geringer aus als in einem nicht-sequentiellen Design, in welchem die gewünschte Power von 90% annähernd erreicht wird. Im Self-Designing ist bei beiden Kombinationsmethoden die empirische Power kleiner als in den übrigen flexiblen Designs, wobei dies in den geringeren mittleren Stichprobenumfängen begründet ist. Insgesamt fallen ASN und  $N_{med}$  kleiner aus als die in einem nicht-sequentiellen Design verwendete Fallzahl  $n_{fest} = 258$ .

Ist die verallgemeinerte Risikodifferenz größer als vermutet (Fall 2d und 2e), resultiert in einem nicht-sequentiellen Design mit  $n_{fest} = 258$  eine empirische Power, welche den vorgegebenen Wert von 90% überschreitet. In den flexiblen Designs ist die empirische Power in den meisten betrachteten Konstellationen größer als der vorgegebene Wert 90%.

Aufgrund der Möglichkeit des vorzeitigen Abbruchs der Studie sind sowohl die Sequenz

der Konfidenzintervalle als auch das finale Intervall konservativ (vgl. Kapitel 2.5.1) und in den Konstellationen, in denen die mittlere Stufenanzahl wesentlich geringer als die geplante Anzahl Sequenzen ist, ist die resultierende empirische Überdeckungswahrscheinlichkeit wesentlich größer als das festgelegte Konfidenzniveau von 95%. In einigen Fällen fällt die empirische Überdeckungswahrscheinlichkeit sehr gering aus, welche durch größere minimale und maximale Stufenumfänge erhöht werden kann. Bei dem Self-Designing mit inverser Normalmethode ist die empirische Überdeckungswahrscheinlichkeit der Sequenz der Konfidenzintervalle teilweise deutlich geringer als 95%, da das Konfidenzniveau in diesem Fall nicht kontrolliert wird.

Insgesamt sind die Ergebnisse der Simulationsstudien flexibler Designs mit ordinaler Zielvariable vergleichbar mit den im binären Fall erhaltenen Resultaten. Als Unterschied ist zu erkennen, dass das empirische Niveau bei Betrachtung einer ordinalen Zielgröße stärker erhöht ist als im binären Fall.

## 4 Zusammenfassung

Basierend auf zwei Verfahren zur Kombination von p–Werten, der gewichteten inversen Normalmethode sowie der verallgemeinerten inversen  $\chi^2$ –Methode, wurden adaptiv gruppensequentielle Designs sowie das Self–Designing klinischer Studien vorgestellt. Diese flexiblen Verfahren bieten die Möglichkeit, im Verlauf der Untersuchung Änderungen am Studiendesign vorzunehmen, wobei in dieser Arbeit der Fokus auf der Anpassung der Fallzahl lag. Zudem kann die klinische Prüfung in den Zwischenauswertungen mit Ablehnung der Nullhypothese vorzeitig beendet werden. Hierbei ist zu beachten, dass bei dem Self–Designing mit inverser Normalmethode nur am Studienende das Testen der Nullhypothese unter Einhaltung des globalen Niveaus erfolgen kann.

Bei den adaptiv gruppensequentiellen Testverfahren ist die Anzahl der Stufen sowie die Gewichtung der Sequenzen vor Studienbeginn fest vorgegeben. Aufgrund des mehrfachen Überprüfens der Nullhypothese ist eine Anpassung der Entscheidungsbereiche erforderlich. Hierfür wurden verschiedene Strategien vorgestellt und anschließend hinsichtlich der Aufteilung des globalen Niveaus miteinander verglichen. Die bekanntesten Verfahren zur Herleitung kritischer Werte stammen von Pocock (1977) sowie von O’Brien und Fleming (1979). Als Spezialfälle sind diese in einer von Wang und Tsiatis (1987) vorgeschlagenen allgemeineren Klasse kritischer Schranken enthalten. Neben diesen Vorgehensweisen zur Adjustierung der Entscheidungsbereiche wurde das sehr einfache Prinzip des sequentiellen Testens von Haybittle (1971) und Peto et al. (1976) betrachtet. Die Überprüfung des Testproblems in den Sequenzen  $j \in \{1, \dots, K - 1\}$  erfolgt hier mit dem Wert  $\sqrt{j} \cdot 3$  ( $\approx \sqrt{j} \cdot \Phi^{-1}(0.9987)$ ) für die Stouffer–Statistik bzw. dem  $F_{\chi^2(\lambda_{\Sigma}(j))}^{-1}(0.9987)$ –Quantil für die ungewichtete  $\chi^2$ –Statistik. Bei diesem Vorgehen kann in der letzten Stufe  $K$  approximativ ein Test zum Niveau  $\alpha$  mit dem kritischen Wert eines nicht–sequentiellen einstufigen Designs durchgeführt werden. Dieses Prinzip wurde leicht modifiziert, indem anstelle des 0.9987–Quantils der entsprechenden Verteilung jeweils das 0.995–Quantil betrachtet wurde. Bei Kombination der Ergebnisse der einzelnen Sequenzen mit der inversen Normalmethode können die kritischen Werte der klassischen gruppensequentiellen Verfahren verwendet werden. Für die inverse  $\chi^2$ –Methode wurden das Pocock– sowie O’Brien/Fleming–Design in Hartung (2006) übertragen. Eine Erweiterung des Wang/Tsiatis– und des Haybittle/Peto–Designs erfolgte für diese Kombinationsmethode in dieser Arbeit. In einem Vergleich für den Verbrauch des globalen Niveaus zeigte sich, dass die verschiedenen

Designs für beide Kombinationsmethoden ein sehr ähnliches Verhalten aufweisen. Insgesamt wird in dem O'Brien/Fleming-Design erst in den letzten Stufen ein Großteil des globalen Niveaus vergeben, bei dem Pocock-Design hingegen bereits zu Beginn der Untersuchung. Das Wang/Tsiatis-Design liegt bei Wahl des Parameters  $\Delta = 0.25$  zwischen diesen beiden Verfahren. Im Haybittle/Peto-Design ist die Wahrscheinlichkeit für ein vorzeitiges Beenden der Studie mit Ablehnung der Nullhypothese in den Sequenzen  $1, \dots, K-1$  sehr gering. In der modifizierten Version dieses Designs kam es zu einer Erhöhung des in diesen Stufen vergebenen Anteils des globalen Niveaus.

Besteht der Wunsch, die Anzahl der Sequenzen vor Studienbeginn nicht festzulegen, stellen das Self-Designing klinischer Studien mit gewichteter inverser Normalmethode und verallgemeinerter inverser  $\chi^2$ -Methode geeignete Verfahren dar (vgl. Hartung (2006)). Hier wird durch die Vergabe der Varianz der finalen Prüfgröße das Ende der Studie gesteuert. Ist die Varianz noch nicht komplett verbraucht, ist die Verteilung der Kombinationsstatistiken nicht bekannt. Mittels der Betrachtung von Martingalen und der charakteristischen Funktion erfolgte der Nachweis der Verteilung der finalen Teststatistiken. Bei Verwendung der inversen  $\chi^2$ -Methode ist die Prüfgröße monoton steigend und somit kann in den Zwischenauswertungen das Testen der Nullhypothese zu dem globalen kritischen Wert erfolgen. Dies ist für die inverse Normalmethode nicht der Fall. Erst nach Vergabe des globalen Gewichts  $\lambda_G$  ist unter Verwendung der finalen  $N(0, \lambda_G)$ -verteilten Prüfgröße die Überprüfung des Testproblems unter Einhaltung des vorgegebenen Niveaus möglich. Die Gewichtung der einzelnen Sequenzen wird im Self-Designing ebenso wie die Fallzahl basierend auf den vorliegenden Beobachtungen ermittelt.

Neben der Überprüfung von Testproblemen ist die Abschätzung der Größe eines Wirkungsunterschieds von Interesse. Hierfür wurden Methoden zur Intervall- und Punktschätzung vorgestellt. Mittels so genannter Repeated Konfidenzintervalle kann in jeder Stufe ein Intervall angegeben werden, welches eine Überdeckungswahrscheinlichkeit größer oder gleich dem geforderten Konfidenzniveau besitzt (vgl. Jennison und Turnbull (1989), Hartung und Knapp (2006)). Erfolgt kein vorzeitiger Abbruch der Studie, hält die Sequenz der Intervalle bei den adaptiv gruppensequentiellen Designs mit inverser Normalmethode sowie dem Pocock-Design mit inverser  $\chi^2$ -Methode das gewünschte Konfidenzniveau genau ein. Für das O'Brien/Fleming-Design sowie das Self-Designing mit inverser  $\chi^2$ -Methode besitzt das Intervall der letzten Stufe, d.h. wenn das globale Gewicht komplett verbraucht

ist, exakt das festgelegte Niveau. Für das Self-Designing mit inverser Normalmethode besteht in Analogie zum einmaligen Testen nur am Studienende die Möglichkeit ein Konfidenzintervall anzugeben, welches das vorgegebene Niveau einhält. Für die Ermittlung zweiseitiger Konfidenzintervalle in den adaptiv gruppensequentiellen Designs mit inverser Normalmethode werden kritische Werte für zweiseitige Tests benötigt. In einem nicht-sequentiellen Design können für die Überprüfung eines zweiseitigen Testproblems zum Niveau  $2\alpha$  die einseitigen kritischen Werte zum Niveau  $\alpha$  verwendet werden. Im Mehrdimensionalen ist dies im Allgemeinen nicht erfüllt (vgl. Wassmer (2001)). Für gebräuchliche Werte des Signifikanzniveaus sowie bei gleicher Gewichtung der einzelnen Sequenzen sind die ein- und zweiseitigen kritischen Werte in der Regel gleich. Bei sehr unterschiedlicher Gewichtung der Stufen können sich diese bereits für kleinere Werte für das Niveau unterscheiden. Für die Ermittlung eines einzelnen Schätzwerts am Studienende wurden in der Meta-Analyse verwendete Verfahren vorgestellt (vgl. Hartung und Knapp (2005)). Angewandt in klinischen Prüfungen mit flexiblem Designs sind diese Punktschätzer in der Regel nicht unverzerrt für den unbekannt Parameter, wobei es verschiedene Verzerrungsquellen gibt. Soll beispielsweise am Ende der Untersuchung eine Schätzung für den interessierenden Parameter angegeben werden und findet ein vorzeitiger Abbruch der Studie Berücksichtigung, erfolgt eine Selektion der Daten. Hierdurch kann es passieren, dass diese Schätzung auf vergleichsweise extremen Daten basiert (vgl. Coburger und Wassmer (2001, 2003)). Vorteilhaft an diesen in der Meta-Analyse verwendeten Schätzern ist, dass sie leicht zu ermitteln sind. Für den Fall, dass lediglich eine Anpassung der Fallzahl und kein vorzeitiger Abbruch der Untersuchung Berücksichtigung findet, wurde ein erwartungstreuer Schätzer hergeleitet.

Die in der Literatur vorgestellten flexiblen Designs basieren in der Regel auf der Annahme stetiger Daten, wobei häufig die Einschränkung auf normalverteilte Zielvariablen mit bekannter Varianz erfolgt. Deshalb wurde in dieser Arbeit das Vorliegen diskreter Beobachtungen untersucht. Hierbei war insbesondere von Interesse, ob Prüfgrößen, welche bereits in nicht-sequentiellen Designs das Testniveau nur asymptotisch einhalten, in flexibel durchgeführten Untersuchungen Anwendung finden können. Bei Vorliegen einer binären Zielvariable wurde für die Messung des Behandlungsunterschieds die Risikodifferenz verwendet. Eine verallgemeinerte Risikodifferenz wurde für den Fall einer ordinalen Zielgröße betrachtet. Es erfolgte für beide Situationen die Ermittlung einer Stichprobenfunktion, welche anschließend auf den sequentiellen Fall übertragen wurde. Die für

den Fall ordinaler Daten auf der Betrachtung der verallgemeinerten Risikodifferenz hergeleitete Fallzahl stimmt bei Vorliegen von nur zwei Kategorien mit einer für die binäre Situation bekannten Stichprobenfunktion überein. Die adaptiv gruppensequentiellen Designs vom Pocock-, O'Brien/Fleming-, Wang/Tsiatis- und Haybittle/Peto-Typ sowie das Self-Designing mit inverser Normal- und  $\chi^2$ -Methode wurden bei Vorliegen einer binären oder ordinalen Zielgröße in Simulationen umfangreich untersucht. Im binären Fall erfolgte das Testen der Nullhypothese in den einzelnen Sequenzen mit Prüfgrößen ohne und mit Stetigkeitskorrektur. Es zeigte sich, dass das liberale bzw. konservative Testverhalten dieser Statistiken in flexiblen Designs verstärkt wird. Bei Betrachtung einer ordinalen Variable ergaben sich mit den im binären Fall erhaltenen Ergebnissen vergleichbare Resultate, wobei das empirische Niveau im ordinalen Fall tendenziell höher ausfiel. Insgesamt können flexible Designs in klinischen Prüfungen mit binärer und ordinaler Zielvariable Anwendung finden, wobei von der Verwendung von Teststatistiken mit Stetigkeitskorrektur abgeraten wird. Die Wahl des flexiblen Designs sollte nach der Zielsetzung der Studie erfolgen. Das mittlere Verhalten der adaptiv gruppensequentiellen Designs ist sehr ähnlich mit Ausnahme des Designs vom Haybittle/Peto-Typ. In den Simulationsstudien ergaben sich für dieses Design sehr hohe mittlere und mediane Fallzahlen bei nur geringfügig größeren Werten für die empirische Power. Bei dem Self-Designing mit inverser Normalmethode besteht der Nachteil, dass die Überprüfung des Testproblems nicht in den einzelnen Sequenzen erfolgen kann. Dies ist erst möglich, wenn das globale Niveau komplett verbraucht ist. Das Self-Designing mit inverser  $\chi^2$ -Methode ermöglicht das Testen der Nullhypothese in jeder Sequenz und stellt insgesamt eine gute Alternative zu den adaptiv gruppensequentiellen Designs dar, wenn die maximale Stufenanzahl nicht festgelegt werden soll.

## Literatur

- [1] Armitage, P., McPherson, C. K., Rowe, B. C. (1969). Repeated significance tests on accumulating data. *Journal of the Royal Statistical Society, Series A*, 132, 235–244.
- [2] Bauer, P. (1989). Multistage testing with adaptive designs. *Biometrie und Informatik in Medizin und Biologie* 20, 130–148.
- [3] Bauer, P. (1992). The choice of sequential boundaries based on the concept of power spending. *Biometrie und Informatik in Medizin und Biologie* 23, 3–15.
- [4] Bauer, P., Kieser, M. (1999). Combining different phases in the development of medical treatments within a single trial. *Statistics in Medicine* 18, 1833–1848.
- [5] Bauer, P., Köhne, K. (1994). Evaluation of experiments with adaptive interim analyses. *Biometrics* 50, 1029–1041.
- [6] Bauer, P., Röhmel, J. (1995). An adaptive method for establishing a dose–response relationship. *Statistics in Medicine* 14, 1595–1607.
- [7] Bock, J. (1998). *Bestimmung des Stichprobenumfangs*. R. Oldenbourg Verlag, München.
- [8] Brannath, W., Posch, M., Bauer, P. (2002). Recursive combination tests. *Journal of the American Statistical Association* 97, 236–244.
- [9] Coburger, S., Wassmer, G. (2001). Conditional point estimation in adaptive group sequential test designs. *Biometrical Journal* 43, 821–833.
- [10] Coburger, S., Wassmer, G. (2003). Sample size reassessment in adaptive clinical trials using a bias corrected estimate. *Biometrical Journal* 45, 812–825.
- [11] Edwardes, M. D., Baltzan, M. (2000). The generalization of the odds ratio, risk ratio and risk difference to  $r \times k$  tables. *Statistics in Medicine* 19, 1901–1914.
- [12] Fahrmeir, L., Kaufmann, H. L., Ost, F. (1985). *Stochastische Prozesse. Eine Einführung in Theorie und Anwendungen*, Hanser Verlag, München.
- [13] Fisher, L. (1932). *Statistical methods for research workers*. 4. Auflage, Oliver and Boyd, London.

- 
- [14] Fisher, L.D. (1998). Self-designing clinical trials. *Statistics in Medicine* 17, 1551–1562.
- [15] Goodman, L. A., Kruskal, W. H. (1972). Measures of association for cross classification, IV: simplification of asymptotic variances. *Journal of the American Statistical Association* 67, 415–421.
- [16] Graybill, F. A., Deal, R. B. (1959). Combining unbiased estimators. *Biometrics* 15, 543–550.
- [17] Hartung, J. (2000). A new class of self-designing clinical trials. Erschienen in: Hasman, A., Blobel, B., Dudeck, J., Engelbrecht, R., Gell. G., Prokosch, H.-U. (Hrsg.). *Medical Infobahn for Europe*. Proceedings of MIE 2000 and GMDS 2000. IOS Press, Amsterdam, 310–314.
- [18] Hartung, J. (2001). A self-designing rule for clinical trials with arbitrary response variables. *Controlled Clinical Trials* 22, 111–116.
- [19] Hartung, J. (2006). Flexible designs by adaptive plans of generalized Pocock- and O'Brien-Fleming-type and by self-designing clinical trials. *Biometrical Journal* 48, 521–536.
- [20] Hartung, J., Elpelt, B., Klösener, K.-H. (1999). *Statistik – Lehr- und Handbuch der angewandten Statistik*. 12. Auflage, R. Oldenbourg Verlag, München.
- [21] Hartung, J., Knapp, G. (2001). A refined method for the meta-analysis of controlled clinical trials with binary outcome. *Statistics in Medicine* 20, 3875–3889.
- [22] Hartung, J., Knapp, G. (2003). A new class of completely self-designing clinical trials. *Biometrical Journal* 45, 3–19.
- [23] Hartung, J., Knapp, G. (2004). Improved tests of homogeneity in randomized controlled multi-center trials with binary outcome. *Far East Journal of Theoretical Statistics* 13, 101-126.
- [24] Hartung, J., Knapp, G. (2005). Models for combining results of different experiments: retrospective and prospective. *American Journal of Mathematical and Management Sciences* 25, 149–188.

- [25] Hartung, J., Knapp, G. (2006). Repeated confidence intervals in self-designing clinical trials and switching between noninferiority and superiority, *Biometrical Journal* 48, 697–709.
- [26] Haybittle, J. L. (1971). Repeated assessment of results in clinical trials of cancer treatment. *British Journal of Radiology* 44, 793–797.
- [27] Jennison, C., Turnbull, B. W. (1984). Repeated confidence intervals for group sequential clinical trials. *Controlled Clinical Trials* 5, 33–45.
- [28] Jennison, C., Turnbull, B. W. (1989). Interim analyses: the repeated confidence interval approach. *Journal of the Royal Statistical Society, Series B*, 51, 305–361.
- [29] Jennison, C., Turnbull, B. W. (2000). *Group sequential methods with applications to clinical trials*. Chapman & Hall, Boca Raton, London, New York, Washington, D.C.
- [30] Lan, K. K. G., DeMets, D. L. (1983). Discrete sequential boundaries for clinical trials. *Biometrika* 70, 659–663.
- [31] Lehmacher, W., Wassmer, G. (1999). Adaptive sample size calculations in group sequential trials. *Biometrics* 55, 1286–1290.
- [32] Mood, A. M., Graybill, F. A., Boes, D. C. (1974). *Introduction to the theory of statistics*. 3. Auflage, McGraw-Hill Book Company, Singapore.
- [33] Müller, P. H. (Hrsg.) (1991). *Lexikon der Stochastik*. Akademie Verlag, Berlin.
- [34] O'Brien, P. C., Fleming, T. R. (1979). A multiple testing procedure for clinical trials. *Biometrics* 35, 549–556.
- [35] Peto, R., Pike, M. C., Armitage, P., Breslow, N. E., Cox, D. R., Howard, S. V., Mantel, N., McPherson, K. Peto, J., Smith, P. G. (1976). Design and analysis of randomized clinical trials requiring prolonged observation for each patient. I. introduction and design. *British Journal of Cancer* 34, 585–612.
- [36] Pocock, S. J. (1977). Group sequential methods in the design and analysis of clinical trials. *Biometrika* 64, 191–199.

- [37] Proschan, M. A., Hunsberger, S. A. (1995). Designed extension of studies based on conditional power. *Biometrics* 51, 1315–1324.
- [38] R Development Core Team (2007). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. URL <http://www.R-project.org>. ISBN 3-900051-07-0.
- [39] Schouten, H. J. A., Mollenaar, W., Van Strik, R., Boomsma, A. (1980). Comparing two independent binomial proportions by a modified chi square test. *Biometrical Journal* 22, 241–248.
- [40] Schumacher, M., Schulgen, G. (2002). *Methodik klinischer Studien*. Springer Verlag, Heidelberg.
- [41] Shen, Y., Fisher, L. (1999). Statistical inference for self-designing clinical trials with a one-sided hypothesis. *Biometrics* 55, 190–197.
- [42] Somers, R. H. (1962). A new asymmetric measure of association for ordinal variables. *American Sociological Review* 27, 799–811.
- [43] Stouffer, S. A., Suchmann, E. A., DeVinney, L. C., Star, S. A., Williams, R. M. Jr. (1949). *The american soldier, volume I: adjustment during army life*. Princeton University Press, New York.
- [44] Tsiatis, A. A., Mehta, C. (2003). On the inefficiency of the adaptive design for monitoring clinical trials. *Biometrika* 90, 367–378.
- [45] Wald, A. (1947). *Sequential analysis*. Wiley, New York.
- [46] Wang, S. K., Tsiatis, A. A. (1987). Approximately optimal one-parameter boundaries for group sequential tests. *Biometrics* 43, 193–200.
- [47] Wassmer, G. (2001). *Statistische Testverfahren für gruppensequentielle und adaptive Pläne in klinischen Studien*. 2. Auflage, Verlag Alexander Mönch, München.

## A Kritische Werte für adaptiv gruppensequentielle Designs

Die in den folgenden Tabellen verzeichneten kritischen Werte wurden mit der statistischen Software R, Version 2.6.2, ermittelt. Bei Verwendung der inversen Normalmethode wurden diese mittels der in dem Paket *Mnormt* enthaltenen Funktion `pnnorm()` berechnet. Für die inverse  $\chi^2$ -Methode wurden die kritischen Werte mittels Simulationen mit 100000 Wiederholungen erhalten.

**Tabelle A.1:** Kritische Werte bei Verwendung der Stouffer-Statistik im Pocock-Design, vgl. (2.18), und im O'Brien/Fleming-Design, vgl. (2.24), sowie bei Verwendung der inversen  $\chi^2$ -Statistik mit Gewichten  $\lambda_i \equiv 1$  sowie der Fisher-Statistik im Pocock-Design, vgl. (2.19)

$K$	Inverse Normalmethode						Inverse $\chi^2$ -Methode								
	Pocock-Design			O'Brien/Fleming-Design			Pocock-Design			Pocock-Design					
	$cv_{Z,PO}(\alpha, K, \{1\})$			$cv_{Z,OBF}(\alpha, K, \{1\})$			$cv_{S,PO}(\alpha, K, \{1\})$			$cv_{S,PO}(\alpha, K, \{2\})$					
	$\alpha$			$\alpha$			$\alpha$			$\alpha$					
	0.05	0.025	0.005	0.05	0.025	0.005	0.05	0.025	0.005	0.05	0.025	0.005	0.005	0.025	0.005
2	1.8754	2.1783	2.7718	2.3730	2.7965	3.6481	6.8321	8.2592	11.5411	10.5312	12.2173	15.9614			
3	1.9922	2.2895	2.8730	2.9611	3.4711	4.4945	9.2459	10.8306	14.3940	14.3973	16.2828	20.4295			
4	2.0674	2.3613	2.9387	3.4662	4.0486	5.2182	11.3979	13.0939	16.9008	17.9278	19.9750	24.4140			
5	2.1217	2.4132	2.9863	3.9151	4.5617	5.8611	13.3818	15.1821	19.1724	21.2641	23.4509	28.1900			
6	2.1636	2.4532	3.0230	4.3230	5.0282	6.4453	15.0117	16.9058	21.0277	24.1557	26.4736	31.4514			
7	2.1973	2.4855	3.0527	4.6997	5.4589	6.9845	16.8594	18.8284	23.1716	27.2996	29.6914	34.9121			
8	2.2254	2.5123	3.0773	5.0513	5.8610	7.4881	18.8037	20.8401	25.2872	30.5823	33.1058	38.4887			
9	2.2492	2.5351	3.0983	5.3822	6.2394	7.9621	20.4871	22.6181	27.2561	33.5456	36.1870	41.7257			
10	2.2698	2.5548	3.1163	5.6956	6.5978	8.4111	22.1321	24.3057	29.0719	36.4346	39.1586	44.8979			
11	2.2880	2.5720	3.1319	5.9942	6.9393	8.8385	23.7555	26.0010	30.8797	39.2679	42.0985	48.0730			
12	2.3042	2.5876	3.1463	6.2793	7.2659	9.2450	25.3539	27.6846	32.6331	42.1202	45.0485	51.1052			
15	2.3431	2.6256	3.1825	7.0722	8.1712	10.3847	29.9342	32.3959	37.6526	50.3838	53.4989	59.9866			

$K$	Inverse Normalmethode			Inverse $\chi^2$ -Methode		
	Pocock-Design		O'Brien/Fleming-Design	Pocock-Design		
	$cv_{Z,PO}(\alpha, K, \{1\})$	$cv_{Z,OFF}(\alpha, K, \{1\})$	$cv_{Z,OFF}(\alpha, K, \{1\})$	$cv_{S,PO}(\alpha, K, \{1\})$	$cv_{S,PO}(\alpha, K, \{2\})$	
	$\alpha$	$\alpha$	$\alpha$	$\alpha$	$\alpha$	$\alpha$
0.05	0.025	0.005	0.05	0.025	0.005	0.05
0.05	0.025	0.005	0.05	0.025	0.005	0.05
20	2.6712	3.2235	8.8415	9.5006	12.0525	37.3482
						45.8510
						63.7001
						67.1008
						74.3415

**Tabelle A.2:** Kritische Werte bei Verwendung der Stouffer-Statistik im Wang/Tsiatis-Design, vgl. (2.31), sowie bei Verwendung der inversen  $\chi^2$ -Statistik mit Gewichten  $\lambda_i \equiv 1$  und  $\lambda_i \equiv 2$ , vgl. (2.32)

$\Delta$	$K$	Inverse Normalmethode			Inverse $\chi^2$ -Methode					
		$cv_{Z,WT}(\alpha, K, \{1\}, \Delta)$		$cv_{S,WT}(\alpha, K, \{1\}, \Delta)$	$cv_{S,WT}(\alpha, K, \{2\}, \Delta)$					
		$\alpha$	$\alpha$	$\alpha$	$\alpha$	$\alpha$	$\alpha$			
0.0	2	2.3730	2.7965	3.6481	0.05	0.025	0.005	0.05	0.025	0.005
	3	2.9611	3.4711	4.4945	5.9915	7.3778	10.5966	9.4877	11.1433	14.8603
	4	3.4662	4.0486	5.2182	7.8147	9.3484	12.8382	12.5916	14.4494	18.5476
	5	3.9151	4.5617	5.8611	9.4877	11.1433	14.8603	15.5073	17.5346	21.9550
	6	4.3230	5.0282	6.4453	11.0705	12.8325	16.7496	18.3070	20.4832	25.1882
	7	4.6997	5.4589	6.9845	12.5916	14.4494	18.5476	21.0261	23.3367	28.2995
	8	5.0513	5.8610	7.4881	14.0671	16.0128	20.2777	23.6848	26.1190	31.3194
	9	5.3822	6.2394	7.9621	15.5073	17.5346	21.9550	26.2962	28.8454	34.2672
	10	5.6956	6.5978	8.4111	16.9190	19.0228	23.5894	28.8693	31.5264	37.1565
	11	5.9942	6.9393	8.8385	18.3070	20.4832	25.1882	31.4104	34.1696	39.9969
	12	6.2793	7.2659	9.2450	19.6751	21.9201	26.7569	33.9244	36.7807	42.7957
					21.0261	23.3367	28.2995	36.4150	39.3641	45.5585

$\Delta$	$K$	Inverse Normalmethode			Inverse $\chi^2$ -Methode							
		$cv_{Z,WT}(\alpha, K, \{1\}, \Delta)$			$cv_{S,WT}(\alpha, K, \{1\}, \Delta)$			$cv_{S,WT}(\alpha, K, \{2\}, \Delta)$				
		$\alpha$			$\alpha$			$\alpha$				
		0.05	0.025	0.005	0.05	0.025	0.005	0.05	0.025	0.005		
	15	7.0722	8.1712	10.3847	24.9958	27.4884	32.8013	43.7730	46.9792	53.6720		
	20	8.8415	9.5006	12.0525	31.4104	34.1696	39.9969	55.7585	59.3417	66.7660		
0.1	2	2.2425	2.6314	3.4136	6.0409	7.4436	10.6360	9.5138	11.1596	14.8639		
	3	2.6943	3.1442	4.0496	7.8781	9.4295	12.9172	12.6144	14.4769	18.5722		
	4	3.0690	3.5692	4.5752	9.5548	11.2147	14.9287	15.5410	17.5661	21.9626		
	5	3.3936	3.9371	5.0304	11.1591	12.9263	16.8695	18.3383	20.4895	25.2112		
	6	3.6823	4.2645	5.4356	12.6697	14.5135	18.6286	21.0607	23.3788	28.3011		
	7	3.9441	4.5614	5.8033	14.1497	16.0922	20.3169	23.7234	26.1689	31.3953		
	8	4.1847	4.8343	6.1413	15.5869	17.6067	22.0127	26.3457	28.9020	34.3124		
	9	4.4081	5.0875	6.4555	17.0173	19.0965	23.7016	28.8986	31.6193	37.1675		
	10	4.6172	5.3249	6.7496	18.3505	20.5149	25.2087	31.4542	34.2159	40.0354		
	11	4.8141	5.5484	7.0269	19.7641	22.0003	26.8312	33.9651	36.8328	42.8859		
	12	5.0007	5.7605	7.2888	21.1240	23.4136	28.4227	36.4542	39.4152	45.6348		
	15	5.5084	6.3358	8.0073	25.0935	27.5425	32.8718	43.8365	47.0760	53.6839		
	20	6.2351	7.1611	9.0362	31.4961	34.2644	40.1045	55.7789	59.3744	66.8250		
0.2	2	2.1287	2.4877	3.2058	6.1505	7.5481	10.7490	9.6099	11.2545	14.9370		
	3	2.4670	2.8639	3.6622	8.0222	9.5842	13.0228	12.7417	14.5920	18.6599		
	4	2.7356	3.1643	4.0273	9.7373	11.3906	15.1332	15.6643	17.6957	22.0698		
	5	2.9615	3.4174	4.3351	11.3406	13.1115	17.0562	18.5074	20.6608	25.3566		
	6	3.1581	3.6377	4.6031	12.8639	14.7122	18.8105	21.2164	23.5283	28.4220		
	7	3.3330	3.8338	4.8423	14.3512	16.2975	20.5191	23.8834	26.3265	31.5594		
	8	3.4911	4.0112	5.0587	15.8253	17.8568	22.2705	26.5110	29.0580	34.4663		
	9	3.6358	4.1731	5.2549	17.2057	19.3177	23.9958	29.1069	31.7802	37.3455		

$\Delta$	$K$	Inverse Normalmethode			Inverse $\chi^2$ -Methode						
		$cv_{Z,WT}(\alpha, K, \{1\}, \Delta)$			$cv_{S,WT}(\alpha, K, \{1\}, \Delta)$			$cv_{S,WT}(\alpha, K, \{2\}, \Delta)$			
		$\alpha$			$\alpha$			$\alpha$			
		0.05	0.025	0.005	0.05	0.025	0.005	0.05	0.025	0.005	
	10	3.7696	4.3240	5.4402	18.6114	20.7804	25.3967	31.6190	34.3870	40.2343	
	11	3.8943	4.4640	5.6110	19.9529	22.1918	27.0197	34.1386	37.0046	43.0613	
	12	4.0113	4.5949	5.7711	21.1240	23.6401	28.5568	36.6309	39.5845	45.8002	
	15	4.3242	4.9463	6.2030	25.3120	27.8111	33.0706	44.0244	47.2598	53.8630	
	20	4.7578	5.4361	6.8062	31.7729	34.5204	40.3748	55.9702	59.5560	66.9959	
0.25	2	2.0777	2.4239	3.1131	6.2254	7.6229	10.8277	9.6899	11.3318	15.0256	
	3	2.3674	2.7411	3.4906	8.1335	9.6889	13.1918	12.8460	14.7028	18.7628	
	4	2.5915	2.9887	3.7873	9.8616	11.5312	15.2725	15.7910	17.8271	22.2043	
	5	2.7767	3.1941	4.0341	11.4861	13.2506	17.2073	18.6449	20.7997	25.4849	
	6	2.9357	3.3707	4.2463	13.0218	14.8614	18.9641	21.3684	23.6721	28.5596	
	7	3.0756	3.5265	4.4343	14.5216	16.4620	20.6989	24.0435	26.4732	31.7231	
	8	3.2011	3.6661	4.6028	15.9667	17.9933	22.3906	26.6740	29.2169	34.6372	
	9	3.3150	3.7931	4.7561	17.4050	19.4875	24.0906	29.2764	31.9464	37.5334	
	10	3.4197	3.9097	4.8971	18.7553	20.9100	25.6072	31.7814	34.5548	40.3705	
	11	3.5164	4.0176	5.0273	20.1747	22.4012	27.2526	34.3121	37.1780	43.2480	
	12	3.6065	4.1180	5.1486	21.5405	23.8433	28.8069	36.8123	39.7596	45.9613	
	15	3.8467	4.3854	5.4753	25.4674	27.9532	33.2187	44.2108	47.4507	54.0288	
	20	4.1709	4.7526	5.9190	31.9556	34.7022	40.5777	56.1484	59.7375	67.1629	
0.3	2	2.0305	2.3651	3.0284	6.3123	7.7221	10.9171	9.7957	11.4327	15.1882	
	3	2.2766	2.6297	3.3345	8.2833	9.8404	13.2894	12.9928	14.8356	18.9331	
	4	2.4616	2.8307	3.5701	10.0207	11.6975	15.4393	15.9891	18.0241	22.4879	
	5	2.6114	2.9943	3.7631	11.6979	13.4669	17.3923	18.8480	21.0270	25.7311	
	6	2.7380	3.1332	3.9273	13.2386	15.0858	19.1637	21.5962	23.8849	28.8500	

$\Delta$	$K$	Inverse Normalmethode			Inverse $\chi^2$ -Methode						
		$cv_{Z,WT}(\alpha, K, \{1\}, \Delta)$			$cv_{S,WT}(\alpha, K, \{1\}, \Delta)$			$cv_{S,WT}(\alpha, K, \{2\}, \Delta)$			
		$\alpha$			$\alpha$			$\alpha$			
		0.05	0.025	0.005	0.05	0.025	0.005	0.05	0.025	0.005	
	7	2.8481	3.2542	4.0710	14.7554	16.6924	20.9397	24.2705	26.6727	31.8297	
	8	2.9458	3.3617	4.1975	16.2013	18.2307	22.6306	26.9249	29.4882	34.8762	
	9	3.0338	3.4587	4.3140	17.6576	19.7396	24.3298	29.5316	32.1783	37.7197	
	10	3.1139	3.5471	4.4194	19.0173	21.1672	25.8729	32.0432	34.8296	40.5982	
	11	3.1876	3.6283	4.5162	20.4338	22.6669	27.5500	34.5787	37.4394	43.4977	
	12	3.2560	3.7038	4.6062	21.7885	24.1147	29.1511	37.0937	40.0277	46.2131	
	15	3.4343	3.9009	4.8442	25.7844	28.2635	33.5845	44.4994	47.7144	54.3193	
	20	3.6767	4.1677	5.1700	32.2556	35.0256	40.8558	56.4533	60.0431	67.4404	
0.4	2	1.9465	2.2625	2.8837	6.5396	7.9557	11.1873	10.0845	11.7686	15.4810	
	3	2.1197	2.4395	3.0709	8.6692	10.2423	13.7555	13.4956	15.3450	19.4592	
	4	2.2412	2.5651	3.2062	10.5565	12.2274	15.9858	16.6343	18.6771	23.0798	
	5	2.3349	2.6624	3.3124	12.3024	14.0913	18.0451	19.6060	21.7343	26.4258	
	6	2.4110	2.7420	3.3999	13.9305	15.7881	19.8764	22.4306	24.7141	29.6619	
	7	2.4752	2.8093	3.4744	15.5158	17.4485	21.7072	25.1855	27.6126	32.8409	
	8	2.5306	2.8674	3.5388	17.0319	19.0525	23.4223	27.8905	30.4242	35.7731	
	9	2.5794	2.9189	3.5967	18.5169	20.5930	25.1810	30.5523	33.1816	38.7619	
	10	2.6230	2.9649	3.6479	19.9256	22.1046	26.7573	33.1073	35.8490	41.6408	
	11	2.6623	3.0064	3.6946	21.3759	23.5743	28.4288	35.6737	38.5197	44.5544	
	12	2.6981	3.0443	3.7370	22.7967	25.0556	30.0134	38.2404	41.1583	47.2656	
	15	2.7886	3.1421	3.8515	26.8451	29.3410	34.6341	45.7323	48.9183	55.5090	
	20	2.9063	3.2684	3.9948	33.4455	36.1927	41.9321	57.8027	61.3292	68.6602	
0.5	2	1.8754	2.1783	2.7718	6.8321	8.2592	11.5411	10.5312	12.2173	15.9614	
	3	1.9922	2.2895	2.8730	9.2459	10.8306	14.3940	14.3973	16.2828	20.4295	

