

Auswertung und Fallzahlberechnung in Dosis-Findungs-Studien mit aktiver Kontrolle mit SAS/STAT und SAS/IML

Hans-Joachim Helms
Abteilung Medizinische Statistik
Humboldtallee 32
37073 Göttingen
Hans-Joachim.Helms@med.uni-goettingen.de

Norbert Benda
Bundesinstitut für Arzneimittel
und Medizinprodukte
Kurt-Georg-Kiesinger-Allee 3
53175 Bonn
Norbert.Benda@bfarm.de

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

Zusammenfassung

Es werden hier in diesem Beitrag Dosis-Findungs-Studien mit aktiver Kontrolle betrachtet. Das Ziel dieser Art von Studien ist das Finden der Zieldosis, welche zum gleichen Wirkungsgrad führt wie die aktive Kontrolle, sowie die Bestimmung eines zugehörigen Konfidenzintervalls.

Es wird für verschiedene Dosis-Wirkungs-Funktionen vorgestellt, wie das Problem der Dosis-Findung in SAS/IML sowie SAS/STAT implementiert werden kann. Weiterhin werden Kriterien für eine Fallzahlplanung vorgestellt und erläutert, wie diese in SAS/IML umgesetzt werden können. Abschließend werden, ergänzend zu den bereits in SAS implementierten Verfahren, zwei weitere Methoden zur Berechnung eines Konfidenzintervalls der Zieldosis diskutiert.

Schlüsselwörter: Dosis-Findungs-Studie, aktive Kontrolle, Fallzahlplanung, PROC NLMIXED

1 Einführung

In einer Dosis-Findungs-Studie mit aktiver Kontrolle wird ein neuartiges Medikament in verschiedenen Dosis-Stufen mit einem bereits bekannt Medikament (der aktiven Kontrolle) verglichen. Ziel dieser Studien ist das Finden der Dosis-Konzentration d^* , welche zum gleichen Wirkungsgrad führt wie die aktive Kontrolle. Dafür muss der Schnittpunkt der linearen oder nichtlinearen Dosis-Wirkungs-Kurve mit dem Wirkungsgrad der aktiven Kontrolle berechnet werden. Von Interesse ist zudem ein Konfidenzintervall für die Dosis d^* , auf dessen Grundlage eine Fallzahlplanung durchgeführt werden kann. Die Bestimmung der Zieldosis ist in Abbildung 1 beispielhaft dargestellt. Da selbst bei

linearen Dosis-Wirkungs-Profilen dieses Schätzproblem nichtlinear ist, müssen geeignete Näherungsverfahren verwendet werden.

