

Fallzahlplanung für klinische Studien mit Zähldaten

Simon Schneider
 Abteilung Medizinische Statistik
 Humboldtallee 32
 37073 Göttingen
 simon.schneider@med.uni-goettingen.de

Heinz Schmidli
 Statistical Methodology Novartis
 Pharma AG
 Lichtstrasse 35
 4056 Basel
 heinz.schmidli@novartis.com

Tim Friede
 Abteilung Medizinische Statistik
 Humboldtallee 32
 37073 Göttingen
 tim.friede@med.uni-goettingen.de

Zusammenfassung

Um die Wirksamkeit und Sicherheit neu entwickelter Medikamente zu überprüfen, wird jährlich eine Vielzahl an klinischen Studien durchgeführt. In der Planungsphase einer klinischen Studie ist die Fallzahlplanung ein wichtiger Baustein, um gute Chancen zu haben, die Studie erfolgreich abschließen zu können. Um einen Stichprobenumfang zu bestimmen, werden zunächst Annahmen über verschiedene Parameter benötigt. Handelt es sich um einen normalverteilten primären Endpunkt, dann müssen sowohl der voraussichtliche Effekt als auch die Varianz des primären Endpunktes spezifiziert werden. Ab Version SAS 9.1 stehen dem Anwender die beiden Prozeduren PROC POWER und PROC GLMPOWER zur Verfügung, die eine Vielzahl an Anwendungsbereichen abdecken. Wir stellen eine Erweiterung dieser Prozeduren in Form des Makros SSECD („Sample size estimation for count data“) vor, welches sich auf Zähldaten als Endpunkt fokussiert. Im Detail sind Fallzahlplanungen für die Poisson- und Negativbinomial-Modelle umgesetzt. Durch den Einsatz des Makros SSECD wird dem Anwender die Möglichkeit eröffnet, eine Fallzahlplanung für Zähldaten durchzuführen und sich diese graphisch darstellen zu lassen.

Schlüsselwörter: Fallzahlplanung, Zähldaten, Poisson-Modell, Negativ-Binomial, Makro SSECD

1 Motivation

In der Planungsphase einer klinischen Studie ist die Fallzahlplanung ein wichtiger Baustein, um eine gute Chance zu haben, die Studie erfolgreich abzuschließen. Die Wichtigkeit der Fallzahlplanung wird aus zwei verschiedenen Sichtweisen begründet.

Wird in der Planungsphase eine zu *kleine* Fallzahl gewählt, so kann es passieren, dass ein klinisch relevanter Effekt nicht nachgewiesen werden kann. In der Statistik wird in diesem Zusammenhang von einer unterpowereten Studie gesprochen. Aus wirtschaftlicher Sicht werden in diesem Szenario Ressourcen unnötig verschwendet.

Die Wahl einer zu *großen* Fallzahl ist einerseits ethisch nicht vertretbar, da der unnötige Patienteneinschluss Risiken beinhaltet. Andererseits kann es zu einem Nachweis von irrelevanten Effekten führen, was aus wirtschaftlicher Sicht abermals als Ressourcenverschwendung zu deuten ist.

Aus diesen ethischen, statistischen und wirtschaftlichen Gesichtspunkten fordern Ethikkommissionen und Zulassungsbehörden eine Fallzahlplanung in der Planungsphase einer klinischen Studie. Um eine initiale Fallzahl zu bestimmen, werden Annahmen über einen erwarteten, klinisch relevanten Effekt und über sogenannte Störparameter benötigt. Im Falle von normalverteilten Endpunkten müsste der voraussichtliche Effekt und die Varianz des Endpunktes spezifiziert werden. Für eine Berechnung der Fallzahl stehen in SAS/STAT dem Anwender die beiden Prozeduren PROC POWER und PROC GLMPOWER mit einer Vielzahl an Anwendungsbereichen zur Verfügung. In dieser Arbeit wird eine Erweiterung der vorhandenen Prozeduren auf Zähldaten anhand des Makros SSECD vorgestellt. Zur Illustration werden Beispiele von klinischen Studien zur Multiplen Sklerose herangezogen, die in Abschnitt 2 eingeführt werden.