$\Delta$	$K$	Inverse Normalmethode			Inverse $\chi^2$ -Methode								
		$cv_{Z,WT}(\alpha, K, \{1\}, \Delta)$			$cv_{S,WT}(\alpha, K, \{1\}, \Delta)$			$cv_{S,WT}(\alpha, K, \{2\}, \Delta)$					
		$\alpha$	$\alpha$	$\alpha$	$\alpha$	$\alpha$	$\alpha$	$\alpha$	$\alpha$	$\alpha$	$\alpha$		
		0.05	0.025	0.005	0.05	0.025	0.005	0.05	0.025	0.005	0.05	0.025	0.005
	4	2.0674	2.3613	2.9387	11.3979	13.0939	16.9008	17.9278	19.9750	24.4140			
	5	2.1217	2.4132	2.9863	13.3818	15.1821	19.1724	21.2641	23.4509	28.1900			
	6	2.1636	2.4532	3.0230	15.0117	16.9058	21.0277	24.1557	26.4736	31.4514			
	7	2.1973	2.4855	3.0527	16.8594	18.8284	23.1716	27.2996	29.6914	34.9121			
	8	2.2254	2.5123	3.0773	18.8037	20.8401	25.2872	30.5823	33.1058	38.4887			
	9	2.2492	2.5351	3.0983	20.4871	22.6181	27.2561	33.5456	36.1870	41.7257			
	10	2.2698	2.5548	3.1163	22.1321	24.3057	29.0719	36.4346	39.1586	44.8979			
	11	2.2880	2.5720	3.1319	23.7555	26.0010	30.8797	39.2679	42.0985	48.0730			
	12	2.3042	2.5876	3.1463	25.3539	27.6846	32.6331	42.1202	45.0485	51.1052			
	15	2.3431	2.6256	3.1825	29.9342	32.3959	37.6526	50.3838	53.4989	59.9866			
	20	2.3918	2.6712	3.2235	37.3482	40.0424	45.8510	63.7001	67.1008	74.3415			
0.6	2	1.8161	2.1110	2.6906	7.2133	8.6665	11.9570	11.1612	12.8892	16.6597			
	3	1.8904	2.1751	2.7367	10.0488	11.6957	15.3261	15.7883	17.7402	22.0624			
	4	1.9332	2.2113	2.7615	12.7112	14.4928	18.3920	20.2469	22.4071	27.0460			
	5	1.9613	2.2347	2.7769	15.2417	17.1429	21.3109	24.6925	27.0344	32.1012			
	6	1.9814	2.2512	2.7871	17.7144	19.7170	24.0839	29.0475	31.5670	36.9406			
	7	1.9966	2.2635	2.7947	20.1679	22.2953	26.9063	33.4483	36.1177	41.7256			
	8	2.0084	2.2730	2.7978	22.5968	24.8016	29.6410	37.8802	40.7064	46.6306			
	9	2.0179	2.2804	2.8048	25.0305	27.3847	32.3335	42.3041	45.2712	51.6356			
	10	2.0256	2.2865	2.8079	27.4425	29.8708	34.9579	46.8354	49.9313	56.5591			
	11	2.0321	2.2915	2.8105	29.8934	32.4200	37.8271	51.3850	54.6880	61.5458			
	12	2.0374	2.2955	2.8125	32.3528	34.9562	40.6039	55.9840	59.4208	66.6348			
	15	2.0499	2.3056	2.8177	39.7566	42.6314	48.7863	70.0139	73.8527	81.8275			

$\Delta$	$K$	Inverse Normalmethode			Inverse $\chi^2$ -Methode								
		$cv_{Z,WT}(\alpha, K, \{1\}, \Delta)$			$cv_{S,WT}(\alpha, K, \{1\}, \Delta)$			$cv_{S,WT}(\alpha, K, \{2\}, \Delta)$					
		$\alpha$	$\alpha$	$\alpha$	$\alpha$	$\alpha$	$\alpha$	$\alpha$	$\alpha$	$\alpha$	$\alpha$		
		0.05	0.025	0.005	0.05	0.025	0.005	0.05	0.025	0.005	0.05	0.025	0.005
0.7	20	2.0640	2.3149	2.8191	52.4684	55.7761	63.5758	94.2835	98.7383	108.0906			
	2	1.7676	2.0590	2.6364	7.6640	9.1596	12.5286	11.9680	13.7540	17.6437			
	3	1.8113	2.0917	2.6529	11.1696	12.8748	16.7004	17.8547	19.9293	24.5214			
	4	1.8327	2.1068	2.6592	14.6635	16.5761	20.7385	24.0120	26.3922	31.4698			
	5	1.8449	2.1149	2.6621	18.2451	20.3285	24.8590	30.5726	33.2380	38.9483			
	6	1.8526	2.1198	2.6634	21.9353	24.1780	29.0479	37.4214	40.3069	46.4741			
	7	1.8577	2.1228	2.6640	25.7866	28.2245	33.4561	44.5594	47.7349	54.3368			
	8	1.8613	2.1247	2.6643	29.7243	32.3447	37.9239	52.0407	55.4507	62.5813			
	9	1.8640	2.1262	2.6649	33.8636	36.6324	42.4529	59.7801	63.4274	71.0412			
	10	1.8658	2.1271	2.6651	38.0722	41.0233	47.2303	67.8716	71.7557	79.7493			
	11	1.8674	2.1278	2.6652	42.4232	45.5335	52.0295	76.1734	80.2837	88.6993			
	12	1.8686	2.1283	2.6653	46.8973	50.1566	57.0652	84.6776	89.0318	97.9504			
	15	1.8706	2.1288	2.6654	60.9554	64.6786	72.1733	111.5724	116.5422	126.6455			
	20	1.8720	2.1293	2.6655	86.3794	90.7693	99.6902	160.5031	166.4274	178.4880			

**Tabelle A.3:** Kritische Werte im (modifizierten) Haybittle/Peto-Design bei Verwendung der Stouffer-Statistik, vgl. (2.38) und (2.40), sowie bei Verwendung der inversen  $\chi^2$ -Statistik mit Gewichten  $\lambda_i \equiv 1$ , vgl. (2.39) und (2.41)

$K$	$cv_{Z,HP}(\alpha, K, \{1\})$			$cv_{Z,HP}^*(\alpha, K, \{1\})$			$cv_{S,HP}(\alpha, K, \{1\})$			$cv_{S,HP}^*(\alpha, K, \{1\})$		
	$\alpha$	$\alpha$	$\alpha$	$\alpha$	$\alpha$	$\alpha$	$\alpha$	$\alpha$	$\alpha$	$\alpha$	$\alpha$	$\alpha$
	0.05	0.025	0.005	0.05	0.025	0.005	0.05	0.025	0.005	0.05	0.025	0.005
2	1.6472	1.9673	2.6443	1.6597	2.0027	-	5.9894	7.3945	10.6017	5.9894	7.3945	-
3	1.6504	1.9751	2.7201	1.6753	2.0458	-	7.8186	9.3538	13.0809	7.8186	9.4487	-
4	1.6537	1.9828	2.8118	1.6905	2.0897	-	9.4854	11.1584	15.3835	9.5100	11.3729	-
5	1.6568	1.9900	2.9383	1.7051	2.1349	-	11.0846	12.8517	17.6606	11.1450	13.2008	-
6	1.6597	1.9969	3.1976	1.7190	2.1822	-	12.5942	14.4699	19.8140	12.7002	14.9847	-
7	1.6625	2.0034	-	1.7322	2.2330	-	14.0715	16.0591	21.9959	14.2169	16.7146	-
8	1.6651	2.0097	-	1.7449	2.2887	-	15.5044	17.5876	24.1883	15.6956	18.4095	-
9	1.6676	2.0155	-	1.7570	2.3529	-	16.9358	19.1359	28.1016	17.1684	20.0991	-
10	1.6698	2.0211	-	1.7686	2.4315	-	18.3680	20.6286	-	18.6498	21.7441	-
11	1.6719	2.0264	-	1.7798	2.5404	-	19.7439	22.1035	-	20.0880	23.6025	-
12	1.6743	2.0314	-	1.7906	2.7922	-	21.0916	23.5068	-	21.4686	25.1780	-
15	1.6800	2.0450	-	1.8200	-	-	25.0741	27.7506	-	25.6082	30.2571	-
20	1.6879	2.0633	-	1.8653	-	-	31.5284	34.5789	-	32.3406	40.4414	-

## B Verteilungen der finalen Teststatistiken im Self-Designing

Bei dem Self-Designing klinischer Studien werden die Gewichtungen der einzelnen Sequenzen im Verlauf der Untersuchung festgelegt. Somit kann die Verteilung der Kombinationsstatistik erst bestimmt werden, wenn das vor Studienbeginn gewählte globale Gewicht  $\lambda_G$  komplett verbraucht ist, d.h. es gilt  $\sum_{i=1}^j \lambda_i = \lambda_G$ . Basierend auf dem Ansatz von Shen und Fisher (1999) erfolgt der Nachweis der Verteilung der finalen Teststatistik zunächst für das Self-Designing mit gewichteter Normalmethode, anschließend in analoger Vorgehensweise für das Self-Designing mit verallgemeinerter inverser  $\chi^2$ -Methode. Da hierfür Martingale sowie die charakteristische Funktion betrachtet werden, werden diese vorab kurz vorgestellt.

Die charakteristische Funktion ist ein wichtiges Hilfsmittel in der Statistik, insbesondere für die Herleitung der Verteilung von Zufallsvariablen. In diesem Zusammenhang werden Funktionen betrachtet, die Werte in der Menge  $\mathbb{C}$  der *komplexen Zahlen* annehmen. Für die Definition der charakteristischen Funktion sei  $g : \mathbb{R} \rightarrow \mathbb{C}$  eine Funktion

$$g = \operatorname{Re}(g) + i \operatorname{Im}(g)$$

mit  $\operatorname{Re}(g)$  Real- und  $\operatorname{Im}(g)$  Imaginärteil. Für die imaginäre Einheit  $i$  gilt:  $i^2 = -1$ .

**Definition B.1** Sei  $X$  eine Zufallsvariable. Dann ist

$$E[g(X)] = E[e^{itX}] = \varphi_X(t)$$

die charakteristische Funktion von  $X$  ( $t \in \mathbb{R}$  fest).

Es gilt:

- (i)  $|\varphi_X(t)| \leq 1$ ,
  - (ii)  $\varphi_X(0) = 1$ ,
  - (iii)  $\varphi_{aX+b}(t) = e^{ibt} \varphi_X(at)$ ,
- (siehe Müller (1991)).

**Satz B.1** Seien  $X$  und  $Y$  Zufallsvariablen mit Verteilungsfunktion  $F$  bzw.  $G$ . Dann sind die folgenden Aussagen äquivalent

- (i)  $F = G$ ,
- (ii)  $\varphi_X = \varphi_Y$ .

**Definition B.2**  $X = \{X_t, t \in \mathbb{N}\}$  heißt Martingal bezüglich  $\{F_t, t \in \mathbb{N}\}$ , falls gilt

- (i)  $E(|X_t|) < \infty, t \in \mathbb{N}$ ,
- (ii)  $E(X_{t+1}|F_t) = E(X_{t+1}|X_t, \dots, X_1) = X_t$ ,

(vgl. Fahrmeir et al. (1985)).

### B.1 Verteilung der finalen Teststatistik $Z_k$

Der Nachweis der Verteilung der finalen Teststatistik  $Z_k$  beruht auf der Ausnutzung der Eigenschaften von Martingalen und der charakteristischen Funktion. Sei

$$Y_j = \exp \left( it \sum_{i=1}^j \sqrt{\lambda_i} \Phi^{-1}(1 - p_i) \right) / \exp \left( -\frac{t^2}{2} \sum_{i=1}^j \lambda_i \right)$$

ein Martingal.

Zu zeigen:

- (i)  $E[|Y_j|] < \infty$
- (ii)  $E[Y_{j+1} | Y_1, \dots, Y_j] = Y_j$ .

Zu (i):

$$\begin{aligned} E[|Y_j|] &= E \left[ \left| \frac{\exp \left( it \sum_{i=1}^j \sqrt{\lambda_i} \Phi^{-1}(1 - p_i) \right)}{\exp \left( -\frac{t^2}{2} \sum_{i=1}^j \lambda_i \right)} \right| \right] = \left| E \left[ \frac{\exp \left( it \sum_{i=1}^j \sqrt{\lambda_i} \Phi^{-1}(1 - p_i) \right)}{\exp \left( -\frac{t^2}{2} \sum_{i=1}^j \lambda_i \right)} \right] \right| \\ &= \frac{\left| E \left[ \exp \left( it \sum_{i=1}^j \sqrt{\lambda_i} \Phi^{-1}(1 - p_i) \right) \right] \right|}{\exp \left( -\frac{t^2}{2} \sum_{i=1}^j \lambda_i \right)} < \infty, \end{aligned}$$

da gilt

$$\left| E \left[ \exp \left( it \sum_{i=1}^j \sqrt{\lambda_i} \Phi^{-1}(1 - p_i) \right) \right] \right| = |\varphi_{\sum_{i=1}^j \sqrt{\lambda_i} \Phi^{-1}(1-p_i)}(t)| \leq 1 \quad \text{und} \quad \sum_{i=1}^j \lambda_i \xrightarrow{j \rightarrow \infty} \lambda_G.$$

Zu (ii):

$$\begin{aligned} E[Y_{j+1} | Y_1, \dots, Y_j] &= E \left[ \frac{\exp \left( it \sum_{i=1}^{j+1} \sqrt{\lambda_i} \Phi^{-1}(1 - p_i) \right)}{\exp \left( -\frac{t^2}{2} \sum_{i=1}^{j+1} \lambda_i \right)} \right] \\ &= \frac{E \left[ \exp \left( it \left( \sum_{i=1}^j \sqrt{\lambda_i} \Phi^{-1}(1 - p_i) + \sqrt{\lambda_{j+1}} \Phi^{-1}(1 - p_{j+1}) \right) \right) \right]}{\exp \left( -\frac{t^2}{2} \sum_{i=1}^{j+1} \lambda_i \right)} \\ &= \frac{\exp \left( it \sum_{i=1}^j \sqrt{\lambda_i} \Phi^{-1}(1 - p_i) \right) E \left[ \exp \left( it \left( \sqrt{\lambda_{j+1}} \Phi^{-1}(1 - p_{j+1}) \right) \right) \right]}{\exp \left( -\frac{t^2}{2} \sum_{i=1}^j \lambda_i \right) \exp \left( -\frac{t^2}{2} \lambda_{j+1} \right)} \\ &\stackrel{H_0}{=} Y_j \frac{\exp \left( i0t - \lambda_{j+1} \frac{t^2}{2} \right)}{\exp \left( -\frac{t^2}{2} \lambda_{j+1} \right)} = Y_j, \end{aligned}$$

da  $\sqrt{\lambda_{j+1}} \Phi^{-1}(1 - p_{j+1})$  gegeben  $\sqrt{\lambda_1} \Phi^{-1}(1 - p_1), \dots, \sqrt{\lambda_j} \Phi^{-1}(1 - p_j)$  unter der Nullhypothese bedingt  $N(0, \lambda_{j+1})$ -verteilt ist und die zugehörige charakteristische Funktion  $E[\exp(it\sqrt{\lambda_{j+1}} \Phi^{-1}(1 - p_{j+1}))] = \exp(-t^2/2 \lambda_{j+1})$  ist.

Durch Nachweis der Eigenschaften (i) und (ii) konnte gezeigt werden, dass  $Y_j$  ein Martingal ist. Da die Gewichte  $\{\lambda_i, i = 1, 2, \dots\}$  nicht-negativ sind, ist  $k = \inf\{j : \sum_{i=1}^j \lambda_i = \lambda_G\}$  eine Stoppzeit. Insbesondere gilt

$$E[Y_k] = E[Y_0],$$

wobei

$$E[Y_0] = E \left[ \frac{\exp(it \cdot 0)}{\exp(-(t^2/2) \cdot 0)} \right] = E[1] = 1$$

ist (vgl. Müller (1991)). Somit gilt

$$E[Y_k] = E \left[ \frac{\exp \left( it \sum_{i=1}^k \sqrt{\lambda_i} \Phi^{-1}(1 - p_i) \right)}{\exp \left( -\frac{t^2}{2} \sum_{i=1}^k \lambda_i \right)} \right] = \frac{E \left[ \exp \left( it \sum_{i=1}^k \sqrt{\lambda_i} \Phi^{-1}(1 - p_i) \right) \right]}{\exp \left( -\frac{t^2}{2} \lambda_G \right)} = 1$$

$$\Leftrightarrow E \left[ \exp \left( it \sum_{i=1}^k \sqrt{\lambda_i} \Phi^{-1}(1 - p_i) \right) \right] = \exp \left( -\frac{t^2}{2} \lambda_G \right) \Leftrightarrow E [\exp (itZ_k)] = \exp \left( -\frac{t^2}{2} \lambda_G \right).$$

Da  $\exp(-t^2/2 \lambda_G)$  sowohl die charakteristische Funktion der  $N(0, \lambda_G)$ -Verteilung als auch von  $Z_k = \sum_{i=1}^k \sqrt{\lambda_i} \Phi^{-1}(1 - p_i)$  ist, ist  $Z_k$  unter der Nullhypothese  $N(0, \lambda_G)$ -verteilt.  $\square$

## B.2 Verteilung der finalen Teststatistik $S_k$

Der Nachweis der Verteilung der finalen Teststatistik  $S_k$  beruht auf der Ausnutzung der Eigenschaften von Martingalen und der charakteristischen Funktion. Sei

$$Y_j = \exp \left( it \sum_{i=1}^j q_i(\lambda_i) \right) / (1 - 2it)^{-\frac{\lambda_{\Sigma}(j)}{2}}$$

ein Martingal.

Zu zeigen:

- (i)  $E[|Y_j|] < \infty$
- (ii)  $E[Y_{j+1} | Y_1, \dots, Y_j] = Y_j$ .

Zu (i):

$$E[|Y_j|] = E \left[ \left| \frac{\exp \left( it \sum_{i=1}^j q_i(\lambda_i) \right)}{(1 - 2it)^{-\frac{\lambda_{\Sigma}(j)}{2}}} \right| \right] \leq \left| \frac{E \left[ \exp \left( it \sum_{i=1}^j q_i(\lambda_i) \right) \right]}{(1 - 2it)^{-\frac{\lambda_{\Sigma}(j)}{2}}} \right|$$

$$= \frac{\left| E \left[ \exp \left( it \sum_{i=1}^j q_i(\lambda_i) \right) \right] \right|}{\left| (1 - 2it)^{-\frac{\lambda_{\Sigma(j)}}{2}} \right|} < \infty,$$

da  $\left| E \left[ \exp \left( it \sum_{i=1}^j q_i(\lambda_i) \right) \right] \right| = |\varphi_{\sum_{i=1}^j q_i(\lambda_i)}(t)| \leq 1$  und  $\sum_{i=1}^j \lambda_i \xrightarrow{j \rightarrow \infty} \lambda_G$ .

Zu (ii):

$$E[Y_{j+1} | Y_1, \dots, Y_j] = E \left[ \frac{\exp \left( it \sum_{i=1}^{j+1} q_i(\lambda_i) \right)}{(1 - 2it)^{-\frac{\lambda_{\Sigma(j+1)}}{2}}} \right]$$

$$= \frac{\exp \left( it \sum_{i=1}^j q_i(\lambda_i) \right)}{(1 - 2it)^{-\frac{\lambda_{\Sigma(j)}}{2}}} \frac{E[\exp(it q_{j+1}(\lambda_{j+1}))]}{(1 - 2it)^{-\frac{\lambda_{j+1}}{2}}} \stackrel{H_0}{=} Y_j,$$

da  $q_{j+1}(\lambda_{j+1})$  gegeben  $q_1(\lambda_1), \dots, q_j(\lambda_j)$  unter der Nullhypothese bedingt  $\chi^2$ -verteilt ist mit  $\lambda_{j+1}$  Freiheitsgraden und die zugehörige charakteristische Funktion  $E[\exp(it q_{j+1}(\lambda_{j+1}))] = (1 - 2it)^{-\frac{\lambda_{j+1}}{2}}$  ist.

Durch Nachweis der Eigenschaften (i) und (ii) konnte gezeigt werden, dass  $Y_j$  ein Martingal ist. Da die Gewichte  $\{\lambda_i, i = 1, 2, \dots\}$  nicht-negativ sind, ist  $k = \inf\{j : \sum_{i=1}^j \lambda_i = \lambda_G\}$  eine Stoppzeit. Insbesondere gilt

$$E[Y_k] = E[Y_0],$$

wobei

$$E[Y_0] = E \left[ \frac{\exp(it \cdot 0)}{(1 - 2it)^0} \right] = E[1] = 1$$

ist (vgl. Müller (1991), S. 246). Somit gilt

$$E \left[ \frac{\exp \left( it \sum_{i=1}^k q_i(\lambda_i) \right)}{(1 - 2it)^{-\frac{\lambda_{\Sigma(k)}}{2}}} \right] = \frac{E \left[ \exp \left( it \sum_{i=1}^k q_i(\lambda_i) \right) \right]}{(1 - 2it)^{-\frac{\lambda_{\Sigma(k)}}{2}}} = 1$$

$$\Leftrightarrow E \left[ \exp \left( it \sum_{i=1}^k q_i(\lambda_i) \right) \right] = (1 - 2it)^{-\frac{\lambda_{\Sigma}(k)}{2}} \Leftrightarrow E [ \exp(itS_k) ] = (1 - 2it)^{-\frac{\lambda_{\Sigma}(k)}{2}}.$$

Da  $(1 - 2it)^{-\frac{\lambda_{\Sigma}(k)}{2}}$  sowohl die charakteristische Funktion der  $\chi^2$ -Verteilung mit  $\lambda_{\Sigma}(k)$  Freiheitsgraden als auch von  $S_k = \sum_{i=1}^k q_i(\lambda_i)$  ist, stammt  $S_k$  unter der Nullhypothese aus einer  $\chi^2(\lambda_{\Sigma}(k))$ -Verteilung.  $\square$

## C Ergebnisse der Simulationsstudien

### C.1 Binäre Zielvariable

In den folgenden Tabellen finden sich die Ergebnisse der Konstellationen II-VIII von Szenario I der Simulationsstudie aus Kapitel 3.1.4 sowohl unter der Null- als auch unter der Alternativhypothese.

#### 1) Unter der Nullhypothese

**Tabelle C.1:** Empirisches Niveau  $\hat{\alpha}\%$  im nicht-sequentiellen Design (NSD) sowie  $\hat{\alpha}\%$ , mittlerer und medianer Umfang (ASN = „Average Sample Number“,  $N_{med}$ ), mittlere Stufenanzahl  $\bar{k}$  und empirische Überdeckungswahrscheinlichkeiten  $(1 - 2\hat{\alpha})\%$  der Sequenz der beobachteten Intervalle  $\{KI_j\}$  sowie des finalen Intervalls  $KI_k$  in flexiblen Designs mit binärer Zielvariable unter der Nullhypothese (Szenario I, Konstellation II)

TS	NSD	Design	Inverse Normalmethode					Inverse $\chi^2$ -Methode						
			$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$ $\{KI_j\}$ $KI_k$	$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$ $\{KI_j\}$ $KI_k$		
1a) $p_1 = p_2 = 0.2, D = 0.0$														
$T_{(1)}$	2.54	PO	2.55	656.65	670	4.94	94.57	96.67	2.65	656.78	670	4.93	94.48	96.20
		OBF	2.37	655.67	670	4.98	95.12	95.60	2.67	654.96	670	4.98	94.73	94.73
		WT	2.38	657.07	670	4.97	95.13	96.04	2.66	656.41	670	4.96	94.71	95.05
		HP	2.57	664.27	670	4.99	94.77	95.01	2.51	652.48	670	4.99	94.93	94.95
		SD	2.46	534.61	546	4.05	93.56	94.97	2.69	534.97	546	4.02	94.79	94.79
$T_{(2)}$	2.40	PO	2.41	657.29	670	4.94	94.95	96.85	2.30	658.07	670	4.94	95.12	96.69
		OBF	2.26	656.25	670	4.98	95.48	95.87	2.50	656.65	670	4.98	94.98	94.98
		WT	2.24	657.58	670	4.97	95.28	96.14	2.34	657.44	670	4.97	95.17	95.50
		HP	2.44	664.93	670	4.99	95.12	95.28	2.32	653.27	670	4.99	95.26	95.28
		SD	2.36	534.93	546	4.05	93.98	95.20	2.36	535.39	546	4.02	95.06	95.06
$T_{(3)}$	1.58	PO	0.77	663.84	670	4.98	98.46	99.05	0.95	664.25	670	4.98	98.22	98.81
		OBF	0.45	662.64	670	5.00	98.88	99.05	0.66	662.32	670	4.99	98.50	98.50
		WT	0.58	663.22	670	4.99	98.69	99.01	0.84	663.37	670	4.99	98.23	98.35
		HP	0.54	667.18	670	5.00	98.84	98.86	0.74	659.35	670	4.99	98.44	98.44
		SD	0.38	539.36	546	4.02	98.27	98.83	0.65	539.37	546	4.01	98.47	98.47
$T_{(4)}$	2.00	PO	1.40	661.51	670	4.96	96.65	98.06	1.45	661.80	670	4.96	96.30	97.55
		OBF	1.04	660.14	670	4.99	96.70	97.16	1.43	659.79	670	4.99	96.22	96.22
		WT	1.13	661.32	670	4.99	96.87	97.57	1.52	660.86	670	4.98	95.98	96.56
		HP	1.26	666.45	670	4.99	96.58	96.70	1.41	656.95	670	4.99	96.04	96.05
		SD	1.11	537.82	546	4.04	95.63	96.52	1.24	537.58	546	4.02	96.13	96.13
1b) $p_1 = p_2 = 0.5, D = 0.0$														
$T_{(1)}$	2.39	PO	2.46	657.07	670	4.94	94.84	96.68	2.82	655.48	670	4.92	94.08	95.88
		OBF	2.62	655.71	670	4.98	94.83	95.30	2.80	654.23	670	4.97	94.18	94.18
		WT	2.59	656.19	670	4.96	94.78	95.76	2.88	654.89	670	4.95	94.08	94.52
		HP	2.32	664.17	670	4.98	94.77	95.12	2.81	652.12	670	4.98	94.20	94.27

TS	NSD $\hat{\alpha}\%$	Design	Inverse Normalmethode						Inverse $\chi^2$ -Methode					
			$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2 \hat{\alpha})\%$ { $KI_j$ } $KI_k$		$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2 \hat{\alpha})\%$ { $KI_j$ } $KI_k$	
$T_{(2)}$	2.39	SD	2.42	534.88	546	4.05	93.37	94.81	2.69	534.95	546	4.02	94.44	94.44
		PO	2.34	657.97	670	4.94	95.19	96.86	2.52	656.83	670	4.93	94.77	96.30
		OBF	2.41	655.92	670	4.99	95.06	95.55	2.52	656.06	670	4.98	94.69	94.69
		WT	2.53	656.56	670	4.97	94.99	95.95	2.68	655.75	670	4.96	94.49	94.96
		HP	2.19	665.01	670	4.99	95.15	95.37	2.45	653.28	670	4.99	94.85	94.86
$T_{(3)}$	1.64	SD	2.55	535.51	546	4.05	93.70	94.87	2.48	535.90	546	4.02	94.95	94.95
		PO	0.97	662.51	670	4.97	97.92	98.76	1.30	661.84	670	4.96	97.06	98.17
		OBF	0.63	661.53	670	5.00	98.49	98.60	0.87	660.77	670	4.99	97.81	97.81
		WT	0.71	662.28	670	4.99	98.33	98.74	1.03	661.47	670	4.98	97.67	97.89
		HP	0.74	666.94	670	4.99	98.19	98.39	1.10	658.35	670	4.99	97.61	97.67
$T_{(4)}$	2.39	SD	0.65	538.90	546	4.03	97.65	98.30	0.84	538.39	546	4.01	97.82	97.82
		PO	1.59	660.32	670	4.96	96.10	97.75	1.83	659.70	670	4.95	95.70	97.05
		OBF	1.41	658.94	670	4.99	96.22	96.56	1.64	658.29	670	4.99	95.41	95.41
		WT	1.51	659.55	670	4.98	96.03	96.96	1.68	658.79	670	4.97	95.43	96.01
		HP	1.35	665.98	670	4.99	96.05	96.25	1.75	655.77	670	4.99	95.49	95.52
1c) $p_1 = p_2 = 0.8, D = 0.0$														
$T_{(1)}$	2.60	SD	1.21	537.07	546	4.03	95.15	96.19	1.48	537.28	546	4.02	95.81	95.81
		PO	2.77	655.62	670	4.93	94.60	96.45	3.02	654.64	670	4.92	94.05	95.93
		OBF	2.52	654.99	670	4.98	94.97	95.33	2.99	654.98	670	4.98	94.29	94.29
		WT	2.54	656.78	670	4.97	94.96	95.80	2.85	655.06	670	4.96	94.39	94.84
		HP	2.67	663.16	670	4.98	94.77	95.07	2.63	652.54	670	4.99	94.64	94.66
$T_{(2)}$	2.48	SD	2.59	534.76	546	4.05	93.62	94.96	2.56	534.97	546	4.02	94.74	94.74
		PO	2.34	657.74	670	4.94	95.13	96.98	2.66	656.36	670	4.93	94.82	96.34
		OBF	2.46	655.81	670	4.98	95.02	95.45	2.72	655.76	670	4.98	94.72	94.72
		WT	2.40	657.54	670	4.97	95.06	95.83	2.73	655.62	670	4.96	94.89	95.24
		HP	2.48	664.18	670	4.99	95.04	95.21	2.33	653.20	670	4.99	95.18	95.18
$T_{(3)}$	1.47	SD	2.42	535.09	546	4.05	93.98	95.19	2.53	535.50	546	4.02	95.07	95.07
		PO	0.89	663.07	670	4.98	98.25	98.94	0.88	663.93	670	4.98	98.14	98.84
		OBF	0.48	662.60	670	5.00	98.80	98.94	0.78	662.13	670	5.00	98.28	98.28
		WT	0.64	662.85	670	4.99	98.62	98.99	0.78	662.64	670	4.99	98.25	98.45
		HP	0.54	667.76	670	5.00	98.78	98.86	0.70	659.36	670	5.00	98.43	98.44
$T_{(4)}$	2.01	SD	0.33	538.97	546	4.02	98.37	98.95	0.54	539.12	546	4.01	98.59	98.59
		PO	1.64	659.98	670	4.96	96.30	97.76	1.55	661.09	670	4.96	96.19	97.63
		OBF	1.24	659.41	670	4.99	96.57	96.93	1.41	658.87	670	4.99	95.81	95.81
		WT	1.29	660.61	670	4.99	96.48	97.31	1.55	660.06	670	4.98	95.89	96.35
		HP	1.27	666.30	670	4.99	96.47	96.58	1.48	656.23	670	4.99	96.01	96.01
		SD	1.00	538.18	546	4.04	95.93	96.84	1.23	537.97	546	4.02	96.35	96.35

**Tabelle C.2:** Empirisches Niveau  $\hat{\alpha}\%$  im nicht-sequentiellen Design (NSD) sowie  $\hat{\alpha}\%$ , mittlerer und medianer Umfang (ASN = „Average Sample Number“,  $N_{med}$ ), mittlere Stufenanzahl  $\bar{k}$  und empirische Überdeckungswahrscheinlichkeiten  $(1 - 2\hat{\alpha})\%$  der Sequenz der beobachteten Intervalle  $\{KI_j\}$  sowie des finalen Intervalls  $KI_k$  in flexiblen Designs mit binärer Zielvariable unter der Nullhypothese (Szenario I, Konstellation III)

TS	NSD	Design	Inverse Normalmethode						Inverse $\chi^2$ -Methode					
			$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$	$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$		
	$\hat{\alpha}\%$													
1a) $p_1 = p_2 = 0.2, D = 0.0$														
$T_{(1)}$	2.54	PO	2.37	307.69	320	4.94	94.68	96.73	2.80	309.37	320	4.93	94.11	95.94
		OBF	2.57	300.09	320	4.98	94.72	95.26	2.66	302.83	320	4.98	94.70	94.70
		WT	2.39	302.55	320	4.97	94.91	95.87	2.71	304.10	320	4.96	94.42	94.89
		HP	2.50	315.53	320	4.99	94.72	94.99	2.59	299.59	320	4.99	94.62	94.65
		SD	2.33	197.08	196	3.46	94.22	95.11	2.66	196.67	196	3.34	94.39	94.39
$T_{(2)}$	2.40	PO	2.04	308.53	320	4.95	95.46	97.22	2.34	309.91	320	4.94	95.11	96.62
		OBF	1.52	301.20	320	4.99	96.19	96.63	1.92	303.47	320	4.99	95.68	95.68
		WT	1.49	303.94	320	4.99	96.02	96.89	1.96	305.11	320	4.97	95.49	95.84
		HP	2.42	315.87	320	4.99	95.29	95.42	2.12	300.94	320	4.99	95.49	95.49
		SD	1.96	197.19	196	3.47	95.00	95.61	2.11	197.42	196	3.35	95.21	95.21
$T_{(3)}$	1.58	PO	0.40	314.02	320	4.99	98.74	99.46	0.67	314.36	320	4.98	98.53	99.12
		OBF	0.11	310.02	320	5.00	99.46	99.56	0.37	309.22	320	5.00	99.07	99.07
		WT	0.22	310.81	320	4.99	99.35	99.57	0.53	310.61	320	4.99	98.76	98.88
		HP	0.41	317.93	320	5.00	99.19	99.23	0.41	305.48	320	4.99	98.97	98.98
		SD	0.15	199.03	196	3.25	98.88	99.18	0.48	198.40	196	3.24	98.73	98.73
$T_{(4)}$	2.00	PO	1.26	311.24	320	4.96	97.10	98.40	1.39	312.58	320	4.96	96.89	97.98
		OBF	0.60	305.89	320	5.00	97.76	98.08	1.10	307.33	320	4.99	96.85	96.85
		WT	0.63	307.88	320	4.99	97.82	98.42	1.07	308.31	320	4.99	96.66	97.20
		HP	0.96	316.97	320	4.99	97.24	97.39	1.00	303.57	320	4.99	96.74	96.76
		SD	0.39	199.09	196	3.34	96.88	97.54	0.91	197.52	196	3.28	96.67	96.67
1b) $p_1 = p_2 = 0.5, D = 0.0$														
$T_{(1)}$	2.39	PO	2.55	307.96	320	4.94	94.55	96.53	3.08	308.98	320	4.92	93.63	95.49
		OBF	3.12	299.77	320	4.97	94.18	94.70	3.40	301.80	320	4.96	93.41	93.41
		WT	3.06	302.21	320	4.96	94.07	95.12	3.28	303.39	320	4.95	93.31	93.83
		HP	2.85	315.32	320	4.98	94.41	94.79	3.31	298.70	320	4.97	93.41	93.47
		SD	2.89	195.92	196	3.42	93.31	94.59	3.21	196.45	196	3.34	93.77	93.77
$T_{(2)}$	2.39	PO	2.17	308.52	320	4.95	95.37	97.01	2.66	309.56	320	4.93	94.49	96.07
		OBF	2.17	301.12	320	4.98	95.14	95.67	2.59	303.03	320	4.98	94.47	94.47
		WT	2.15	302.50	320	4.97	95.33	96.25	2.48	304.43	320	4.96	94.43	94.96
		HP	2.39	315.58	320	4.99	95.10	95.35	2.60	299.52	320	4.98	94.64	94.66
		SD	2.54	196.54	196	3.43	94.32	95.09	2.36	197.63	196	3.35	94.83	94.83
$T_{(3)}$	1.64	PO	0.82	312.62	320	4.98	98.05	99.02	1.33	312.91	320	4.96	97.34	98.29
		OBF	0.24	307.63	320	5.00	99.05	99.29	0.89	308.17	320	4.99	97.87	97.87
		WT	0.47	309.50	320	4.99	98.75	99.15	1.02	309.21	320	4.98	97.70	97.96
		HP	0.64	317.63	320	4.99	98.56	98.85	0.94	304.33	320	4.99	97.90	97.93
		SD	0.35	199.08	196	3.29	98.26	98.79	0.77	196.69	196	3.23	97.95	97.95
$T_{(4)}$	2.39	PO	1.55	310.39	320	4.96	96.50	97.99	1.92	311.41	320	4.95	95.72	97.15
		OBF	1.01	305.07	320	4.99	96.85	97.32	1.75	305.21	320	4.98	95.55	95.55
		WT	1.19	306.01	320	4.98	96.82	97.66	1.77	306.84	320	4.97	95.48	96.05
		HP	1.45	316.57	320	4.99	96.54	96.74	1.74	301.81	320	4.99	95.57	95.58

TS	NSD $\hat{\alpha}\%$	Design	Inverse Normalmethode						Inverse $\chi^2$ -Methode					
			$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$ { $KI_j$ } $KI_k$		$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$ { $KI_j$ } $KI_k$	
		SD	1.02	197.93	196	3.35	95.85	96.78	1.48	198.43	196	3.31	95.65	95.65
1c) $p_1 = p_2 = 0.8, D = 0.0$														
$T_{(1)}$	2.60	PO	2.67	306.92	320	4.94	94.62	96.53	3.10	308.27	320	4.92	93.92	95.85
		OBF	2.50	299.80	320	4.98	94.87	95.26	2.75	301.78	320	4.98	94.47	94.47
		WT	2.47	301.78	320	4.97	94.88	95.73	2.75	303.44	320	4.96	94.64	95.16
		HP	2.73	315.26	320	4.99	94.51	94.79	2.62	299.32	320	4.99	94.52	94.53
		SD	2.88	196.91	196	3.47	93.50	94.53	2.74	196.65	196	3.37	94.77	94.77
$T_{(2)}$	2.48	PO	2.16	307.85	320	4.95	95.36	97.07	2.39	309.41	320	4.94	95.12	96.72
		OBF	1.84	299.66	320	4.99	95.83	96.17	1.99	303.40	320	4.99	95.60	95.60
		WT	1.68	302.75	320	4.98	96.07	96.83	2.09	304.68	320	4.97	95.54	95.99
		HP	2.53	315.69	320	4.99	95.09	95.22	1.95	299.99	320	4.99	95.63	95.63
		SD	2.22	197.02	196	3.47	94.75	95.47	1.95	197.65	196	3.38	95.71	95.71
$T_{(3)}$	1.47	PO	0.52	313.55	320	4.98	98.76	99.38	0.58	314.07	320	4.98	98.55	99.19
		OBF	0.08	309.37	320	5.00	99.60	99.67	0.44	308.78	320	5.00	98.91	98.91
		WT	0.14	310.99	320	5.00	99.34	99.64	0.53	310.20	320	4.99	98.75	98.92
		HP	0.32	318.11	320	5.00	99.25	99.38	0.41	306.12	320	5.00	98.93	98.94
		SD	0.18	198.96	196	3.26	98.66	99.12	0.43	198.67	196	3.26	98.84	98.84
$T_{(4)}$	2.01	PO	1.32	310.99	320	4.96	96.83	98.25	1.48	311.90	320	4.96	96.63	97.78
		OBF	0.56	305.49	320	5.00	97.65	98.02	0.99	306.18	320	4.99	96.83	96.83
		WT	0.66	307.60	320	4.99	97.63	98.27	1.05	307.46	320	4.98	96.84	97.19
		HP	0.95	317.00	320	4.99	97.49	97.64	1.03	303.30	320	4.99	96.80	96.80
		SD	0.60	199.06	196	3.35	96.85	97.52	0.99	197.70	196	3.30	96.53	96.53

**Tabelle C.3:** Empirisches Niveau  $\hat{\alpha}\%$  im nicht-sequentiellen Design (NSD) sowie  $\hat{\alpha}\%$ , mittlerer und medianer Umfang (ASN = „Average Sample Number“,  $N_{med}$ ), mittlere Stufenanzahl  $\bar{k}$  und empirische Überdeckungswahrscheinlichkeiten  $(1 - 2\hat{\alpha})\%$  der Sequenz der beobachteten Intervalle  $\{KI_j\}$  sowie des finalen Intervalls  $KI_k$  in flexiblen Designs mit binärer Zielvariable unter der Nullhypothese (Szenario I, Konstellation IV)