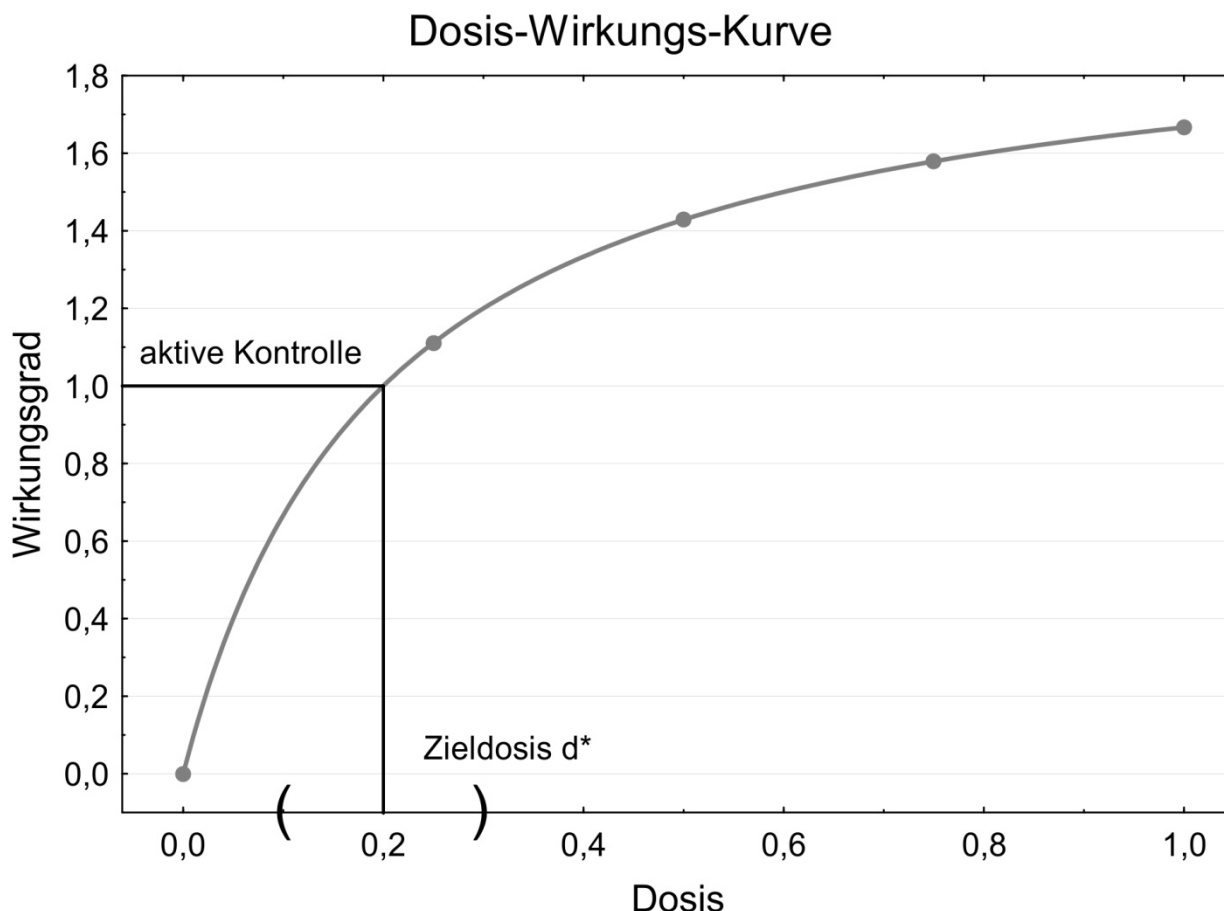


Abbildung 1: Bestimmung der Zieldosis in einer Dosis-Findungs-Studie mit aktiver Kontrolle.

2 Statistisches Modell

Es werden k Gruppen mit den Dosisstufen d_1, \dots, d_k und eine aktive Kontrolle (ak) untersucht. Die Messwerte der j -ten Person in der i -ten Dosisstufe können wie folgt dargestellt werden:

$$Y_{ij} = f_{\theta}(d_i) + \epsilon_{ij} \quad i = 1, \dots, k \quad j = 1, \dots, n,$$

wobei $f_{\theta}(d_i)$ eine Erwartungswertfunktion mit den Parametervektor θ ist, welche von der Dosis d_i abhängt. Die Messwerte der aktiven Kontrolle lassen sich als

$$Y_{ak,j} = \mu + \epsilon_{ak,j} \quad j = 1, \dots, n_{ak} = r \cdot n,$$

mit Erwartungswert μ schreiben. Sei ϵ der multivariat normalverteilte Vektor des Versuchsfehlers mit Kovarianzmatrix $\sigma^2 \cdot I$. Weiterhin sei $n_d = k \cdot n$ die Fallzahl aller Dosisgruppen und $N = n_d + n_{ak}$ die Gesamtfallzahl. Im Folgenden wird das lineare Modell sowie das nichtlineare Emax-Modell betrachtet.

2.1 Das lineare Modell

Wird eine lineare Dosis-Wirkungs-Kurve angenommen, so vereinfacht sich die Erwartungswertfunktion zu $f_\theta(d_i) = \theta_0 + \theta_1 d_i$ und das statistische Modell kann in Vektorschreibweise dargestellt werden:

$$Y_d = X\theta + \epsilon_d, X = \begin{pmatrix} 1 & \dots & 1 \\ d_1 & \dots & d_k \end{pmatrix}'$$

$$Y_{ak} = \mu 1_{n_{ak}} + \epsilon_{ak}.$$

Um die Zieldosis \widehat{d}^* zu bestimmen muss die Gleichung $f_\theta(d^*) = \theta_0 + \theta_1 d^* = \mu$ gelöst werden. Im Allgemeinen sind aber die Parameter der Dosis-Wirkungs-Kurve θ_0 und θ_1 sowie der Erwartungswert der aktiven Kontrolle μ unbekannt und müssen aus den Daten geschätzt werden. Daher werden im Folgenden die Schätzer dieser Parameter dargestellt.