2 Beispiel: Klinische Studien in Multipler Sklerose

Multiple Sklerose (MS) ist eine chronisch-entzündliche Erkrankung des zentralen Nervensystems. Diese tritt hauptsächlich bei jungen Erwachsenen auf und beginnt mit einer remittierenden schubförmigen Phase. In klinischen Phase II Studien ist ein häufig verwendeter primärer Endpunkt die jährliche Schubrate (ARR) bzw. die mittels Magnetresonanztomographie gewonnenen, neu entstandenen kumulierten Läsionen in Phase III Studien.

In der Planungsphase einer klinischen Studie, die als primären Endpunkt die jährliche Schubrate hat, ist eine Annahme über die Kontrollschubrate zur Bestimmung der initialen Fallzahl notwendig. Andere parametrische Modellierungen, die in Abschnitt 3 vorgestellt werden, benötigen zudem eine Annahme über die Dispersionsparameter. Mit diesen Annahmen wird eine klinische Studie für eine voraussichtliche Alternative zu einem gegebenen Signifikanzniveau gepowert. Um in einem späteren Abschnitt die Funktionsweise des Makros SSECD demonstrieren zu können, werden zwei fiktive Studienplanungsphasen vorgestellt, die auf historischen Daten aus abgeschlossenen Experimenten stammen.

- Die erste fiktive Studie wird anhand der Ergebnisse aus Rudick *et al.* (2006) geplant, wobei die Nachbeobachtung für jeden Patienten ein Jahr beträgt. In dem kontrollierten und randomisierten Versuch wurden 1171 Patienten balanciert auf eine Add-on Placebo ($n = 582$) und auf eine Kombinationstherapiegruppe ($n = 589$) verteilt. Die beobachtete Kontrollschubrate beträgt $\lambda_0 = 0,89$ und der beobachtete Effekt $\theta = 0,47$ (Verhältnis der beiden Raten).
- Die zweite Studie wird mittels der Resultate aus Comi *et al.* (2012) geplant. Hierbei beträgt die Nachbeobachtung für jeden Patienten ein Jahr. In der genannten Studie wurde eine Therapiegruppe gegen einen Placebo-Arm getestet. Als Kontrollschubrate ergibt sich $\lambda_0 = 0,39$. Zusätzlich wurde ein

Quotient aus Varianz / Erwartungswert der Kontrollrate von 1,8 und ein Effekt von $\theta = 0,77$ beobachtet.

Studie 1 (2) soll in der Planungsphase bei einem zweiseitigen Signifikanzniveau von $\alpha = 5\%$ für einen klinisch relevanten Effekt θ^* auf $1 - \beta = 90\%$ (80%) gepowert werden. Bevor in Abschnitt 5 die Fallzahlen mithilfe des Makros SSECD bestimmt werden, werden die verschiedenen Modelle für Zähldaten und die speziellen Fallzahlformeln vorgestellt.

3 Statistische Modelle für Zähldaten

Im Folgenden wird eine klinische Studie, in der N Patienten zufällig im Verhältnis 1:k einer Kontroll- ($i = 0$) oder einer Behandlungsgruppe ($i = 1$) zugeordnet werden, zugrunde gelegt. Die Anzahl an Events (Zähldaten) für Patient $j = 1, \dots, n_i$ in Behandlungsgruppe i werden mit y_{ij} bezeichnet. In den nächsten Abschnitten werden verschiedene Modellierungen der Zähldaten y_{ij} und die zugehörige Fallzahlformel vorgestellt.