TS	NSD $\hat{\alpha}\%$	Design	Inverse Normalmethode						Inverse $\chi^2$ -Methode					
			$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$ { $KI_j$ } $KI_k$		$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$ { $KI_j$ } $KI_k$	
1a) $p_1 = p_2 = 0.2, D = 0.0$														
$T_{(1)}$	2.54	PO	2.48	633.91	670	4.94	94.70	96.67	2.60	635.35	668	4.93	94.37	96.25
		OBF	2.31	614.28	670	4.98	95.14	95.58	2.53	617.17	668	4.98	94.84	94.84
		WT	2.29	619.56	670	4.97	95.05	96.00	2.44	621.29	668	4.97	94.77	95.19
		HP	2.71	652.35	670	4.98	94.60	94.86	2.55	608.44	668	4.99	94.70	94.73
		SD	2.44	537.58	546	4.64	93.88	95.04	2.43	538.21	544	4.50	94.91	94.91
$T_{(2)}$	2.40	PO	1.99	635.52	670	4.95	95.48	97.24	2.18	636.74	668	4.94	95.30	96.83
		OBF	1.49	615.52	670	4.99	96.26	96.66	2.00	617.81	668	4.99	95.60	95.60
		WT	1.55	621.99	670	4.98	96.00	96.92	1.88	622.68	668	4.97	95.51	95.92
		HP	2.46	654.40	670	4.99	95.05	95.23	2.15	611.77	668	4.99	95.48	95.49
		SD	1.88	536.61	546	4.63	95.00	95.82	1.97	540.60	544	4.50	95.55	95.55
$T_{(3)}$	1.58	PO	0.53	648.54	670	4.98	98.57	99.30	0.83	646.93	668	4.98	98.29	98.91
		OBF	0.25	637.75	670	5.00	99.18	99.31	0.48	632.37	668	5.00	98.64	98.64
		WT	0.28	639.58	670	4.99	99.09	99.43	0.61	635.69	668	4.99	98.41	98.58
		HP	0.47	660.08	670	4.99	98.84	98.89	0.52	622.71	668	4.99	98.58	98.58

TS	NSD $\hat{\alpha}\%$	Design	Inverse Normalmethode						Inverse $\chi^2$ -Methode					
			$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2 \hat{\alpha})\%$ $\{KI_j\}$ $KI_k$		$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2 \hat{\alpha})\%$ $\{KI_j\}$ $KI_k$	
		SD	0.16	556.80	546	4.45	98.85	99.06	0.50	549.42	544	4.41	98.66	98.66
$T_{(4)}$	2.00	PO	1.20	642.07	670	4.97	96.69	98.23	1.41	642.03	668	4.96	96.56	97.69
		OBF	0.63	628.96	670	5.00	97.31	97.63	1.13	626.12	668	4.99	96.49	96.49
		WT	0.70	632.41	670	4.99	97.20	97.98	1.16	629.55	668	4.98	96.33	96.79
		HP	1.07	657.18	670	4.99	96.57	96.75	1.08	617.93	668	4.99	96.26	96.27
		SD	0.52	553.05	546	4.54	96.66	97.36	1.01	546.89	544	4.46	96.50	96.50
1b) $p_1 = p_2 = 0.5, D = 0.0$														
$T_{(1)}$	2.39	PO	2.80	632.05	670	4.93	94.72	96.36	3.13	633.70	668	4.91	93.67	95.57
		OBF	3.16	609.66	670	4.97	94.08	94.57	3.34	613.21	668	4.96	93.79	93.79
		WT	3.25	616.14	670	4.95	93.84	94.91	3.28	618.72	668	4.95	93.88	94.36
		HP	2.68	652.85	670	4.98	94.60	94.98	3.41	606.46	668	4.97	93.59	93.65
		SD	2.95	531.99	546	4.57	92.80	94.28	3.46	535.81	544	4.48	93.50	93.50
$T_{(2)}$	2.39	PO	2.15	634.11	670	4.94	95.44	97.01	2.69	634.93	668	4.93	94.58	96.14
		OBF	2.60	609.20	670	4.98	94.96	95.35	2.56	616.11	668	4.98	94.83	94.83
		WT	2.26	618.52	670	4.97	95.12	96.03	2.63	620.98	668	4.96	94.76	95.17
		HP	2.53	653.55	670	4.99	95.04	95.32	2.82	610.58	668	4.98	94.52	94.52
		SD	2.42	533.28	546	4.58	94.10	95.12	2.58	538.54	544	4.50	94.62	94.62
$T_{(3)}$	1.64	PO	0.92	644.05	670	4.97	98.20	98.84	1.35	643.59	668	4.96	97.10	98.17
		OBF	0.50	630.85	670	5.00	98.73	98.82	1.09	627.45	668	4.99	97.75	97.75
		WT	0.67	634.61	670	4.99	98.46	98.81	1.01	631.43	668	4.98	97.73	97.92
		HP	0.77	659.14	670	4.99	98.18	98.38	1.35	619.75	668	4.98	97.35	97.37
		SD	0.49	552.78	546	4.48	98.14	98.57	0.89	550.42	544	4.45	97.83	97.83
$T_{(4)}$	2.39	PO	1.65	638.75	670	4.95	96.22	97.79	1.91	638.91	668	4.95	95.34	96.90
		OBF	1.36	621.20	670	4.99	96.22	96.56	1.74	621.03	668	4.98	95.52	95.52
		WT	1.26	627.80	670	4.98	96.31	97.24	1.77	625.67	668	4.97	95.40	95.96
		HP	1.32	656.80	670	4.99	96.09	96.35	1.89	614.19	668	4.98	95.13	95.14
		SD	1.22	544.65	546	4.53	95.30	96.26	1.81	545.43	544	4.47	95.23	95.23
1c) $p_1 = p_2 = 0.8, D = 0.0$														
$T_{(1)}$	2.60	PO	2.64	631.32	670	4.94	94.70	96.55	2.92	632.79	668	4.92	94.22	96.11
		OBF	2.45	613.07	670	4.98	95.04	95.41	2.58	614.16	668	4.98	94.55	94.55
		WT	2.45	617.67	670	4.97	95.09	95.89	2.63	618.32	668	4.96	94.76	95.20
		HP	2.71	652.30	670	4.99	94.47	94.81	2.57	607.92	668	4.99	94.73	94.74
		SD	2.40	537.04	546	4.66	94.02	95.19	2.61	537.08	544	4.52	94.74	94.74
$T_{(2)}$	2.48	PO	2.05	633.53	670	4.95	95.46	97.20	2.34	634.70	668	4.94	95.11	96.75
		OBF	1.79	611.33	670	4.99	95.75	96.18	1.97	615.75	668	4.99	95.65	95.65
		WT	1.61	620.25	670	4.98	96.26	97.05	2.08	620.15	668	4.98	95.48	95.91
		HP	2.62	653.62	670	4.99	94.97	95.10	1.94	610.57	668	4.99	95.72	95.72
		SD	1.92	536.25	546	4.65	95.11	95.87	1.99	540.09	544	4.52	95.49	95.49
$T_{(3)}$	1.47	PO	0.59	647.25	670	4.98	98.60	99.28	0.74	645.56	668	4.98	98.32	98.94
		OBF	0.19	636.33	670	5.00	99.28	99.38	0.56	631.43	668	5.00	98.64	98.64
		WT	0.22	640.18	670	5.00	99.15	99.47	0.65	634.41	668	4.99	98.47	98.63
		HP	0.50	660.63	670	5.00	98.81	98.92	0.44	623.75	668	5.00	98.67	98.68
		SD	0.13	556.33	546	4.45	98.97	99.28	0.41	549.77	544	4.43	98.77	98.77
$T_{(4)}$	2.01	PO	1.33	641.22	670	4.96	96.53	98.03	1.33	641.48	668	4.97	96.40	97.76
		OBF	0.70	626.30	670	5.00	97.09	97.47	1.21	624.26	668	4.99	96.03	96.03
		WT	0.70	631.51	670	4.99	97.05	97.84	1.20	627.39	668	4.99	96.23	96.64
		HP	1.09	657.25	670	4.99	96.55	96.70	1.13	617.26	668	4.99	96.32	96.32

TS	NSD	Design	Inverse Normalmethode					Inverse $\chi^2$ -Methode						
			$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$	$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$		
	$\hat{\alpha}\%$						$\{KI_j\}$	$KI_k$				$\{KI_j\}$	$KI_k$	
		SD	0.54	552.53	546	4.55	96.84	97.32	0.88	547.08	544	4.49	96.48	96.48

**Tabelle C.4:** Empirisches Niveau  $\hat{\alpha}\%$  im nicht-sequentiellen Design (NSD) sowie  $\hat{\alpha}\%$ , mittlerer und medianer Umfang (ASN = „Average Sample Number“,  $N_{med}$ ), mittlere Stufenanzahl  $\bar{k}$  und empirische Überdeckungswahrscheinlichkeiten  $(1 - 2\hat{\alpha})\%$  der Sequenz der beobachteten Intervalle  $\{KI_j\}$  sowie des finalen Intervalls  $KI_k$  in flexiblen Designs mit binärer Zielvariable unter der Nullhypothese (Szenario I, Konstellation V)

TS	NSD	Design	Inverse Normalmethode					Inverse $\chi^2$ -Methode						
			$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$	$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$		
	$\hat{\alpha}\%$						$\{KI_j\}$	$KI_k$				$\{KI_j\}$	$KI_k$	
1a) $p_1 = p_2 = 0.2, D = 0.0$														
$T_{(1)}$	2.54	PO	2.64	303.73	308	2.97	94.71	96.20	2.77	303.47	308	2.96	94.41	95.78
		OBF	2.53	304.32	308	2.99	95.00	95.24	2.56	303.96	308	2.99	94.60	94.60
		WT	2.66	304.06	308	2.98	94.64	95.32	2.69	303.93	308	2.98	94.49	94.84
		HP	2.52	306.39	308	3.00	94.90	94.99	2.61	304.65	308	3.00	94.70	94.70
		SD	2.46	209.97	208	2.05	94.43	94.67	2.64	209.10	208	2.04	94.58	94.58
$T_{(2)}$	2.40	PO	2.34	304.24	308	2.97	95.22	96.60	2.35	304.12	308	2.97	95.22	96.34
		OBF	2.46	304.31	308	2.99	95.06	95.29	2.44	304.24	308	2.99	94.86	94.86
		WT	2.46	304.12	308	2.98	94.89	95.52	2.38	304.18	308	2.98	95.04	95.40
		HP	2.36	306.51	308	3.00	95.15	95.22	2.38	304.91	308	3.00	95.19	95.19
		SD	2.35	210.05	208	2.05	94.74	94.91	2.66	208.97	208	2.04	94.61	94.61
$T_{(3)}$	1.58	PO	1.09	305.66	308	2.98	97.82	98.58	1.19	305.79	308	2.98	97.66	98.29
		OBF	0.75	305.56	308	3.00	98.37	98.48	0.93	305.84	308	3.00	98.01	98.01
		WT	0.88	305.71	308	2.99	98.14	98.43	1.03	305.85	308	2.99	97.76	97.95
		HP	0.84	307.04	308	3.00	98.31	98.34	1.04	306.37	308	3.00	97.93	97.93
		SD	1.06	209.69	208	2.04	97.60	97.67	1.13	209.24	208	2.03	97.51	97.51
$T_{(4)}$	2.00	PO	1.61	305.12	308	2.98	96.60	97.61	1.83	304.88	308	2.98	96.24	97.17
		OBF	1.33	305.01	308	3.00	96.69	96.95	1.60	305.05	308	2.99	96.13	96.13
		WT	1.54	305.09	308	2.99	96.49	97.01	1.78	304.96	308	2.98	96.06	96.31
		HP	1.34	306.84	308	3.00	96.76	96.82	1.68	305.67	308	3.00	96.37	96.37
		SD	1.57	209.80	208	2.04	96.15	96.28	1.70	209.19	208	2.03	95.83	95.83
1b) $p_1 = p_2 = 0.5, D = 0.0$														
$T_{(1)}$	2.39	PO	2.85	303.83	308	2.97	94.62	96.01	2.72	304.08	308	2.97	94.66	95.99
		OBF	2.74	303.95	308	2.99	94.86	95.02	2.94	304.13	308	2.99	94.31	94.31
		WT	2.82	303.97	308	2.98	94.82	95.40	2.83	304.29	308	2.98	94.46	94.83
		HP	2.75	306.09	308	2.99	94.84	94.92	2.94	304.74	308	2.99	94.40	94.40
		SD	2.51	210.41	208	2.06	94.79	94.94	2.57	208.86	208	2.03	94.80	94.80
$T_{(2)}$	2.39	PO	2.77	303.88	308	2.97	94.75	96.14	2.45	304.48	308	2.97	95.26	96.34
		OBF	2.63	304.06	308	2.99	95.10	95.26	2.79	304.34	308	2.99	94.60	94.60
		WT	2.71	304.14	308	2.98	95.01	95.60	2.58	304.49	308	2.98	94.90	95.22
		HP	2.62	306.20	308	3.00	95.12	95.20	2.78	304.88	308	3.00	94.70	94.70
		SD	2.32	210.09	208	2.05	95.07	95.16	2.49	208.97	208	2.03	95.03	95.03
$T_{(3)}$	1.64	PO	1.15	305.73	308	2.99	97.65	98.40	1.57	305.37	308	2.98	96.81	97.85
		OBF	0.98	305.36	308	3.00	97.91	98.00	1.41	305.48	308	2.99	97.38	97.38
		WT	1.25	305.43	308	2.99	97.62	97.90	1.44	305.61	308	2.99	97.43	97.57
		HP	1.15	306.85	308	3.00	97.74	97.79	1.40	306.01	308	3.00	97.42	97.42

TS	NSD	Design	Inverse Normalmethode					Inverse $\chi^2$ -Methode						
			$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$	$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$		
	$\hat{\alpha}\%$						$\{KI_j\}$	$KI_k$				$\{KI_j\}$	$KI_k$	
		SD	1.23	209.61	208	2.04	97.28	97.37	1.29	209.59	208	2.03	97.22	97.22
$T_{(4)}$	2.39	PO	2.06	304.58	308	2.97	95.81	97.10	2.09	304.82	308	2.97	95.80	96.94
		OBF	1.72	304.77	308	3.00	96.43	96.58	2.03	304.90	308	2.99	95.74	95.74
		WT	1.92	304.87	308	2.99	96.21	96.69	2.05	305.00	308	2.98	95.72	96.12
		HP	1.84	306.53	308	3.00	96.34	96.42	2.05	305.48	308	3.00	95.80	95.80
		SD	1.77	209.45	208	2.04	95.88	96.00	1.96	209.30	208	2.03	95.68	95.68
1c) $p_1 = p_2 = 0.8, D = 0.0$														
$T_{(1)}$	2.60	PO	2.73	303.62	308	2.97	94.89	96.26	2.90	303.48	308	2.96	94.32	95.80
		OBF	2.63	304.11	308	2.99	94.74	95.05	2.71	304.03	308	2.99	94.87	94.87
		WT	2.60	304.13	308	2.98	94.89	95.58	2.79	304.00	308	2.98	94.70	95.10
		HP	2.48	306.26	308	3.00	95.02	95.16	2.74	304.68	308	3.00	94.73	94.73
		SD	2.60	210.15	208	2.05	94.63	94.85	2.59	209.32	208	2.04	94.90	94.90
$T_{(2)}$	2.48	PO	2.50	303.92	308	2.97	95.20	96.55	2.53	304.07	308	2.97	95.12	96.30
		OBF	2.40	304.26	308	2.99	95.00	95.35	2.57	304.24	308	2.99	95.05	95.05
		WT	2.50	304.20	308	2.98	95.10	95.78	2.67	304.25	308	2.98	95.01	95.40
		HP	2.26	306.38	308	3.00	95.36	95.50	2.55	304.71	308	3.00	95.02	95.02
		SD	2.49	210.20	208	2.05	94.87	95.07	2.71	209.03	208	2.04	94.78	94.78
$T_{(3)}$	1.47	PO	1.04	305.60	308	2.98	97.88	98.74	1.17	305.84	308	2.98	97.68	98.35
		OBF	0.77	305.55	308	3.00	98.33	98.49	0.92	305.67	308	3.00	98.02	98.02
		WT	0.78	305.70	308	2.99	98.31	98.68	1.04	305.83	308	2.99	97.90	98.11
		HP	0.85	307.10	308	3.00	98.31	98.33	0.96	306.27	308	3.00	98.00	98.00
		SD	1.02	209.60	208	2.04	97.57	97.72	1.05	209.33	208	2.03	97.71	97.71
$T_{(4)}$	2.01	PO	1.71	304.83	308	2.98	96.50	97.56	1.85	304.96	308	2.98	96.19	97.20
		OBF	1.38	305.06	308	3.00	96.85	97.15	1.62	305.01	308	2.99	96.23	96.23
		WT	1.55	304.97	308	2.99	96.68	97.27	1.81	305.16	308	2.99	96.35	96.59
		HP	1.45	306.71	308	3.00	96.58	96.71	1.67	305.47	308	3.00	96.23	96.23
		SD	1.53	209.83	208	2.04	96.23	96.43	1.73	209.29	208	2.04	96.05	96.05

**Tabelle C.5:** Empirisches Niveau  $\hat{\alpha}\%$  im nicht-sequentiellen Design (NSD) sowie  $\hat{\alpha}\%$ , mittlerer und medianer Umfang (ASN = „Average Sample Number“,  $N_{med}$ ), mittlere Stufenanzahl  $\bar{k}$  und empirische Überdeckungswahrscheinlichkeiten  $(1 - 2\hat{\alpha})\%$  der Sequenz der beobachteten Intervalle  $\{KI_j\}$  sowie des finalen Intervalls  $KI_k$  in flexiblen Designs mit binärer Zielvariable unter der Nullhypothese (Szenario I, Konstellation VI)

TS	NSD	Design	Inverse Normalmethode					Inverse $\chi^2$ -Methode						
			$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$	$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$		
	$\hat{\alpha}\%$						$\{KI_j\}$	$KI_k$				$\{KI_j\}$	$KI_k$	
1a) $p_1 = p_2 = 0.2, D = 0.0$														
$T_{(1)}$	2.54	PO	2.78	487.56	498	2.97	94.60	96.09	2.67	488.34	498	2.96	94.56	95.98
		OBF	2.49	486.87	498	2.99	95.15	95.38	2.51	487.56	498	2.99	94.77	94.77
		WT	2.62	486.92	498	2.98	94.81	95.44	2.61	488.01	498	2.98	94.61	94.99
		HP	2.39	492.86	498	3.00	94.96	95.09	2.55	489.75	498	3.00	94.94	94.94
		SD	2.69	488.27	498	3.01	93.97	94.72	2.60	488.84	498	3.00	94.70	94.70
$T_{(2)}$	2.40	PO	2.51	488.21	498	2.97	95.07	96.42	2.40	489.04	498	2.97	95.13	96.26
		OBF	2.38	487.02	498	2.99	95.24	95.43	2.36	488.15	498	2.99	94.95	94.95
		WT	2.35	487.21	498	2.98	95.21	95.80	2.39	488.13	498	2.98	94.97	95.32
		HP	2.33	493.19	498	3.00	95.15	95.27	2.38	489.75	498	3.00	95.25	95.25

TS	NSD $\hat{\alpha}\%$	Design	Inverse Normalmethode						Inverse $\chi^2$ -Methode					
			$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$		$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$	
							$\{KI_j\}$	$KI_k$					$\{KI_j\}$	$KI_k$
		SD	2.56	488.18	498	3.01	94.42	95.13	2.48	488.98	498	3.00	94.93	94.93
$T_{(3)}$	1.58	PO	1.19	491.49	498	2.98	97.67	98.41	1.30	492.39	498	2.98	97.45	98.11
		OBF	0.77	490.13	498	3.00	98.12	98.24	1.09	491.46	498	3.00	97.62	97.62
		WT	1.03	490.80	498	2.99	97.89	98.16	1.21	491.80	498	2.99	97.52	97.72
		HP	0.85	494.75	498	3.00	98.20	98.24	1.19	492.49	498	3.00	97.58	97.58
		SD	0.87	490.60	498	3.00	97.63	98.09	1.04	491.64	498	3.00	97.78	97.78
$T_{(4)}$	2.00	PO	1.74	490.09	498	2.98	96.32	97.39	1.90	490.76	498	2.97	96.02	97.04
		OBF	1.45	489.00	498	3.00	96.40	96.59	1.59	489.82	498	2.99	95.90	95.90
		WT	1.55	489.42	498	2.99	96.35	96.83	1.78	489.77	498	2.98	95.86	96.10
		HP	1.38	493.90	498	3.00	96.50	96.59	1.65	491.10	498	3.00	95.90	95.90
		SD	1.55	489.59	498	3.01	95.58	96.10	1.57	490.31	498	3.00	95.65	95.65
1b) $p_1 = p_2 = 0.5, D = 0.0$														
$T_{(1)}$	2.39	PO	2.68	487.52	498	2.97	94.66	96.10	2.62	488.33	498	2.96	94.52	95.77
		OBF	2.60	485.00	498	2.99	94.84	95.08	2.90	486.72	498	2.98	94.28	94.28
		WT	2.63	486.01	498	2.98	94.77	95.52	2.93	487.37	498	2.98	94.16	94.46
		HP	2.51	492.34	498	2.99	94.87	95.03	2.80	488.48	498	2.99	94.26	94.26
		SD	2.50	487.37	498	3.02	94.24	95.01	2.55	488.26	498	3.00	94.57	94.57
$T_{(2)}$	2.39	PO	2.54	487.78	498	2.97	94.84	96.23	2.42	488.87	498	2.97	94.88	96.03
		OBF	2.43	485.56	498	2.99	94.95	95.21	2.82	487.17	498	2.98	94.30	94.30
		WT	2.51	486.40	498	2.98	94.89	95.58	2.77	487.67	498	2.98	94.42	94.77
		HP	2.52	492.32	498	2.99	94.99	95.15	2.77	488.62	498	2.99	94.34	94.34
		SD	2.45	487.28	498	3.01	94.29	95.03	2.49	488.75	498	3.00	94.73	94.73
$T_{(3)}$	1.64	PO	1.50	490.13	498	2.98	97.03	97.91	1.76	490.82	498	2.97	96.46	97.46
		OBF	1.05	488.61	498	3.00	97.53	97.76	1.44	489.79	498	2.99	96.91	96.91
		WT	1.31	489.21	498	2.99	97.27	97.77	1.51	489.99	498	2.99	96.75	96.98
		HP	1.24	493.87	498	3.00	97.49	97.56	1.47	491.10	498	3.00	96.85	96.85
		SD	1.06	489.94	498	3.01	97.00	97.56	1.41	490.63	498	3.00	97.01	97.01
$T_{(4)}$	2.39	PO	2.22	488.26	498	2.97	95.42	96.81	2.17	489.63	498	2.97	95.26	96.44
		OBF	1.75	487.18	498	2.99	96.05	96.30	2.12	488.30	498	2.99	94.98	94.98
		WT	1.90	487.85	498	2.99	95.64	96.40	2.04	488.67	498	2.98	95.20	95.62
		HP	1.83	493.10	498	2.99	95.92	96.08	2.12	489.79	498	2.99	94.93	94.93
		SD	1.61	488.89	498	3.01	95.28	95.96	1.88	489.77	498	3.00	95.31	95.31
1c) $p_1 = p_2 = 0.8, D = 0.0$														
$T_{(1)}$	2.60	PO	2.67	487.80	498	2.97	94.96	96.35	2.80	488.21	498	2.96	94.53	95.94
		OBF	2.46	486.10	498	2.99	94.96	95.27	2.68	487.31	498	2.99	94.94	94.94
		WT	2.47	487.48	498	2.98	95.21	95.90	2.74	487.83	498	2.98	94.84	95.17
		HP	2.47	492.54	498	3.00	95.22	95.36	2.66	488.96	498	3.00	94.92	94.92
		SD	2.48	487.81	498	3.01	94.10	94.94	2.68	489.04	498	3.00	94.71	94.71
$T_{(2)}$	2.48	PO	2.36	488.57	498	2.97	95.30	96.71	2.45	489.41	498	2.97	95.22	96.39
		OBF	2.19	486.65	498	2.99	95.23	95.57	2.53	488.03	498	2.99	95.20	95.20
		WT	2.39	487.59	498	2.98	95.28	95.99	2.57	488.24	498	2.98	95.08	95.39
		HP	2.39	492.74	498	3.00	95.34	95.47	2.60	489.01	498	3.00	94.99	94.99
		SD	2.26	488.18	498	3.01	94.60	95.28	2.59	489.01	498	3.00	94.89	94.89
$T_{(3)}$	1.47	PO	1.11	491.27	498	2.98	97.73	98.58	1.26	492.49	498	2.98	97.50	98.22
		OBF	0.85	489.88	498	3.00	98.13	98.31	1.01	490.90	498	3.00	97.79	97.79
		WT	0.87	490.62	498	2.99	98.05	98.47	1.13	491.45	498	2.99	97.65	97.85
		HP	0.88	494.73	498	3.00	98.18	98.21	1.07	492.64	498	3.00	97.66	97.66

TS	NSD	Design	Inverse Normalmethode					Inverse $\chi^2$ -Methode						
			$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	(1 - 2 $\hat{\alpha}$ )% { $KI_j$ } $KI_k$	$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	(1 - 2 $\hat{\alpha}$ )% { $KI_j$ } $KI_k$		
		SD	0.85	490.11	498	3.00	97.50	98.04	1.09	491.78	498	3.00	97.72	97.72
$T_{(4)}$	2.01	PO	1.66	490.12	498	2.98	96.38	97.49	1.94	490.81	498	2.97	95.92	96.94
		OBF	1.34	488.72	498	3.00	96.65	96.91	1.78	489.60	498	2.99	96.03	96.03
		WT	1.50	489.19	498	2.99	96.50	97.11	1.87	489.96	498	2.99	96.04	96.36
		HP	1.55	493.58	498	3.00	96.29	96.40	1.71	490.68	498	3.00	95.85	95.85
		SD	1.39	488.68	498	3.01	95.83	96.53	1.61	490.43	498	3.00	95.91	95.91

**Tabelle C.6:** Empirisches Niveau  $\hat{\alpha}\%$  im nicht-sequentiellen Design (NSD) sowie  $\hat{\alpha}\%$ , mittlerer und medianer Umfang (ASN = „Average Sample Number“,  $N_{med}$ ), mittlere Stufenanzahl  $\bar{k}$  und empirische Überdeckungswahrscheinlichkeiten (1 - 2 $\hat{\alpha}$ )% der Sequenz der beobachteten Intervalle { $KI_j$ } sowie des finalen Intervalls  $KI_k$  in flexiblen Designs mit binärer Zielvariable unter der Nullhypothese (Szenario I, Konstellation VII)

TS	NSD	Design	Inverse Normalmethode					Inverse $\chi^2$ -Methode						
			$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	(1 - 2 $\hat{\alpha}$ )% { $KI_j$ } $KI_k$	$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	(1 - 2 $\hat{\alpha}$ )% { $KI_j$ } $KI_k$		
1a) $p_1 = p_2 = 0.2, D = 0.0$														
$T_{(1)}$	2.54	PO	2.56	296.96	308	2.97	94.92	96.29	2.82	298.52	308	2.96	94.37	95.75
		OBF	2.51	293.67	308	2.99	94.79	95.00	2.61	296.14	308	2.99	94.57	94.57
		WT	2.64	294.01	308	2.98	94.53	95.19	2.59	296.80	308	2.98	94.59	94.94
		HP	2.50	303.76	308	3.00	94.97	95.05	2.64	298.69	308	3.00	94.69	94.69
		SD	2.60	213.71	208	2.27	94.67	94.84	2.56	214.62	208	2.22	94.52	94.52
$T_{(2)}$	2.40	PO	2.23	297.55	308	2.97	95.45	96.75	2.36	299.16	308	2.97	95.25	96.35
		OBF	2.11	294.28	308	2.99	95.49	95.68	2.28	296.85	308	2.99	95.13	95.13
		WT	1.95	295.13	308	2.99	95.47	96.06	2.26	297.40	308	2.98	95.11	95.42
		HP	2.46	303.83	308	3.00	95.21	95.26	2.29	299.06	308	3.00	95.24	95.24
		SD	2.33	214.05	208	2.27	95.11	95.23	2.29	214.93	208	2.22	94.88	94.88
$T_{(3)}$	1.58	PO	1.00	300.30	308	2.99	97.92	98.66	1.10	301.69	308	2.98	97.77	98.40
		OBF	0.52	297.47	308	3.00	98.59	98.73	0.82	299.61	308	3.00	98.12	98.12
		WT	0.63	298.29	308	3.00	98.37	98.69	0.94	300.25	308	2.99	97.91	98.07
		HP	0.81	305.15	308	3.00	98.19	98.23	0.88	301.32	308	3.00	98.04	98.04
		SD	0.63	215.45	208	2.22	98.12	98.14	1.03	214.73	208	2.19	97.66	97.66
$T_{(4)}$	2.00	PO	1.60	298.92	308	2.98	96.56	97.57	1.84	300.27	308	2.97	96.06	97.09
		OBF	1.08	295.99	308	3.00	96.95	97.19	1.53	298.05	308	2.99	96.10	96.10
		WT	1.09	297.07	308	2.99	97.05	97.54	1.66	298.75	308	2.99	96.12	96.38
		HP	1.32	304.52	308	3.00	96.82	96.90	1.60	300.09	308	3.00	96.39	96.39
		SD	1.11	214.81	208	2.25	96.66	96.76	1.47	215.08	208	2.21	95.89	95.89
1b) $p_1 = p_2 = 0.5, D = 0.0$														
$T_{(1)}$	2.39	PO	2.90	297.33	308	2.97	94.55	95.94	2.67	299.12	308	2.97	94.68	96.00
		OBF	2.93	293.59	308	2.99	94.56	94.72	3.13	296.27	308	2.98	93.98	93.98
		WT	2.84	294.45	308	2.98	94.65	95.26	2.91	297.18	308	2.98	94.26	94.65
		HP	2.76	303.57	308	2.99	94.81	94.90	3.05	298.56	308	2.99	94.15	94.15
		SD	2.77	213.76	208	2.28	94.30	94.73	2.63	214.25	208	2.23	94.80	94.80
$T_{(2)}$	2.39	PO	2.65	297.62	308	2.97	94.87	96.24	2.53	299.35	308	2.97	95.16	96.24
		OBF	2.48	294.01	308	2.99	95.29	95.44	2.71	296.70	308	2.99	94.66	94.66
		WT	2.69	294.78	308	2.98	94.98	95.59	2.53	297.60	308	2.98	94.88	95.21
		HP	2.58	303.67	308	3.00	95.13	95.21	2.74	298.94	308	3.00	94.65	94.65

TS	NSD	Design	Inverse Normalmethode						Inverse $\chi^2$ -Methode					
			$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$		$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$	
	$\hat{\alpha}\%$						$\{KI_j\}$	$KI_k$					$\{KI_j\}$	$KI_k$
		SD	2.60	214.43	208	2.28	94.76	94.91	2.40	214.77	208	2.23	95.13	95.13
$T_{(3)}$	1.64	PO	1.05	299.99	308	2.99	97.75	98.49	1.51	301.26	308	2.98	96.88	97.91
		OBF	0.81	296.63	308	3.00	98.08	98.16	1.35	299.10	308	2.99	97.43	97.43
		WT	0.96	297.75	308	2.99	97.89	98.17	1.41	299.70	308	2.99	97.44	97.59
		HP	1.03	304.95	308	3.00	97.86	97.91	1.36	300.80	308	3.00	97.44	97.44
		SD	0.88	214.72	208	2.24	97.75	97.84	1.27	216.59	208	2.22	97.21	97.21
$T_{(4)}$	2.39	PO	2.01	298.47	308	2.97	95.77	97.06	2.05	300.26	308	2.97	95.77	96.90
		OBF	1.69	295.25	308	3.00	96.36	96.54	2.00	297.72	308	2.99	95.58	95.58
		WT	1.89	296.18	308	2.99	96.10	96.64	2.04	298.67	308	2.98	95.58	95.94
		HP	1.76	304.27	308	3.00	96.38	96.46	1.94	299.90	308	3.00	95.78	95.78
		SD	1.48	213.55	208	2.25	96.05	96.19	1.92	215.74	208	2.23	95.64	95.64
1c) $p_1 = p_2 = 0.8, D = 0.0$														
$T_{(1)}$	2.60	PO	2.57	296.83	308	2.97	95.07	96.44	2.82	298.48	308	2.96	94.38	95.90
		OBF	2.61	293.02	308	2.99	94.83	95.14	2.80	296.18	308	2.99	94.82	94.82
		WT	2.52	294.48	308	2.98	95.03	95.65	2.63	296.53	308	2.98	94.82	95.25
		HP	2.38	303.39	308	3.00	95.04	95.19	2.72	297.89	308	3.00	94.61	94.61
		SD	2.58	214.28	208	2.29	94.83	94.99	2.72	214.70	208	2.24	94.74	94.74
$T_{(2)}$	2.48	PO	2.19	297.15	308	2.97	95.53	96.88	2.37	299.23	308	2.97	95.38	96.50
		OBF	2.17	294.44	308	3.00	95.30	95.63	2.39	296.86	308	2.99	95.36	95.36
		WT	2.26	294.78	308	2.99	95.40	96.03	2.41	297.28	308	2.98	95.24	95.58
		HP	2.18	303.55	308	3.00	95.30	95.46	2.43	298.12	308	3.00	95.20	95.20
		SD	2.44	214.73	208	2.29	95.18	95.25	2.52	215.09	208	2.24	95.05	95.05
$T_{(3)}$	1.47	PO	0.93	299.81	308	2.99	97.95	98.82	1.13	301.16	308	2.98	97.74	98.40
		OBF	0.58	297.21	308	3.00	98.54	98.72	0.77	298.69	308	3.00	98.13	98.13
		WT	0.54	298.16	308	3.00	98.52	98.87	0.91	299.37	308	2.99	97.95	98.17
		HP	0.70	305.31	308	3.00	98.37	98.41	0.91	301.28	308	3.00	98.06	98.06
		SD	0.62	215.69	208	2.23	98.28	98.30	0.97	215.12	208	2.20	97.84	97.84
$T_{(4)}$	2.01	PO	1.44	298.46	308	2.98	96.76	97.80	1.77	300.28	308	2.98	96.20	97.27
		OBF	1.22	295.94	308	3.00	96.90	97.20	1.52	298.07	308	2.99	96.24	96.24
		WT	1.09	296.86	308	2.99	96.97	97.60	1.68	298.46	308	2.99	96.38	96.64
		HP	1.54	304.31	308	3.00	96.45	96.55	1.56	299.93	308	3.00	96.32	96.32
		SD	1.20	215.15	208	2.26	96.67	96.73	1.70	215.24	208	2.22	96.09	96.09

**Tabelle C.7:** Empirisches Niveau  $\hat{\alpha}\%$  im nicht-sequentiellen Design (NSD) sowie  $\hat{\alpha}\%$ , mittlerer und medianer Umfang (ASN = „Average Sample Number“,  $N_{med}$ ), mittlere Stufenanzahl  $\bar{k}$  und empirische Überdeckungswahrscheinlichkeiten  $(1 - 2\hat{\alpha})\%$  der Sequenz der beobachteten Intervalle  $\{KI_j\}$  sowie des finalen Intervalls  $KI_k$  in flexiblen Designs mit binärer Zielvariable unter der Nullhypothese (Szenario I, Konstellation VIII)

TS	NSD	Design	Inverse Normalmethode						Inverse $\chi^2$ -Methode					
			$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$		$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$	
	$\hat{\alpha}\%$						$\{KI_j\}$	$KI_k$					$\{KI_j\}$	$KI_k$
1a) $p_1 = p_2 = 0.2, D = 0.0$														
$T_{(1)}$	2.54	PO	2.57	472.41	498	2.97	94.92	96.29	2.72	476.22	498	2.96	94.57	95.92
		OBF	2.51	465.05	498	2.99	94.91	95.12	2.52	469.83	498	2.99	94.69	94.69
		WT	2.59	466.14	498	2.98	94.86	95.44	2.53	471.53	498	2.98	94.68	95.05
		HP	2.43	485.57	498	3.00	94.99	95.11	2.58	475.13	498	3.00	94.86	94.86