2.1.1 Schätzer im linearen Modell

Die folgenden Resultate stammen aus der Theorie der linearen Modelle und können unter anderem bei Mattai und Provost (1992) nachgelesen werden.

Der Erwartungswert der aktiven Kontrolle lässt sich durch den Mittelwert der aktiven Kontrolle schätzen:

$$\hat{\mu} = \bar{Y}_{ak} \sim N\left(\mu, \frac{\sigma^2}{n_{ak}}\right)$$

Die Parameterschätzer der Dosis-Wirkungs-Funktion lassen sich wie folgt darstellen:

$$\hat{\theta} = \begin{pmatrix} \hat{\theta}_0 \\ \hat{\theta}_1 \end{pmatrix} \sim N(\theta, \sigma^2 (X'X)^{-1})$$

$$\hat{\theta} = (X'X)^{-1} X'Y_d.$$

Weiterhin ist die Varianz des Messfehlers unbekannt und muss geschätzt werden:

$$\widehat{\sigma^2} = \frac{1}{n_d - 2} Y_d' (I_{n_d} - X'(X'X)^{-1}X) Y_d \sim \frac{\sigma^2}{n_d - 2} \cdot \chi_{n_d - 2}^2.$$

Werden diese Schätzer in die Maximum-Likelihood-Gleichung der Zieldosis eingesetzt, so ergibt sich ein Schätzer für die Zieldosis d^* :

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

Somit steht ein Schätzer zur Bestimmung der Zieldosis d^* zur Verfügung, für welchen in den nächsten Abschnitten die Implementierung in SAS/IML sowie die Konstruktion von Konfidenzintervallen vorgestellt wird.

2.1.2 Die Implementierung der Schätzung der Zioldosis in SAS/IML

Wird zur Dateneingabe dieselbe Notation wie im vorgestellten linearen Modell verwendet, so können die notwendigen Parameterschätzer unter Ausnutzung der Matrix-Schreibweise in SAS/IML wie folgt implementiert werden.

```
***** Schätzen der Parameter *****;
mu_hut = SUM(y_ak)/n_ak;
theta_hut = xx_inv * x` * y_d;
theta0_hut = theta_hut[1];
theta1_hut = theta_hut[2];
sigma_2_hut = y_d` * (I(n_d) - x * xx_inv * x`) * y_d / (n_d -
rang_x);

*** Schätzen der Zioldosis d_stern *****;
d_stern_hut = (mu_hut - theta0_hut) / theta1_hut;
```

2.1.3 Berechnung des Konfidenzintervalls der Zioldosis

Im Folgenden wird ein Verfahren zur Berechnung eines Konfidenzintervalls der Zioldosis d^* und die entsprechende Implementierung in SAS/IML oder in SAS/STAT mit Hilfe der Prozedur NLMXED dargestellt.

2.1.3.1 Approximation des Standardfehlers durch die Δ -Regel

Da alle Parameterschätzer, welche zur Bestimmung der Zioldosis d^* benötigt werden, normalverteilt sind, kann mit Hilfe der Δ -Regel gezeigt werden, dass auch der Schätzer der Zioldosis \widehat{d}^* asymptotisch normalverteilt ist mit Varianz δ^2 , wobei:

$$\delta^2 = Var(\widehat{d}^*) = \frac{\sigma^2}{\theta_1^2} \left[\frac{(\sum [d_i - d^*]^2)}{(k(\sum d_i^2) - (\sum d_i)^2)} + \frac{1}{r} \right].$$

Weiterhin sei $\widehat{\delta}^2$ der Plug-In Schätzer der unbekanntem Varianz δ^2 , dann kann für die Zioldosis ein Konfidenzintervall wie folgt konstruiert werden:

$$(1 - \alpha) - KI: \widehat{d}^* \pm u_{1-\frac{\alpha}{2}} \cdot \sqrt{\widehat{\delta}^2}.$$

Im Folgenden wird die Implementierung des Konfidenzintervalls in SAS/IML und alternativ in der Prozedur NLMIXED in SAS/STAT vorgestellt.