3.1 Poisson-Modell

In diesem Abschnitt wird angenommen, dass die Beobachtungen y_{ij} aus einer Poissonverteilung mit dem Parameter $t_{ij}\lambda_i$ kommen. Hierbei ist t_{ij} die Follow-up Zeit für den Patienten j in der Behandlungsgruppe i . Das statistische Modell ist somit durch

$$y_{ij} \sim Pois(t_{ij}\lambda_i)$$

beschrieben. In klinischen Studien ist oftmals die relative Risikoreduktion $\theta = \lambda_1 / \lambda_0$, die den Behandlungseffekt definiert, von Hauptinteresse. Um die Hypothese $H_0: \theta \geq 1$ gegen die Alternative $H_1: \theta < 1$ zu testen, wird die approximative Normalverteilung des logarithmierten Schätzers für das relative Risiko mittels Waldtest ausgenutzt:

$$Z = \frac{\log(\hat{\theta}) - \log(\theta)}{\frac{1}{\sum y_{0j}} + \frac{1}{\sum y_{1j}}} \sim N(0,1)$$

Hierbei ist der Schätzer für den Behandlungseffekt wie folgt definiert: $\hat{\theta} = \frac{\frac{\sum y_{1j}}{\sum t_{1j}}}{\frac{\sum y_{0j}}{\sum t_{0j}}}$

Zur Varianzschätzung wird die Delta-Methode angewendet.

Eine Fallzahlformel für die Kontrollgruppe, die eine Power von $1 - \beta$ für einen angenommenen Effekt θ^* und eine angenommene Rate von λ_0^* sichert, ist von Keene *et al.* (2007) gegeben und von Friede und Schmidli (2010a) für unbalancierte Designs ausgeweitet zu

$$n_0 = \frac{(z_{1-\alpha} + z_{1-\beta})^2}{\log(\theta^*)^2} \frac{1 + k\theta^*}{k\theta^*} \frac{1}{\lambda^* t}, \quad (\text{Gl. 1})$$

wobei z_γ das γ -Quantil der Standardnormalverteilung und t die Studienlänge darstellen. Um die Fallzahl für einen zweiseitigen Hypothesentest zu bestimmen, wird in (Gl.1) $\alpha = \alpha/2$ gesetzt.

3.2 Quasi-Likelihood Ansatz

Das in Abschnitt 2.1 vorgestellte Poisson-Modell impliziert die Annahme der Gleichheit von Erwartungswert und Varianz der Zähldaten ($\text{Var}(y_{ij}) = t_{ij}\lambda_i$). Wie in Abschnitt 2 angedeutet, ist diese Annahme für viele Phase II bzw. Phase III Studien in Multiple Sklerose nicht immer angemessen. Eine Möglichkeit diese sogenannte Überdispersion (größere Variabilität als im statistischen Modell angenommen) zu erlauben und zu schätzen bietet ein Quasi-Likelihood Ansatz (Friede und Schmidli, 2010a).

Das statistische Modell aus Abschnitt 3.1 wird im Folgenden zugrunde gelegt, jedoch kann die Varianz durch die Einführung eines Dispersionsparameters $\sigma^2 \geq 1$ vergrößert werden. Die neue Varianz der Zähldaten ergibt sich somit zu: $\text{Var}(y_{ij}) = \sigma^2 t_{ij}\lambda_i$, wobei der Dispersionsparameter in den Gruppen gleich ist.

Die Schätzung des zusätzlichen Parameters geschieht mittels Momentenmethode und ist in Friede und Schmidli (2010a) detailliert beschrieben. Der Quasi-Likelihood Ansatz bietet eine benutzerfreundliche und unkomplizierte Umsetzung hinsichtlich Teststatistik und Fallzahlformel, da der Dispersionsparameter lediglich den Standardfehler in der Teststatistik vergrößert ($Z^{od} = Z/\sigma$) und somit als Faktor in die Fallzahlformel eingeht. Eine modifizierte Fallzahlformel für die Kontrollgruppe, die eine Power von $1 - \beta$ für einen angenommenen Effekt θ^* , eine angenommene Rate von λ_0^* und einen Dispersionsparameter σ^{2*} sichert, ist von Friede und Schmidli (2010a) gegeben:

$$n_0 = \frac{(z_{1-\alpha} + z_{1-\beta})^2}{\log(\theta^*)^2} \frac{1 + k\theta^*}{k\theta^*} \frac{1}{\lambda^* t} \sigma^{2*} \quad (\text{Gl. 2})$$