TS	NSD $\hat{\alpha}\%$	Design	Inverse Normalmethode						Inverse $\chi^2$ -Methode					
			$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2 \hat{\alpha})\%$ $\{KI_j\}$ $KI_k$		$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2 \hat{\alpha})\%$ $\{KI_j\}$ $KI_k$	
$T_{(2)}$	2.40	SD	2.61	468.33	498	3.13	94.41	94.97	2.40	474.75	498	3.10	95.02	95.02
		PO	2.26	473.39	498	2.97	95.43	96.71	2.34	476.83	498	2.97	95.24	96.33
		OBF	2.07	466.08	498	2.99	95.54	95.72	2.25	471.48	498	2.99	95.08	95.08
		WT	1.91	468.28	498	2.99	95.64	96.23	2.21	472.24	498	2.98	95.10	95.42
		HP	2.34	485.55	498	3.00	95.25	95.34	2.26	475.26	498	3.00	95.34	95.34
$T_{(3)}$	1.58	SD	2.22	468.78	498	3.13	94.96	95.43	2.07	475.45	498	3.10	95.47	95.47
		PO	1.06	479.38	498	2.99	97.85	98.56	1.23	481.69	498	2.98	97.54	98.17
		OBF	0.57	473.83	498	3.00	98.32	98.51	0.95	476.54	498	3.00	97.76	97.76
		WT	0.68	475.29	498	3.00	98.23	98.56	1.10	477.86	498	2.99	97.64	97.84
		HP	0.86	488.81	498	3.00	98.02	98.06	1.06	479.99	498	3.00	97.77	97.77
$T_{(4)}$	2.00	SD	0.58	477.78	498	3.10	98.02	98.31	0.97	480.14	498	3.09	97.82	97.82
		PO	1.61	476.48	498	2.98	96.39	97.45	1.88	479.32	498	2.97	96.02	97.05
		OBF	1.05	470.46	498	3.00	96.65	96.90	1.51	473.53	498	2.99	95.96	95.96
		WT	1.11	472.64	498	2.99	96.84	97.31	1.57	475.23	498	2.99	96.00	96.22
		HP	1.35	487.10	498	3.00	96.60	96.69	1.56	477.57	498	3.00	96.06	96.06
$T_{(1)}$	2.39	SD	1.18	473.41	498	3.12	96.21	96.69	1.36	477.43	498	3.09	96.01	96.01
		PO	2.99	472.44	498	2.96	94.36	95.72	2.93	476.17	498	2.96	94.21	95.55
		OBF	2.74	464.46	498	2.99	94.70	94.92	3.03	470.12	498	2.98	93.93	93.93
		WT	2.79	466.44	498	2.98	94.63	95.20	2.91	472.32	498	2.98	94.10	94.52
		HP	2.47	485.07	498	2.99	94.82	94.94	2.93	474.17	498	2.99	94.24	94.24
$T_{(2)}$	2.39	SD	2.67	468.29	498	3.14	94.30	94.88	2.96	473.56	498	3.09	94.22	94.22
		PO	2.65	473.11	498	2.97	94.80	96.16	2.53	477.02	498	2.97	94.88	96.11
		OBF	2.36	465.01	498	2.99	95.23	95.45	2.71	470.90	498	2.99	94.47	94.47
		WT	2.61	467.17	498	2.98	94.92	95.41	2.67	472.97	498	2.98	94.54	94.99
		HP	2.46	485.04	498	2.99	94.90	95.01	2.76	474.85	498	2.99	94.46	94.46
$T_{(3)}$	1.64	SD	2.51	468.43	498	3.14	94.56	95.05	2.78	474.37	498	3.09	94.41	94.41
		PO	1.26	478.05	498	2.98	97.27	98.16	1.74	480.40	498	2.97	96.49	97.50
		OBF	1.08	471.58	498	3.00	97.53	97.71	1.48	474.67	498	2.99	97.01	97.01
		WT	1.05	474.18	498	2.99	97.51	97.93	1.58	476.03	498	2.99	96.82	97.05
		HP	1.10	487.46	498	3.00	97.66	97.72	1.41	478.50	498	3.00	97.10	97.10
$T_{(4)}$	2.39	SD	0.79	476.11	498	3.11	97.56	97.85	1.46	479.12	498	3.08	97.01	97.01
		PO	2.19	474.96	498	2.97	95.28	96.71	2.15	478.63	498	2.97	95.24	96.52
		OBF	1.80	467.78	498	2.99	95.80	96.00	2.09	472.88	498	2.99	95.11	95.11
		WT	1.97	470.58	498	2.98	95.59	96.26	2.30	474.34	498	2.98	95.01	95.51
		HP	1.74	486.36	498	3.00	95.97	96.11	2.00	476.60	498	3.00	95.12	95.12
$T_{(1)}$	2.60	SD	1.40	472.85	498	3.13	95.84	96.29	2.11	475.73	498	3.09	95.19	95.19
		PO	2.57	472.33	498	2.97	95.03	96.43	2.70	475.86	498	2.96	94.60	96.01
		OBF	2.45	463.95	498	2.99	95.01	95.36	2.63	470.42	498	2.99	95.04	95.04
		WT	2.43	467.08	498	2.99	95.17	95.81	2.64	470.90	498	2.98	94.81	95.20
		HP	2.44	484.50	498	2.99	95.12	95.25	2.64	473.12	498	3.00	94.84	94.84
$T_{(2)}$	2.48	SD	2.55	466.67	498	3.14	94.49	95.17	2.56	474.64	498	3.10	94.97	94.97
		PO	2.15	473.24	498	2.97	95.48	96.91	2.32	476.97	498	2.97	95.36	96.47
		OBF	2.13	466.16	498	3.00	95.41	95.73	2.33	471.37	498	2.99	95.48	95.48
		WT	2.20	467.56	498	2.99	95.49	96.12	2.35	472.04	498	2.98	95.29	95.60
		HP	2.28	484.79	498	3.00	95.28	95.43	2.51	473.40	498	3.00	95.13	95.13

1b)  $p_1 = p_2 = 0.5, D = 0.0$

1c)  $p_1 = p_2 = 0.8, D = 0.0$

TS	NSD $\hat{\alpha}\%$	Design	Inverse Normalmethode						Inverse $\chi^2$ -Methode					
			$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$ { $KI_j$ } $KI_k$		$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$ { $KI_j$ } $KI_k$	
$T_{(3)}$	1.47	SD	2.42	467.44	498	3.14	94.86	95.38	2.48	475.25	498	3.10	95.18	95.18
		PO	0.98	478.56	498	2.98	97.81	98.71	1.20	480.76	498	2.98	97.57	98.27
		OBF	0.64	473.39	498	3.00	98.35	98.54	0.84	475.11	498	3.00	97.96	97.96
		WT	0.59	475.05	498	3.00	98.33	98.75	0.98	476.39	498	2.99	97.75	97.96
		HP	0.72	489.15	498	3.00	98.25	98.30	1.02	480.00	498	3.00	97.74	97.74
$T_{(4)}$	2.01	SD	0.52	477.38	498	3.10	98.07	98.42	0.83	479.85	498	3.09	97.90	97.90
		PO	1.43	475.86	498	2.98	96.65	97.72	1.86	479.17	498	2.98	95.92	96.98
		OBF	1.14	470.59	498	3.00	96.74	97.01	1.70	473.63	498	2.99	96.12	96.12
		WT	1.06	472.16	498	2.99	96.81	97.44	1.76	474.46	498	2.99	96.07	96.39
		HP	1.40	487.07	498	3.00	96.54	96.68	1.62	477.19	498	3.00	95.98	95.98
SD	1.04	472.90	498	3.12	96.30	96.69	1.48	477.50	498	3.09	96.09	96.09		

## 2) Unter der Alternativhypothese

**Tabelle C.8:** Empirische Power  $(1 - \hat{\beta})\%$  im nicht-sequentiellen Design (NSD) sowie  $(1 - \hat{\beta})\%$ , mittlerer und medianer Umfang (ASN = „Average Sample Number“,  $N_{med}$ ), mittlere Stufenanzahl  $\bar{k}$  und empirische Überdeckungswahrscheinlichkeiten  $(1 - 2\hat{\alpha})\%$  der Sequenz der beobachteten Intervalle  $\{KI_j\}$  sowie des finalen Intervalls  $KI_k$  in flexiblen Designs mit binärer Zielvariable unter der Alternativhypothese (Szenario I, Konstellation II)

TS	NSD $(1 - \hat{\beta})\%$	Design	Inverse Normalmethode						Inverse $\chi^2$ -Methode					
			$(1 - \hat{\beta})\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$ { $KI_j$ } $KI_k$		$(1 - \hat{\beta})\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$ { $KI_j$ } $KI_k$	
2a) $p_1 = 0.3, p_2 = 0.2, D = 0.1$														
$T_{(1)}$	61.24	PO	71.16	461.14	422	3.92	96.24	97.91	66.19	477.67	640	3.98	95.66	97.04
		OBF	77.29	456.64	432	4.40	97.12	97.55	73.59	467.77	504	4.44	96.38	96.38
		WT	76.67	455.39	422	4.14	97.04	97.84	72.96	469.03	528	4.19	96.42	96.71
		HP	80.33	499.74	490	4.48	97.08	97.25	74.27	474.70	502	4.71	96.23	96.28
		SD	64.47	401.40	470	4.24	93.85	94.80	66.89	407.21	500	3.85	96.35	96.35
$T_{(2)}$	59.87	PO	70.51	465.51	555	3.95	96.61	98.09	64.72	484.86	652	4.02	95.76	97.17
		OBF	76.93	460.41	450	4.43	97.46	97.87	72.31	472.97	524	4.48	96.79	96.79
		WT	76.92	460.25	422	4.17	97.24	97.93	71.70	476.03	562	4.25	96.61	96.89
		HP	79.98	504.13	500	4.52	97.11	97.25	73.05	482.26	526	4.75	96.50	96.53
		SD	65.01	406.35	476	4.25	95.21	95.98	66.16	410.45	502	3.89	96.80	96.80
$T_{(3)}$	53.20	PO	50.80	538.38	670	4.40	98.48	99.21	48.43	545.10	670	4.39	97.98	98.61
		OBF	54.31	532.17	608	4.76	99.22	99.33	53.95	534.05	642	4.73	98.60	98.60
		WT	55.84	531.14	626	4.59	98.98	99.40	55.33	536.68	655	4.55	98.52	98.77
		HP	59.79	555.63	663	4.76	98.83	98.90	55.23	539.59	624	4.87	98.42	98.43
		SD	39.49	455.00	542	4.19	98.68	99.09	46.69	451.30	546	3.97	98.64	98.64
$T_{(4)}$	57.41	PO	61.61	502.29	648	4.17	97.57	98.77	57.06	516.05	670	4.21	96.87	97.93
		OBF	67.13	496.87	562	4.60	98.10	98.43	64.31	502.88	602	4.60	97.08	97.08
		WT	67.45	492.99	582	4.37	97.82	98.40	64.66	506.33	618	4.40	97.23	97.48
		HP	69.81	530.01	582	4.65	97.86	97.95	65.25	510.33	590	4.80	96.78	96.80

TS	NSD (1 - β̂)%	Design	Inverse Normalmethode						Inverse χ <sup>2</sup> -Methode					
			(1 - β̂)%	ASN	N <sub>med</sub>	k̄	(1 - 2 α̂)%	(1 - β̂)%	ASN	N <sub>med</sub>	k̄	(1 - 2 α̂)%		
							{KI <sub>j</sub> }	KI <sub>k</sub>					{KI <sub>j</sub> }	KI <sub>k</sub>
		SD	50.72	433.10	512	4.22	96.30	96.98	57.20	430.45	530	3.93	97.16	97.16
2b) p <sub>1</sub> = 0.9, p <sub>2</sub> = 0.8, D = 0.1														
T <sub>(1)</sub>	61.24	PO	88.22	378.93	298	3.47	96.15	98.43	85.11	393.23	392	3.50	95.86	97.56
		OBF	89.88	382.02	354	4.15	96.72	97.25	88.14	391.35	366	4.22	96.38	96.38
		WT	91.16	375.41	342	3.77	96.42	97.61	88.85	392.21	370	3.86	96.19	96.75
		HP	92.62	429.87	422	4.16	96.46	96.99	89.12	405.49	410	4.62	96.23	96.33
		SD	80.42	346.92	328	4.29	94.37	95.31	82.89	348.85	313	3.73	95.94	95.94
T <sub>(2)</sub>	59.87	PO	87.74	385.11	345	3.52	96.20	98.45	84.15	404.28	416	3.59	95.95	97.70
		OBF	89.35	385.55	358	4.20	96.83	97.32	86.64	399.14	380	4.28	96.75	96.75
		WT	90.68	381.40	350	3.82	96.98	97.94	88.04	399.59	384	3.91	96.69	97.12
		HP	92.81	435.90	428	4.22	96.81	97.33	88.50	414.88	422	4.69	96.47	96.61
		SD	80.94	349.41	330	4.30	95.44	96.14	82.55	355.59	334	3.80	96.44	96.44
T <sub>(3)</sub>	53.20	PO	68.62	486.25	592	4.18	98.86	99.70	67.24	498.48	628	4.18	98.25	99.26
		OBF	67.16	485.76	510	4.68	99.47	99.58	68.36	483.56	520	4.65	98.98	98.98
		WT	71.58	480.63	518	4.44	99.34	99.65	71.42	485.65	540	4.41	98.90	99.07
		HP	71.21	504.67	490	4.65	99.43	99.53	71.32	491.77	528	4.85	98.81	98.83
		SD	50.17	426.17	468	4.23	98.87	99.12	61.08	416.75	486	3.94	98.75	98.75
T <sub>(4)</sub>	57.41	PO	80.36	431.72	422	3.82	97.36	98.97	76.87	448.91	422	3.87	97.12	98.31
		OBF	80.57	435.92	422	4.45	97.82	98.21	78.56	440.71	422	4.47	97.61	97.61
		WT	82.53	429.56	422	4.15	97.97	98.66	80.70	441.26	422	4.18	97.51	97.90
		HP	82.61	469.07	436	4.43	97.92	98.13	81.18	452.51	428	4.77	97.45	97.48
		SD	62.37	391.05	416	4.28	96.67	97.30	71.31	386.44	426	3.87	97.26	97.26
2c) p <sub>1</sub> = 0.7, p <sub>2</sub> = 0.5, D = 0.2														
T <sub>(1)</sub>	89.91	PO	99.68	213.79	174	2.44	96.09	97.76	99.51	218.86	174	2.42	96.38	97.37
		OBF	98.20	229.29	206	3.49	96.81	97.20	98.27	226.20	186	3.39	95.21	95.21
		WT	99.63	221.78	198	2.92	97.21	98.06	99.22	223.31	174	2.91	96.86	97.16
		HP	99.52	272.29	286	3.00	97.27	97.42	99.32	239.44	214	3.78	94.15	94.29
		SD	92.05	219.25	186	4.05	90.50	91.43	97.10	212.64	174	3.15	95.04	95.04
T <sub>(2)</sub>	89.63	PO	99.66	218.18	174	2.49	96.62	98.03	99.44	223.82	174	2.47	96.32	97.29
		OBF	97.99	232.96	208	3.56	96.79	97.22	97.94	232.41	192	3.52	95.95	95.95
		WT	99.49	225.09	202	2.96	97.40	98.05	98.97	229.06	186	2.99	97.13	97.35
		HP	99.57	276.61	292	3.07	97.83	97.95	99.13	249.02	224	3.97	94.59	94.71
		SD	93.69	220.82	186	4.12	92.89	93.63	96.85	217.04	180	3.25	95.85	95.85
T <sub>(3)</sub>	87.47	PO	98.62	264.46	242	2.89	98.33	99.08	98.39	270.46	244	2.84	98.01	98.68
		OBF	91.87	288.25	270	4.04	98.79	98.99	92.47	279.01	254	3.89	96.87	96.87
		WT	97.12	271.30	248	3.41	98.86	99.29	96.27	274.05	244	3.42	98.11	98.27
		HP	96.60	315.18	298	3.49	98.96	99.03	96.31	291.01	270	4.21	96.73	96.79
		SD	78.19	269.53	242	4.18	97.36	97.80	88.70	257.68	220	3.48	97.70	97.70
T <sub>(4)</sub>	88.99	PO	99.43	235.41	174	2.63	97.01	98.23	99.07	246.79	174	2.66	96.42	97.38
		OBF	95.99	255.60	232	3.75	97.20	97.60	95.75	250.34	216	3.66	95.95	95.95
		WT	98.90	244.36	218	3.14	97.61	98.22	97.93	250.70	216	3.19	97.21	97.45
		HP	98.24	293.17	298	3.26	98.12	98.20	98.18	265.25	242	4.07	95.13	95.18
		SD	84.56	244.61	208	4.15	94.15	94.76	93.33	236.75	186	3.34	96.33	96.33
2d) p <sub>1</sub> = 0.3, p <sub>2</sub> = 0.1, D = 0.2														
T <sub>(1)</sub>	98.45	PO	100.00	151.26	146	1.99	96.54	98.44	100.00	152.92	174	1.96	95.71	97.31
		OBF	99.09	165.52	160	3.20	96.84	97.36	99.10	161.76	161	3.12	96.35	96.35
		WT	99.94	158.98	152	2.51	97.74	98.70	99.75	158.16	158	2.53	97.03	97.44
		HP	99.98	201.63	174	2.43	98.55	98.86	99.64	171.87	168	3.56	95.14	95.31

TS	NSD (1 - β̂)%	Design	Inverse Normalmethode						Inverse χ <sup>2</sup> -Methode							
			(1 - β̂)%	ASN	N <sub>med</sub>	k̄	(1 - 2 α̂)%	{KI <sub>j</sub> }	KI <sub>k</sub>	(1 - β̂)%	ASN	N <sub>med</sub>	k̄	(1 - 2 α̂)%	{KI <sub>j</sub> }	KI <sub>k</sub>
		SD	94.32	159.22	150	3.91	92.23	93.16	97.89	150.03	140	2.89	96.16	96.16		
T <sub>(2)</sub>	98.31	PO	100.00	152.72	148	2.01	96.85	98.56	100.00	156.39	174	2.01	96.63	97.82		
		OBF	98.46	167.88	162	3.31	97.12	97.62	98.36	166.56	170	3.25	96.31	96.31		
		WT	99.91	160.93	154	2.54	97.94	98.84	99.45	161.52	164	2.61	97.52	97.83		
		HP	99.95	205.66	174	2.50	98.39	98.65	99.50	178.57	174	3.73	95.94	96.15		
		SD	95.18	160.74	154	3.96	94.50	95.29	98.05	155.06	150	3.04	96.69	96.69		
T <sub>(3)</sub>	97.56	PO	99.89	185.38	174	2.39	98.68	99.48	99.76	194.57	174	2.43	98.36	99.01		
		OBF	89.91	231.14	212	3.99	99.13	99.28	90.45	209.63	184	3.82	98.25	98.25		
		WT	97.68	204.75	180	3.15	99.25	99.65	96.61	202.86	174	3.12	99.08	99.16		
		HP	97.92	239.64	240	2.94	99.36	99.45	96.03	220.64	206	4.10	98.15	98.20		
		SD	71.33	219.69	192	4.11	98.48	98.94	87.08	196.90	174	3.33	98.68	98.68		
T <sub>(4)</sub>	97.91	PO	99.97	167.11	174	2.17	97.34	98.92	99.94	173.17	174	2.18	97.17	98.26		
		OBF	96.56	194.47	180	3.59	98.04	98.33	95.46	184.85	174	3.48	97.41	97.41		
		WT	99.22	177.63	174	2.83	98.57	99.10	98.42	177.75	174	2.86	98.12	98.33		
		HP	99.42	219.20	193	2.67	98.55	98.79	98.38	195.88	174	3.86	96.48	96.57		
		SD	79.64	186.59	180	4.04	95.74	96.37	90.97	174.42	174	3.15	97.43	97.43		
2e) p <sub>1</sub> = 0.6, p <sub>2</sub> = 0.3, D = 0.3																
T <sub>(1)</sub>	99.85	PO	100.00	115.98	106	1.71	97.09	98.08	100.00	115.12	102	1.65	96.57	97.55		
		OBF	99.59	125.86	110	2.85	95.63	95.85	99.62	120.53	106	2.60	93.82	93.82		
		WT	99.99	119.84	98	2.17	98.39	98.67	99.89	118.49	98	2.12	97.25	97.39		
		HP	100.00	146.74	154	1.96	99.17	99.34	99.94	122.81	106	2.77	93.41	93.55		
		SD	94.81	123.62	110	3.56	88.02	88.91	98.11	115.58	96	2.53	94.01	94.01		
T <sub>(2)</sub>	99.85	PO	100.00	117.67	108	1.73	97.21	98.15	100.00	118.87	114	1.71	97.09	97.75		
		OBF	99.31	127.43	114	2.94	97.01	97.36	99.31	123.58	110	2.76	94.29	94.29		
		WT	99.94	121.69	104	2.19	98.73	99.01	99.84	121.01	106	2.18	97.43	97.59		
		HP	100.00	151.43	156	2.06	99.33	99.50	99.86	128.85	116	3.12	93.90	93.98		
		SD	95.53	125.87	110	3.64	91.15	91.92	98.72	118.59	104	2.63	95.84	95.84		
T <sub>(3)</sub>	99.69	PO	100.00	131.46	120	1.93	98.83	99.23	100.00	131.54	126	1.83	98.48	98.89		
		OBF	94.73	158.38	148	3.52	99.02	99.14	95.61	145.57	138	3.18	98.81	98.81		
		WT	99.44	142.20	132	2.57	99.09	99.26	98.40	141.89	140	2.62	98.75	98.80		
		HP	99.82	168.41	174	2.26	99.54	99.61	98.49	149.27	142	3.31	98.76	98.83		
		SD	82.47	151.68	138	3.78	96.52	97.12	92.17	137.96	128	2.87	97.78	97.78		
T <sub>(4)</sub>	99.85	PO	100.00	121.24	114	1.77	97.51	98.29	100.00	125.10	118	1.78	97.22	97.76		
		OBF	98.33	139.45	124	3.17	96.77	96.98	98.38	131.96	124	2.91	94.43	94.43		
		WT	99.87	128.16	116	2.29	98.90	99.12	99.45	129.54	124	2.38	97.40	97.47		
		HP	99.97	158.37	160	2.15	99.51	99.66	99.63	136.76	126	3.19	93.34	93.47		
		SD	89.36	136.58	120	3.67	91.99	92.83	96.46	126.68	116	2.69	95.54	95.54		
2f) p <sub>1</sub> = 0.9, p <sub>2</sub> = 0.6, D = 0.3																
T <sub>(1)</sub>	99.99	PO	100.00	89.07	50	1.45	97.44	98.52	100.00	89.58	50	1.44	96.92	97.61		
		OBF	99.55	99.52	82	2.68	95.45	95.89	99.84	93.96	70	2.43	94.42	94.42		
		WT	100.00	93.59	76	1.91	98.22	98.79	99.97	91.08	66	1.88	96.80	97.15		
		HP	100.00	114.41	108	1.72	99.30	99.71	99.94	95.41	78	2.64	93.48	93.76		
		SD	95.31	100.70	84	3.43	87.96	89.06	98.72	91.38	70	2.36	94.45	94.45		
T <sub>(2)</sub>	99.99	PO	100.00	89.87	50	1.46	97.56	98.57	100.00	91.20	50	1.46	97.14	97.76		
		OBF	99.46	100.84	84	2.80	97.62	97.89	99.62	96.58	78	2.61	94.96	94.96		
		WT	100.00	94.86	78	1.94	98.54	99.02	99.91	93.64	72	1.97	97.17	97.33		
		HP	100.00	116.91	112	1.77	99.45	99.76	99.92	99.66	88	2.87	95.37	95.64		

TS	NSD (1 - β̂)%	Design	Inverse Normalmethode					Inverse χ²-Methode						
			(1 - β̂)%	ASN	N <sub>med</sub>	k̄	(1 - 2 α̂)% {KI <sub>j</sub> } KI <sub>k</sub>	(1 - β̂)%	ASN	N <sub>med</sub>	k̄	(1 - 2 α̂)% {KI <sub>j</sub> } KI <sub>k</sub>		
		SD	96.14	101.66	86	3.50	92.45	93.26	98.68	93.40	74	2.51	95.79	95.79
T <sub>(3)</sub>	99.97	PO	100.00	101.39	88	1.65	98.64	99.36	100.00	104.09	86	1.67	98.57	99.00
		OBF	93.25	135.07	122	3.56	99.55	99.63	94.77	117.48	108	3.16	99.30	99.30
		WT	99.26	113.82	100	2.40	99.51	99.77	98.46	112.71	104	2.41	98.89	98.95
		HP	99.89	131.98	130	2.01	99.71	99.86	98.19	121.77	112	3.31	99.31	99.36
		SD	75.63	130.16	114	3.68	97.80	98.49	89.74	112.25	100	2.80	97.98	97.98
T <sub>(4)</sub>	99.99	PO	100.00	94.38	50	1.54	98.21	99.02	100.00	95.98	50	1.53	97.88	98.48
		OBF	98.43	113.45	98	3.07	97.30	97.47	98.22	104.09	88	2.77	95.74	95.74
		WT	99.86	102.15	86	2.14	98.74	99.22	99.58	100.90	84	2.16	97.70	97.82
		HP	99.98	122.66	114	1.85	99.47	99.77	99.57	106.50	94	2.99	94.96	95.12
		SD	85.62	113.42	96	3.57	94.04	94.99	92.91	100.60	84	2.60	96.15	96.15

**Tabelle C.9:** Empirische Power (1 - β̂)% im nicht-sequentiellen Design (NSD) sowie (1 - β̂)%, mittlerer und medianer Umfang (ASN = „Average Sample Number“, N<sub>med</sub>), mittlere Stufenanzahl k̄ und empirische Überdeckungswahrscheinlichkeiten (1 - 2 α̂)% der Sequenz der beobachteten Intervalle {KI<sub>j</sub>} sowie des finalen Intervalls KI<sub>k</sub> in flexiblen Designs mit binärer Zielvariable unter der Alternativhypothese (Szenario I, Konstellation III)

TS	NSD (1 - β̂)%	Design	Inverse Normalmethode					Inverse χ²-Methode						
			(1 - β̂)%	ASN	N <sub>med</sub>	k̄	(1 - 2 α̂)% {KI <sub>j</sub> } KI <sub>k</sub>	(1 - β̂)%	ASN	N <sub>med</sub>	k̄	(1 - 2 α̂)% {KI <sub>j</sub> } KI <sub>k</sub>		
2a) p <sub>1</sub> = 0.3, p <sub>2</sub> = 0.2, D = 0.1														
T <sub>(1)</sub>	61.24	PO	34.36	244.23	280	4.41	95.80	97.77	32.54	252.82	300	4.42	94.61	96.28
		OBF	35.96	232.51	248	4.80	95.92	96.45	36.12	237.55	260	4.75	95.41	95.41
		WT	36.14	235.58	262	4.65	96.03	96.91	35.98	240.05	264	4.61	95.35	95.79
		HP	49.20	271.33	320	4.77	96.40	96.62	37.00	241.38	264	4.85	95.18	95.21
		SD	31.65	181.05	196	4.53	92.77	93.91	31.39	181.55	196	4.13	94.30	94.30
T <sub>(2)</sub>	59.87	PO	31.75	248.87	284	4.49	96.14	97.98	29.79	256.77	302	4.47	95.11	96.80
		OBF	32.08	236.60	248	4.89	96.03	96.53	31.44	243.03	264	4.82	96.10	96.10
		WT	31.49	240.91	264	4.78	96.41	97.42	30.99	245.23	266	4.71	95.77	96.07
		HP	48.56	273.95	320	4.81	96.81	97.05	32.86	246.45	264	4.88	96.06	96.10
		SD	29.62	184.00	196	4.61	94.62	95.55	26.63	187.07	196	4.23	95.31	95.31
T <sub>(3)</sub>	53.20	PO	11.17	278.64	320	4.80	98.58	99.29	14.28	278.60	320	4.75	98.15	98.87
		OBF	8.70	267.72	280	4.97	99.01	99.20	14.33	264.54	290	4.93	98.47	98.47
		WT	9.63	272.05	290	4.93	99.03	99.33	14.28	267.93	298	4.88	98.45	98.61
		HP	21.59	295.32	320	4.94	99.37	99.46	15.20	266.14	278	4.94	98.34	98.35
		SD	6.66	199.31	196	4.02	98.55	98.86	11.27	197.90	196	4.02	98.21	98.21
T <sub>(4)</sub>	57.41	PO	20.86	264.17	300	4.65	97.23	98.61	22.35	266.89	316	4.59	96.61	97.89
		OBF	18.34	253.79	266	4.93	97.72	97.98	22.77	253.22	276	4.87	96.94	96.94
		WT	18.90	256.92	274	4.85	97.72	98.31	22.51	255.84	284	4.78	96.95	97.27
		HP	34.12	285.50	320	4.88	98.17	98.28	23.96	256.15	270	4.91	96.93	96.94
		SD	14.62	194.19	196	4.30	96.26	97.34	18.71	191.69	196	4.09	96.29	96.29
2b) p <sub>1</sub> = 0.9, p <sub>2</sub> = 0.8, D = 0.1														
T <sub>(1)</sub>	61.24	PO	46.19	224.57	246	4.27	95.50	97.98	43.75	233.03	264	4.21	94.96	96.99
		OBF	46.35	215.20	214	4.81	96.09	96.67	45.17	221.92	222	4.76	95.54	95.54
		WT	46.34	217.54	222	4.64	95.98	97.20	44.47	222.84	226	4.58	95.54	96.07
		HP	66.02	251.23	282	4.69	95.95	96.35	47.93	224.83	234	4.87	95.48	95.53

TS	NSD (1 - β̂)%	Design	Inverse Normalmethode						Inverse χ <sup>2</sup> -Methode					
			(1 - β̂)%	ASN	N <sub>med</sub>	k̄	(1 - 2 α̂)%	(1 - β̂)%	ASN	N <sub>med</sub>	k̄	(1 - 2 α̂)%		
							{KI <sub>j</sub> }	KI <sub>k</sub>					{KI <sub>j</sub> }	KI <sub>k</sub>
		SD	42.35	176.60	196	4.96	94.21	95.06	40.91	179.66	196	4.50	95.00	95.00
T <sub>(2)</sub>	59.87	PO	41.22	232.00	256	4.42	95.91	98.24	39.21	238.58	268	4.34	95.37	97.45
		OBF	42.57	220.65	216	4.91	96.16	96.80	40.21	227.55	226	4.84	96.35	96.35
		WT	41.89	225.11	226	4.79	96.22	97.40	39.64	229.63	234	4.68	96.11	96.69
		HP	64.84	254.91	286	4.78	96.38	96.72	43.32	230.95	238	4.91	96.00	96.11
		SD	39.39	181.55	196	5.03	95.49	96.19	35.58	185.81	196	4.59	95.45	95.45
T <sub>(3)</sub>	53.20	PO	12.07	271.71	292	4.80	99.10	99.78	15.38	270.77	306	4.76	98.63	99.44
		OBF	9.40	262.66	266	4.99	99.42	99.51	16.07	255.44	270	4.95	98.92	98.92
		WT	9.87	266.19	272	4.95	99.35	99.58	16.11	258.35	276	4.88	98.84	98.98
		HP	26.81	289.06	320	4.95	99.49	99.61	17.38	257.64	268	4.96	98.85	98.87
		SD	6.98	202.12	196	4.15	98.90	99.23	12.90	200.44	196	4.27	98.75	98.75
T <sub>(4)</sub>	57.41	PO	25.05	251.98	276	4.59	97.61	98.91	27.21	255.08	288	4.55	97.37	98.40
		OBF	21.64	245.33	254	4.95	98.04	98.39	27.62	241.72	250	4.90	97.51	97.51
		WT	20.99	247.68	262	4.89	98.06	98.68	27.30	244.47	258	4.81	97.45	97.80
		HP	43.70	273.73	310	4.88	98.31	98.47	30.20	245.10	258	4.93	97.17	97.20
		SD	16.64	196.35	196	4.52	97.29	97.97	22.83	192.99	196	4.39	97.31	97.31
2c) p <sub>1</sub> = 0.7, p <sub>2</sub> = 0.5, D = 0.2														
T <sub>(1)</sub>	89.91	PO	80.28	158.43	146	3.40	94.59	96.29	79.35	163.07	156	3.29	94.63	96.03
		OBF	77.76	157.92	152	4.30	91.42	92.05	79.28	158.10	152	4.13	91.76	91.76
		WT	78.93	156.21	152	3.93	92.98	94.04	79.53	159.28	154	3.83	93.00	93.32
		HP	91.40	184.17	180	3.98	95.84	96.01	81.91	164.36	160	4.34	91.37	91.40
		SD	73.20	142.82	140	5.13	86.57	88.83	75.63	140.92	138	4.29	88.93	88.93
T <sub>(2)</sub>	89.63	PO	77.96	164.66	162	3.54	95.39	97.05	76.35	169.77	170	3.44	95.14	96.32
		OBF	73.94	165.59	158	4.53	93.56	94.07	74.55	168.48	160	4.37	92.13	92.13
		WT	74.56	164.49	158	4.20	93.78	94.72	74.24	168.62	162	4.06	93.23	93.42
		HP	91.12	187.61	184	4.11	96.79	96.99	78.15	172.89	170	4.47	92.44	92.49
		SD	71.40	149.42	160	5.35	92.77	93.65	70.62	152.55	172	4.64	90.99	90.99
T <sub>(3)</sub>	87.47	PO	51.21	203.21	218	4.13	98.38	98.98	57.86	198.94	212	3.90	97.37	97.95
		OBF	41.18	203.62	208	4.81	98.36	98.59	53.45	194.56	190	4.60	94.74	94.74
		WT	44.50	203.53	208	4.60	97.94	98.37	53.92	195.71	192	4.43	95.49	95.69
		HP	72.02	221.09	232	4.52	99.07	99.15	58.23	198.35	202	4.64	97.19	97.21
		SD	36.31	180.72	196	5.04	97.82	98.40	49.94	173.20	196	4.73	97.45	97.45
T <sub>(4)</sub>	88.99	PO	67.80	180.09	182	3.74	95.96	97.24	68.71	182.53	180	3.66	95.41	96.49
		OBF	59.81	182.56	180	4.63	93.89	94.16	66.82	177.75	170	4.44	92.85	92.85
		WT	61.87	181.64	182	4.33	94.79	95.38	67.32	178.69	174	4.18	94.06	94.30
		HP	83.03	203.25	206	4.30	97.44	97.57	70.59	182.96	182	4.54	93.17	93.21
		SD	54.02	164.05	184	5.17	94.36	95.31	61.94	159.80	184	4.64	91.41	91.41
2d) p <sub>1</sub> = 0.3, p <sub>2</sub> = 0.1, D = 0.2														
T <sub>(1)</sub>	98.45	PO	91.53	122.35	104	2.94	94.43	97.30	91.44	127.77	112	2.87	93.99	95.83
		OBF	88.05	131.80	124	4.26	92.57	93.24	89.47	133.26	130	4.08	92.89	92.89
		WT	88.65	128.09	118	3.75	93.30	94.92	89.66	131.65	124	3.64	93.62	94.13
		HP	97.17	142.94	132	3.57	95.22	95.50	91.56	135.41	128	4.30	92.28	92.40
		SD	87.12	126.05	118	5.67	90.04	91.53	88.64	127.77	114	4.81	91.15	91.15
T <sub>(2)</sub>	98.31	PO	89.73	130.05	112	3.10	95.25	97.88	88.46	136.88	120	3.10	94.99	96.80
		OBF	84.35	140.94	142	4.55	95.35	95.89	84.61	144.13	152	4.34	93.93	93.93
		WT	85.40	139.04	140	4.10	95.30	96.37	85.48	142.88	152	3.92	94.74	95.10
		HP	97.00	147.59	138	3.82	96.21	96.49	88.55	144.84	142	4.43	93.65	93.72

TS	NSD (1 - β̂)%	Design	Inverse Normalmethode						Inverse χ <sup>2</sup> -Methode							
			(1 - β̂)%	ASN	N <sub>med</sub>	k̄	(1 - 2 α̂)%	{KI <sub>j</sub> }	KI <sub>k</sub>	(1 - β̂)%	ASN	N <sub>med</sub>	k̄	(1 - 2 α̂)%	{KI <sub>j</sub> }	KI <sub>k</sub>
		SD	85.62	134.79	126	5.96	94.45	95.16	85.13	141.48	148	5.22	93.46	93.46		
T <sub>(3)</sub>	97.56	PO	59.09	181.23	196	3.96	98.69	99.43	67.96	176.95	180	3.85	97.97	98.65		
		OBF	42.71	195.12	198	4.88	99.31	99.48	62.51	175.10	170	4.68	97.05	97.05		
		WT	47.24	191.76	198	4.61	99.07	99.36	63.88	174.57	172	4.38	97.52	97.64		
		HP	75.01	193.88	196	4.48	99.33	99.46	67.76	178.59	178	4.66	98.58	98.63		
		SD	39.99	185.23	196	5.45	98.90	99.26	58.18	172.26	190	5.38	98.69	98.69		
T <sub>(4)</sub>	97.91	PO	78.26	151.62	156	3.45	96.90	98.44	81.25	152.85	158	3.40	96.44	97.57		
		OBF	64.87	166.96	164	4.71	96.42	96.79	75.90	156.31	154	4.47	94.83	94.83		
		WT	67.40	163.46	164	4.35	96.44	97.07	77.45	155.79	156	4.15	95.65	95.85		
		HP	86.62	168.83	166	4.12	97.74	97.97	80.57	159.41	160	4.51	94.85	94.90		
		SD	60.99	162.79	176	5.76	96.21	96.86	71.68	153.56	170	5.26	93.63	93.63		
2e) p <sub>1</sub> = 0.6, p <sub>2</sub> = 0.3, D = 0.3																
T <sub>(1)</sub>	99.85	PO	97.54	92.13	66	2.33	94.36	95.82	98.08	93.61	62	2.19	94.63	95.81		
		OBF	95.92	101.14	70	3.61	87.40	87.98	97.48	98.46	70	3.26	87.26	87.26		
		WT	96.74	96.27	68	3.01	90.62	91.53	97.55	96.40	68	2.85	89.73	89.98		
		HP	99.07	107.01	98	2.74	96.66	96.80	98.13	99.38	74	3.39	86.41	86.44		
		SD	92.36	99.81	86	4.95	85.04	86.33	96.10	94.54	74	3.81	82.25	82.25		
T <sub>(2)</sub>	99.85	PO	97.31	95.78	68	2.42	95.80	96.95	97.24	98.90	70	2.36	95.48	96.01		
		OBF	94.07	109.35	94	3.96	95.65	96.04	95.13	108.63	94	3.62	87.19	87.19		
		WT	95.27	104.45	88	3.33	95.74	96.35	95.81	105.75	92	3.11	90.25	90.34		
		HP	98.98	111.41	102	3.00	97.38	97.51	96.61	108.26	92	3.70	90.58	90.59		
		SD	92.54	106.40	92	5.31	90.94	91.82	95.15	105.71	88	4.31	94.33	94.33		
T <sub>(3)</sub>	99.69	PO	86.47	125.25	112	3.07	98.26	98.76	91.52	118.97	102	2.76	97.70	98.15		
		OBF	74.14	143.94	152	4.46	99.15	99.29	85.56	129.27	128	3.99	98.71	98.71		
		WT	78.44	137.83	144	3.94	98.66	98.96	86.63	128.08	126	3.66	97.99	98.12		
		HP	91.66	133.92	126	3.50	98.90	98.97	88.62	128.80	124	3.90	98.41	98.46		
		SD	70.71	141.47	146	5.51	97.28	97.91	83.27	128.55	122	4.69	96.62	96.62		
T <sub>(4)</sub>	99.85	PO	94.57	103.70	72	2.54	96.24	97.24	95.07	107.91	86	2.58	94.91	95.41		
		OBF	87.15	122.18	114	4.12	95.20	95.55	91.99	114.69	106	3.73	87.29	87.29		
		WT	90.42	115.07	102	3.43	94.89	95.51	93.40	112.46	102	3.29	90.36	90.49		
		HP	96.38	120.63	112	3.19	97.97	98.08	94.23	115.02	102	3.78	87.95	87.97		
		SD	83.60	119.60	110	5.31	92.51	93.73	90.97	112.19	98	4.37	93.36	93.36		
2f) p <sub>1</sub> = 0.9, p <sub>2</sub> = 0.6, D = 0.3																
T <sub>(1)</sub>	99.99	PO	99.46	72.71	50	1.86	94.58	97.15	99.49	75.63	50	1.86	94.01	95.50		
		OBF	98.07	85.60	68	3.46	88.97	90.03	98.91	83.55	68	3.06	88.30	88.30		
		WT	98.78	79.99	62	2.69	91.17	93.11	99.19	80.84	58	2.52	90.87	91.54		
		HP	99.54	85.19	74	2.36	96.20	96.67	99.21	83.60	74	3.24	86.63	86.72		
		SD	94.77	87.10	76	4.88	84.11	86.45	97.79	81.49	68	3.67	84.06	84.06		
T <sub>(2)</sub>	99.99	PO	99.16	76.43	50	1.97	96.19	98.04	99.05	80.35	50	2.00	95.20	96.04		
		OBF	96.79	94.02	79	3.86	96.47	97.02	97.78	93.58	74	3.43	90.14	90.14		
		WT	97.84	88.83	68	3.07	96.39	97.51	98.07	90.63	68	2.87	92.85	92.98		
		HP	99.63	88.07	80	2.56	97.17	97.51	98.27	90.83	74	3.41	93.35	93.45		
		SD	94.86	93.81	82	5.29	92.23	93.21	97.09	93.64	74	4.26	95.13	95.13		
T <sub>(3)</sub>	99.97	PO	90.68	104.94	70	2.66	98.33	99.02	94.65	102.96	74	2.60	98.04	98.56		
		OBF	73.71	138.45	148	4.53	99.65	99.78	88.62	116.19	114	3.99	99.14	99.14		
		WT	81.02	127.04	132	3.80	99.45	99.74	91.34	113.46	112	3.47	99.06	99.17		
		HP	90.32	114.80	110	3.37	99.31	99.45	90.90	113.82	110	3.85	99.11	99.15		