2.1.3.2 Implementierung der Δ -Regel in SAS/IML

Wird für die Parameterschätzer die vorgestellte Notation verwendet, so lässt sich das 95% Konfidenzintervall der Zioldosis wie folgt implementieren:

```
*** Geschätzter Gradient *****;
g1_hut = -1/theta1_hut;
g2_hut = -(mu_hut-theta0_hut)/(theta1_hut*theta1_hut);
g3_hut = 1/theta1_hut;
grad_d_stern_hut =g1_hut//g2_hut//g3_hut;
```

```

*** Schätzung der Varianz *****;
var_d_stern_hut = sigma_2_hut * grad_d_stern_hut` * sigm_inv *
grad_d_stern_hut;

**** KI *****;
ki_o = d_stern_hut +(1.96*SQRT(var_d_stern_hut));
ki_u = d_stern_hut -(1.96*SQRT(var_d_stern_hut));

```

2.1.3.3 Implementierung der Δ-Regel in SAS/STAT

Alternativ zur Implementierung in SAS/IML können Schätzer und Konfidenzintervall der Zieldosis auch durch **Anwenden** der Prozedur NLMIXED bestimmt werden. Dafür wird d^* als ein Extraparameter definiert und separat im estimate Statement geschätzt. Weiterhin muss der aktiven Kontrolle der Mittelwert μ und den Dosisgruppen die korrespondierende Erwartungswertfunktion zugewiesen werden. Dies ist mittels einer if Abfrage in der NLMIXED Prozedur möglich, wenn der aktiven Kontrolle in der Variable dose ein negativer Wert zugeordnet wird.

```

proc nlmixed data=daten_linear;
    if dose>=0 then f=theta0+theta1*dose;
    else f=mu;
*** dose<0: aktive Kontrolle *****;
    model y ~ N(f,sigma**2);
    estimate 'D_STERN' (mu-theta0)/theta1;
run;

```

Tabelle 1: SAS Output von PROC NLMIXED

Additional Estimates								
Label	Estimate	Standard Error	DF	t Value	Pr > t	Alpha	Lower	Upper
D_STERN	0.4756	0.06206	280	6.31	<.0001	0.05	0.3595	0.5838

Mittels der Prozedur NLMIXED wird ein Schätzer für die Zieldosis sowie ein entsprechendes Konfidenzintervall über die Δ -Regel ausgegeben.

2.2 Ein nichtlineares Modell: Das Emax-Modell

Im Folgenden wird beispielhaft das Emax-Modell als nichtlineares Modell betrachtet. Im Emax-Modell lässt sich die Erwartungswertfunktion als $f_{\theta}(d_i) = \theta_0 + \theta_1 d_i / (\theta_2 + d_i)$ schreiben und die Messwerte der Dosisgruppen lassen sich wie folgt darstellen:

$$Y_{ij} = \theta_0 + \frac{\theta_1 d_i}{(\theta_2 + d_i)} + \epsilon_{ij} \quad i = 1, \dots, k \quad j = 1, \dots, n.$$

Dann definiert θ_0 den minimalen Behandlungseffekt, θ_1 den maximalen Behandlungseffekt bei einer „unendlich“ großen Dosis und θ_2 die Dosis welche zu 50% des maximalen Behandlungseffektes führt (ED_{50}). Die Messwerte der aktiven Kontrolle lassen sich analog zum linearen Modell beschreiben. Weiterhin soll auch in diesem Modell die Zieldosis d^* bestimmt werden, für welche gilt.

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

2.2.1 Schätzung der Zieldosis im Emax-Modell

Wird die oben definierte Gleichung nach d^* aufgelöst, kann ein Schätzer für die Zieldosis konstruiert werden:

$$d^* = \frac{\theta_2}{\left(\frac{\theta_1}{(\mu - \theta_0)} - 1\right)}. \quad (1)$$

Allerdings sind auch im nichtlinearen Modell die Parameter im Allgemeinen unbekannt und müssen geschätzt werden. Leider ist dies im Emax-Modell nicht in geschlossener Form möglich und so müssen alle Parameterschätzer iterativ bestimmt werden. Weiterhin kann mit Hilfe der Δ -Regel ein Konfidenzintervall konstruiert werden. Die entsprechende Implementierung in SAS/STAT wird im Folgenden dargestellt.