Die Herleitung der Formel ergibt sich umgehend aus der Herleitung der Fallzahlformel in Abschnitt 3.1 und der Anmerkung, dass der Dispersionsparameter lediglich als Faktor in die Teststatistik eingeht. Der vorgestellte Quasi-Likelihood Ansatz ist robust gegenüber Missspezifikationen in der Modellwahl (McCullagh und Nelder, 1989).

3.3 Negativ-Binomialmodell

Ein weiteres und oft genutztes Modell zur Erzeugung von überdispensierten Zähldaten ist das Negativ-Binomialmodell. Dass dieses Modell sich für die Beschreibung von Krankheitsverläufen bei Multipler Sklerose eignet, zeigen z.B. Sormani *et al.* (1999).

Analog zu Abschnitt 3.1 wird angenommen, dass die Beobachtungen y_{ij} aus einer Poissonverteilung mit Parameter $t_{ij}\lambda_i v_{ij}$ stammen, wobei v_{ij} einen Interindividualeffekt darstellt. Dieser zufällige Effekt spiegelt eine Heterogenität im Patientenkollektiv wider

und wird im Folgenden durch eine Gammaverteilung modelliert. Das statistische Modell ist wie folgt gegeben:

$$y_{ij}|v_{ij} \sim \text{Pois}(t_{ij}\lambda_i v_{ij}) \quad \text{mit} \quad v_{ij} \sim \Gamma(1/\varphi, \varphi)$$

Für den Erwartungswert und die Varianz der Zählraten gelten $E(y_{ij}) = t_{ij}\lambda_i = \mu_i$ bzw. $\text{Var}(y_{ij}) = \mu_i(1 + \varphi\mu_i)$. Folglich ist die Überdispersion durch den Faktor $1 + \varphi\mu_i$ beschrieben. Eine Fallzahlformel für den approximativen Waldtest wird durch Friede und Schmidli (2010b) gegeben. Eine gewünschte Power von $1 - \beta$ sichert die Fallzahl für die Kontrollgruppe n_0 für einen angenommenen Effekt θ^* , eine angenommene Rate von λ_0^* und einen Shapeparameter φ^* :

$$n_0 = \frac{(z_{1-\alpha} + z_{1-\beta})^2}{\log(\theta^*)^2} \left(\frac{1 + k\theta^*}{k\theta^*} \frac{1}{\lambda_0^* t} + \frac{k+1}{k} \varphi^* \right) \quad (\text{Gl. 3})$$

3.4 Nicht-Unterlegenheitsstudien

Der Nachweis der Wirksamkeit einer Therapie wird in der nahen Vergangenheit immer öfter durch Nicht-Unterlegenheitsstudien gezeigt. Diese haben den Nachweis von einer höchstens irrelevant schlechteren Wirksamkeit einer neuen Medikation zu einem Standardmedikament zum Ziel. Für die Hypothesenformulierung wird eine positive Nicht-Unterlegenheitsmarge δ eingeführt. Die modifizierte Hypothese lautet $H_0: \theta \geq \delta$ gegen die Alternativhypothese $H_1: \theta < \delta$. Durch die neue Formulierung ergeben sich Implikationen für die Fallzahlformeln in (Gl.1), (Gl.2) und (Gl.3). Der Quotient $\log(\theta^*)^2$ wird für die Bestimmung der Fallzahl für Nicht-Unterlegenheitsstudien durch $\log(\theta^*/\delta)^2$ ersetzt. Das vorgestellte Makro SSECD enthält die Möglichkeit eine Marge zu definieren, wobei ein $\delta = 1$ das Testen von Überlegenheit impliziert. Powerberechnungen für Nicht-Unterlegenheitsstudien für weitere Testprozeduren (z.B. Score Test, Likelihood-Ratio Test) werden von Miede und Cohrs (2005) bereitgestellt.