TS	NSD	Design	Inverse Normalmethode					Inverse $\chi^2$ -Methode						
			$(1 - \hat{\beta})\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2 \hat{\alpha})\%$	$(1 - \hat{\beta})\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2 \hat{\alpha})\%$		
		$(1 - \hat{\beta})\%$					$\{KI_j\}$				$\{KI_j\}$		$KI_k$	
		SD	68.18	140.10	148	5.65	98.29	98.71	85.78	121.43	116	4.93	97.97	97.97
$T_{(4)}$	99.99	PO	97.26	85.82	50	2.20	96.86	98.22	98.02	86.96	50	2.18	96.11	96.74
		OBF	90.06	111.39	104	4.11	96.94	97.38	95.34	100.31	86	3.56	90.87	90.87
		WT	93.08	102.99	86	3.34	96.62	97.31	96.45	97.36	74	3.06	92.99	93.17
		HP	96.20	98.38	90	2.88	97.92	98.07	96.24	98.75	86	3.58	91.74	91.78
		SD	82.41	112.84	102	5.42	94.47	95.47	90.46	101.79	82	4.41	95.04	95.04

**Tabelle C.10:** Empirische Power  $(1 - \hat{\beta})\%$  im nicht-sequentiellen Design (NSD) sowie  $(1 - \hat{\beta})\%$ , mittlerer und medianer Umfang (ASN = „Average Sample Number“,  $N_{med}$ ), mittlere Stufenanzahl  $\bar{k}$  und empirische Überdeckungswahrscheinlichkeiten  $(1 - 2\hat{\alpha})\%$  der Sequenz der beobachteten Intervalle  $\{KI_j\}$  sowie des finalen Intervalls  $KI_k$  in flexiblen Designs mit binärer Zielvariable unter der Alternativhypothese (Szenario I, Konstellation IV)

TS	NSD	Design	Inverse Normalmethode					Inverse $\chi^2$ -Methode						
			$(1 - \hat{\beta})\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2 \hat{\alpha})\%$	$(1 - \hat{\beta})\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2 \hat{\alpha})\%$		
		$(1 - \hat{\beta})\%$					$\{KI_j\}$				$\{KI_j\}$		$KI_k$	
2a) $p_1 = 0.3, p_2 = 0.2, D = 0.1$														
$T_{(1)}$	61.24	PO	55.47	433.27	482	4.30	95.54	97.41	55.36	444.74	494	4.26	95.29	96.75
		OBF	56.18	403.17	426	4.76	95.64	96.12	59.98	409.66	422	4.67	95.54	95.54
		WT	56.92	411.22	436	4.59	95.60	96.50	60.10	412.40	430	4.48	95.72	96.03
		HP	75.12	460.36	480	4.68	96.31	96.46	60.92	417.10	440	4.82	95.21	95.31
		SD	58.68	385.50	434	5.80	92.07	93.48	60.87	392.97	440	5.30	94.56	94.56
$T_{(2)}$	59.87	PO	53.21	442.12	490	4.38	96.02	97.69	53.23	453.76	508	4.30	95.58	96.98
		OBF	52.98	411.44	430	4.86	96.29	96.75	55.55	421.02	434	4.75	96.01	96.01
		WT	53.45	420.80	440	4.72	96.22	97.01	55.25	427.18	440	4.60	96.11	96.33
		HP	74.64	465.21	486	4.74	96.69	96.87	58.33	426.20	448	4.85	95.97	96.02
		SD	57.46	389.78	432	5.87	94.70	95.78	57.07	409.22	450	5.43	95.24	95.24
$T_{(3)}$	53.20	PO	27.52	519.14	568	4.72	98.40	99.09	35.41	508.03	575	4.60	98.05	98.69
		OBF	25.94	495.38	514	4.96	98.50	98.65	37.80	472.27	496	4.87	98.15	98.15
		WT	26.28	499.43	524	4.90	98.68	98.96	37.28	478.99	508	4.78	97.93	98.19
		HP	50.92	532.19	598	4.90	98.84	98.91	39.22	479.31	500	4.92	98.25	98.26
		SD	24.93	498.02	546	5.60	98.81	99.10	37.89	470.71	546	5.43	97.86	97.86
$T_{(4)}$	57.41	PO	40.58	479.65	540	4.55	97.15	98.29	45.54	479.16	548	4.44	96.58	97.66
		OBF	38.05	456.93	472	4.90	96.85	97.32	46.92	442.56	458	4.80	96.18	96.18
		WT	38.51	462.73	486	4.80	97.09	97.72	47.82	450.53	474	4.68	96.67	96.94
		HP	62.47	499.78	554	4.82	97.38	97.45	49.67	450.93	472	4.88	96.64	96.66
		SD	38.61	452.41	502	5.73	95.93	97.13	47.11	439.55	502	5.41	95.96	95.96
2b) $p_1 = 0.9, p_2 = 0.8, D = 0.1$														
$T_{(1)}$	61.24	PO	71.88	370.09	379	4.11	95.82	98.24	72.70	377.06	384	3.97	95.29	97.28
		OBF	69.40	348.73	342	4.76	96.25	96.83	72.49	354.96	344	4.65	96.07	96.07
		WT	70.27	354.13	350	4.56	96.13	97.30	72.37	361.53	356	4.43	95.89	96.39
		HP	87.69	385.88	376	4.56	96.36	96.82	75.07	361.75	356	4.81	95.89	95.94
		SD	73.65	335.95	338	6.23	93.85	95.06	74.43	352.21	358	5.66	95.01	95.01
$T_{(2)}$	59.87	PO	68.87	382.32	394	4.25	96.12	98.52	69.67	390.68	404	4.11	95.70	97.66
		OBF	66.29	358.66	348	4.87	96.51	97.00	68.98	368.34	360	4.74	96.61	96.61
		WT	66.89	367.08	360	4.71	96.73	97.72	68.22	372.29	366	4.54	96.53	96.95
		HP	87.66	391.16	380	4.66	96.57	97.00	71.99	372.15	370	4.86	96.30	96.36

TS	NSD (1 - β̂)%	Design	Inverse Normalmethode						Inverse χ <sup>2</sup> -Methode					
			(1 - β̂)%	ASN	N <sub>med</sub>	k̄	(1 - 2 α̂)%	(1 - β̂)%	ASN	N <sub>med</sub>	k̄	(1 - 2 α̂)%		
							{KI <sub>j</sub> }	KI <sub>k</sub>					{KI <sub>j</sub> }	KI <sub>k</sub>
		SD	71.34	346.15	348	6.31	95.68	96.50	71.02	369.48	378	5.80	95.86	95.86
T <sub>(3)</sub>	53.20	PO	36.00	490.95	520	4.74	98.84	99.64	48.41	469.10	503	4.57	98.43	99.22
		OBF	31.06	473.67	478	4.98	99.29	99.40	48.33	436.28	443	4.89	98.79	98.79
		WT	33.40	480.78	488	4.93	99.16	99.49	48.34	441.85	450	4.78	98.85	98.94
		HP	60.41	485.22	498	4.91	99.31	99.42	51.14	440.64	454	4.93	98.68	98.69
		SD	33.66	498.70	536	5.90	98.91	99.13	48.66	449.81	482	5.78	98.71	98.71
T <sub>(4)</sub>	57.41	PO	53.08	436.79	468	4.47	97.25	98.76	59.81	429.06	458	4.33	97.18	98.29
		OBF	47.09	424.01	436	4.93	97.61	98.06	58.13	400.72	404	4.81	97.48	97.48
		WT	47.70	428.28	442	4.83	97.88	98.50	59.51	406.97	416	4.67	97.32	97.69
		HP	73.24	439.57	436	4.80	97.77	97.97	62.90	408.17	422	4.89	97.00	97.01
		SD	49.52	435.59	460	6.10	96.95	97.69	59.74	410.64	438	5.84	96.63	96.63
2c) p <sub>1</sub> = 0.7, p <sub>2</sub> = 0.5, D = 0.2														
T <sub>(1)</sub>	89.91	PO	94.95	213.65	174	3.18	94.36	95.93	96.03	219.75	174	3.03	94.77	95.92
		OBF	91.02	218.87	183	4.21	91.23	91.60	93.76	215.18	180	3.98	91.95	91.95
		WT	92.24	215.70	180	3.81	92.34	93.26	94.65	216.93	174	3.64	92.55	92.89
		HP	97.00	228.31	206	3.73	95.79	95.93	95.18	216.49	182	4.23	90.66	90.69
		SD	89.92	209.46	180	6.02	84.81	87.44	94.09	205.32	174	5.02	86.83	86.83
T <sub>(2)</sub>	89.63	PO	93.21	224.93	174	3.32	95.12	96.71	95.00	229.09	174	3.15	95.29	96.31
		OBF	88.29	230.37	194	4.47	93.59	93.95	90.54	232.03	194	4.21	92.33	92.33
		WT	90.34	227.61	186	4.07	94.26	95.13	91.80	232.72	192	3.86	93.19	93.37
		HP	97.00	232.86	212	3.87	96.76	96.92	92.91	229.89	196	4.38	92.02	92.07
		SD	89.76	222.20	186	6.25	91.54	92.65	91.53	228.34	186	5.42	89.17	89.17
T <sub>(3)</sub>	87.47	PO	80.23	301.61	298	3.91	97.86	98.64	88.76	280.07	256	3.58	97.36	97.92
		OBF	68.82	314.66	324	4.77	98.39	98.69	80.93	278.29	260	4.47	93.84	93.84
		WT	72.80	312.00	322	4.50	97.45	98.01	82.82	281.38	260	4.24	95.00	95.15
		HP	86.74	287.28	272	4.32	98.71	98.80	84.58	278.76	258	4.54	96.90	96.93
		SD	71.01	328.31	336	6.39	97.67	98.24	81.61	287.81	270	5.76	97.36	97.36
T <sub>(4)</sub>	88.99	PO	88.98	254.91	214	3.53	95.78	96.92	92.39	251.67	208	3.38	95.60	96.48
		OBF	81.38	269.62	258	4.58	93.35	93.63	87.71	249.30	218	4.30	92.59	92.59
		WT	83.33	265.11	246	4.21	94.27	94.87	88.81	248.78	214	3.97	93.51	93.59
		HP	91.70	255.69	236	4.06	96.96	97.11	89.26	249.67	218	4.45	92.77	92.83
		SD	81.38	269.39	244	6.25	93.33	94.50	87.27	250.21	212	5.57	88.93	88.93
2d) p <sub>1</sub> = 0.3, p <sub>2</sub> = 0.1, D = 0.2														
T <sub>(1)</sub>	98.45	PO	98.48	151.03	106	2.76	94.39	97.32	98.33	155.34	124	2.64	94.33	96.12
		OBF	94.67	165.91	130	4.18	92.19	92.86	95.94	166.15	144	3.93	93.34	93.34
		WT	96.35	161.17	128	3.62	93.50	94.93	96.75	165.49	146	3.48	94.17	94.49
		HP	98.37	160.58	148	3.30	95.86	96.20	96.51	162.89	136	4.20	92.60	92.75
		SD	93.00	156.76	124	6.22	88.37	90.39	95.45	159.68	136	5.22	90.45	90.45
T <sub>(2)</sub>	98.31	PO	97.30	160.38	130	2.92	95.58	97.76	97.14	167.30	152	2.87	95.08	96.77
		OBF	92.06	177.81	152	4.47	95.71	96.12	93.00	181.26	166	4.16	94.35	94.35
		WT	93.58	175.91	150	3.97	95.17	96.40	94.30	178.05	166	3.72	95.00	95.40
		HP	98.30	165.22	152	3.50	96.30	96.76	94.77	174.29	152	4.31	94.30	94.51
		SD	91.95	169.70	146	6.53	94.52	95.61	93.29	181.86	172	5.67	93.57	93.57
T <sub>(3)</sub>	97.56	PO	84.61	250.48	234	3.74	98.56	99.27	90.74	225.57	194	3.53	98.13	98.68
		OBF	71.00	287.49	292	4.85	99.29	99.43	82.38	229.89	202	4.53	97.15	97.15
		WT	75.01	280.49	288	4.52	98.99	99.29	85.03	228.63	200	4.17	97.20	97.34
		HP	82.21	229.43	214	4.26	99.14	99.26	84.64	224.45	201	4.57	98.58	98.60

TS	NSD (1 - β̂)%	Design	Inverse Normalmethode						Inverse χ <sup>2</sup> -Methode					
			(1 - β̂)%	ASN	N <sub>med</sub>	k̄	(1 - 2 α̂)%	(1 - β̂)%	ASN	N <sub>med</sub>	k̄	(1 - 2 α̂)%		
							{KI <sub>j</sub> }	KI <sub>k</sub>					{KI <sub>j</sub> }	KI <sub>k</sub>
		SD	72.91	312.24	312	6.83	98.87	99.26	81.15	246.69	230	6.18	98.98	98.98
T <sub>(4)</sub>	99.91	PO	94.25	197.27	174	3.23	96.90	98.56	95.11	193.64	174	3.14	96.68	97.63
		OBF	83.79	230.26	210	4.64	96.60	96.95	88.71	199.85	174	4.30	95.20	95.20
		WT	87.02	223.47	198	4.23	96.33	97.08	90.50	200.46	174	3.97	95.42	95.69
		HP	90.80	192.09	174	3.83	97.45	97.70	91.03	193.86	174	4.38	95.03	95.13
		SD	80.61	238.61	218	6.68	95.99	97.05	85.33	208.51	182	5.89	92.92	92.92
2e) p <sub>1</sub> = 0.6, p <sub>2</sub> = 0.3, D = 0.3														
T <sub>(1)</sub>	99.85	PO	99.58	107.33	66	2.25	94.68	95.95	99.73	107.63	62	2.08	94.91	96.01
		OBF	98.31	119.49	72	3.56	87.64	88.25	98.86	114.69	70	3.20	87.42	87.42
		WT	98.97	113.73	68	2.94	91.04	91.90	99.01	113.73	68	2.78	90.24	90.54
		HP	99.15	117.15	98	2.57	96.92	97.11	98.97	111.38	76	3.32	86.25	86.35
		SD	93.80	113.42	86	5.26	83.93	85.47	97.46	107.19	74	4.03	81.90	81.90
T <sub>(2)</sub>	99.85	PO	99.40	111.29	68	2.32	95.82	96.98	99.47	114.92	72	2.24	95.60	96.20
		OBF	96.53	128.53	94	3.91	95.79	96.18	97.48	128.09	94	3.53	87.55	87.55
		WT	98.07	123.54	88	3.25	95.94	96.50	97.85	125.54	96	3.04	91.14	91.19
		HP	99.20	121.20	100	2.86	97.60	97.74	98.13	121.84	92	3.64	90.83	90.86
		SD	94.55	121.83	92	5.63	90.90	91.89	96.83	123.44	92	4.54	94.85	94.85
T <sub>(3)</sub>	99.69	PO	95.99	154.24	120	2.93	98.21	98.64	97.56	139.40	102	2.56	97.61	98.01
		OBF	87.87	186.70	164	4.40	99.24	99.35	92.26	157.50	140	3.89	98.89	98.89
		WT	91.07	177.73	158	3.83	98.65	98.94	93.49	155.78	138	3.55	98.37	98.47
		HP	93.18	146.90	140	3.31	98.85	98.94	93.56	149.40	130	3.82	98.58	98.64
		SD	81.58	186.57	170	6.13	97.00	97.69	89.37	158.92	145	5.15	96.81	96.81
T <sub>(4)</sub>	99.85	PO	99.07	122.54	72	2.40	96.44	97.32	98.86	125.88	82	2.43	95.00	95.53
		OBF	94.18	150.96	118	4.05	95.24	95.44	95.82	136.08	104	3.64	87.78	87.78
		WT	96.62	142.45	102	3.34	94.81	95.36	97.45	134.23	104	3.20	90.28	90.36
		HP	97.02	132.29	118	3.00	98.15	98.31	96.98	131.85	104	3.71	88.14	88.21
		SD	88.86	146.29	114	5.77	92.32	93.58	93.90	133.26	102	4.68	93.14	93.14
2f) p <sub>1</sub> = 0.9, p <sub>2</sub> = 0.6, D = 0.3														
T <sub>(1)</sub>	99.99	PO	99.90	79.52	50	1.83	95.01	97.32	99.88	82.78	50	1.82	94.39	95.73
		OBF	98.66	93.65	68	3.42	89.63	90.76	99.19	92.36	68	3.02	88.58	88.58
		WT	99.51	87.75	62	2.64	91.53	93.24	99.61	89.44	58	2.49	91.19	91.90
		HP	99.67	89.72	74	2.24	97.00	97.44	99.27	88.94	74	3.20	87.20	87.34
		SD	94.97	92.66	76	5.02	83.97	86.52	97.71	87.83	68	3.78	82.63	82.63
T <sub>(2)</sub>	99.99	PO	99.73	83.50	50	1.91	96.36	98.00	99.51	88.71	50	1.95	95.27	96.25
		OBF	97.52	103.91	80	3.82	96.83	97.39	98.48	103.73	74	3.37	90.22	90.22
		WT	98.49	98.89	68	3.05	96.79	97.71	98.66	102.06	68	2.87	92.70	92.79
		HP	99.59	92.71	80	2.44	97.77	98.24	98.54	97.34	76	3.39	93.63	93.78
		SD	94.93	99.25	82	5.41	92.34	93.39	97.58	101.85	74	4.34	95.32	95.32
T <sub>(3)</sub>	99.97	PO	97.45	124.24	70	2.57	98.48	99.01	97.41	117.43	74	2.54	98.16	98.62
		OBF	86.76	172.60	154	4.48	99.69	99.76	91.72	131.41	116	3.90	99.26	99.26
		WT	92.10	159.08	138	3.72	99.35	99.61	94.33	128.89	116	3.40	98.97	99.09
		HP	91.45	121.47	110	3.20	99.40	99.52	92.22	124.36	110	3.82	99.13	99.20
		SD	78.14	178.84	170	6.22	98.65	99.03	87.76	136.01	126	5.15	98.15	98.15
T <sub>(4)</sub>	99.99	PO	99.21	98.38	50	2.16	97.17	98.27	99.27	97.37	50	2.12	96.19	96.81
		OBF	94.59	129.96	102	4.06	97.06	97.46	96.57	112.94	86	3.51	91.00	91.00
		WT	97.03	121.48	88	3.30	96.53	97.28	97.69	111.21	78	3.05	93.30	93.46
		HP	96.36	104.73	90	2.76	98.20	98.43	97.02	107.38	88	3.57	91.81	91.93

TS	NSD (1 - β̂)%	Design	Inverse Normalmethode					Inverse χ <sup>2</sup> -Methode						
			(1 - β̂)%	ASN	N <sub>med</sub>	k̄	(1 - 2 α̂)%	(1 - β̂)%	ASN	N <sub>med</sub>	k̄	(1 - 2 α̂)%		
							{KI <sub>j</sub> }	KI <sub>k</sub>					{KI <sub>j</sub> }	KI <sub>k</sub>
		SD	85.71	130.61	106	5.69	94.68	95.95	91.19	112.17	84	4.58	95.12	95.12

**Tabelle C.11:** Empirische Power (1-β̂)% im nicht-sequentiellen Design (NSD) sowie (1-β̂)%, mittlerer und medianer Umfang (ASN = „Average Sample Number“, N<sub>med</sub>), mittlere Stufenanzahl k̄ und empirische Überdeckungswahrscheinlichkeiten (1-2α̂)% der Sequenz der beobachteten Intervalle {KI<sub>j</sub>} sowie des finalen Intervalls KI<sub>k</sub> in flexiblen Designs mit binärer Zielvariable unter der Alternativhypothese (Szenario I, Konstellation V)

TS	NSD (1 - β̂)%	Design	Inverse Normalmethode					Inverse χ <sup>2</sup> -Methode						
			(1 - β̂)%	ASN	N <sub>med</sub>	k̄	(1 - 2 α̂)%	(1 - β̂)%	ASN	N <sub>med</sub>	k̄	(1 - 2 α̂)%		
							{KI <sub>j</sub> }	KI <sub>k</sub>					{KI <sub>j</sub> }	KI <sub>k</sub>
2a) p <sub>1</sub> = 0.3, p <sub>2</sub> = 0.2, D = 0.1														
T <sub>(1)</sub>	61.24	PO	43.63	261.08	308	2.62	95.95	97.56	39.89	262.26	308	2.62	95.26	96.82
		OBF	48.01	266.51	308	2.84	96.69	97.00	44.29	266.54	308	2.79	96.28	96.28
		WT	47.22	264.82	308	2.74	96.65	97.38	44.00	265.97	308	2.71	95.83	96.35
		HP	50.06	276.44	308	2.88	96.73	96.83	45.23	268.98	308	2.90	96.26	96.26
		SD	36.69	208.26	208	2.33	94.63	94.90	36.25	204.73	208	2.19	95.19	95.19
T <sub>(2)</sub>	59.87	PO	42.61	263.43	308	2.65	96.15	97.63	38.30	265.55	308	2.65	95.63	96.95
		OBF	47.18	267.48	308	2.85	97.14	97.45	43.19	267.55	308	2.81	96.50	96.50
		WT	46.63	265.82	308	2.74	96.98	97.63	42.80	267.53	308	2.73	96.18	96.59
		HP	49.46	277.40	308	2.89	97.00	97.09	44.03	270.16	308	2.91	96.63	96.63
		SD	36.32	208.78	208	2.34	95.44	95.61	35.05	205.10	208	2.20	95.58	95.58
T <sub>(3)</sub>	53.20	PO	29.18	276.39	308	2.76	98.41	98.99	26.08	279.26	308	2.77	97.77	98.42
		OBF	30.64	278.58	308	2.92	98.92	99.03	29.22	278.19	308	2.88	98.41	98.41
		WT	31.10	277.94	308	2.85	98.78	99.04	29.34	279.32	308	2.83	98.20	98.41
		HP	32.39	285.39	308	2.93	98.90	98.95	29.74	281.31	308	2.95	98.47	98.47
		SD	23.31	212.46	208	2.27	98.10	98.17	23.60	209.80	208	2.19	97.77	97.77
T <sub>(4)</sub>	57.41	PO	36.16	269.57	308	2.70	97.17	98.38	33.01	271.00	308	2.70	96.73	97.79
		OBF	38.60	273.25	308	2.89	97.97	98.13	36.54	272.81	308	2.84	97.25	97.25
		WT	38.83	271.98	308	2.80	97.78	98.31	36.22	273.46	308	2.78	97.20	97.52
		HP	41.11	281.30	308	2.91	97.81	97.89	37.43	275.41	308	2.93	97.38	97.38
		SD	29.49	210.82	208	2.31	96.22	96.43	29.16	207.56	208	2.20	96.29	96.29
2b) p <sub>1</sub> = 0.9, p <sub>2</sub> = 0.8, D = 0.1														
T <sub>(1)</sub>	61.24	PO	61.27	241.38	308	2.46	95.99	97.69	56.02	243.35	308	2.45	95.52	96.91
		OBF	65.09	249.50	302	2.77	96.56	96.89	61.45	248.37	308	2.69	96.22	96.22
		WT	64.79	246.60	308	2.61	96.48	97.37	60.43	248.05	308	2.58	96.03	96.34
		HP	66.98	260.84	308	2.80	96.87	97.03	61.81	252.65	308	2.86	96.51	96.51
		SD	52.11	205.97	208	2.47	95.01	95.20	50.14	200.46	208	2.26	95.73	95.73
T <sub>(2)</sub>	59.87	PO	59.97	244.58	308	2.49	96.45	97.92	54.45	246.80	308	2.49	95.73	97.13
		OBF	64.15	250.82	302	2.78	96.90	97.25	60.44	250.26	308	2.72	96.41	96.41
		WT	63.68	248.28	308	2.63	96.71	97.50	59.10	250.14	308	2.61	96.38	96.67
		HP	66.35	262.33	308	2.82	97.11	97.22	61.26	254.29	308	2.89	96.68	96.68
		SD	51.66	206.69	208	2.47	95.72	95.84	48.66	201.66	208	2.27	96.32	96.32
T <sub>(3)</sub>	53.20	PO	40.37	265.70	308	2.69	98.61	99.27	36.79	269.02	308	2.70	98.22	98.71
		OBF	40.68	268.14	308	2.90	99.03	99.21	39.74	269.12	308	2.86	98.71	98.71
		WT	42.36	267.13	308	2.81	99.05	99.39	40.63	269.02	308	2.77	98.60	98.83
		HP	43.72	275.99	308	2.92	99.15	99.29	40.56	271.15	308	2.95	98.78	98.78

TS	NSD (1 - β̂)%	Design	Inverse Normalmethode						Inverse χ <sup>2</sup> -Methode							
			(1 - β̂)%	ASN	N <sub>med</sub>	k̄	(1 - 2 α̂)%	{KI <sub>j</sub> }	KI <sub>k</sub>	(1 - β̂)%	ASN	N <sub>med</sub>	k̄	(1 - 2 α̂)%	{KI <sub>j</sub> }	KI <sub>k</sub>
		SD	31.10	214.24	208	2.38	98.66	98.69	31.28	209.53	208	2.27	98.31	98.31		
T <sub>(4)</sub>	57.41	PO	51.35	254.38	308	2.58	97.45	98.63	46.57	256.10	308	2.57	96.93	98.06		
		OBF	52.77	259.51	308	2.84	98.22	98.47	50.23	259.44	308	2.78	97.50	97.50		
		WT	53.70	257.21	308	2.72	97.90	98.46	50.24	259.90	308	2.69	97.14	97.48		
		HP	55.02	269.16	308	2.88	98.16	98.27	50.97	263.33	308	2.92	97.68	97.68		
		SD	39.97	210.65	208	2.42	97.06	97.11	39.57	205.52	208	2.27	97.14	97.14		
2c) p <sub>1</sub> = 0.7, p <sub>2</sub> = 0.5, D = 0.2																
T <sub>(1)</sub>	89.91	PO	91.21	182.35	184	1.97	95.95	97.20	88.79	182.71	184	1.93	95.87	96.96		
		OBF	90.96	193.21	184	2.44	96.66	96.89	90.51	190.15	184	2.29	96.42	96.42		
		WT	92.56	186.88	184	2.15	96.65	97.29	90.82	188.93	184	2.11	96.67	96.86		
		HP	92.43	206.90	192	2.42	97.18	97.24	90.70	193.94	184	2.54	95.91	95.91		
		SD	80.43	184.49	208	2.79	92.91	93.15	81.30	174.84	208	2.23	95.37	95.37		
T <sub>(2)</sub>	89.63	PO	90.63	183.33	184	1.97	96.38	97.53	88.26	186.05	184	1.97	96.00	97.00		
		OBF	90.41	194.08	184	2.46	96.96	97.19	89.96	191.71	184	2.30	96.73	96.73		
		WT	92.06	189.58	184	2.19	96.92	97.46	90.11	191.15	184	2.14	96.69	96.90		
		HP	92.34	208.19	194	2.44	97.24	97.30	90.12	196.09	184	2.58	96.09	96.09		
		SD	80.21	185.47	208	2.80	94.22	94.41	80.59	176.77	208	2.28	95.91	95.91		
T <sub>(3)</sub>	87.47	PO	84.44	200.65	184	2.14	98.10	98.78	80.70	201.91	184	2.11	97.72	98.38		
		OBF	79.37	211.26	206	2.63	98.31	98.42	80.26	208.75	186	2.49	97.98	97.98		
		WT	84.57	205.27	184	2.34	98.56	98.96	82.24	207.49	184	2.31	98.17	98.28		
		HP	83.21	221.16	210	2.57	98.63	98.70	82.29	212.19	190	2.68	97.38	97.38		
		SD	65.28	197.04	208	2.72	97.28	97.58	69.74	187.36	208	2.37	97.33	97.33		
T <sub>(4)</sub>	88.99	PO	88.46	190.30	184	2.03	97.46	98.39	84.92	194.69	184	2.05	96.91	97.67		
		OBF	86.35	201.50	184	2.53	97.73	97.88	86.27	198.98	184	2.36	97.13	97.13		
		WT	89.06	196.60	184	2.26	97.90	98.30	87.53	197.88	184	2.19	97.43	97.67		
		HP	88.79	213.84	200	2.50	97.78	97.84	86.95	203.19	190	2.62	96.75	96.75		
		SD	72.85	190.80	208	2.76	95.32	95.48	76.40	181.91	208	2.32	96.38	96.38		
2d) p <sub>1</sub> = 0.3, p <sub>2</sub> = 0.1, D = 0.2																
T <sub>(1)</sub>	98.45	PO	98.80	140.04	84	1.58	95.86	97.56	98.21	141.14	84	1.56	95.88	96.96		
		OBF	97.20	154.42	136	2.21	96.52	96.97	97.89	149.40	124	1.96	96.61	96.61		
		WT	98.68	148.20	122	1.84	96.57	97.56	98.44	147.62	114	1.75	96.61	97.00		
		HP	97.89	169.37	184	2.06	96.59	96.70	98.29	153.65	136	2.27	96.28	96.33		
		SD	91.24	162.17	154	3.03	93.67	93.87	93.77	150.79	142	2.24	95.56	95.56		
T <sub>(2)</sub>	98.31	PO	98.69	142.56	84	1.61	96.01	97.60	98.02	144.65	84	1.60	96.29	97.37		
		OBF	97.07	155.27	136	2.23	96.82	97.22	97.68	151.78	128	2.03	96.66	96.66		
		WT	98.60	149.54	130	1.86	96.69	97.68	98.18	150.33	120	1.79	96.59	97.04		
		HP	98.27	170.52	184	2.09	96.97	97.08	98.26	155.86	138	2.35	96.57	96.63		
		SD	91.65	163.38	158	3.04	95.40	95.57	93.18	152.63	144	2.28	96.01	96.01		
T <sub>(3)</sub>	97.56	PO	95.76	161.68	148	1.81	98.20	98.94	94.52	166.72	160	1.83	97.93	98.74		
		OBF	86.54	179.94	184	2.50	98.68	98.86	90.13	172.43	166	2.28	98.70	98.70		
		WT	93.86	170.50	170	2.13	98.78	99.14	94.20	170.45	174	2.02	98.54	98.80		
		HP	90.59	187.00	184	2.34	98.95	99.04	91.91	175.64	176	2.53	98.25	98.28		
		SD	71.73	186.04	208	2.99	98.16	98.30	81.97	171.96	200	2.46	98.20	98.20		
T <sub>(4)</sub>	97.91	PO	97.76	150.76	130	1.70	97.22	98.47	96.64	151.46	84	1.66	97.10	98.08		
		OBF	93.60	167.22	158	2.37	97.76	98.01	95.01	161.08	150	2.10	97.59	97.59		
		WT	96.91	159.03	148	1.99	97.77	98.39	96.80	159.64	150	1.90	97.53	97.82		
		HP	94.42	178.17	184	2.21	97.98	98.06	96.05	164.77	158	2.43	97.35	97.39		

TS	NSD (1 - β̂)%	Design	Inverse Normalmethode						Inverse χ <sup>2</sup> -Methode					
			(1 - β̂)%	ASN	N <sub>med</sub>	k̄	(1 - 2 α̂)%	(1 - β̂)%	ASN	N <sub>med</sub>	k̄	(1 - 2 α̂)%		
							{KI <sub>j</sub> }	KI <sub>k</sub>					{KI <sub>j</sub> }	KI <sub>k</sub>
		SD	80.32	175.10	180	3.03	96.46	96.64	87.69	161.89	164	2.39	97.01	97.01
2e) p <sub>1</sub> = 0.6, p <sub>2</sub> = 0.3, D = 0.3														
T <sub>(1)</sub>	99.85	PO	99.89	115.10	84	1.35	96.26	97.69	99.78	115.19	84	1.32	96.49	97.49
		OBF	98.82	125.08	104	1.92	96.30	96.62	99.30	120.34	84	1.65	96.31	96.31
		WT	99.78	120.02	84	1.56	97.45	98.06	99.74	119.46	84	1.49	97.18	97.45
		HP	99.40	137.92	138	1.71	97.42	97.51	99.50	122.72	96	1.89	94.49	94.55
		SD	93.50	138.32	122	3.02	89.61	89.98	98.10	125.56	94	1.93	94.82	94.82
T <sub>(2)</sub>	99.85	PO	99.88	115.13	84	1.35	96.66	97.74	99.76	117.77	84	1.36	96.48	97.53
		OBF	98.58	126.06	106	1.96	96.76	96.99	99.35	121.70	84	1.66	97.03	97.03
		WT	99.77	121.26	84	1.58	97.60	98.15	99.72	120.50	84	1.49	97.27	97.55
		HP	99.50	139.17	140	1.73	97.88	97.95	99.58	124.12	96	1.95	95.28	95.33
		SD	94.14	139.40	122	3.07	91.95	92.33	98.25	127.05	98	1.98	95.57	95.57
T <sub>(3)</sub>	99.69	PO	99.52	124.79	84	1.47	97.90	98.62	99.39	124.04	84	1.41	98.02	98.58
		OBF	95.20	139.17	122	2.12	99.01	99.08	96.69	132.26	100	1.87	98.64	98.64
		WT	98.80	130.38	104	1.70	98.71	99.02	98.82	131.76	84	1.66	98.52	98.63
		HP	97.09	146.74	152	1.86	98.93	98.95	97.30	134.35	106	2.10	98.89	98.90
		SD	81.01	155.07	142	3.11	96.71	97.10	93.01	137.84	118	2.15	97.41	97.41
T <sub>(4)</sub>	99.85	PO	99.80	117.39	84	1.36	97.56	98.42	99.60	119.92	84	1.37	97.14	97.69
		OBF	97.57	131.85	114	2.03	96.87	97.03	98.58	125.98	90	1.75	97.05	97.05
		WT	99.53	124.48	84	1.61	98.10	98.48	99.61	123.45	84	1.51	97.71	97.87
		HP	98.70	141.28	146	1.76	98.22	98.29	99.10	128.38	96	1.96	95.98	96.03
		SD	87.93	146.45	132	3.08	92.65	93.17	96.38	131.81	106	2.09	95.29	95.29
2f) p <sub>1</sub> = 0.9, p <sub>2</sub> = 0.6, D = 0.3														
T <sub>(1)</sub>	99.99	PO	100.00	95.86	84	1.14	96.85	97.76	100.00	95.64	84	1.12	96.68	97.51
		OBF	99.43	104.75	90	1.67	96.59	96.90	99.85	99.06	84	1.35	97.52	97.52
		WT	99.99	99.52	84	1.29	97.84	98.37	100.00	97.89	84	1.22	97.77	97.96
		HP	99.91	113.48	84	1.39	98.40	98.55	99.90	101.84	84	1.55	96.08	96.16
		SD	95.53	118.55	104	3.01	88.88	89.68	99.26	104.34	90	1.70	95.30	95.30
T <sub>(2)</sub>	99.99	PO	100.00	96.71	84	1.15	97.08	97.98	100.00	96.85	84	1.14	96.82	97.65
		OBF	99.32	105.19	90	1.69	97.66	97.94	99.82	100.02	84	1.41	97.66	97.66
		WT	100.00	99.86	84	1.30	98.05	98.60	99.99	98.66	84	1.24	97.79	97.99
		HP	99.93	114.58	84	1.41	98.52	98.67	99.88	102.81	84	1.61	96.97	97.06
		SD	95.65	119.17	106	3.03	92.32	92.86	99.29	105.73	90	1.74	96.48	96.48
T <sub>(3)</sub>	99.97	PO	99.99	101.32	84	1.21	98.37	99.13	99.99	102.79	84	1.22	98.36	98.85
		OBF	95.26	119.43	102	1.95	99.14	99.30	98.47	108.19	84	1.57	99.15	99.15
		WT	99.66	107.53	84	1.45	99.02	99.40	99.84	106.40	84	1.36	98.78	98.91
		HP	98.51	121.71	84	1.55	99.14	99.24	98.60	110.58	84	1.80	99.31	99.37
		SD	76.86	138.08	122	3.11	97.24	97.57	93.84	115.51	90	1.97	98.18	98.18
T <sub>(4)</sub>	99.99	PO	100.00	98.69	84	1.18	97.68	98.50	100.00	98.36	84	1.16	97.73	98.34
		OBF	98.62	111.36	90	1.80	97.47	97.69	99.29	103.25	84	1.45	98.13	98.13
		WT	99.91	103.23	84	1.37	98.43	98.94	99.94	101.57	84	1.29	98.29	98.43
		HP	99.42	118.16	84	1.48	98.78	98.87	99.59	106.12	84	1.70	96.91	96.94
		SD	83.56	126.92	112	3.09	93.72	94.42	96.66	109.41	90	1.83	96.54	96.54

**Tabelle C.12:** Empirische Power  $(1-\hat{\beta})\%$  im nicht-sequentiellen Design (NSD) sowie  $(1-\hat{\beta})\%$ , mittlerer und medianer Umfang (ASN = „Average Sample Number“,  $N_{med}$ ), mittlere Stufenanzahl  $\bar{k}$  und empirische Überdeckungswahrscheinlichkeiten  $(1-2\hat{\alpha})\%$  der Sequenz der beobachteten Intervalle  $\{KI_j\}$  sowie des finalen Intervalls  $KI_k$  in flexiblen Designs mit binärer Zielvariable unter der Alternativhypothese (Szenario I, Konstellation VI)