2.2.2 Implementierung in SAS/STAT

Analog zum linearen Modell wird d^* als ein Extraparameter definiert und separat im `estimate` Statement nach der obigen Formel (1) geschätzt. Weiterhin muss der aktiven Kontrolle der Mittelwert μ und den Dosisgruppen die korrespondierende Erwartungswertfunktion zugewiesen werden. Dies ist mittels einer `if` Abfrage in der NLMIXED Prozedur möglich, wenn der aktiven Kontrolle in der Variable `dose` ein negativer Wert zugeordnet wird.

```
proc nlmixed data=daten_emax;
    if dose>=0 then f=theta0+theta1*dose/(theta2+dose);
    else f=mu;
*** dose< 0: aktive Kontrolle *****;
    model y ~ N(f,sigma**2;
    estimate 'D_STERN' theta2/(theta1/(mu-theta0)-1);
run;
```

Tabelle 2: SAS Output von PROC NL MIXED

Additional Estimates								
Label	Estimate	Standard Error	DF	t Value	Pr > t	Alpha	Lower	Upper
D_STERN	0.1927	0.07311	280	2.64	0.0089	0.05	0.04875	0.3366

Mittels der Prozedur NL MIXED wird ein Schätzer für die Zieldosis sowie ein entsprechendes Konfidenzintervall über die Δ -Regel ausgegeben.

3 Fallzahlplanung

Nachdem beschrieben wurde, wie ein Punktschätzer für die Zieldosis d^* sowie ein zugehöriges Konfidenzintervall in SAS/IML oder SAS/STAT implementiert werden kann, bleibt die Frage: Wie und nach welchen Kriterien kann eine Fallzahl für eine Dosis-Findungs-Studie mit aktiver Kontrolle durchgeführt werden? Ein Überblick über verschiedene Kriterien zur Fallzahlplanung wird unter anderem bei Beal (1989) oder Jiroutek (2003) gegeben. Im Folgenden werden zwei verschiedene Kriterien und entsprechende Implementierungsansätze vorgestellt.

3.1 Erwartete Intervallbreite

Die geplante Fallzahl soll so gewählt werden, dass das Konfidenzintervall der Zieldosis d^* in Erwartung nicht breiter ist als eine Konstante $2 \cdot c > 0$. Um dieses Verfahren näher zu erläutern wird wieder das lineare Modell betrachtet. Dann kann die benötigte Fallzahl ($n^{(E)}$) so gewählt werden, dass:

$$n^{(E)} \geq \frac{\sigma^2}{\theta_1^2} \left[\frac{(\sum [d_i - d^*]^2)}{(k(\sum d_i^2) - (\sum d_i)^2)} + \frac{1}{r} \right] \left(\frac{u_1 - \frac{\alpha}{2}}{c} \right)^2.$$

Mit dieser ermittelten Fallzahl ist das Konfidenzintervall der Zieldosis in Erwartung nicht breiter als $2 \cdot c > 0$. Eine Implementierung der Formel ist in SAS/IML eins zu eins möglich, falls die notwendigen Variablen $\sigma^2, \theta_0, \theta_1$, sowie der Erwartungswert der aktiven Kontrolle μ gegeben sind. Dann kann die geplante Fallzahl für gewähltes α und c direkt berechnet werden.

3.2 Quantil der Intervallbreite

Bei diesem Kriterium soll die geplante Fallzahl so gewählt werden, dass das Konfidenzintervall der Zieldosis d^* mit einer Wahrscheinlichkeit von $\gamma\%$ nicht breiter ist als eine Konstante $2 \cdot c > 0$.

Das Ziel ist es das obere $\gamma\%$ -Quantil der Verteilung des Varianzschätzers der Zieldosis zu bestimmen und dieses zur Fallzahlplanung zu verwenden. Dafür wird allerdings die Verteilung des Varianzschätzers der Zieldosis benötigt, welche sich bereits im linearen

Fall nicht bestimmen lässt. Somit muss die geplante Fallzahl iterativ durch Simulationen bestimmt werden, was sich in SAS/IML realisieren lässt.