4 Implementierung

Die in Abschnitt 3 vorgestellten Fallzahlformeln sind in dem vorgestellten Makro SSECD¹ implementiert. Das Makro „Sample size estimation for count data“ wird wie folgt aufgerufen:

```
%ssecd(   alpha=,
          power=,
          theta_a=,
          lambda0=,
          margin=,
```

¹ Das SAS-Makro SSECD wird auf Nachfrage vom Autor verschickt.

```
t=,
phi=,
sigma2=,
method=)
```

Hierbei stellt `alpha` das Signifikanzniveau und `power` die gewünschte Power $1 - \beta$ dar. Der angenommene Effekt ist mit `theta_a` und die angenommene Ereignisrate in der Kontrollgruppe mit `lambda0` gegeben. `Margin` bezeichnet die Nicht-Unterlegenheitsmarge δ . Wird auf Überlegenheit getestet, so muss diese gleich eins gesetzt werden. Des Weiteren steht `phi` für den Shapeparameter φ der Gammaverteilung und `sigma2` bezeichnet den eingeführten Dispersionsparameter σ^2 aus Abschnitt 3.2. Um sich Fallzahlen für das Poisson-Modell aus Abschnitt 3.1 berechnen zu lassen, muss `method=1` gesetzt werden. Für den Quasi-Likelihood Ansatz muss `method=2` und anderenfalls (Negativ-Binomialmodell) als `method=3` definiert werden.

Die Ausgabe wird im nächsten Abschnitt an einem Beispiel vorgestellt.

5 Auswertung der Beispiele

In diesem Abschnitt werden die Ergebnisse aus den in Abschnitt 2 vorgestellten Studien aufgegriffen (Tabelle 1). Der beobachtete Quotient aus Varianz zu Erwartungswert in Comi *et al.* (2012) beträgt $\sigma^2 = 1,8$. Wird der Dispersionsparameter in das parametrische Negativ-Binomialmodell übertragen, so ergibt sich ein Shape-Parameter von $\varphi = 2,05$.

Tabelle 1: Charakteristika der historischen Studien zur Fallzahlplanung der fiktiven Experimente

Studie	Kontrollrate	Beob. Effekt	σ^2	φ
Rudick <i>et al.</i>	0,81	0,47	-	-
Comi <i>et al.</i>	0,39	0,77	1,8	2,05

Die erste fiktive Ein-Jahres Studie soll bei einem zweiseitigen Signifikanzniveau von $\alpha = 5\%$ für einen klinisch relevanten Effekt von 0,66 auf $1 - \beta = 90\%$ gepowert werden. Werden die Bezeichnungen des Makros SSECDC aus Abschnitt 4 beachtet, so ergibt sich der folgende Programmcode:

```
%ssecdc( alpha=0.025,
power=0.90,
theta_a=0.66,
lambda0=0.81,
margin=1,
t=1,
phi=,
sigma2=,
method=1)
```

Zu beachten ist, dass `method=1` gewählt ist, somit keine Dispersion angenommen wird. Die Fallzahl ergibt sich aus dem Poisson-Modell in Abschnitt 3.1. Die Makrovariablen `phi` und `sigma2` können leer gelassen werden, da diese zur Berechnung ignoriert werden. Als Ergebnis wird folgender Output geliefert (Tab. 2):

Tabelle 2: Output des Makros SSECD für Fallzahlplanung nach Rudick et al. (2006).

	alpha	k	Effect	Margin	Control rate	n Group1	n Group2
Results	0.025	1	0.66	1	0.81	190	190

Um eine Power von 90 % bei einem angenommenen Effekt von $\theta = 0,66$ sichern zu können, wird in einem Poisson-Modell eine Fallzahl von 190 Personen pro Gruppe benötigt (Tab. 2).