TS	NSD	Design	Inverse Normalmethode					Inverse $\chi^2$ -Methode						
			$(1-\hat{\beta})\%$	ASN	$N_{med}$	$\bar{k}$	$(1-2\hat{\alpha})\%$	$(1-\hat{\beta})\%$	ASN	$N_{med}$	$\bar{k}$	$(1-2\hat{\alpha})\%$		
			$\{KI_j\}$					$KI_k$						
2a) $p_1 = 0.3, p_2 = 0.2, D = 0.1$														
$T_{(1)}$	61.24	PO	59.51	388.37	498	2.59	95.78	97.41	56.98	392.77	498	2.57	95.24	96.68
		OBF	64.21	384.78	460	2.81	96.48	96.74	62.69	389.16	498	2.75	96.34	96.34
		WT	63.62	385.71	480	2.70	96.49	97.16	61.66	392.27	498	2.66	95.88	96.29
		HP	67.00	403.22	498	2.86	96.70	96.78	63.85	394.77	498	2.89	96.15	96.15
		SD	63.26	386.91	436	3.13	94.42	95.12	63.95	389.81	482	2.86	96.38	96.38
$T_{(2)}$	59.87	PO	58.44	391.27	498	2.60	96.00	97.47	56.15	396.34	498	2.59	95.42	96.75
		OBF	63.44	386.30	460	2.82	96.88	97.14	62.00	391.90	498	2.77	96.37	96.37
		WT	63.29	387.60	486	2.71	96.61	97.26	61.07	393.09	498	2.67	96.29	96.71
		HP	66.77	404.63	498	2.87	96.86	96.94	62.83	396.96	498	2.90	96.58	96.58
		SD	62.70	387.78	440	3.13	95.32	95.86	63.17	393.73	496	2.87	96.72	96.72
$T_{(3)}$	53.20	PO	45.34	414.16	498	2.72	98.19	98.88	43.81	420.68	498	2.72	97.57	98.36
		OBF	47.44	408.90	498	2.90	98.72	98.89	47.67	414.32	498	2.85	98.06	98.06
		WT	48.07	410.51	498	2.82	98.36	98.66	47.94	417.47	498	2.79	97.85	98.01
		HP	51.12	423.84	498	2.92	98.71	98.74	49.29	418.32	498	2.94	98.10	98.10
		SD	45.09	409.17	484	3.07	98.23	98.48	49.84	415.58	498	2.92	98.13	98.13
$T_{(4)}$	57.41	PO	52.69	402.53	498	2.66	96.86	97.97	50.63	406.25	498	2.64	96.45	97.59
		OBF	55.57	396.30	484	2.86	97.59	97.75	55.61	402.17	498	2.80	96.99	96.99
		WT	55.50	398.59	498	2.77	97.49	97.97	55.47	403.77	498	2.73	96.94	97.22
		HP	59.45	414.64	498	2.90	97.70	97.73	56.16	407.77	498	2.92	97.28	97.28
		SD	53.67	397.74	452	3.10	96.20	96.72	56.46	404.17	498	2.90	97.11	97.11
2b) $p_1 = 0.9, p_2 = 0.8, D = 0.1$														
$T_{(1)}$	61.24	PO	78.72	340.36	408	2.39	96.03	97.81	76.26	347.75	448	2.38	95.68	97.09
		OBF	80.50	342.09	364	2.72	96.62	96.91	80.18	345.52	386	2.63	96.60	96.60
		WT	81.54	339.89	376	2.54	96.52	97.39	79.74	347.89	402	2.51	96.11	96.52
		HP	83.19	361.71	368	2.76	96.80	96.95	80.77	351.33	394	2.82	96.68	96.69
		SD	79.31	345.41	350	3.17	94.71	95.46	80.97	349.37	382	2.80	96.40	96.40
$T_{(2)}$	59.87	PO	78.03	344.12	414	2.42	96.57	98.05	75.53	353.39	461	2.41	95.83	97.28
		OBF	79.75	343.69	366	2.73	96.96	97.28	79.46	348.66	392	2.66	96.72	96.72
		WT	80.89	342.31	380	2.56	96.74	97.51	78.78	351.25	408	2.54	96.40	96.74
		HP	82.99	364.20	372	2.77	97.02	97.11	80.16	353.55	398	2.85	96.90	96.91
		SD	79.40	346.93	356	3.18	95.72	96.34	80.53	351.59	384	2.81	96.73	96.73
$T_{(3)}$	53.20	PO	62.21	384.77	478	2.63	98.43	99.18	61.38	392.64	498	2.62	98.04	98.56
		OBF	61.63	377.43	416	2.86	98.78	98.94	64.62	382.32	462	2.79	98.57	98.57
		WT	63.88	378.65	432	2.76	98.91	99.29	65.24	385.05	472	2.70	98.39	98.60
		HP	64.78	389.73	444	2.89	99.09	99.16	64.87	387.47	467	2.93	98.67	98.67
		SD	56.49	378.47	400	3.08	98.60	98.84	63.65	383.24	440	2.89	98.55	98.55
$T_{(4)}$	57.41	PO	71.06	362.58	444	2.51	97.36	98.56	69.63	369.50	492	2.49	96.95	97.98
		OBF	71.82	359.36	388	2.79	97.78	98.04	71.89	365.54	426	2.72	97.23	97.23
		WT	73.11	360.74	404	2.66	97.63	98.24	72.75	368.54	442	2.62	97.06	97.35
		HP	74.08	377.28	408	2.84	98.07	98.18	73.08	369.59	438	2.88	97.61	97.62

TS	NSD (1 - β̂)%	Design	Inverse Normalmethode						Inverse χ <sup>2</sup> -Methode					
			(1 - β̂)%	ASN	N <sub>med</sub>	k̄	(1 - 2 α̂)%	(1 - β̂)%	ASN	N <sub>med</sub>	k̄	(1 - 2 α̂)%		
							{KI <sub>j</sub> }	KI <sub>k</sub>					{KI <sub>j</sub> }	KI <sub>k</sub>
		SD	66.89	364.85	380	3.14	96.77	97.21	71.56	365.53	406	2.86	97.39	97.39
2c) p <sub>1</sub> = 0.7, p <sub>2</sub> = 0.5, D = 0.2														
T <sub>(1)</sub>	89.91	PO	98.31	222.00	201	1.88	95.86	97.20	98.23	227.36	226	1.84	95.72	96.88
		OBF	95.94	227.94	208	2.36	97.13	97.41	97.08	228.41	227	2.19	96.96	96.96
		WT	98.04	223.47	212	2.08	97.14	97.68	98.10	229.84	236	2.02	96.90	97.12
		HP	97.30	251.60	250	2.31	97.21	97.28	97.85	234.94	250	2.47	95.61	95.62
		SD	92.41	239.58	228	3.25	91.74	92.20	96.82	233.10	238	2.46	96.20	96.20
T <sub>(2)</sub>	89.63	PO	98.22	223.01	202	1.88	96.23	97.45	97.95	231.99	231	1.87	95.73	96.92
		OBF	95.44	229.57	210	2.39	97.22	97.47	97.00	231.04	234	2.22	97.41	97.41
		WT	97.88	226.31	214	2.12	97.33	97.77	97.85	232.58	240	2.05	96.87	97.13
		HP	97.48	253.26	250	2.34	97.30	97.39	97.62	236.59	250	2.50	96.06	96.06
		SD	92.63	241.55	230	3.27	93.69	94.21	96.48	235.69	240	2.50	96.79	96.79
T <sub>(3)</sub>	87.47	PO	96.04	249.96	250	2.05	97.84	98.60	95.94	256.36	250	2.00	97.48	98.19
		OBF	89.11	256.57	244	2.55	98.76	98.86	92.03	255.16	250	2.38	98.05	98.05
		WT	94.68	251.39	246	2.26	98.73	99.08	94.71	256.73	250	2.21	98.25	98.48
		HP	92.23	270.88	250	2.48	98.61	98.68	93.55	257.91	250	2.61	97.11	97.12
		SD	80.66	268.18	256	3.22	97.14	97.53	90.32	260.26	250	2.62	97.75	97.75
T <sub>(4)</sub>	88.99	PO	97.58	233.78	234	1.94	97.23	98.17	97.14	243.32	250	1.94	97.03	97.80
		OBF	93.21	241.62	224	2.46	97.81	97.96	94.74	242.53	250	2.28	97.46	97.46
		WT	96.68	236.32	228	2.17	97.82	98.18	96.72	242.79	250	2.10	97.61	97.76
		HP	95.26	261.13	250	2.40	97.78	97.84	95.80	246.40	250	2.54	96.13	96.15
		SD	86.48	252.38	256	3.25	94.15	94.54	94.34	246.34	250	2.54	97.08	97.08
2d) p <sub>1</sub> = 0.3, p <sub>2</sub> = 0.1, D = 0.2														
T <sub>(1)</sub>	98.45	PO	99.90	157.83	84	1.53	96.17	97.79	99.93	160.43	84	1.50	95.96	96.94
		OBF	97.99	168.45	136	2.15	96.93	97.38	98.87	166.95	126	1.91	97.00	97.00
		WT	99.59	164.61	128	1.80	97.08	97.88	99.60	166.22	118	1.70	97.04	97.30
		HP	99.00	193.07	198	1.96	97.03	97.20	99.22	169.55	134	2.19	96.25	96.35
		SD	94.12	182.59	154	3.32	92.88	93.41	98.12	170.73	142	2.31	96.56	96.56
T <sub>(2)</sub>	98.31	PO	99.85	159.74	84	1.56	96.35	97.79	99.88	164.41	84	1.54	96.33	97.36
		OBF	97.75	169.72	136	2.18	97.31	97.65	98.59	167.66	128	1.97	97.03	97.03
		WT	99.57	165.54	130	1.81	97.14	97.93	99.49	168.35	122	1.73	97.01	97.31
		HP	99.08	193.98	198	1.98	97.37	97.54	99.35	173.34	138	2.27	96.78	96.89
		SD	94.95	184.43	158	3.34	94.66	95.20	97.87	172.51	144	2.34	96.83	96.83
T <sub>(3)</sub>	97.56	PO	99.43	183.90	148	1.74	98.24	98.98	99.41	190.43	160	1.73	98.05	98.70
		OBF	90.56	201.07	182	2.43	98.63	98.71	93.70	191.95	168	2.19	98.52	98.52
		WT	97.11	190.69	168	2.05	98.71	99.08	97.82	190.78	174	1.93	98.56	98.73
		HP	93.44	212.29	218	2.21	98.85	98.89	95.26	194.48	176	2.44	98.19	98.22
		SD	73.88	217.99	206	3.31	98.15	98.61	89.80	201.13	186	2.56	98.44	98.44
T <sub>(4)</sub>	97.91	PO	99.77	169.87	130	1.63	97.40	98.45	99.77	171.55	84	1.58	97.17	98.02
		OBF	95.42	183.81	158	2.30	97.86	98.12	96.83	179.78	154	2.04	97.52	97.52
		WT	98.55	176.38	148	1.93	97.88	98.49	98.91	178.93	152	1.83	97.80	97.99
		HP	96.73	203.25	208	2.10	97.93	98.03	97.91	183.77	158	2.35	97.14	97.19
		SD	81.78	201.25	182	3.34	96.11	96.51	93.60	185.89	164	2.45	97.62	97.62
2e) p <sub>1</sub> = 0.6, p <sub>2</sub> = 0.3, D = 0.3														
T <sub>(1)</sub>	99.85	PO	100.00	123.68	84	1.33	96.28	97.45	100.00	124.92	84	1.30	96.48	97.33
		OBF	99.00	132.12	104	1.89	96.69	96.87	99.42	127.61	84	1.62	96.82	96.82
		WT	99.79	126.92	84	1.54	97.84	98.29	99.88	126.80	84	1.45	97.33	97.52
		HP	99.60	152.09	136	1.65	97.96	98.03	99.66	131.03	96	1.85	94.79	94.82

TS	NSD (1 - β̂)%	Design	Inverse Normalmethode					Inverse χ <sup>2</sup> -Methode						
			(1 - β̂)%	ASN	N <sub>med</sub>	k̄	(1 - 2 α̂)% {KI <sub>j</sub> } KI <sub>k</sub>	(1 - β̂)%	ASN	N <sub>med</sub>	k̄	(1 - 2 α̂)% {KI <sub>j</sub> } KI <sub>k</sub>		
		SD	93.58	145.20	122	3.18	88.13	88.64	99.12	132.10	94	1.94	95.68	95.68
T <sub>(2)</sub>	99.85	PO	100.00	123.99	84	1.33	96.66	97.53	100.00	127.64	84	1.33	96.53	97.39
		OBF	98.78	132.67	106	1.93	96.92	97.06	99.46	128.70	84	1.63	97.33	97.33
		WT	99.81	128.12	84	1.56	97.95	98.35	99.92	128.45	84	1.46	97.52	97.74
		HP	99.75	153.13	138	1.66	98.33	98.39	99.65	132.58	96	1.90	95.75	95.78
		SD	94.43	146.28	122	3.22	91.60	92.12	99.17	134.07	98	1.99	96.09	96.09
T <sub>(3)</sub>	99.69	PO	99.94	133.82	84	1.43	98.12	98.70	99.98	134.57	84	1.38	97.93	98.42
		OBF	95.79	147.53	120	2.08	98.72	98.82	97.33	140.73	100	1.83	98.71	98.71
		WT	99.32	138.48	102	1.66	98.91	99.15	99.24	140.30	84	1.62	98.58	98.72
		HP	98.19	160.29	150	1.78	99.00	99.03	97.87	143.78	104	2.06	98.98	98.99
		SD	82.62	165.66	142	3.28	96.80	97.16	94.18	147.29	116	2.17	97.64	97.64
T <sub>(4)</sub>	99.85	PO	100.00	126.16	84	1.34	97.60	98.40	100.00	129.72	84	1.34	96.87	97.35
		OBF	98.04	139.42	112	2.00	97.20	97.32	98.88	134.21	90	1.71	97.31	97.31
		WT	99.74	131.92	84	1.58	98.24	98.62	99.85	131.87	84	1.47	97.75	97.88
		HP	99.26	155.14	140	1.69	98.56	98.59	99.43	136.71	96	1.90	96.36	96.39
		SD	88.68	154.97	132	3.27	91.86	92.51	97.40	139.58	104	2.09	95.36	95.36
2f) p <sub>1</sub> = 0.9, p <sub>2</sub> = 0.6, D = 0.3														
T <sub>(1)</sub>	99.99	PO	100.00	99.57	84	1.14	96.47	97.43	100.00	99.70	84	1.12	96.56	97.31
		OBF	99.36	107.68	90	1.67	96.54	96.87	99.84	102.90	84	1.36	97.48	97.48
		WT	100.00	103.26	84	1.30	97.57	98.11	99.99	101.75	84	1.22	97.64	97.78
		HP	99.92	121.50	84	1.38	98.88	99.11	99.88	105.59	84	1.53	96.36	96.51
		SD	95.53	121.21	106	3.08	88.23	89.32	99.38	106.70	90	1.70	95.71	95.71
T <sub>(2)</sub>	99.99	PO	100.00	100.47	84	1.16	96.76	97.59	100.00	101.14	84	1.14	96.66	97.43
		OBF	99.40	108.09	90	1.69	97.68	97.92	99.84	103.91	84	1.41	97.69	97.69
		WT	99.99	103.45	84	1.30	98.00	98.47	99.99	103.12	84	1.25	97.79	97.95
		HP	99.95	122.81	84	1.40	98.94	99.16	99.89	106.79	84	1.60	97.23	97.40
		SD	95.84	122.34	106	3.11	92.77	93.47	99.35	107.78	90	1.74	96.90	96.90
T <sub>(3)</sub>	99.97	PO	100.00	105.40	84	1.21	98.24	98.88	100.00	108.07	84	1.22	98.30	98.68
		OBF	95.82	124.53	102	1.94	99.27	99.40	98.45	111.99	84	1.56	99.18	99.18
		WT	99.74	111.88	84	1.45	99.03	99.35	99.87	111.06	84	1.37	98.74	98.88
		HP	99.11	129.86	84	1.52	99.46	99.54	98.45	115.15	84	1.78	99.44	99.55
		SD	76.34	142.63	122	3.19	97.88	98.25	93.93	118.19	90	1.97	98.54	98.54
T <sub>(4)</sub>	99.99	PO	100.00	102.53	84	1.18	97.34	98.10	100.00	103.26	84	1.16	97.48	98.04
		OBF	98.81	114.77	90	1.80	97.53	97.76	99.33	107.43	84	1.45	98.14	98.14
		WT	99.95	107.40	84	1.38	98.27	98.66	99.97	106.31	84	1.30	98.25	98.35
		HP	99.58	126.39	84	1.47	99.22	99.32	99.56	110.06	84	1.68	97.37	97.50
		SD	83.41	129.93	112	3.16	94.38	95.10	96.70	111.86	90	1.84	96.84	96.84

**Tabelle C.13:** Empirische Power  $(1-\hat{\beta})\%$  im nicht-sequentiellen Design (NSD) sowie  $(1-\hat{\beta})\%$ , mittlerer und medianer Umfang (ASN = „Average Sample Number“,  $N_{med}$ ), mittlere Stufenanzahl  $\bar{k}$  und empirische Überdeckungswahrscheinlichkeiten  $(1-2\hat{\alpha})\%$  der Sequenz der beobachteten Intervalle  $\{KI_j\}$  sowie des finalen Intervalls  $KI_k$  in flexiblen Designs mit binärer Zielvariable unter der Alternativhypothese (Szenario I, Konstellation VII)

TS	NSD	Design	Inverse Normalmethode						Inverse $\chi^2$ -Methode					
			$(1-\hat{\beta})\%$	ASN	$N_{med}$	$\bar{k}$	$(1-2\hat{\alpha})\%$	$(1-\hat{\beta})\%$	ASN	$N_{med}$	$\bar{k}$	$(1-2\hat{\alpha})\%$		
						$\{KI_j\}$						$\{KI_j\}$		
						$KI_k$						$KI_k$		
2a) $p_1 = 0.3, p_2 = 0.2, D = 0.1$														
$T_{(1)}$	61.24	PO	37.80	246.03	290	2.69	95.88	97.40	36.60	249.00	308	2.65	95.18	96.67
		OBF	40.56	244.62	256	2.90	96.47	96.76	39.77	248.11	270	2.83	96.24	96.24
		WT	39.84	244.99	268	2.81	96.64	97.35	39.24	248.75	280	2.76	95.81	96.34
		HP	47.26	264.59	308	2.90	96.50	96.62	40.76	251.72	292	2.92	95.87	95.87
		SD	35.90	210.46	208	3.01	94.23	94.40	35.66	211.10	208	2.78	94.85	94.85
$T_{(2)}$	59.87	PO	36.29	248.35	292	2.71	96.09	97.59	34.68	252.92	308	2.69	95.63	96.85
		OBF	38.84	246.43	250	2.93	96.75	97.05	37.57	251.17	276	2.87	96.41	96.41
		WT	38.09	247.46	269	2.84	96.86	97.58	36.77	251.82	280	2.79	96.11	96.53
		HP	46.70	265.91	308	2.91	96.79	96.90	39.18	253.96	294	2.93	96.47	96.47
		SD	34.76	213.19	208	3.02	95.45	95.61	33.67	214.39	208	2.81	95.70	95.70
$T_{(3)}$	53.20	PO	21.33	261.20	308	2.81	98.24	98.94	22.18	265.49	308	2.80	97.62	98.34
		OBF	19.43	259.14	282	2.97	98.70	98.82	23.90	260.14	292	2.91	98.25	98.25
		WT	20.37	259.65	296	2.91	98.50	98.89	23.89	262.17	304	2.87	98.09	98.28
		HP	28.78	276.37	308	2.95	98.85	98.90	24.83	265.05	308	2.96	98.25	98.25
		SD	16.57	225.43	208	2.91	98.12	98.45	20.60	220.70	208	2.76	97.83	97.83
$T_{(4)}$	57.41	PO	28.82	254.82	308	2.76	97.26	98.39	29.59	257.46	308	2.73	96.56	97.62
		OBF	28.74	252.39	267	2.94	97.59	97.75	31.13	255.18	284	2.88	97.20	97.20
		WT	28.61	253.31	279	2.87	97.46	97.99	30.64	256.49	294	2.83	97.01	97.35
		HP	37.92	271.18	308	2.93	97.73	97.81	32.61	259.41	308	2.94	97.14	97.14
		SD	24.33	220.04	208	2.97	96.53	96.83	26.97	216.98	208	2.78	96.44	96.44
2b) $p_1 = 0.9, p_2 = 0.8, D = 0.1$														
$T_{(1)}$	61.24	PO	52.99	227.47	242	2.56	95.92	97.66	51.39	230.19	258	2.51	95.55	96.79
		OBF	54.32	229.30	222	2.88	96.46	96.75	54.73	232.45	234	2.77	96.13	96.13
		WT	54.56	228.51	226	2.73	96.34	97.18	53.23	232.21	238	2.67	95.93	96.27
		HP	63.84	247.34	288	2.85	96.55	96.74	56.02	236.80	248	2.88	96.16	96.17
		SD	48.90	207.55	208	3.22	95.28	95.41	48.57	207.88	208	2.95	95.57	95.57
$T_{(2)}$	59.87	PO	50.55	231.09	242	2.61	96.47	97.96	48.57	235.28	258	2.57	95.88	97.16
		OBF	52.97	231.85	222	2.91	96.67	96.96	52.80	235.02	236	2.81	96.39	96.39
		WT	52.61	231.48	226	2.77	96.70	97.52	51.33	236.29	240	2.72	96.26	96.65
		HP	62.95	249.23	294	2.87	96.86	96.99	54.27	239.44	250	2.91	96.63	96.64
		SD	47.24	210.86	208	3.22	96.14	96.28	46.10	211.62	208	2.98	96.35	96.35
$T_{(3)}$	53.20	PO	28.24	251.14	276	2.77	98.55	99.33	30.37	255.02	286	2.75	98.30	98.82
		OBF	24.64	250.38	240	2.97	99.09	99.29	32.32	250.32	254	2.91	98.74	98.74
		WT	25.94	250.31	248	2.90	98.85	99.26	32.48	251.81	262	2.84	98.51	98.78
		HP	38.65	266.46	308	2.95	99.12	99.21	33.44	254.24	276	2.96	98.77	98.77
		SD	19.35	225.79	218	3.04	98.75	98.91	26.88	223.31	208	2.96	98.51	98.51
$T_{(4)}$	57.41	PO	39.79	240.44	256	2.68	97.49	98.69	40.51	243.54	278	2.63	97.07	98.11
		OBF	36.85	240.91	230	2.94	97.82	98.01	42.78	242.52	246	2.84	97.43	97.43
		WT	37.85	241.26	236	2.84	97.85	98.38	42.34	243.55	250	2.77	97.10	97.40
		HP	50.30	257.83	308	2.92	97.99	98.09	43.85	246.99	262	2.93	97.59	97.59

TS	NSD (1 - β̂)%	Design	Inverse Normalmethode						Inverse χ <sup>2</sup> -Methode					
			(1 - β̂)%	ASN	N <sub>med</sub>	k̄	(1 - 2 α̂)%	(1 - β̂)%	ASN	N <sub>med</sub>	k̄	(1 - 2 α̂)%		
							{KI <sub>j</sub> }	KI <sub>k</sub>					{KI <sub>j</sub> }	KI <sub>k</sub>
		SD	31.54	220.96	214	3.15	97.46	97.60	36.05	216.89	208	2.97	97.22	97.22
2c) p <sub>1</sub> = 0.7, p <sub>2</sub> = 0.5, D = 0.2														
T <sub>(1)</sub>	89.91	PO	83.72	172.78	182	2.12	95.24	96.62	84.55	174.30	184	2.04	95.31	96.55
		OBF	83.06	180.51	184	2.61	94.70	94.92	85.61	178.57	184	2.41	95.07	95.07
		WT	84.05	175.04	184	2.34	95.41	95.92	85.71	177.87	184	2.25	95.47	95.70
		HP	89.12	190.95	184	2.53	96.72	96.78	86.15	182.46	188	2.61	94.92	94.93
		SD	78.28	178.99	194	3.53	89.38	90.26	81.31	174.10	206	2.92	93.27	93.27
T <sub>(2)</sub>	89.63	PO	83.35	175.05	184	2.14	95.86	97.03	83.16	178.18	184	2.10	95.50	96.67
		OBF	81.72	184.28	192	2.68	94.93	95.15	84.48	182.66	186	2.47	95.77	95.77
		WT	83.23	180.57	184	2.41	95.68	96.14	84.13	182.20	184	2.31	95.68	95.88
		HP	89.03	193.03	184	2.57	96.90	96.93	84.69	185.21	190	2.66	95.25	95.25
		SD	77.40	183.67	206	3.56	91.81	92.76	79.73	180.56	208	3.01	93.70	93.70
T <sub>(3)</sub>	87.47	PO	70.23	193.05	214	2.32	97.25	98.11	73.54	192.78	216	2.23	97.56	98.20
		OBF	63.00	203.60	214	2.80	97.40	97.51	72.41	197.92	214	2.63	97.05	97.05
		WT	67.76	198.09	214	2.56	97.46	97.75	73.92	199.05	214	2.48	97.30	97.42
		HP	77.24	208.90	212	2.68	98.48	98.54	74.13	201.72	214	2.74	97.20	97.20
		SD	55.45	204.16	208	3.49	97.63	98.20	66.06	195.86	208	3.11	95.57	95.57
T <sub>(4)</sub>	88.99	PO	78.23	181.53	192	2.19	96.74	97.74	78.54	185.00	184	2.17	96.55	97.21
		OBF	73.74	192.04	214	2.72	95.40	95.52	79.52	188.98	210	2.52	96.05	96.05
		WT	76.01	187.28	214	2.46	96.34	96.72	80.09	187.71	205	2.35	96.36	96.52
		HP	83.94	200.83	196	2.63	97.52	97.57	80.71	192.37	194	2.70	95.76	95.76
		SD	66.59	191.49	208	3.51	93.10	94.53	74.23	186.56	208	3.05	94.25	94.25
2d) p <sub>1</sub> = 0.3, p <sub>2</sub> = 0.1, D = 0.2														
T <sub>(1)</sub>	98.45	PO	95.35	137.68	84	1.76	95.38	97.13	95.80	138.72	84	1.70	95.35	96.48
		OBF	92.98	151.81	136	2.46	95.16	95.53	95.30	148.14	122	2.14	96.16	96.16
		WT	94.70	146.30	116	2.09	95.80	96.61	95.50	146.33	92	1.95	96.11	96.56
		HP	96.14	152.53	134	2.26	96.31	96.45	95.96	149.17	122	2.39	96.05	96.11
		SD	88.90	159.66	148	3.67	91.30	92.31	92.74	151.74	136	2.88	94.30	94.30
T <sub>(2)</sub>	98.31	PO	94.94	141.63	84	1.81	95.70	97.30	94.89	144.56	84	1.78	95.90	97.06
		OBF	92.35	156.94	144	2.54	96.16	96.45	94.58	153.03	140	2.25	96.64	96.64
		WT	94.09	151.38	144	2.15	96.30	97.08	94.64	151.75	138	2.02	96.74	97.06
		HP	96.05	153.93	136	2.30	96.82	96.95	95.41	153.78	135	2.47	96.28	96.33
		SD	88.19	165.31	160	3.70	94.84	95.35	91.73	159.70	158	2.99	95.33	95.33
T <sub>(3)</sub>	97.56	PO	82.52	164.04	174	2.06	98.40	99.04	87.84	166.25	182	2.03	98.09	98.72
		OBF	68.35	184.42	214	2.75	98.57	98.75	85.05	171.91	182	2.47	98.20	98.20
		WT	74.81	175.89	200	2.42	98.71	99.04	87.41	170.67	184	2.25	98.22	98.42
		HP	83.50	177.02	174	2.55	99.00	99.07	86.02	172.81	176	2.62	98.44	98.46
		SD	62.30	201.97	214	3.66	98.74	98.97	77.71	180.02	200	3.17	97.89	97.89
T <sub>(4)</sub>	97.91	PO	89.42	151.44	98	1.92	97.06	98.26	92.71	150.41	84	1.83	96.92	97.79
		OBF	83.19	170.04	178	2.64	96.89	97.09	90.90	160.65	158	2.30	97.21	97.21
		WT	86.12	162.59	168	2.29	97.13	97.74	92.10	159.40	166	2.11	97.37	97.55
		HP	89.75	165.07	154	2.44	97.82	97.93	91.93	162.54	156	2.54	96.78	96.81
		SD	75.92	183.61	200	3.69	95.81	96.76	85.37	168.75	174	3.08	96.09	96.09
2e) p <sub>1</sub> = 0.6, p <sub>2</sub> = 0.3, D = 0.3														
T <sub>(1)</sub>	99.85	PO	98.65	112.69	84	1.48	95.09	96.55	99.13	113.34	84	1.43	95.74	96.74
		OBF	97.21	123.95	90	2.10	92.67	93.06	98.62	119.26	84	1.78	94.56	94.56
		WT	98.29	118.76	84	1.73	94.36	95.04	99.03	118.23	84	1.62	95.12	95.40
		HP	97.88	122.80	98	1.89	96.59	96.67	98.71	120.39	90	2.01	92.77	92.83

TS	NSD (1 - β̂)%	Design	Inverse Normalmethode					Inverse χ <sup>2</sup> -Methode						
			(1 - β̂)%	ASN	N <sub>med</sub>	k̄	(1 - 2 α̂)% {KI <sub>j</sub> } KI <sub>k</sub>	(1 - β̂)%	ASN	N <sub>med</sub>	k̄	(1 - 2 α̂)% {KI <sub>j</sub> } KI <sub>k</sub>		
		SD	92.37	132.91	108	3.46	85.18	87.32	97.98	121.19	90	2.33	90.81	90.81
T <sub>(2)</sub>	99.85	PO	98.71	114.63	84	1.50	96.35	97.43	99.02	116.76	84	1.48	95.77	96.75
		OBF	96.88	127.68	90	2.18	96.57	96.78	98.47	122.66	84	1.82	95.52	95.52
		WT	98.12	122.36	84	1.79	97.02	97.47	98.77	121.75	84	1.66	95.80	95.99
		HP	98.01	124.00	102	1.93	97.15	97.24	98.68	123.43	96	2.07	94.00	94.06
		SD	92.46	137.88	118	3.53	92.16	92.58	97.51	127.57	96	2.46	91.70	91.70
T <sub>(3)</sub>	99.69	PO	95.09	127.28	84	1.67	97.93	98.64	97.82	124.31	84	1.55	97.86	98.35
		OBF	89.65	144.19	130	2.33	99.31	99.40	94.99	134.30	96	2.03	98.75	98.75
		WT	92.81	136.39	90	1.95	98.85	99.16	96.86	133.80	84	1.85	98.70	98.78
		HP	93.12	134.98	116	2.09	98.86	98.88	95.61	134.92	96	2.21	98.70	98.71
		SD	80.55	162.15	154	3.60	97.35	97.61	91.84	142.15	116	2.67	97.70	97.70
T <sub>(4)</sub>	99.85	PO	97.92	117.11	84	1.51	97.29	98.20	98.52	119.44	84	1.50	96.66	97.23
		OBF	94.39	134.06	110	2.23	95.96	96.10	97.54	126.92	90	1.90	95.34	95.34
		WT	96.46	126.33	84	1.81	96.69	97.05	98.39	124.25	84	1.66	96.23	96.39
		HP	96.31	128.15	106	1.97	97.29	97.37	97.70	127.45	96	2.08	94.38	94.44
		SD	86.63	147.00	128	3.56	92.71	93.42	95.73	131.95	96	2.56	91.29	91.29
2f) p <sub>1</sub> = 0.9, p <sub>2</sub> = 0.6, D = 0.3														
T <sub>(1)</sub>	99.99	PO	99.86	95.12	84	1.20	95.76	97.26	99.95	94.73	84	1.17	96.36	97.23
		OBF	99.10	105.73	90	1.79	93.56	94.18	99.78	100.26	84	1.44	96.23	96.23
		WT	99.72	100.09	84	1.40	95.22	96.31	99.93	98.84	84	1.30	96.60	96.89
		HP	99.39	101.93	84	1.53	96.79	96.98	99.79	101.13	84	1.64	94.77	94.87
		SD	94.95	116.97	102	3.32	86.16	87.91	98.89	104.25	90	1.96	92.64	92.64
T <sub>(2)</sub>	99.99	PO	99.84	96.81	84	1.23	96.70	97.83	99.94	97.11	84	1.21	96.42	97.38
		OBF	98.82	108.33	90	1.86	97.50	97.71	99.71	102.96	84	1.52	96.97	96.97
		WT	99.71	102.68	84	1.44	97.76	98.22	99.87	101.83	84	1.35	97.22	97.42
		HP	99.27	102.91	84	1.58	97.51	97.68	99.70	103.18	84	1.71	96.05	96.15
		SD	94.86	120.97	104	3.38	93.21	93.63	98.84	108.75	90	2.07	94.78	94.78
T <sub>(3)</sub>	99.97	PO	98.23	104.79	84	1.34	98.22	99.03	99.65	105.71	84	1.33	98.25	98.68
		OBF	92.43	128.89	104	2.15	99.49	99.63	98.17	113.33	84	1.71	99.24	99.24
		WT	96.41	116.25	84	1.64	99.16	99.50	99.28	111.65	84	1.52	98.96	99.05
		HP	95.44	113.29	84	1.79	99.05	99.15	97.93	113.81	84	1.90	99.25	99.32
		SD	79.02	150.30	136	3.49	98.41	98.55	93.19	121.76	90	2.33	98.74	98.74
T <sub>(4)</sub>	99.99	PO	99.47	99.71	84	1.27	97.25	98.21	99.88	99.08	84	1.23	97.35	98.00
		OBF	97.58	117.41	90	1.99	97.37	97.51	99.11	106.56	84	1.56	97.27	97.27
		WT	98.89	108.67	84	1.54	98.11	98.50	99.72	105.30	84	1.41	97.61	97.72
		HP	97.63	107.71	84	1.68	98.07	98.17	99.02	107.39	84	1.80	95.81	95.84
		SD	85.58	133.28	112	3.42	94.65	95.40	95.90	112.81	90	2.16	94.84	94.84

**Tabelle C.14:** Empirische Power  $(1-\hat{\beta})\%$  im nicht-sequentiellen Design (NSD) sowie  $(1-\hat{\beta})\%$ , mittlerer und medianer Umfang (ASN = „Average Sample Number“,  $N_{med}$ ), mittlere Stufenanzahl  $\bar{k}$  und empirische Überdeckungswahrscheinlichkeiten  $(1-2\hat{\alpha})\%$  der Sequenz der beobachteten Intervalle  $\{KI_j\}$  sowie des finalen Intervalls  $KI_k$  in flexiblen Designs mit binärer Zielvariable unter der Alternativhypothese (Szenario I, Konstellation VIII)

TS	NSD	Design	Inverse Normalmethode						Inverse $\chi^2$ -Methode					
			$(1-\hat{\beta})\%$	ASN	$N_{med}$	$\bar{k}$	$(1-2\hat{\alpha})\%$		$(1-\hat{\beta})\%$	ASN	$N_{med}$	$\bar{k}$	$(1-2\hat{\alpha})\%$	
2a) $p_1 = 0.3, p_2 = 0.2, D = 0.1$														
$T_{(1)}$	61.24	PO	50.43	363.21	402	2.67	95.67	97.25	51.45	369.24	422	2.63	95.09	96.49
		OBF	53.69	351.86	362	2.89	96.33	96.50	55.19	358.57	378	2.81	96.15	96.15
		WT	53.26	354.28	370	2.80	96.21	96.78	54.75	362.30	392	2.73	95.81	96.20
		HP	62.87	381.09	442	2.90	96.10	96.19	56.41	364.05	392	2.91	96.24	96.24
		SD	55.83	355.52	370	3.53	94.04	94.53	57.33	363.93	394	3.29	95.33	95.33
$T_{(2)}$	59.87	PO	49.42	365.66	404	2.69	96.22	97.55	50.08	373.99	424	2.66	95.34	96.55
		OBF	52.05	355.47	362	2.92	96.81	97.02	53.41	362.93	384	2.84	96.50	96.50
		WT	52.09	359.14	372	2.82	96.54	97.04	52.42	366.86	394	2.77	96.20	96.63
		HP	62.58	382.30	444	2.90	96.37	96.46	55.57	366.57	398	2.92	96.83	96.83
		SD	54.99	359.04	370	3.53	95.20	95.69	55.54	370.75	400	3.33	96.11	96.11
$T_{(3)}$	53.20	PO	32.38	393.68	438	2.81	98.12	98.88	36.74	395.18	448	2.77	97.65	98.34
		OBF	32.06	383.75	382	2.96	98.68	98.77	39.96	381.04	400	2.89	98.07	98.07
		WT	32.62	386.35	392	2.91	98.52	98.77	39.64	386.97	414	2.85	97.86	98.06
		HP	45.69	406.83	498	2.95	98.64	98.67	40.77	388.50	426	2.95	98.06	98.06
		SD	32.87	397.78	408	3.48	98.33	98.82	41.21	393.97	428	3.36	98.04	98.04
$T_{(4)}$	57.41	PO	41.66	378.93	417	2.74	96.92	97.99	44.43	382.25	432	2.70	96.37	97.48
		OBF	42.04	369.48	368	2.94	97.30	97.40	47.74	370.09	392	2.86	96.89	96.89
		WT	42.11	372.99	382	2.86	97.41	97.86	46.65	375.84	400	2.80	96.87	97.16
		HP	53.80	394.17	466	2.93	97.23	97.28	48.69	377.40	410	2.94	97.27	97.27
		SD	43.59	379.53	390	3.50	96.45	97.08	48.74	381.05	412	3.35	96.75	96.75
2b) $p_1 = 0.9, p_2 = 0.8, D = 0.1$														
$T_{(1)}$	61.24	PO	69.50	323.24	352	2.54	95.94	97.67	69.74	326.85	358	2.47	95.54	96.90
		OBF	69.29	316.80	338	2.86	96.23	96.58	72.34	322.68	342	2.74	96.33	96.33
		WT	69.89	317.61	340	2.71	96.42	97.32	72.01	325.26	347	2.63	95.91	96.36
		HP	78.73	338.05	354	2.83	96.55	96.75	73.58	326.39	352	2.86	96.32	96.34
		SD	70.63	320.16	340	3.64	94.67	95.08	72.91	328.01	348	3.33	96.28	96.28
$T_{(2)}$	59.87	PO	67.78	328.75	352	2.58	96.52	97.99	67.90	334.79	366	2.53	95.85	97.26
		OBF	68.07	320.85	338	2.90	96.59	96.93	70.91	325.82	344	2.78	96.68	96.68
		WT	68.74	321.53	340	2.75	96.59	97.43	70.20	330.99	352	2.68	96.28	96.69
		HP	78.42	340.39	358	2.85	96.76	96.92	72.27	329.77	352	2.89	96.76	96.78
		SD	69.91	324.31	344	3.65	95.58	96.01	70.90	334.07	352	3.37	96.84	96.84
$T_{(3)}$	53.20	PO	45.70	367.74	380	2.75	98.53	99.30	51.69	369.33	398	2.72	98.15	98.63
		OBF	41.17	363.02	352	2.96	98.82	99.01	53.67	354.35	362	2.87	98.63	98.63
		WT	43.36	365.51	360	2.89	98.87	99.28	54.26	357.52	368	2.80	98.41	98.64
		HP	57.75	374.84	406	2.95	98.99	99.05	55.16	359.35	378	2.95	98.59	98.60
		SD	41.10	383.30	386	3.58	98.56	98.98	52.97	368.91	388	3.43	98.63	98.63
$T_{(4)}$	57.41	PO	57.15	346.78	362	2.65	97.31	98.58	60.90	350.01	378	2.59	97.02	98.02
		OBF	53.72	341.15	342	2.92	97.68	97.87	62.61	340.24	356	2.81	97.31	97.31
		WT	55.89	345.15	348	2.82	97.54	98.18	62.75	344.48	360	2.73	97.07	97.40
		HP	67.70	356.96	382	2.90	97.89	97.97	63.89	344.10	364	2.91	97.51	97.53