4 Diskussion

Zur Bestimmung der Zieldosis d^* sowie eines zugehörigen Konfidenzintervalls kann zum einen direkt PROC NLMIXED verwendet werden, welche zur Bestimmung eines Konfidenzintervalls auf die Methode der Δ -Regel zurückgreift oder es kann die Δ -Regel direkt in SAS/IML mittels Matrixschreibweise implementiert werden. Dies kann sinnvoll sein, falls Ergebnisse der Δ -Regel (z.B. zur Validierung) direkt in SAS/IML weiter verwendet werden sollen. Falls nicht auf die asymptotischen Resultate der Δ -Regel zurückgegriffen werden soll, um ein Konfidenzintervall zu bestimmen, so kann auch die Methode von Fisch und Strehlau (1993) oder ein parametrischer Bootstrap verwendet werden.

4.1 Methode nach Fisch und Strehlau (1993)

Alternativ zur bisher vorgestellten Methode lässt sich ein Konfidenzintervall von d^* auch nach der Methode von Fisch und Strehlau (1993) berechnen, welche für Kalibrierungsprobleme vorgeschlagen wurde.

Dabei wird die Maximum-Likelihood Gleichung des Schätzers der Zieldosis \widehat{d}^* in die folgende Form umgeschrieben:

$$\hat{\mu} - \hat{\theta}_0 - \hat{\theta}_1 d = 0 \quad (2)$$

und die Dosis d als fest, aber unbekannt angesehen.

Danach wird eine Verteilung für die linke Seite aus (2) hergeleitet und für diese Hilfsvariable ($W(d)$) ein Konfidenzintervall konstruiert, welches nicht von der unbekannt Dosis d abhängt. Wird dann, wie in Abbildung 2 zu sehen, die Dosis so verschoben, dass die Konfidenzintervall-Grenzen der Hilfsvariablen $W(d)$ erreicht werden, kann dadurch ein Konfidenzintervall für die Zieldosis d^* bestimmt werden.