Für die fiktive Planung der zweiten Studie wird die Fallzahl zunächst mittels Quasi-Likelihood-Ansatz bestimmt, um der Dispersion gerecht zu werden. Folgender Code erfüllt die Bedingungen aus Tabelle 1, wobei ein klinisch relevanter Effekt als $\theta^* = 0.75$ für die kommende Studie definiert wird:

```
%ssecd( alpha=0.025,
         power=0.80,
         theta_a=0.75,
         lambda0=0.39,
         margin=1,
         t=1,
         phi=,
         sigma2=1.8,
         method=2)
```

Tabelle 3: Resultierende Fallzahlen bei Annahme eines Quasi-Likelihood-Ansatzes für die fiktive Studienplanung 2.

	alpha	k	Effect	Margin	Control rate	Sigma	n Group1	n Group2
Results	0.025	1	0.75	1	0.39	1.8	1022	1022

Wird ein Negativ-Binomialmodell vorausgesetzt, so ändert sich die Methode zu `method=3` und der Shape-Parameter der Gammaverteilung wird zur initialen Fallzahlplanung benötigt.

```
%ssecd( alpha=0.025,
         power=0.80,
         theta_a=0.75,
         lambda0=0.39,
```

```
margin=1,
t=1,
phi=2.05,
sigma2=,
method=3)
```

Tabelle 4: Resultierende Fallzahlen bei Annahme eines parametrischen Negativ-Binomialmodells für die fiktive Studienplanung 2.

	alpha	k	Shape	Effect	Margin	Control rate	n Group1	n Group2
Results	0.025	1	2.05	0.75	1	0.39	958	958

Tabellen 3 und 4 liefern die Fallzahlen für den Quasi-Likelihood-Ansatz und das Negativ-Binomialmodell. Beide Szenarien liefern Fallzahlen von ca. 1000 Patienten pro Gruppe, die eine Power von 80 % für den angenommenen Effekt von $\theta^* = 0,75$ sichern.