TS	NSD (1 - β̂)%	Design	Inverse Normalmethode						Inverse χ <sup>2</sup> -Methode					
			(1 - β̂)%	ASN	N <sub>med</sub>	k̄	(1 - 2 α̂)%	(1 - β̂)%	ASN	N <sub>med</sub>	k̄	(1 - 2 α̂)%		
							{KI <sub>j</sub> }	KI <sub>k</sub>					{KI <sub>j</sub> }	KI <sub>k</sub>
		SD	54.24	357.07	366	3.62	97.00	97.43	62.02	350.95	370	3.41	97.43	97.43
2c) p <sub>1</sub> = 0.7, p <sub>2</sub> = 0.5, D = 0.2														
T <sub>(1)</sub>	89.91	PO	93.81	219.48	182	2.08	95.25	96.55	95.29	220.13	202	1.98	95.44	96.55
		OBF	90.83	222.82	194	2.58	94.74	94.98	94.05	221.59	210	2.38	95.29	95.29
		WT	92.83	220.27	202	2.32	95.27	95.90	94.49	225.39	222	2.23	95.25	95.48
		HP	94.11	227.36	206	2.49	96.92	97.01	94.16	223.31	202	2.58	94.67	94.67
		SD	88.51	228.00	198	3.72	89.09	90.15	93.60	224.59	206	3.07	92.68	92.68
T <sub>(2)</sub>	89.63	PO	93.48	221.14	188	2.10	95.78	96.94	94.59	227.15	220	2.05	95.32	96.40
		OBF	89.88	226.74	202	2.65	95.07	95.29	93.11	227.51	220	2.42	96.12	96.12
		WT	92.70	225.55	208	2.38	95.50	96.13	93.11	230.75	230	2.28	95.44	95.63
		HP	94.15	228.93	206	2.52	97.10	97.17	93.78	228.10	206	2.63	95.34	95.34
		SD	87.40	232.35	208	3.75	91.96	93.07	92.46	233.62	218	3.16	93.45	93.45
T <sub>(3)</sub>	87.47	PO	85.34	252.45	250	2.28	97.49	98.23	90.56	250.92	250	2.18	97.39	98.14
		OBF	77.28	263.01	258	2.77	97.44	97.61	86.23	251.38	250	2.57	97.01	97.01
		WT	81.42	259.44	258	2.53	97.34	97.72	88.08	254.02	250	2.43	97.41	97.66
		HP	86.60	253.37	250	2.63	98.59	98.68	87.25	253.00	250	2.71	97.19	97.19
		SD	73.38	279.31	272	3.77	97.36	98.09	82.86	260.75	250	3.30	95.70	95.70
T <sub>(4)</sub>	88.99	PO	90.58	233.90	218	2.15	96.71	97.73	92.78	238.46	248	2.12	96.85	97.60
		OBF	85.10	242.15	238	2.69	95.49	95.69	90.01	237.66	240	2.47	96.26	96.26
		WT	87.86	238.45	230	2.43	96.15	96.60	91.88	239.59	250	2.31	96.47	96.68
		HP	90.60	240.44	226	2.58	97.31	97.41	90.87	237.69	224	2.65	95.80	95.81
		SD	80.74	253.76	250	3.76	92.95	94.56	88.43	243.80	230	3.20	94.13	94.13
2d) p <sub>1</sub> = 0.3, p <sub>2</sub> = 0.1, D = 0.2														
T <sub>(1)</sub>	98.45	PO	98.52	160.40	84	1.74	95.10	97.05	99.17	162.00	84	1.66	95.08	96.22
		OBF	95.86	174.10	134	2.43	95.45	95.89	97.82	170.70	122	2.12	96.05	96.05
		WT	97.86	170.08	116	2.08	95.75	96.66	98.65	170.06	90	1.92	96.02	96.55
		HP	96.71	165.87	134	2.20	96.51	96.74	98.08	168.94	120	2.37	95.74	95.78
		SD	92.46	181.86	148	3.78	91.02	92.20	96.34	174.03	136	2.93	94.61	94.61
T <sub>(2)</sub>	98.31	PO	98.51	165.69	84	1.79	95.37	97.08	98.90	169.77	84	1.75	95.70	96.97
		OBF	94.79	178.90	144	2.52	96.45	96.82	97.47	177.00	138	2.23	96.38	96.38
		WT	97.35	175.37	142	2.13	96.26	97.19	97.78	176.76	134	2.01	96.38	96.84
		HP	96.72	167.58	136	2.24	96.78	96.98	97.63	172.98	132	2.45	96.35	96.40
		SD	91.32	186.72	158	3.80	94.16	95.08	95.30	184.42	158	3.05	95.38	95.38
T <sub>(3)</sub>	97.56	PO	92.10	200.69	170	2.03	98.25	99.03	95.90	201.00	184	2.00	97.82	98.60
		OBF	81.13	230.54	230	2.71	98.51	98.66	90.83	202.67	182	2.44	97.95	97.95
		WT	87.02	220.65	212	2.40	98.48	98.79	93.86	202.96	188	2.23	98.08	98.26
		HP	86.91	196.47	174	2.49	98.71	98.80	90.85	198.51	176	2.59	98.45	98.47
		SD	76.26	253.60	250	3.83	98.62	99.08	85.87	216.63	204	3.27	97.93	97.93
T <sub>(4)</sub>	97.91	PO	96.67	180.73	94	1.89	96.64	97.80	98.22	179.18	84	1.80	96.80	97.70
		OBF	90.98	204.87	180	2.62	97.15	97.41	95.27	187.78	158	2.27	97.28	97.28
		WT	93.96	199.09	172	2.28	96.84	97.49	96.47	187.52	166	2.10	97.13	97.32
		HP	91.60	181.17	154	2.38	97.54	97.67	95.41	186.02	156	2.52	96.81	96.85
		SD	83.50	218.82	200	3.81	95.79	96.88	90.87	197.28	178	3.14	96.47	96.47
2e) p <sub>1</sub> = 0.6, p <sub>2</sub> = 0.3, D = 0.3														
T <sub>(1)</sub>	99.85	PO	99.65	123.81	84	1.46	95.07	96.41	99.90	125.41	84	1.42	95.81	96.69
		OBF	98.34	134.72	90	2.08	92.71	93.01	99.35	130.68	84	1.76	94.46	94.46
		WT	99.34	130.85	84	1.72	94.58	95.15	99.66	129.72	84	1.60	95.18	95.35
		HP	98.24	129.11	100	1.86	96.84	96.89	99.30	129.13	92	2.01	92.95	93.01

TS	NSD (1 - β̂)%	Design	Inverse Normalmethode					Inverse χ <sup>2</sup> -Methode						
			(1 - β̂)%	ASN	N <sub>med</sub>	k̄	(1 - 2 α̂)% {KI <sub>j</sub> } KI <sub>k</sub>	(1 - β̂)%	ASN	N <sub>med</sub>	k̄	(1 - 2 α̂)% {KI <sub>j</sub> } KI <sub>k</sub>		
		SD	93.40	143.73	108	3.51	84.74	86.84	98.53	130.74	90	2.36	90.54	90.54
T <sub>(2)</sub>	99.85	PO	99.73	126.23	84	1.48	96.50	97.38	99.84	129.54	84	1.47	95.82	96.64
		OBF	97.63	138.32	90	2.18	96.43	96.59	99.13	133.58	84	1.80	95.59	95.59
		WT	99.15	134.25	84	1.78	97.21	97.58	99.52	133.89	84	1.65	96.05	96.16
		HP	98.49	130.14	102	1.90	97.19	97.26	99.05	131.67	96	2.05	94.14	94.21
		SD	93.44	148.45	118	3.58	91.90	92.39	98.24	137.98	96	2.48	91.89	91.89
T <sub>(3)</sub>	99.69	PO	98.64	144.78	84	1.63	98.03	98.68	99.49	139.81	84	1.53	98.01	98.40
		OBF	93.65	165.66	130	2.32	99.14	99.23	96.19	148.76	96	2.02	98.63	98.63
		WT	97.30	156.54	90	1.93	98.92	99.17	98.21	148.59	84	1.82	98.81	98.86
		HP	93.86	143.07	116	2.06	98.92	98.95	96.67	145.99	96	2.20	98.75	98.77
		SD	84.26	184.08	156	3.66	97.04	97.47	93.36	155.94	116	2.69	97.64	97.64
T <sub>(4)</sub>	99.85	PO	99.59	129.99	84	1.49	97.34	98.12	99.70	133.59	84	1.49	96.75	97.25
		OBF	96.37	149.50	110	2.22	95.88	95.99	98.47	139.52	90	1.88	95.22	95.22
		WT	98.62	142.63	84	1.80	96.79	97.22	99.35	137.31	84	1.64	96.32	96.48
		HP	96.94	135.55	104	1.95	97.52	97.58	98.20	137.87	96	2.08	94.41	94.48
		SD	88.18	162.33	128	3.61	92.19	92.99	96.47	143.32	96	2.58	91.53	91.53
2f) p <sub>1</sub> = 0.9, p <sub>2</sub> = 0.6, D = 0.3														
T <sub>(1)</sub>	99.99	PO	99.98	100.29	84	1.22	95.35	97.07	99.99	99.37	84	1.18	96.25	97.10
		OBF	99.10	110.49	90	1.80	93.50	94.00	99.80	105.19	84	1.45	96.01	96.01
		WT	99.89	105.04	84	1.40	94.87	95.88	99.94	103.75	84	1.31	96.56	96.88
		HP	99.33	104.99	84	1.53	96.90	97.12	99.75	104.91	84	1.65	94.69	94.81
		SD	95.07	119.78	102	3.33	86.32	88.02	99.01	106.91	90	1.97	93.07	93.07
T <sub>(2)</sub>	99.99	PO	99.95	102.08	84	1.24	96.43	97.67	00.00	102.32	84	1.21	96.23	97.12
		OBF	98.75	112.35	90	1.86	97.32	97.50	99.69	108.18	84	1.53	96.76	96.76
		WT	99.84	107.98	84	1.45	97.56	98.02	99.89	107.30	84	1.37	97.14	97.37
		HP	99.26	105.34	84	1.56	97.68	97.92	99.71	106.74	84	1.71	96.21	96.34
		SD	95.01	124.92	104	3.40	93.18	93.67	98.90	111.97	90	2.08	95.06	95.06
T <sub>(3)</sub>	99.97	PO	99.60	114.59	84	1.35	98.28	98.96	99.84	113.41	84	1.34	98.26	98.63
		OBF	94.78	142.24	104	2.13	99.61	99.66	98.08	120.20	84	1.72	99.16	99.16
		WT	98.86	130.13	84	1.65	99.18	99.37	99.49	118.69	84	1.52	98.96	99.03
		HP	95.86	117.98	84	1.77	99.32	99.39	97.93	119.49	84	1.91	99.26	99.36
		SD	80.98	163.69	138	3.51	98.44	98.56	93.36	127.49	90	2.35	98.97	98.97
T <sub>(4)</sub>	99.99	PO	99.92	107.09	84	1.28	97.22	98.24	99.95	105.62	84	1.25	97.42	98.06
		OBF	98.11	125.62	90	1.98	97.32	97.44	99.04	113.05	84	1.58	97.18	97.18
		WT	99.48	117.18	84	1.54	97.90	98.30	99.73	111.90	84	1.43	97.58	97.77
		HP	97.63	111.08	84	1.68	98.12	98.28	99.03	111.92	84	1.81	96.26	96.39
		SD	86.36	140.17	112	3.43	94.52	95.21	96.01	117.61	90	2.18	95.46	95.46

## C.2 Ordinale Zielvariable

In den folgenden Tabellen finden sich die Ergebnisse der Konstellationen II-VIII der Simulationsstudie aus Kapitel 3.2.4 sowohl unter der Null- als auch unter der Alternativhypothese.

### 1) Unter der Nullhypothese

**Tabelle C.15:** Empirisches Niveau  $\hat{\alpha}\%$  im nicht-sequentiellen Design (NSD) sowie  $\hat{\alpha}\%$ , mittlerer und medianer Umfang (ASN = „Average Sample Number“,  $N_{med}$ ), mittlere Stufenanzahl  $\bar{k}$  und empirische Überdeckungswahrscheinlichkeiten  $(1 - 2\hat{\alpha})\%$  der Sequenz der beobachteten Intervalle  $\{KI_j\}$  sowie des finalen Intervalls  $KI_k$  in flexiblen Designs mit ordinaler Zielvariable unter der Nullhypothese (Konstellation II)

NSD	Design	Inverse Normalmethode					Inverse $\chi^2$ -Methode						
		$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$	$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$		
						$\{KI_j\}$	$KI_k$				$KI_k$	$\{KI_j\}$	
1a) $\vec{p}_1 = \vec{p}_2 = (0.25, 0.25, 0.25, 0.25)$ , $D_G = 0.0$													
2.66	PO	3.08	689.79	706	4.92	93.88	95.96	3.04	689.79	706	4.92	93.54	95.56
	OBF	2.94	690.22	706	4.97	94.15	94.59	3.14	688.86	706	4.97	93.46	93.46
	WT	3.10	690.40	706	4.96	94.01	95.05	3.19	689.18	706	4.95	93.48	93.95
	HP	2.78	698.32	706	4.98	94.06	94.45	3.27	685.88	706	4.98	93.29	93.32
	SD	2.84	565.16	576	4.06	93.40	94.66	3.03	564.88	576	4.02	93.97	93.97
1b) $\vec{p}_1 = \vec{p}_2 = (0.70, 0.10, 0.10, 0.10)$ , $D_G = 0.0$													
2.37	PO	3.13	689.54	706	4.92	93.80	95.87	3.17	688.79	706	4.91	93.67	95.68
	OBF	2.73	690.88	706	4.98	94.48	94.99	3.09	689.58	706	4.97	93.82	93.82
	WT	2.84	691.27	706	4.96	94.34	95.24	3.29	688.66	706	4.95	93.84	94.22
	HP	2.75	700.02	706	4.99	94.31	94.60	3.06	687.21	706	4.98	93.93	93.95
	SD	2.59	564.90	576	4.05	93.37	94.80	2.92	564.41	576	4.02	94.20	94.20
1c) $\vec{p}_1 = \vec{p}_2 = (0.10, 0.10, 0.10, 0.70)$ , $D_G = 0.0$													
2.69	PO	2.98	690.01	706	4.92	94.29	96.23	2.99	690.03	706	4.92	94.15	95.83
	OBF	2.61	691.06	706	4.98	94.62	95.13	2.78	689.46	706	4.97	94.24	94.24
	WT	2.98	690.62	706	4.96	94.31	95.23	2.80	690.52	706	4.95	94.31	94.80
	HP	2.58	699.32	706	4.98	94.67	95.02	2.68	687.58	706	4.98	94.23	94.28
	SD	2.75	565.60	576	4.06	93.16	94.46	2.74	565.57	576	4.02	94.63	94.63
1d) $\vec{p}_1 = \vec{p}_2 = (0.05, 0.15, 0.75, 0.05)$ , $D_G = 0.0$													
2.49	PO	3.26	689.39	706	4.92	93.90	95.95	3.08	689.41	706	4.91	93.93	95.69
	OBF	2.95	689.64	706	4.98	94.35	94.88	3.21	688.12	706	4.97	94.00	94.00
	WT	3.10	690.69	706	4.96	94.15	95.15	3.07	688.87	706	4.95	94.02	94.51
	HP	2.91	698.64	706	4.98	94.47	94.80	3.10	686.95	706	4.98	93.57	93.58
	SD	3.01	564.86	576	4.06	92.80	94.29	2.84	565.02	576	4.02	94.19	94.19
1e) $\vec{p}_1 = \vec{p}_2 = (0.15, 0.75, 0.05, 0.05)$ , $D_G = 0.0$													
2.49	PO	3.28	688.14	706	4.91	93.84	95.91	3.18	689.04	706	4.91	93.40	95.39
	OBF	2.55	690.28	706	4.98	94.55	95.09	3.21	687.69	706	4.97	93.67	93.67
	WT	2.81	691.24	706	4.96	94.11	95.25	3.32	688.86	706	4.95	93.58	93.99
	HP	2.85	697.82	706	4.98	94.57	94.79	3.47	684.50	706	4.97	93.05	93.12
	SD	2.75	565.02	576	4.06	92.77	94.14	3.11	564.55	576	4.02	93.86	93.86

**Tabelle C.16:** Empirisches Niveau  $\hat{\alpha}\%$  im nicht-sequentiellen Design (NSD) sowie  $\hat{\alpha}\%$ , mittlerer und medianer Umfang (ASN = „Average Sample Number“,  $N_{med}$ ), mittlere Stufenanzahl  $\bar{k}$  und empirische Überdeckungswahrscheinlichkeiten  $(1 - 2\hat{\alpha})\%$  der Sequenz der beobachteten Intervalle  $\{KI_j\}$  sowie des finalen Intervalls  $KI_k$  in flexiblen Designs mit ordinaler Zielvariable unter der Nullhypothese (Konstellation III)

NSD	Design	Inverse Normalmethode						Inverse $\chi^2$ -Methode					
		$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$ $\{KI_j\}$ $KI_k$		$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$ $KI_k$ $\{KI_j\}$	
1a) $\vec{p}_1 = \vec{p}_2 = (0.25, 0.25, 0.25, 0.25)$ , $D_G = 0.0$													
2.66	PO	3.41	327.42	342	4.91	93.05	95.39	3.54	328.93	342	4.90	92.75	94.61
	OBF	3.86	318.74	342	4.95	92.55	93.12	4.26	320.87	342	4.95	91.29	91.29
	WT	3.40	321.76	342	4.94	93.01	94.16	3.90	322.82	342	4.93	91.57	92.09
	HP	3.05	336.62	342	4.98	94.09	94.46	4.50	316.09	342	4.95	90.55	90.59
	SD	3.45	207.64	210	3.33	91.62	92.87	4.01	207.19	210	3.26	91.80	91.80
1b) $\vec{p}_1 = \vec{p}_2 = (0.70, 0.10, 0.10, 0.10)$ , $D_G = 0.0$													
2.37	PO	3.43	327.74	342	4.91	93.55	95.60	3.30	329.20	342	4.91	93.20	95.15
	OBF	3.51	319.60	342	4.96	93.44	93.94	4.04	321.17	342	4.96	92.34	92.34
	WT	3.76	321.21	342	4.94	92.82	93.93	3.80	323.22	342	4.94	92.57	93.09
	HP	2.99	336.53	342	4.98	93.98	94.36	4.22	318.13	342	4.97	92.06	92.12
	SD	3.15	208.15	210	3.35	92.31	93.74	3.96	207.18	210	3.29	92.38	92.38
1c) $\vec{p}_1 = \vec{p}_2 = (0.10, 0.10, 0.10, 0.70)$ , $D_G = 0.0$													
2.69	PO	3.61	327.28	342	4.91	93.26	95.44	3.57	328.69	342	4.91	93.14	95.07
	OBF	3.75	319.27	342	4.96	93.05	93.64	3.75	321.93	342	4.96	92.69	92.69
	WT	3.83	321.13	342	4.94	92.95	94.02	3.91	323.15	342	4.94	92.65	93.00
	HP	2.90	336.57	342	4.98	94.10	94.52	4.11	318.50	342	4.97	91.83	91.87
	SD	3.26	208.30	210	3.38	92.36	93.67	3.85	207.19	210	3.26	92.68	92.68
1d) $\vec{p}_1 = \vec{p}_2 = (0.05, 0.15, 0.75, 0.05)$ , $D_G = 0.0$													
2.49	PO	3.62	327.47	342	4.91	93.28	95.24	3.82	328.40	342	4.90	92.88	94.67
	OBF	3.74	318.47	342	4.96	93.40	93.95	4.11	321.61	342	4.95	92.22	92.22
	WT	3.56	321.73	342	4.94	93.51	94.39	3.97	323.21	342	4.94	92.77	93.16
	HP	3.41	335.64	342	4.97	93.78	94.16	4.15	318.10	342	4.97	91.83	91.89
	SD	3.16	207.98	210	3.37	92.30	93.78	3.72	207.57	210	3.27	92.85	92.85
1e) $\vec{p}_1 = \vec{p}_2 = (0.15, 0.75, 0.05, 0.05)$ , $D_G = 0.0$													
2.49	PO	3.76	326.55	342	4.90	92.97	95.22	4.19	327.75	342	4.90	92.24	94.20
	OBF	3.80	318.64	342	4.96	93.00	93.63	4.62	319.57	342	4.95	91.33	91.33
	WT	3.96	321.55	342	4.94	92.76	93.99	4.43	321.37	342	4.92	91.93	92.43
	HP	3.43	336.50	342	4.98	93.44	93.77	4.57	318.08	342	4.97	91.36	91.41
	SD	3.47	207.56	210	3.38	91.89	93.56	4.13	207.12	210	3.30	92.40	92.40

**Tabelle C.17:** Empirisches Niveau  $\hat{\alpha}\%$  im nicht-sequentiellen Design (NSD) sowie  $\hat{\alpha}\%$ , mittlerer und medianer Umfang (ASN = „Average Sample Number“,  $N_{med}$ ), mittlere Stufenanzahl  $\bar{k}$  und empirische Überdeckungswahrscheinlichkeiten  $(1 - 2\hat{\alpha})\%$  der Sequenz der beobachteten Intervalle  $\{KI_j\}$  sowie des finalen Intervalls  $KI_k$  in flexiblen Designs mit ordinaler Zielvariable unter der Nullhypothese (Konstellation IV)

NSD	Design	Inverse Normalmethode						Inverse $\chi^2$ -Methode					
		$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$ $\{KI_j\}$ $KI_k$		$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$ $KI_k$ $\{KI_j\}$	
1a) $\vec{p}_1 = \vec{p}_2 = (0.25, 0.25, 0.25, 0.25)$ , $D_G = 0.0$													
2.66	PO	3.37	663.84	706	4.91	93.41	95.58	3.51	667.54	706	4.90	92.93	94.96
	OBF	3.86	643.53	706	4.96	92.51	92.90	4.08	647.37	706	4.95	91.75	91.75
	WT	3.68	648.62	706	4.94	92.83	93.85	4.04	651.33	706	4.93	92.04	92.40
	HP	3.12	686.49	706	4.98	93.76	94.11	4.44	639.87	706	4.95	91.34	91.36
	SD	3.97	561.75	576	4.54	90.69	92.11	4.42	562.76	576	4.44	91.34	91.34
1b) $\vec{p}_1 = \vec{p}_2 = (0.70, 0.10, 0.10, 0.10)$ , $D_G = 0.0$													
2.37	PO	3.18	664.26	706	4.92	93.63	95.72	3.37	668.48	706	4.91	93.65	95.50
	OBF	3.42	645.51	706	4.96	93.46	94.08	3.74	647.56	706	4.96	92.78	92.78
	WT	3.37	651.47	706	4.95	93.58	94.66	3.29	654.95	706	4.95	93.53	93.99
	HP	2.64	687.61	706	4.98	94.50	94.75	3.93	639.20	706	4.96	92.88	92.91
	SD	3.15	561.60	576	4.60	91.58	93.53	3.67	564.89	576	4.46	92.89	92.89
1c) $\vec{p}_1 = \vec{p}_2 = (0.10, 0.10, 0.10, 0.70)$ , $D_G = 0.0$													
2.69	PO	3.44	662.81	706	4.91	93.88	95.80	3.37	666.94	706	4.91	93.68	95.44
	OBF	3.54	644.08	706	4.96	93.50	94.12	3.87	647.70	706	4.95	92.87	92.87
	WT	3.61	648.86	706	4.94	93.34	94.44	3.64	654.10	706	4.94	93.02	93.45
	HP	2.82	687.30	706	4.98	94.36	94.74	4.18	639.49	706	4.97	92.44	92.48
	SD	3.05	564.50	576	4.57	92.49	93.94	3.90	563.21	576	4.48	92.77	92.77
1d) $\vec{p}_1 = \vec{p}_2 = (0.05, 0.15, 0.75, 0.05)$ , $D_G = 0.0$													
2.49	PO	3.67	663.42	706	4.90	93.18	95.24	3.35	668.57	706	4.91	93.31	95.30
	OBF	3.57	641.78	706	4.96	93.72	94.19	3.79	647.82	706	4.96	92.78	92.78
	WT	3.27	650.00	706	4.95	93.77	94.74	4.09	649.44	706	4.93	92.29	92.89
	HP	3.04	685.64	706	4.98	94.03	94.32	3.87	639.82	706	4.97	92.73	92.76
	SD	3.29	563.34	576	4.58	92.14	93.78	3.98	563.43	576	4.46	92.84	92.84
1e) $\vec{p}_1 = \vec{p}_2 = (0.15, 0.75, 0.05, 0.05)$ , $D_G = 0.0$													
2.49	PO	3.76	663.17	706	4.91	93.19	95.41	4.06	665.17	706	4.89	92.70	94.56
	OBF	3.61	644.19	706	4.96	93.30	93.86	3.76	649.18	706	4.96	92.20	92.20
	WT	3.94	648.58	706	4.93	92.88	93.86	4.33	650.50	706	4.93	92.41	92.93
	HP	2.98	685.80	706	4.98	94.09	94.45	4.20	636.60	706	4.97	92.13	92.19
	SD	3.20	561.32	576	4.59	91.82	93.75	3.67	563.34	576	4.47	92.97	92.97

**Tabelle C.18:** Empirisches Niveau  $\hat{\alpha}\%$  im nicht-sequentiellen Design (NSD) sowie  $\hat{\alpha}\%$ , mittlerer und medianer Umfang (ASN = „Average Sample Number“,  $N_{med}$ ), mittlere Stufenanzahl  $\bar{k}$  und empirische Überdeckungswahrscheinlichkeiten  $(1 - 2\hat{\alpha})\%$  der Sequenz der beobachteten Intervalle  $\{KI_j\}$  sowie des finalen Intervalls  $KI_k$  in flexiblen Designs mit ordinaler Zielvariable unter der Nullhypothese (Konstellation V)

NSD	Design	Inverse Normalmethode						Inverse $\chi^2$ -Methode					
		$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$ $\{KI_j\}$ $KI_k$		$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$ $KI_k$ $\{KI_j\}$	
1a) $\vec{p}_1 = \vec{p}_2 = (0.25, 0.25, 0.25, 0.25)$ , $D_G = 0.0$													
2.66	PO	2.90	314.76	320	2.96	94.22	95.83	3.04	314.69	320	2.96	93.87	95.20
	OBF	2.88	315.22	320	2.99	94.22	94.52	3.16	315.19	320	2.98	93.39	93.39
	WT	3.01	315.03	320	2.98	93.96	94.79	2.85	315.25	320	2.97	93.84	94.18
	HP	2.72	317.62	320	2.99	94.42	94.64	3.08	316.06	320	2.99	93.88	93.88
	SD	2.63	217.79	216	2.05	94.33	94.47	2.64	217.18	216	2.04	94.23	94.23
1b) $\vec{p}_1 = \vec{p}_2 = (0.70, 0.10, 0.10, 0.10)$ , $D_G = 0.0$													
2.37	PO	3.08	314.55	320	2.96	94.10	95.68	2.93	315.05	320	2.96	94.15	95.49
	OBF	2.88	315.43	320	2.99	94.53	94.69	2.86	315.21	320	2.99	94.16	94.16
	WT	2.92	315.33	320	2.98	94.18	94.99	3.08	315.42	320	2.97	93.95	94.34
	HP	3.00	317.78	320	2.99	94.36	94.49	3.07	315.73	320	2.99	94.21	94.21
	SD	2.57	218.08	216	2.06	94.80	95.02	3.07	217.16	216	2.04	93.92	93.92
1c) $\vec{p}_1 = \vec{p}_2 = (0.10, 0.10, 0.10, 0.70)$ , $D_G = 0.0$													
2.69	PO	2.87	314.88	320	2.96	94.45	95.93	2.93	314.95	320	2.96	94.34	95.50
	OBF	2.74	315.40	320	2.99	94.39	94.70	3.08	315.16	320	2.98	93.98	93.98
	WT	2.64	315.48	320	2.98	94.63	95.28	2.92	315.41	320	2.98	94.38	94.71
	HP	2.78	317.97	320	3.00	94.44	94.59	2.90	316.15	320	3.00	94.15	94.15
	SD	2.61	217.82	216	2.05	94.41	94.57	2.76	217.07	216	2.04	94.67	94.67
1d) $\vec{p}_1 = \vec{p}_2 = (0.05, 0.15, 0.75, 0.05)$ , $D_G = 0.0$													
2.49	PO	3.08	314.90	320	2.96	94.12	95.65	3.00	315.12	320	2.96	94.16	95.40
	OBF	3.08	315.19	320	2.99	94.10	94.36	3.00	315.20	320	2.98	94.06	94.06
	WT	3.20	315.34	320	2.98	93.91	94.65	2.94	315.53	320	2.98	94.07	94.38
	HP	2.75	317.76	320	2.99	94.60	94.74	2.84	316.08	320	2.99	94.04	94.04
	SD	2.57	218.12	216	2.05	94.50	94.68	2.79	216.94	216	2.04	94.29	94.29
1e) $\vec{p}_1 = \vec{p}_2 = (0.15, 0.75, 0.05, 0.05)$ , $D_G = 0.0$													
2.49	PO	3.04	314.72	320	2.96	93.84	95.46	2.72	315.27	320	2.96	94.39	95.44
	OBF	3.23	315.31	320	2.99	93.95	94.33	3.27	315.15	320	2.98	93.43	93.43
	WT	2.99	315.76	320	2.98	93.82	94.73	3.16	315.66	320	2.98	93.32	93.70
	HP	3.20	317.63	320	2.99	93.87	94.09	3.06	316.13	320	2.99	93.93	93.93
	SD	2.79	218.02	216	2.05	94.38	94.51	2.76	217.27	216	2.04	94.09	94.09

**Tabelle C.19:** Empirisches Niveau  $\hat{\alpha}\%$  im nicht-sequentiellen Design (NSD) sowie  $\hat{\alpha}\%$ , mittlerer und medianer Umfang (ASN = „Average Sample Number“,  $N_{med}$ ), mittlere Stufenanzahl  $\bar{k}$  und empirische Überdeckungswahrscheinlichkeiten  $(1 - 2\hat{\alpha})\%$  der Sequenz der beobachteten Intervalle  $\{KI_j\}$  sowie des finalen Intervalls  $KI_k$  in flexiblen Designs mit ordinaler Zielvariable unter der Nullhypothese (Konstellation VI)

NSD	Design	Inverse Normalmethode						Inverse $\chi^2$ -Methode					
		$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$ $\{KI_j\}$ $KI_k$		$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$ $KI_k$ $\{KI_j\}$	
1a) $\vec{p}_1 = \vec{p}_2 = (0.25, 0.25, 0.25, 0.25)$ , $D_G = 0.0$													
2.66	PO	3.00	503.74	516	2.96	94.15	95.93	2.98	504.78	516	2.96	94.15	95.49
	OBF	2.92	502.73	516	2.99	94.36	94.63	3.03	503.67	516	2.98	94.11	94.11
	WT	2.96	503.13	516	2.98	94.18	95.02	3.15	503.95	516	2.97	94.18	94.47
	HP	2.91	509.45	516	2.99	94.41	94.55	3.01	505.79	516	2.99	93.95	93.95
	SD	2.68	504.65	516	3.02	93.86	94.61	2.71	505.21	516	3.00	94.44	94.44
1b) $\vec{p}_1 = \vec{p}_2 = (0.70, 0.10, 0.10, 0.10)$ , $D_G = 0.0$													
2.37	PO	3.01	504.18	516	2.96	94.35	95.89	2.91	505.22	516	2.96	94.48	95.75
	OBF	2.75	502.05	516	2.99	94.53	94.77	2.87	503.52	516	2.98	94.36	94.36
	WT	2.78	503.31	516	2.98	94.57	95.26	2.86	504.69	516	2.97	94.42	94.91
	HP	2.62	510.08	516	3.00	94.88	95.05	2.75	505.98	516	3.00	94.30	94.30
	SD	2.59	505.04	516	3.02	93.88	94.82	2.75	505.13	516	3.00	94.73	94.73
1c) $\vec{p}_1 = \vec{p}_2 = (0.10, 0.10, 0.10, 0.70)$ , $D_G = 0.0$													
2.69	PO	2.83	504.11	516	2.96	94.57	96.04	2.93	504.96	516	2.96	94.36	95.60
	OBF	2.69	502.94	516	2.99	94.72	95.01	2.98	504.22	516	2.98	94.24	94.24
	WT	2.72	504.14	516	2.98	94.67	95.36	2.79	504.73	516	2.98	94.47	94.90
	HP	2.56	510.35	516	3.00	94.88	95.06	2.86	506.31	516	3.00	94.39	94.39
	SD	2.58	505.30	516	3.02	94.09	94.86	2.52	505.57	516	3.00	94.73	94.73
1d) $\vec{p}_1 = \vec{p}_2 = (0.05, 0.15, 0.75, 0.05)$ , $D_G = 0.0$													
2.49	PO	2.89	504.47	516	2.96	94.34	95.75	2.83	505.44	516	2.96	94.38	95.70
	OBF	2.98	502.36	516	2.99	94.21	94.53	2.84	504.39	516	2.99	94.23	94.23
	WT	3.06	503.57	516	2.98	94.08	94.85	2.91	504.86	516	2.98	94.29	94.71
	HP	2.75	509.50	516	2.99	94.76	94.87	2.69	506.15	516	2.99	94.35	94.35
	SD	2.68	504.96	516	3.01	93.83	94.61	2.65	505.10	516	3.00	94.60	94.60
1e) $\vec{p}_1 = \vec{p}_2 = (0.15, 0.75, 0.05, 0.05)$ , $D_G = 0.0$													
2.49	PO	3.02	503.62	516	2.96	93.92	95.59	2.85	505.43	516	2.96	94.31	95.71
	OBF	2.95	502.64	516	2.99	94.23	94.53	3.30	504.38	516	2.98	93.93	93.93
	WT	2.81	503.49	516	2.98	94.23	95.14	3.16	504.46	516	2.97	93.98	94.42
	HP	2.71	509.58	516	2.99	94.18	94.30	2.98	505.56	516	2.99	94.04	94.04
	SD	2.62	505.01	516	3.02	93.84	94.66	2.86	505.11	516	3.00	94.09	94.09

**Tabelle C.20:** Empirisches Niveau  $\hat{\alpha}\%$  im nicht-sequentiellen Design (NSD) sowie  $\hat{\alpha}\%$ , mittlerer und medianer Umfang (ASN = „Average Sample Number“,  $N_{med}$ ), mittlere Stufenanzahl  $\bar{k}$  und empirische Überdeckungswahrscheinlichkeiten  $(1 - 2\hat{\alpha})\%$  der Sequenz der beobachteten Intervalle  $\{KI_j\}$  sowie des finalen Intervalls  $KI_k$  in flexiblen Designs mit ordinaler Zielvariable unter der Nullhypothese (Konstellation VII)

NSD	Design	Inverse Normalmethode						Inverse $\chi^2$ -Methode					
		$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$ $\{KI_j\}$ $KI_k$		$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$ $KI_k$ $\{KI_j\}$	
1a) $\vec{p}_1 = \vec{p}_2 = (0.25, 0.25, 0.25, 0.25)$ , $D_G = 0.0$													
2.66	PO	2.94	307.61	320	2.96	93.87	95.60	3.25	309.03	320	2.96	93.32	94.92
	OBF	3.45	302.91	320	2.98	93.32	93.71	3.45	306.34	320	2.98	93.14	93.14
	WT	3.33	304.24	320	2.98	93.64	94.46	3.05	307.24	320	2.97	93.52	93.84
	HP	2.94	314.44	320	2.99	94.52	94.67	3.38	308.75	320	2.99	93.25	93.25
	SD	3.16	222.12	216	2.29	93.30	93.71	3.44	222.14	216	2.23	93.41	93.41
1b) $\vec{p}_1 = \vec{p}_2 = (0.70, 0.10, 0.10, 0.10)$ , $D_G = 0.0$													
2.37	PO	3.08	307.39	320	2.96	93.80	95.52	3.37	309.13	320	2.96	93.52	94.99
	OBF	3.00	303.80	320	2.99	93.92	94.31	3.16	306.94	320	2.98	93.61	93.61
	WT	3.07	305.24	320	2.98	94.07	94.90	3.17	307.41	320	2.97	93.73	94.09
	HP	3.05	314.44	320	2.99	94.27	94.35	3.35	308.57	320	2.99	93.94	93.94
	SD	3.00	222.56	216	2.30	93.79	94.02	3.06	221.47	216	2.21	93.82	93.82
1c) $\vec{p}_1 = \vec{p}_2 = (0.10, 0.10, 0.10, 0.70)$ , $D_G = 0.0$													
2.69	PO	3.02	307.45	320	2.96	94.24	95.80	3.00	309.67	320	2.96	94.29	95.48
	OBF	3.27	303.77	320	2.99	93.88	94.19	3.19	307.14	320	2.98	93.80	93.80
	WT	3.33	305.02	320	2.98	93.93	94.62	3.51	306.92	320	2.97	93.42	94.00
	HP	2.61	314.87	320	3.00	94.56	94.71	3.26	309.59	320	2.99	93.74	93.74
	SD	2.94	221.96	216	2.29	93.93	94.32	3.15	221.59	216	2.22	93.78	93.78
1d) $\vec{p}_1 = \vec{p}_2 = (0.05, 0.15, 0.75, 0.05)$ , $D_G = 0.0$													
2.49	PO	3.39	306.70	320	2.96	93.55	95.26	3.29	308.92	320	2.96	93.60	94.88
	OBF	3.03	303.85	320	2.99	93.99	94.41	3.05	306.82	320	2.98	93.56	93.56
	WT	3.06	304.66	320	2.98	93.98	94.81	3.42	307.25	320	2.97	93.38	93.93
	HP	2.74	314.83	320	2.99	94.54	94.69	3.16	309.13	320	2.99	93.72	93.72
	SD	2.60	222.24	216	2.29	94.08	94.51	3.51	221.77	216	2.23	93.38	93.38
1e) $\vec{p}_1 = \vec{p}_2 = (0.15, 0.75, 0.05, 0.05)$ , $D_G = 0.0$													
2.49	PO	3.36	307.50	320	2.96	93.63	95.34	3.48	308.74	320	2.95	93.30	94.77
	OBF	3.50	303.43	320	2.99	93.36	93.76	3.89	305.73	320	2.98	93.09	93.09
	WT	3.37	305.37	320	2.98	93.68	94.54	3.44	307.23	320	2.97	93.26	93.70
	HP	2.83	314.64	320	2.99	94.02	94.19	3.38	309.09	320	2.99	93.51	93.51
	SD	3.11	222.26	216	2.30	93.71	94.18	3.36	221.63	216	2.23	93.51	93.51