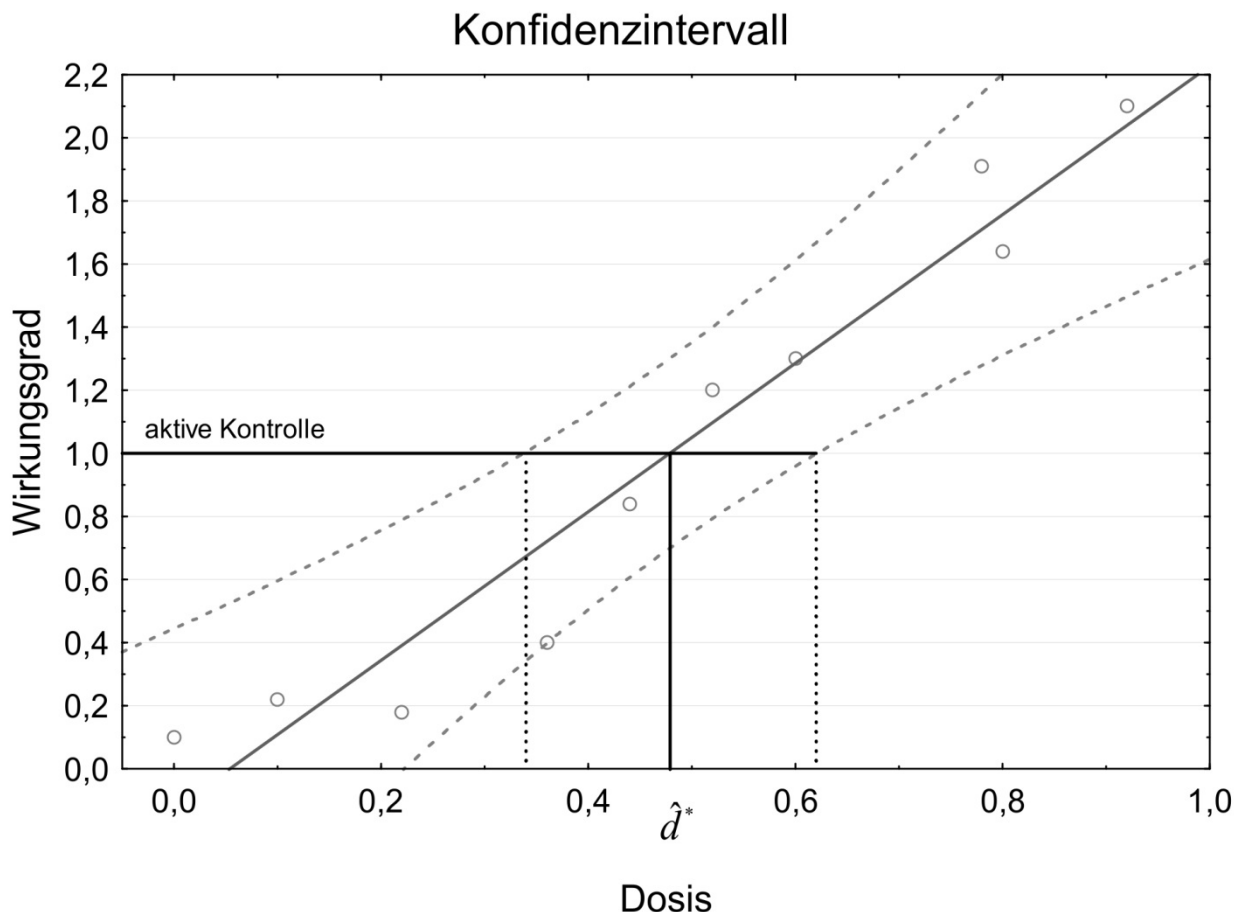


Abbildung 2: Inverse Bestimmung des Konfidenzintervalls der Zieldosis d^* nach der Methode von Fisch und Strehlau (1993)

Für Informationen zur expliziten Herleitung siehe Fisch und Strehlau (1993) sowie Helms et al. (2012). Diese Methode kann in SAS/IML mit Hilfe einer iterativen Schleife implementiert werden. Somit steht ein Verfahren zur Verfügung, welches auf eine exakte Verteilung zurückgreift, wobei ein Konfidenzintervall für die Zieldosis nur numerisch bestimmt werden kann.

4.2 Parametrischer Bootstrap

Da alle Parameterschätzer, welche zur Schätzung der Zieldosis d^* benötigt werden, mit entsprechender Verteilung bekannt sind ($\hat{\theta}$, $\hat{\mu}$ und $\widehat{\sigma^2}$), kann ein parametrischer Bootstrap (z.B. Efron, 1979) zur Konstruktion eines Konfidenzintervalls verwendet werden. Dieser kann in SAS/IML leicht implementiert werden, da lediglich die Parameterschätzer und nicht alle Messwerte „gebootstrap“ werden müssen.

Wird allerdings ein nichtlineares Modell angenommen, so können die notwendigen Parameterschätzer nur noch iterativ bestimmt werden und es sollte dann auf die Prozedur NLMIXED zurückgegriffen werden, um diese Schätzer zu ermitteln. Für Informationen zur expliziten Herleitung siehe Helms et al. (2012).

4.3 Ausblick

Die Methode von Fisch und Strehlau (1993) sowie der parametrische Bootstrap sind zur Bestimmung der Konfidenzintervalle der Zieldosis d^* auf die Parameterschätzer des angenommenen Modells angewiesen. In SAS 9.3 ist es möglich diese Parameterschätzer innerhalb einer SAS/IML Umgebung mit Hilfe der Prozedur NLMIXED zu bestimmen.

Literatur

- [1] S. Beal: Sample Size Determination for Confidence Intervals on the Population Mean and on the Difference Between Two Population Means. *Biometrics*, Vol. 45, No. 3: 969-977, 1989.
- [2] N. Benda, T. Friede: Sample Size Calculation and Recalculation in Dose Finding Studies with Active Control Sample Size Reassessment in Inverse Regression. 55. *Biometrisches Kolloquium*, Hannover, 2009.
- [3] B. Efron: Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics*. 7, No. 1, 1–26, 1979.
- [4] R. Fisch, G. Strehlau: A Simplified Approach to Calibration Confidence Sets. *The American Statistician*, 47 No. 3: 168-171, 1993.
- [5] H. J. Helms, N. Benda, T. Friede: Analysis and sample size calculation in dose-finding studies with active control. 58. *Biometrisches Kolloquium*, Berlin, 2012.
- [6] M. Jiroutek, K. Muller, L. Kupper, P. Stewart: A New Method for Choosing Sample Size for Confidence Interval-Based Inferences. *Biometrics*, Vol. 59, No. 3: 580-590, 2003.
- [7] A. M. Mattai, S. B. Provost: Quadratic forms in random variables: Theory and applications. Marcel Dekker, INC., New York, 1992.