Unter Verwendung der Funktion PROC GPLOT können die Fallzahlen gegen verschiedene Effekte bzw. Kontrollraten abgetragen oder für eine feste Fallzahl Powerkurven zur graphischen Illustration angefertigt werden. Als Beispiel wird auf die erste fiktive Studie zurückgegriffen. Wird für verschiedene Effekte eine Fallzahl benötigt, die eine Power von 90% sichert, so kann mithilfe der Prozedur GPLOT eine Fallzahlkurve erstellt werden. Für die Effekte $\theta = 0.5, \dots, 0.9$ ist dies in Abbildung 1 dargestellt. Folgender Programmschritt wurde verwendet:

```
TITLE 'Fallzahl gegen Effektgröße';
FOOTNOTE1 h=1 j=1 'KSFE 2013 Ulm'
           j=r 'Schneider et al.';

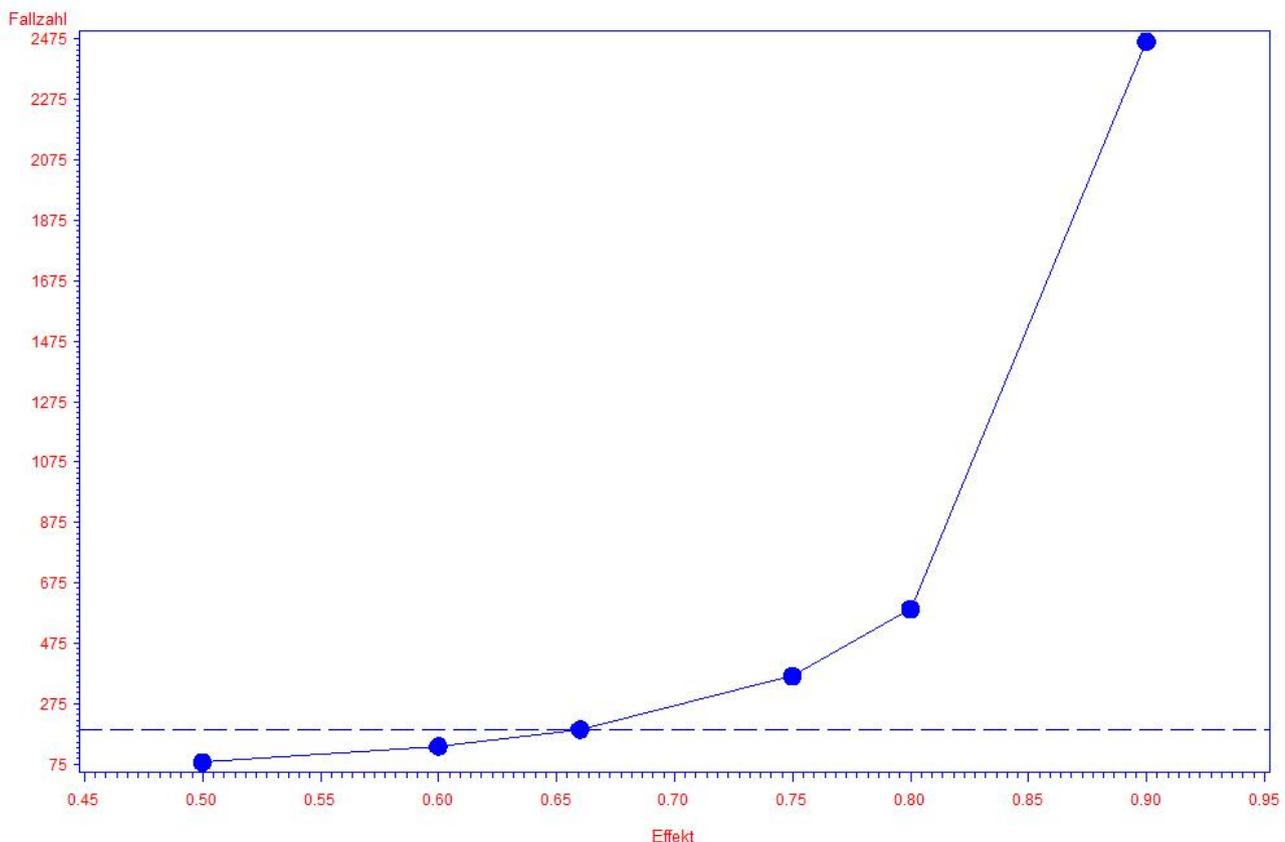
SYMBOL11 color=blue
          interpol=join
          value=dot
          height=2;

PROC GPLOT data=rudick;
  plot Fallzahl*Effekt / haxis=0.45 to 0.95 by 0.05
                        vaxis=75 to 2500 by 200
                        hminor=10
                        vminor=10

                        vref=190 lvref=2
                        cvref=blue
                        caxis=blue
                        ctext=red;

RUN;
QUIT;
```

Fallzahl gegen Effektgröße



KSFE 2013 Ulm

Schneider et al.

Abbildung 1: Fallzahl für verschiedene Effekte, die eine Power von 90% bei konstanter Kontrollrate, sichert.