**Tabelle C.21:** Empirisches Niveau  $\hat{\alpha}\%$  im nicht-sequentiellen Design (NSD) sowie  $\hat{\alpha}\%$ , mittlerer und medianer Umfang (ASN = „Average Sample Number“,  $N_{med}$ ), mittlere Stufenanzahl  $\bar{k}$  und empirische Überdeckungswahrscheinlichkeiten  $(1 - 2\hat{\alpha})\%$  der Sequenz der beobachteten Intervalle  $\{KI_j\}$  sowie des finalen Intervalls  $KI_k$  in flexiblen Designs mit ordinaler Zielvariable unter der Nullhypothese (Konstellation VIII)

NSD	Design	Inverse Normalmethode						Inverse $\chi^2$ -Methode					
		$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$ $\{KI_j\}$ $KI_k$		$\hat{\alpha}\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$ $KI_k$ $\{KI_j\}$	
1a) $\vec{p}_1 = \vec{p}_2 = (0.25, 0.25, 0.25, 0.25)$ , $D_G = 0.0$													
2.66	PO	3.21	486.64	516	2.96	93.93	95.57	2.98	491.38	516	2.96	94.08	95.48
	OBF	3.04	478.30	516	2.99	94.12	94.40	3.05	483.19	516	2.98	93.82	93.82
	WT	3.09	480.89	516	2.98	94.01	94.79	3.31	484.88	516	2.97	93.84	94.15
	HP	3.03	500.73	516	2.99	94.21	94.34	3.24	488.39	516	2.99	93.81	93.81
	SD	3.30	482.22	516	3.13	92.62	93.50	3.57	488.26	516	3.09	93.12	93.12
1b) $\vec{p}_1 = \vec{p}_2 = (0.70, 0.10, 0.10, 0.10)$ , $D_G = 0.0$													
2.37	PO	3.24	487.34	516	2.96	93.81	95.58	3.10	492.04	516	2.96	94.00	95.44
	OBF	3.07	477.42	516	2.99	94.05	94.39	3.23	485.12	516	2.98	94.01	94.01
	WT	3.08	481.10	516	2.98	94.10	94.87	3.46	486.08	516	2.97	93.53	93.97
	HP	2.71	500.91	516	2.99	94.68	94.79	2.82	490.63	516	2.99	94.34	94.34
	SD	3.04	483.68	516	3.14	93.31	94.16	3.07	489.21	516	3.09	94.39	94.39
1c) $\vec{p}_1 = \vec{p}_2 = (0.10, 0.10, 0.10, 0.70)$ , $D_G = 0.0$													
2.69	PO	3.00	487.91	516	2.96	94.27	95.81	3.01	492.06	516	2.96	94.34	95.58
	OBF	3.20	479.44	516	2.99	94.08	94.36	3.19	486.54	516	2.98	93.87	93.87
	WT	3.31	482.84	516	2.98	93.90	94.66	3.38	486.52	516	2.97	93.75	94.27
	HP	2.51	501.60	516	3.00	94.86	95.01	3.24	490.93	516	2.99	93.77	93.77
	SD	3.01	484.21	516	3.13	93.41	94.18	3.44	488.95	516	3.09	93.61	93.61
1d) $\vec{p}_1 = \vec{p}_2 = (0.05, 0.15, 0.75, 0.05)$ , $D_G = 0.0$													
2.49	PO	3.26	486.91	516	2.96	93.75	95.42	3.06	490.98	516	2.96	93.99	95.34
	OBF	2.90	480.05	516	2.99	94.23	94.61	3.22	484.90	516	2.98	93.63	93.63
	WT	3.03	481.70	516	2.98	94.04	94.97	3.26	486.13	516	2.97	93.59	94.14
	HP	2.86	501.30	516	2.99	94.64	94.77	3.06	490.03	516	2.99	93.86	93.86
	SD	3.14	482.42	516	3.14	93.02	94.14	3.27	487.97	516	3.09	93.36	93.36
1e) $\vec{p}_1 = \vec{p}_2 = (0.15, 0.75, 0.05, 0.05)$ , $D_G = 0.0$													
2.49	PO	3.17	487.63	516	2.96	94.04	95.73	3.19	491.46	516	2.96	93.79	95.32
	OBF	3.24	479.35	516	2.99	93.62	93.96	3.41	485.41	516	2.98	93.50	93.50
	WT	3.25	483.14	516	2.98	93.72	94.59	3.39	486.12	516	2.97	93.67	94.08
	HP	2.90	501.67	516	2.99	94.19	94.38	3.05	489.61	516	2.99	93.76	93.76
	SD	2.89	482.88	516	3.13	93.36	94.32	3.43	488.02	516	3.10	93.69	93.69

## 2) Unter der Alternativhypothese

**Tabelle C.22:** Empirische Power  $(1 - \hat{\beta})\%$  im nicht-sequentiellen Design (NSD) sowie  $(1 - \hat{\beta})\%$ , mittlerer und medianer Umfang (ASN = „Average Sample Number“,  $N_{med}$ ), mittlere Stufenanzahl  $\bar{k}$  und empirische Überdeckungswahrscheinlichkeiten  $(1 - 2\hat{\alpha})\%$  der Sequenz der beobachteten Intervalle  $\{KI_j\}$  sowie des finalen Intervalls  $KI_k$  in flexiblen Designs mit ordinaler Zielvariable unter der Alternativhypothese (Konstellation II)

NSD	Design	Inverse Normalmethode					Inverse $\chi^2$ -Methode						
		$(1 - \hat{\beta})\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$	$(1 - \hat{\beta})\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$		
		$\{KI_j\}$					$KI_k$						
2a) $\vec{p}_1 = (0.58, 0.30, 0.10, 0.02)$ , $\vec{p}_2 = (0.46, 0.28, 0.16, 0.10)$ , $D_G = 0.171$													
75.27	PO	96.97	304.14	290	2.88	95.75	97.48	95.42	319.37	308	2.93	95.13	96.68
	OBF	96.31	323.94	300	3.79	95.66	96.25	95.40	322.09	300	3.73	92.38	92.38
	WT	97.72	309.48	286	3.27	96.77	97.70	96.75	318.96	294	3.32	95.55	95.92
	HP	97.26	367.43	386	3.56	96.35	96.56	96.84	333.49	318	4.10	88.83	88.94
	SD	86.46	301.90	260	4.21	85.19	86.34	92.93	294.99	251	3.41	91.75	91.75
2b) $\vec{p}_1 = (0.22, 0.40, 0.34, 0.04)$ , $\vec{p}_2 = (0.15, 0.35, 0.28, 0.22)$ , $D_G = 0.201$													
83.47	PO	99.20	251.45	188	2.55	96.09	97.54	98.48	262.04	188	2.58	95.46	96.70
	OBF	97.62	273.76	244	3.55	94.76	95.24	97.99	266.94	226	3.45	92.24	92.24
	WT	99.33	260.10	232	3.00	96.74	97.62	99.00	265.07	222	3.02	95.25	95.58
	HP	99.09	314.36	318	3.19	96.47	96.69	98.85	276.56	252	3.82	88.78	88.96
	SD	89.78	261.75	214	4.13	83.31	84.38	95.61	253.05	194	3.26	90.62	90.62
2c) $\vec{p}_1 = (0.18, 0.44, 0.32, 0.06)$ , $\vec{p}_2 = (0.10, 0.35, 0.40, 0.15)$ , $D_G = 0.213$													
89.10	PO	99.77	221.47	188	2.36	95.93	97.61	99.55	232.51	188	2.39	95.72	96.93
	OBF	98.38	245.61	220	3.42	94.92	95.46	98.69	236.15	194	3.30	92.26	92.26
	WT	99.77	229.98	188	2.82	96.91	97.70	99.39	234.39	188	2.85	95.23	95.50
	HP	99.17	281.29	296	2.93	96.48	96.70	99.32	248.72	220	3.64	89.42	89.56
	SD	90.76	234.48	200	4.08	82.28	83.69	96.63	224.87	188	3.11	91.10	91.10
2d) $\vec{p}_1 = (0.45, 0.30, 0.20, 0.05)$ , $\vec{p}_2 = (0.25, 0.34, 0.26, 0.15)$ , $D_G = 0.255$													
96.57	PO	100.00	169.71	180	1.99	96.35	97.89	99.99	172.81	188	1.99	95.82	96.85
	OBF	99.08	186.68	182	3.11	94.75	95.07	99.18	180.25	182	2.96	91.06	91.06
	WT	99.91	177.32	174	2.47	97.56	98.32	99.73	178.29	184	2.49	95.73	95.98
	HP	99.87	221.18	188	2.45	97.92	98.15	99.76	187.54	188	3.24	89.17	89.34
	SD	92.35	182.89	174	3.89	78.52	79.71	97.43	171.43	168	2.82	89.49	89.49
2e) $\vec{p}_1 = (0.52, 0.36, 0.08, 0.04)$ , $\vec{p}_2 = (0.30, 0.30, 0.20, 0.20)$ , $D_G = 0.336$													
99.94	PO	100.00	107.75	58	1.50	96.55	97.35	100.00	108.17	58	1.48	95.97	96.78
	OBF	99.70	120.91	98	2.64	91.94	92.24	99.82	111.77	84	2.32	89.72	89.72
	WT	100.00	113.63	92	1.93	97.87	98.17	99.99	110.11	82	1.89	95.87	95.93
	HP	100.00	135.96	126	1.78	98.84	99.06	99.97	113.43	84	2.44	88.92	89.33
	SD	94.08	122.74	100	3.47	74.39	75.52	98.17	110.05	84	2.35	87.57	87.57

**Tabelle C.23:** Empirische Power  $(1 - \hat{\beta})\%$  im nicht-sequentiellen Design (NSD) sowie  $(1 - \hat{\beta})\%$ , mittlerer und medianer Umfang (ASN = „Average Sample Number“,  $N_{med}$ ), mittlere Stufenanzahl  $\bar{k}$  und empirische Überdeckungswahrscheinlichkeiten  $(1 - 2\hat{\alpha})\%$  der Sequenz der beobachteten Intervalle  $\{KI_j\}$  sowie des finalen Intervalls  $KI_k$  in flexiblen Designs mit ordinaler Zielvariable unter der Alternativhypothese (Konstellation III)

NSD	Design	Inverse Normalmethode						Inverse $\chi^2$ -Methode					
		$(1 - \hat{\beta})\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$	$(1 - \hat{\beta})\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$		
						$\{KI_j\}$					$KI_k$	$\{KI_j\}$	
2a) $\vec{p}_1 = (0.58, 0.30, 0.10, 0.02)$ , $\vec{p}_2 = (0.46, 0.28, 0.16, 0.10)$ , $D_G = 0.171$													
75.27	PO	67.32	198.19	196	3.66	92.68	94.62	64.89	206.37	210	3.65	92.04	93.85
	OBF	65.25	192.95	190	4.35	86.52	87.10	68.74	194.79	188	4.21	84.09	84.09
	WT	65.81	192.84	190	4.08	88.39	89.61	68.37	195.18	190	3.98	85.38	86.42
	HP	79.80	232.98	248	4.27	94.35	94.58	70.82	197.04	198	4.36	84.06	84.12
	SD	57.50	165.29	190	4.58	77.29	79.15	61.70	161.10	194	3.95	82.74	82.74
2b) $\vec{p}_1 = (0.22, 0.40, 0.34, 0.04)$ , $\vec{p}_2 = (0.15, 0.35, 0.28, 0.22)$ , $D_G = 0.201$													
83.47	PO	77.08	174.51	162	3.35	91.60	93.80	75.30	184.95	182	3.37	91.04	92.56
	OBF	73.47	173.85	166	4.16	85.16	85.90	77.96	173.95	164	3.99	83.47	83.47
	WT	74.81	173.26	166	3.85	87.00	88.57	77.69	173.10	164	3.72	83.93	84.58
	HP	87.71	208.09	210	4.02	93.59	93.73	80.18	176.53	172	4.11	82.66	82.70
	SD	66.85	153.40	162	4.62	75.05	76.91	72.53	148.83	152	3.83	80.84	80.84
2c) $\vec{p}_1 = (0.18, 0.44, 0.32, 0.06)$ , $\vec{p}_2 = (0.10, 0.35, 0.40, 0.15)$ , $D_G = 0.213$													
89.10	PO	82.87	161.40	134	3.17	91.84	93.76	81.22	169.26	152	3.17	91.46	92.98
	OBF	77.96	162.94	160	4.07	84.33	85.06	82.87	159.77	154	3.84	83.30	83.30
	WT	79.56	160.83	148	3.71	86.45	87.98	81.90	161.04	152	3.58	85.13	85.67
	HP	91.01	192.75	190	3.84	93.86	94.08	85.63	162.88	154	3.97	81.85	81.92
	SD	71.56	147.44	140	4.69	74.77	76.46	78.06	141.40	124	3.82	80.63	80.63
2d) $\vec{p}_1 = (0.45, 0.30, 0.20, 0.05)$ , $\vec{p}_2 = (0.25, 0.34, 0.26, 0.15)$ , $D_G = 0.255$													
96.57	PO	91.15	131.40	106	2.74	92.00	93.94	91.60	135.80	120	2.69	91.48	92.81
	OBF	87.20	136.88	124	3.81	81.40	82.43	91.22	133.62	124	3.56	80.83	80.83
	WT	89.03	133.97	116	3.36	85.08	86.72	91.05	132.33	122	3.22	83.95	84.41
	HP	95.87	156.72	144	3.39	92.71	92.97	93.26	134.64	114	3.66	79.18	79.32
	SD	82.57	129.44	110	4.74	68.97	70.51	89.11	121.89	98	3.71	75.83	75.83
2e) $\vec{p}_1 = (0.52, 0.36, 0.08, 0.04)$ , $\vec{p}_2 = (0.30, 0.30, 0.20, 0.20)$ , $D_G = 0.336$													
99.94	PO	98.79	86.73	58	1.89	92.57	94.29	98.99	88.88	58	1.87	92.38	93.55
	OBF	96.49	98.24	71	3.21	78.69	79.52	98.69	92.76	70	2.80	79.79	79.79
	WT	97.44	92.18	70	2.56	84.94	86.24	98.83	90.14	64	2.39	84.64	84.85
	HP	99.24	102.97	86	2.41	93.97	94.13	99.11	92.25	70	2.80	78.93	79.34
	SD	92.37	98.59	82	4.39	62.76	64.17	97.20	88.82	70	3.12	71.82	71.82

**Tabelle C.24:** Empirische Power  $(1 - \hat{\beta})\%$  im nicht-sequentiellen Design (NSD) sowie  $(1 - \hat{\beta})\%$ , mittlerer und medianer Umfang (ASN = „Average Sample Number“,  $N_{med}$ ), mittlere Stufenanzahl  $\bar{k}$  und empirische Überdeckungswahrscheinlichkeiten  $(1 - 2\hat{\alpha})\%$  der Sequenz der beobachteten Intervalle  $\{KI_j\}$  sowie des finalen Intervalls  $KI_k$  in flexiblen Designs mit ordinaler Zielvariable unter der Alternativhypothese (Konstellation IV)

NSD	Design	Inverse Normalmethode					Inverse $\chi^2$ -Methode						
		$(1 - \hat{\beta})\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$	$(1 - \hat{\beta})\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$		
	$(1 - \hat{\beta})\%$				$\{KI_j\}$	$KI_k$					$KI_k$	$\{KI_j\}$	
2a) $\vec{p}_1 = (0.58, 0.30, 0.10, 0.02)$ , $\vec{p}_2 = (0.46, 0.28, 0.16, 0.10)$ , $D_G = 0.171$													
75.27	PO	87.44	296.26	240	3.47	92.67	94.62	88.79	304.24	260	3.40	92.24	93.82
	OBF	82.98	291.97	264	4.26	85.38	86.04	88.47	286.94	246	4.09	83.56	83.56
	WT	84.50	291.12	260	3.98	87.99	89.14	89.16	288.83	246	3.82	85.57	86.48
	HP	93.55	322.42	303	4.07	93.93	94.02	91.00	289.94	254	4.29	83.12	83.19
	SD	82.85	285.77	226	5.76	72.15	74.36	90.75	271.67	205	4.86	78.04	78.04
2b) $\vec{p}_1 = (0.22, 0.40, 0.34, 0.04)$ , $\vec{p}_2 = (0.15, 0.35, 0.28, 0.22)$ , $D_G = 0.201$													
83.47	PO	92.22	251.39	188	3.21	91.43	93.60	94.34	250.48	188	3.08	91.58	93.14
	OBF	86.81	252.20	202	4.09	84.24	85.20	93.13	239.86	194	3.85	83.25	83.25
	WT	89.31	249.15	200	3.74	86.13	87.46	93.56	244.22	194	3.58	83.81	84.46
	HP	95.34	269.24	244	3.78	93.77	93.98	94.36	237.83	198	4.01	81.62	81.75
	SD	86.46	245.77	200	5.68	69.71	72.25	94.12	224.19	188	4.56	76.70	76.70
2c) $\vec{p}_1 = (0.18, 0.44, 0.32, 0.06)$ , $\vec{p}_2 = (0.10, 0.35, 0.40, 0.15)$ , $D_G = 0.213$													
89.10	PO	94.92	219.64	187	3.01	92.00	94.02	96.48	224.29	188	2.91	91.81	93.30
	OBF	90.54	230.25	194	4.00	84.25	85.07	94.98	216.57	188	3.73	83.23	83.23
	WT	92.02	226.35	188	3.61	87.16	88.72	95.35	219.03	188	3.43	85.56	86.18
	HP	96.43	238.18	212	3.59	93.69	93.93	96.06	216.74	188	3.87	82.66	82.74
	SD	88.98	217.21	168	5.60	68.62	70.89	95.93	202.12	160	4.46	76.51	76.51
2d) $\vec{p}_1 = (0.45, 0.30, 0.20, 0.05)$ , $\vec{p}_2 = (0.25, 0.34, 0.26, 0.15)$ , $D_G = 0.255$													
96.57	PO	98.21	162.62	104	2.56	92.05	93.89	98.70	167.76	122	2.50	92.02	93.27
	OBF	94.57	179.68	128	3.74	80.67	81.65	97.83	168.89	126	3.44	80.93	80.93
	WT	96.32	173.34	116	3.26	84.42	86.11	97.82	167.11	122	3.10	84.87	85.33
	HP	97.74	181.79	168	3.15	92.90	93.13	98.23	164.26	116	3.56	79.06	79.21
	SD	92.17	170.62	114	5.40	63.96	66.07	97.09	152.49	100	4.13	71.65	71.65
2e) $\vec{p}_1 = (0.52, 0.36, 0.08, 0.04)$ , $\vec{p}_2 = (0.30, 0.30, 0.20, 0.20)$ , $D_G = 0.336$													
99.94	PO	99.71	95.34	58	1.82	92.86	94.47	99.89	98.50	58	1.80	92.57	93.41
	OBF	98.70	113.15	72	3.17	79.30	80.16	99.62	105.80	70	2.74	80.00	80.00
	WT	99.22	106.34	70	2.51	85.52	86.84	99.71	102.94	64	2.36	84.84	84.96
	HP	99.36	110.71	84	2.28	94.68	94.95	99.70	100.79	70	2.76	79.42	79.76
	SD	94.19	108.94	82	4.62	60.01	62.00	97.68	95.02	70	3.22	70.44	70.44

**Tabelle C.25:** Empirische Power  $(1 - \hat{\beta})\%$  im nicht-sequentiellen Design (NSD) sowie  $(1 - \hat{\beta})\%$ , mittlerer und medianer Umfang (ASN = „Average Sample Number“,  $N_{med}$ ), mittlere Stufenanzahl  $\bar{k}$  und empirische Überdeckungswahrscheinlichkeiten  $(1 - 2\hat{\alpha})\%$  der Sequenz der beobachteten Intervalle  $\{KI_j\}$  sowie des finalen Intervalls  $KI_k$  in flexiblen Designs mit ordinaler Zielvariable unter der Alternativhypothese (Konstellation V)

NSD	Design	Inverse Normalmethode					Inverse $\chi^2$ -Methode						
		$(1 - \hat{\beta})\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$	$(1 - \hat{\beta})\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$		
						$\{KI_j\}$	$KI_k$				$KI_k$	$\{KI_j\}$	
2a) $\vec{p}_1 = (0.58, 0.30, 0.10, 0.02)$ , $\vec{p}_2 = (0.46, 0.28, 0.16, 0.10)$ , $D_G = 0.171$													
75.27	PO	77.42	221.98	190	2.23	95.51	97.00	73.06	223.34	190	2.22	94.94	96.21
	OBF	78.88	232.14	248	2.61	95.96	96.29	77.33	230.66	250	2.51	94.65	94.65
	WT	79.51	228.65	242	2.41	96.69	97.32	76.36	229.68	252	2.37	95.59	95.87
	HP	80.98	246.31	262	2.65	96.91	96.93	77.09	235.12	260	2.72	92.53	92.55
	SD	64.61	203.59	216	2.61	90.75	91.17	65.95	197.13	216	2.25	92.84	92.84
2b) $\vec{p}_1 = (0.22, 0.40, 0.34, 0.04)$ , $\vec{p}_2 = (0.15, 0.35, 0.28, 0.22)$ , $D_G = 0.201$													
83.47	PO	86.40	201.26	190	2.06	95.82	97.35	83.20	204.42	190	2.07	95.32	96.56
	OBF	86.08	212.24	190	2.50	95.33	95.63	85.50	211.07	190	2.37	94.46	94.46
	WT	87.53	208.71	190	2.27	96.59	97.07	85.79	209.90	190	2.21	95.88	96.26
	HP	88.13	228.23	220	2.52	96.55	96.64	86.18	215.92	196	2.61	91.58	91.59
	SD	74.01	197.37	216	2.70	89.61	90.11	75.21	187.48	216	2.24	92.58	92.58
2c) $\vec{p}_1 = (0.18, 0.44, 0.32, 0.06)$ , $\vec{p}_2 = (0.10, 0.35, 0.40, 0.15)$ , $D_G = 0.213$													
89.10	PO	90.67	189.41	190	1.97	95.71	97.01	87.86	191.93	190	1.96	95.24	96.34
	OBF	90.19	199.49	190	2.42	95.41	95.69	89.53	198.76	190	2.29	94.36	94.36
	WT	91.32	195.99	190	2.17	95.95	96.67	90.08	195.69	190	2.10	95.95	96.29
	HP	91.87	215.03	200	2.42	96.57	96.70	90.65	201.49	190	2.52	90.96	90.96
	SD	78.58	191.55	216	2.77	88.54	88.99	80.86	181.08	216	2.24	92.29	92.29
2d) $\vec{p}_1 = (0.45, 0.30, 0.20, 0.05)$ , $\vec{p}_2 = (0.25, 0.34, 0.26, 0.15)$ , $D_G = 0.255$													
96.57	PO	97.54	155.75	140	1.68	96.02	97.22	96.55	158.85	86	1.68	95.59	96.53
	OBF	95.88	167.35	150	2.21	94.92	95.16	96.51	164.40	148	2.02	95.14	95.14
	WT	97.62	162.14	144	1.90	96.77	97.36	97.12	162.24	142	1.83	96.05	96.38
	HP	97.29	182.22	190	2.13	97.01	97.09	97.17	168.78	158	2.29	92.10	92.10
	SD	87.69	173.85	178	2.93	85.96	86.53	91.44	161.37	160	2.17	91.61	91.61
2e) $\vec{p}_1 = (0.52, 0.36, 0.08, 0.04)$ , $\vec{p}_2 = (0.30, 0.30, 0.20, 0.20)$ , $D_G = 0.336$													
99.94	PO	99.98	109.55	86	1.26	96.40	97.50	99.91	110.05	86	1.24	96.13	96.91
	OBF	99.17	120.31	96	1.79	94.88	95.10	99.65	114.50	86	1.52	95.56	95.56
	WT	99.91	113.99	86	1.43	97.60	98.21	99.95	113.62	86	1.36	96.91	97.14
	HP	99.69	129.97	86	1.56	97.51	97.56	99.71	116.81	86	1.72	93.15	93.18
	SD	92.97	134.02	116	2.99	78.28	78.95	98.35	120.25	92	1.82	91.69	91.69

**Tabelle C.26:** Empirische Power  $(1 - \hat{\beta})\%$  im nicht-sequentiellen Design (NSD) sowie  $(1 - \hat{\beta})\%$ , mittlerer und medianer Umfang (ASN = „Average Sample Number“,  $N_{med}$ ), mittlere Stufenanzahl  $\bar{k}$  und empirische Überdeckungswahrscheinlichkeiten  $(1 - 2\hat{\alpha})\%$  der Sequenz der beobachteten Intervalle  $\{KI_j\}$  sowie des finalen Intervalls  $KI_k$  in flexiblen Designs mit ordinaler Zielvariable unter der Alternativhypothese (Konstellation VI)

NSD	Design	Inverse Normalmethode						Inverse $\chi^2$ -Methode					
		$(1 - \hat{\beta})\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$	$(1 - \hat{\beta})\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$		
$(1 - \hat{\beta})\%$						$\{KI_j\}$	$KI_k$				$KI_k$	$\{KI_j\}$	
2a) $\vec{p}_1 = (0.58, 0.30, 0.10, 0.02)$ , $\vec{p}_2 = (0.46, 0.28, 0.16, 0.10)$ , $D_G = 0.171$													
75.27	PO	82.24	279.62	264	2.31	93.66	95.16	90.05	303.26	258	2.14	95.23	96.58
	OBF	80.45	278.05	280	2.70	91.01	91.38	90.66	302.50	258	2.42	94.77	94.77
	WT	81.95	276.94	273	2.51	92.11	93.00	91.45	302.52	258	2.28	95.57	95.98
	HP	86.99	294.01	274	2.69	95.34	95.51	91.51	306.18	264	2.67	91.64	91.65
	SD	79.39	283.95	274	3.64	82.77	83.89	90.37	305.88	258	2.64	94.13	94.13
2b) $\vec{p}_1 = (0.22, 0.40, 0.34, 0.04)$ , $\vec{p}_2 = (0.15, 0.35, 0.28, 0.22)$ , $D_G = 0.201$													
83.47	PO	89.44	246.81	244	2.17	93.45	95.04	95.63	262.82	258	1.96	95.22	96.33
	OBF	87.05	248.29	238	2.61	90.14	90.63	95.28	265.88	258	2.29	94.51	94.51
	WT	88.61	250.55	249	2.41	91.69	92.51	96.07	262.86	258	2.11	95.85	96.14
	HP	91.83	260.43	254	2.57	95.37	95.48	95.87	269.37	258	2.54	90.55	90.58
	SD	83.99	253.75	238	3.65	79.85	81.52	94.98	267.75	258	2.51	93.51	93.51
2c) $\vec{p}_1 = (0.18, 0.44, 0.32, 0.06)$ , $\vec{p}_2 = (0.10, 0.35, 0.40, 0.15)$ , $D_G = 0.213$													
89.10	PO	93.19	227.25	188	2.08	93.87	95.52	97.78	238.41	252	1.85	95.28	96.33
	OBF	89.33	231.46	206	2.55	90.32	90.76	96.85	238.51	236	2.19	94.66	94.66
	WT	91.63	230.03	206	2.32	92.15	93.00	97.69	239.45	244	2.01	95.83	96.18
	HP	93.54	236.82	210	2.49	95.05	95.13	97.15	246.47	258	2.45	90.78	90.84
	SD	86.96	234.24	204	3.66	79.34	80.57	96.14	243.63	246	2.46	93.11	93.11
2d) $\vec{p}_1 = (0.45, 0.30, 0.20, 0.05)$ , $\vec{p}_2 = (0.25, 0.34, 0.26, 0.15)$ , $D_G = 0.255$													
96.57	PO	97.40	179.19	92	1.81	93.57	95.10	99.65	184.27	86	1.60	95.87	96.67
	OBF	94.53	190.70	140	2.38	88.86	89.37	98.75	187.02	150	1.96	95.41	95.41
	WT	96.13	185.01	108	2.08	91.23	92.10	99.59	188.28	148	1.77	96.33	96.63
	HP	95.88	185.77	146	2.25	95.28	95.37	99.08	190.68	156	2.21	92.47	92.49
	SD	89.96	194.72	144	3.63	75.35	77.26	97.90	190.99	158	2.25	92.61	92.61
2e) $\vec{p}_1 = (0.52, 0.36, 0.08, 0.04)$ , $\vec{p}_2 = (0.30, 0.30, 0.20, 0.20)$ , $D_G = 0.336$													
99.94	PO	99.80	116.49	86	1.35	94.21	95.33	100.00	117.59	86	1.23	96.12	96.91
	OBF	98.49	130.39	92	1.93	89.04	89.48	99.57	119.90	86	1.51	95.87	95.87
	WT	99.42	123.93	86	1.58	92.38	93.05	99.97	119.63	86	1.35	97.03	97.27
	HP	98.65	120.94	86	1.70	95.76	95.84	99.84	124.87	86	1.69	93.95	93.98
	SD	93.13	137.09	104	3.35	71.57	72.59	98.70	124.09	92	1.82	92.46	92.46

**Tabelle C.27:** Empirische Power  $(1 - \hat{\beta})\%$  im nicht-sequentiellen Design (NSD) sowie  $(1 - \hat{\beta})\%$ , mittlerer und medianer Umfang (ASN = „Average Sample Number“,  $N_{med}$ ), mittlere Stufenanzahl  $\bar{k}$  und empirische Überdeckungswahrscheinlichkeiten  $(1 - 2\hat{\alpha})\%$  der Sequenz der beobachteten Intervalle  $\{KI_j\}$  sowie des finalen Intervalls  $KI_k$  in flexiblen Designs mit ordinaler Zielvariable unter der Alternativhypothese (Konstellation VII)

NSD	Design	Inverse Normalmethode					Inverse $\chi^2$ -Methode						
		$(1 - \hat{\beta})\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$	$(1 - \hat{\beta})\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$		
						$\{KI_j\}$	$KI_k$				$KI_k$	$\{KI_j\}$	
2a) $\vec{p}_1 = (0.58, 0.30, 0.10, 0.02)$ , $\vec{p}_2 = (0.46, 0.28, 0.16, 0.10)$ , $D_G = 0.171$													
75.27	PO	68.77	207.27	224	2.33	94.52	96.01	67.36	212.92	230	2.32	93.80	95.00
	OBF	68.88	210.25	222	2.70	91.76	92.12	70.82	213.17	222	2.58	91.36	91.36
	WT	69.73	208.36	222	2.52	92.71	93.55	70.12	213.19	223	2.47	92.49	92.79
	HP	76.64	229.92	242	2.71	95.64	95.70	71.41	218.28	226	2.74	91.15	91.16
	SD	62.50	198.06	216	3.30	84.16	84.85	66.67	193.61	216	2.85	88.27	88.27
2b) $\vec{p}_1 = (0.22, 0.40, 0.34, 0.04)$ , $\vec{p}_2 = (0.15, 0.35, 0.28, 0.22)$ , $D_G = 0.201$													
83.47	PO	79.21	188.90	190	2.18	93.99	95.36	78.23	194.30	190	2.16	93.89	95.07
	OBF	77.21	195.12	218	2.63	90.33	90.78	80.74	193.55	202	2.45	89.98	89.98
	WT	78.67	192.68	211	2.41	91.62	92.62	81.12	194.37	208	2.32	91.33	91.72
	HP	84.78	212.33	212	2.60	95.79	95.91	81.10	199.64	200	2.63	89.48	89.48
	SD	72.36	185.72	216	3.36	81.85	83.12	77.19	182.08	216	2.81	86.40	86.40
2c) $\vec{p}_1 = (0.18, 0.44, 0.32, 0.06)$ , $\vec{p}_2 = (0.10, 0.35, 0.40, 0.15)$ , $D_G = 0.213$													
89.10	PO	84.42	176.57	184	2.08	94.09	95.76	83.90	179.93	190	2.04	94.09	95.14
	OBF	82.12	184.85	190	2.56	90.21	90.70	85.38	183.29	190	2.37	89.50	89.50
	WT	83.87	181.79	190	2.33	92.19	92.91	85.64	183.34	190	2.23	91.09	91.47
	HP	88.74	197.79	190	2.51	95.52	95.63	86.93	187.40	190	2.57	89.50	89.50
	SD	76.22	180.51	204	3.42	81.00	82.11	82.75	174.42	198	2.79	86.63	86.63
2d) $\vec{p}_1 = (0.45, 0.30, 0.20, 0.05)$ , $\vec{p}_2 = (0.25, 0.34, 0.26, 0.15)$ , $D_G = 0.255$													
96.57	PO	93.70	147.53	92	1.81	93.99	95.32	93.76	150.53	86	1.78	94.15	95.11
	OBF	90.80	158.72	138	2.39	89.34	89.84	94.41	156.77	128	2.16	91.12	91.12
	WT	92.28	154.87	113	2.09	92.24	93.15	94.59	155.22	106	1.98	92.76	93.10
	HP	95.03	166.01	148	2.27	95.24	95.32	94.56	159.71	130	2.35	89.37	89.39
	SD	85.53	161.52	144	3.49	75.98	77.61	91.74	154.46	126	2.67	85.01	85.01
2e) $\vec{p}_1 = (0.52, 0.36, 0.08, 0.04)$ , $\vec{p}_2 = (0.30, 0.30, 0.20, 0.20)$ , $D_G = 0.336$													
99.94	PO	99.35	107.58	86	1.35	94.54	95.73	99.62	108.17	86	1.32	95.15	95.99
	OBF	97.76	118.82	92	1.92	89.08	89.56	99.48	113.87	86	1.64	92.54	92.54
	WT	98.83	113.35	86	1.58	92.84	93.75	99.67	113.00	86	1.48	93.80	94.12
	HP	98.58	117.40	86	1.72	95.49	95.60	99.44	114.47	86	1.79	90.45	90.46
	SD	92.31	128.27	104	3.33	71.88	72.88	98.55	115.01	92	2.09	85.33	85.33

**Tabelle C.28:** Empirische Power  $(1 - \hat{\beta})\%$  im nicht-sequentiellen Design (NSD) sowie  $(1 - \hat{\beta})\%$ , mittlerer und medianer Umfang (ASN = „Average Sample Number“,  $N_{med}$ ), mittlere Stufenanzahl  $\bar{k}$  und empirische Überdeckungswahrscheinlichkeiten  $(1 - 2\hat{\alpha})\%$  der Sequenz der beobachteten Intervalle  $\{KI_j\}$  sowie des finalen Intervalls  $KI_k$  in flexiblen Designs mit ordinaler Zielvariable unter der Alternativhypothese (Konstellation VIII)

NSD	Design	Inverse Normalmethode					Inverse $\chi^2$ -Methode						
		$(1 - \hat{\beta})\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$	$(1 - \hat{\beta})\%$	ASN	$N_{med}$	$\bar{k}$	$(1 - 2\hat{\alpha})\%$		
						$\{KI_j\}$	$KI_k$				$KI_k$	$\{KI_j\}$	
2a) $\vec{p}_1 = (0.58, 0.30, 0.10, 0.02)$ , $\vec{p}_2 = (0.46, 0.28, 0.16, 0.10)$ , $D_G = 0.171$													
75.27	PO	82.72	283.37	260	2.30	94.01	95.70	84.62	287.37	258	2.26	94.27	95.36
	OBF	80.47	278.93	270	2.69	91.38	91.72	85.52	280.30	258	2.54	91.21	91.21
	WT	83.46	278.76	260	2.50	92.77	93.59	85.80	284.80	258	2.42	92.49	92.92
	HP	87.42	300.24	280	2.69	95.90	96.01	86.38	287.87	264	2.71	90.76	90.79
	SD	79.81	284.16	270	3.63	83.95	84.82	85.97	281.44	258	3.12	87.52	87.52
2b) $\vec{p}_1 = (0.22, 0.40, 0.34, 0.04)$ , $\vec{p}_2 = (0.15, 0.35, 0.28, 0.22)$ , $D_G = 0.201$													
83.47	PO	90.32	246.79	241	2.15	93.77	95.25	91.85	255.19	258	2.12	94.12	95.40
	OBF	87.24	249.69	238	2.60	90.40	90.89	91.64	251.41	252	2.44	90.30	90.30
	WT	89.01	253.41	256	2.41	91.94	92.77	91.91	251.99	258	2.29	91.86	92.33
	HP	91.96	259.37	252	2.56	95.49	95.62	91.51	254.78	252	2.61	89.41	89.43
	SD	84.13	255.01	242	3.64	80.80	82.36	91.41	249.84	238	3.02	87.10	87.10
2c) $\vec{p}_1 = (0.18, 0.44, 0.32, 0.06)$ , $\vec{p}_2 = (0.10, 0.35, 0.40, 0.15)$ , $D_G = 0.213$													
89.10	PO	93.85	227.32	194	2.06	94.42	95.90	95.04	234.72	226	2.03	93.76	95.06
	OBF	90.00	234.18	208	2.55	90.67	91.08	94.06	231.18	214	2.35	90.19	90.19
	WT	92.18	231.42	208	2.30	92.03	92.92	94.58	232.34	220	2.20	91.95	92.19
	HP	93.94	238.71	214	2.48	95.14	95.32	94.68	232.11	206	2.53	89.47	89.49
	SD	86.79	234.45	206	3.65	80.04	81.30	93.78	227.45	198	2.95	86.52	86.52
2d) $\vec{p}_1 = (0.45, 0.30, 0.20, 0.05)$ , $\vec{p}_2 = (0.25, 0.34, 0.26, 0.15)$ , $D_G = 0.255$													
96.57	PO	97.76	178.94	92	1.80	94.12	95.65	98.81	183.48	86	1.76	94.43	95.52
	OBF	94.55	188.98	138	2.36	89.43	89.84	97.37	184.81	124	2.12	91.60	91.60
	WT	96.57	184.18	112	2.07	91.85	92.61	97.98	186.80	111	1.96	93.69	93.97
	HP	96.53	187.88	148	2.24	95.29	95.37	97.80	185.97	128	2.34	89.44	89.48
	SD	90.50	195.80	148	3.63	75.02	76.98	96.72	182.97	124	2.74	84.87	84.87
2e) $\vec{p}_1 = (0.52, 0.36, 0.08, 0.04)$ , $\vec{p}_2 = (0.30, 0.30, 0.20, 0.20)$ , $D_G = 0.336$													
99.94	PO	99.78	115.81	86	1.34	94.38	95.59	99.99	117.86	86	1.32	94.89	95.81
	OBF	98.70	130.66	92	1.92	89.40	89.94	99.65	122.94	86	1.63	92.72	92.72
	WT	99.49	123.50	86	1.57	92.88	93.55	99.85	122.16	86	1.48	93.70	93.97
	HP	98.93	121.94	86	1.69	96.18	96.27	99.61	121.16	86	1.80	90.83	90.91
	SD	93.29	136.97	104	3.36	71.87	72.89	98.54	120.99	92	2.10	85.30	85.30