6 Ausblick

Da in der Planungsphase einer klinischen Studie oftmals eine erhebliche Unsicherheit bezüglich des angenommenen Effekts beziehungsweise in der Annahme der Kontrollrate besteht, sind adaptive Designs immer populärer und gewinnen immer weiter an Anwendung (Nicholas *et al.*, 2011). Klinische Studien mit einer internen Pilotstudie (IPS) gehören zu den attraktivsten Umsetzungsmöglichkeiten, da aus wirtschaftlicher Sichtweise Ressourcen und aus ethischer Sichtweise Patienten geschont werden. In der IPS kann die Fallzahl an den vorhandenen Patientendaten verblindet oder entblindet adjustiert werden. Einen Vergleich der zur Auswahl stehenden verblindeten Methoden für Zähldaten präsentiert Schneider *et al.* (2012). Diese Methoden können in einer weiteren Entwicklung in das SAS-Programm implementiert werden, so dass verschiedene Szenarien und Ausgänge simuliert und bewertet werden können.

Des Weiteren ist zu erwähnen, dass die Güte der vorgestellten Fallzahlapproximationen bei unbalancierten Designs erheblich an Exaktheit verliert. Für das Poisson-Modell ist diese Eigenschaft bei Ng und Tang (2005) in detaillierten Simulationen zu finden. Festzustellen ist, dass bei nicht zu extremer Unbalanciertheit (2:1, 3:1) die Fallzahlformeln

noch gute Ergebnisse liefern. Werden extremere Situationen betrachtet, so sollten exakte Berechnungen vorgenommen werden (Miede und Cohrs, 2005).

Literatur

- [1] O.N. Keene, M.R.K. Jones, P.W. Lane, J. Anderson: Analysis of exacerbation rates in asthma and chronic obstructive pulmonary disease: example from the TRISTAN study. *Pharmaceutical Statistics*, 6, 89-97, 2007.
- [2] T. Friede, H. Schmidli: Blinded sample size reestimation with count data: Methods and applications in multiple sclerosis. *Statistics in Medicine*, 29, 1145-1156, 2010a.
- [3] P. McCullagh, J.A. Nelder: *Generalized Linear Models* (2nd edn). Chapman & Hall, London, 1989.
- [4] M.P. Sormani, P. Bruzzi, D.H. Miller, C. Gasperini, F. Barkhof, M. Filippi: Modelling MRI enhancing lesion counts in multiple sclerosis using a negative binomial model: implications for clinical trials. *Journal of Neurological Sciences*, 163, 74-80, 1999.
- [5] T. Friede, H. Schmidli: Blinded sample size reestimation with negative binomial counts in superiority and non-inferiority trials. *Methods of Information in Medicine*, 49, 618-624, 2010b.
- [6] R.A. Rudick, W.H. Stuart, P.A. Calabresi, C. Confavreux, S.L. Galetta, E-W. Radue, F.D. Lublin, B. Weinstock-Guttman, D.R. Wynn, F. Lynn, M.A. Panzara, A.W. Sandrock for the SENTINEL Investigators: Natalizumab plus interferon beta-1a for relapsing multiple sclerosis. *The New England Journal of Medicine*, 354, 911-923, 2006.
- [7] G. Comi, D. Jeffrey, L. Kappos, X. Montalban, A. Boyko, M.A. Rocca, M. Filippi: Placebo-controlled trial of oral laquinimod for multiple sclerosis. *The New England Journal of Medicine*, 366, 1000-1009, 2012.
- [8] C. Miede, J. Mueller-Cohrs: Power calculation for non-inferiority trials comparing two Poisson-distributions. *PhUSE*, 2005.
- [9] R. Nicholas, S. Straube, H. Schmidli, S. Schneider, T. Friede: Trends in annualized relapse rates in relapsing remitting multiple sclerosis and consequences for clinical trial design. *Multiple Sclerosis Journal*, 17, 1211-1217, 2011.
- [10] S. Schneider, H. Schmidli, T. Friede: Robustness of methods for blinded sample size reestimation with overdispersed count data. (Submitted)
- [11] H.K.T. Ng, M.-L. Tang: Testing the equality of two Poisson means using the rate ratio. *Statistics in Medicine*, 24, 955-965, 2005.