

Berechnung von Konfidenzintervallen für das relative Risiko in SAS

Daniela Zöller
Institut für Medizinische
Biometrie, Epidemiologie und
Informatik (IMBEI)
Johannes-Gutenberg-
Universität Mainz
D-51101 Mainz
Daniela.Zoeller@uni-mainz.de

Ralf Bender
Institut für Qualität und
Wirtschaftlichkeit im
Gesundheitswesen (IQWiG)
Im Mediapark 8
D-50670 Köln
Ralf.Bender@iqwig.de

Zusammenfassung

Das relative Risiko (RR) ist ein wichtiges Effektmaß in der Medizinischen Statistik und der Epidemiologie. Zur Berechnung von Konfidenzintervallen für das RR stehen eine Reihe verschiedener Methoden zur Verfügung. Für viele Situationen in der Praxis sind insbesondere Methoden zur Erstellung exakter Konfidenzintervalle interessant, welche die Einhaltung des nominalen Niveaus garantieren. Eine Möglichkeit zur Berechnung exakter Konfidenzintervalle für das RR ist in SAS ab Version 9.3 in der Prozedur PROC FREQ implementiert. Diese beruht auf der Invertierung zweier einseitiger unbedingter exakter Tests zum Niveau $\alpha/2$, wodurch man ein exaktes Konfidenzintervall zum Niveau $1 - \alpha$ erhält. Dabei kann sowohl das nicht standardisierte RR, als auch die Miettinen-Nurminen-Score-Statistik als Ordnungsstatistik verwendet werden. Letztere ist auch als Farrington-Manning-Score-Statistik bekannt und wurde von Chan und Zhang (1999) für die Verwendung in unbedingten exakten Test für das RR empfohlen. In diesem Beitrag wird – nach einem groben Überblick über die zugrundeliegende Theorie – erläutert, wie PROC FREQ zur Berechnung exakter Konfidenzintervalle für das RR verwendet werden kann. Zudem wird eine kurze Analyse der praktischen Anwendbarkeit bezogen auf die Laufzeit der Prozedur dargestellt. Aufgrund der Tatsache, dass die benötigte Rechenzeit sehr hoch sein kann, wird kurz auf alternative nicht exakte Methoden zur Erstellung von Konfidenzintervallen für das RR eingegangen, welche kürzere Laufzeiten aufweisen und beispielsweise mit Hilfe der Prozedur PROC FCMP implementiert werden können. Die hier verwendeten Methoden sind die Miettinen-Nurminen-Score-Methode und die MOVER-R-Methode, die sich in Untersuchungen als verlässliche Methoden mit angemessenem Informationsgehalt erwiesen haben.

Schlüsselwörter: Relatives Risiko (RR), Epidemiologie, Konfidenzintervalle, exakte Methoden, asymptotische Methoden, PROC FREQ, PROC FCMP

1 Einleitung

In der Medizinischen Statistik ist das relative Risiko (RR) eines der zentralen Effektmaße bei der Analyse von Vierfeldertafeln (Agresti, 2002; Newcombe, 2013). Das RR wird sowohl in klinischen als auch in epidemiologischen Studien sehr häufig verwendet,

um den Effekt einer Therapie oder einer Exposition auf eine binäre Zielvariable im Vergleich mit einer Kontrolltherapie bzw. unter Nicht-Exposition zu beschreiben. Standardmäßig werden in der Praxis neben dem Punktschätzer auch Konfidenzintervalle für das RR berechnet. Dabei existieren sowohl exakte als auch asymptotische Verfahren um Konfidenzintervalle für das RR zu berechnen. Eine Möglichkeit zur Konstruktion von exakten $(1-\alpha)$ -Konfidenzintervallen ist die Invertierung zweier einseitiger unbedingter exakter Tests zum Niveau $\alpha/2$ (Chan und Zhang, 1999). Diese Möglichkeit ist seit Version 9.3 in PROC FREQ in SAS implementiert, wobei zwei verschiedene Ordnungsstatistiken und somit zwei verschiedene exakte Konfidenzintervalle zur Auswahl stehen. Exakte Konfidenzintervalle garantieren zwar ein hohes Konfidenzniveau, allerdings kann die Berechnung dieser Intervalle sehr aufwändig sein, wodurch die Anwendbarkeit in der Praxis eingeschränkt ist. Alternativ kann auf asymptotische Konfidenzintervalle zurückgegriffen werden. Diese können zwar ein antikonservatives Verhalten aufweisen, sind jedoch leicht berechenbar und haben im Allgemeinen einen höheren Informationsgehalt als exakte Verfahren. In SAS ist lediglich die gewöhnliche Wald-Methode (Agresti und Min, 2005; Koopman, 1984), beruhend auf der Normalapproximation des logarithmierten RR, in der Prozedur PROC FREQ als „asymptotisches Konfidenzintervall“ implementiert. In diesem Beitrag soll aus diesem Grund eine Möglichkeit zur Implementierung asymptotischer Verfahren mit Hilfe der Prozedur PROC FCMP, welche ab Version 9.2 Teil von BASE SAS ist, aufgezeigt werden. Veranschaulicht wird dies mit Hilfe zweier asymptotischer Methoden zur Bestimmung von Konfidenzintervallen: der Miettinen-Nurminen-Score-Methode (Miettinen und Nurminen, 1985) und der MOVER-R-Methode (Zou und Donner, 2008; Newcombe, 2013). Diese beiden Methoden liefern erfahrungsgemäß verlässliche Intervalle mit einem angemessenen Informationsgehalt.

2 Theoretische Grundlagen und Methoden

2.1 Wahrscheinlichkeiten und das RR in einer Vierfeldertafel

		Ereignis		Summe
		eingetreten	nicht eingetreten	
Gruppe	exponiert	A	$C = n_1 - A$	n_1
	nicht-exponiert	B	$D = n_2 - B$	n_2
Summe		$K := A + B$	$C + D$	$N = n_1 + n_2$

Abbildung 1: Absolute Häufigkeiten in einer Vierfeldertafel

Um den Effekt einer Therapie oder einer Exposition auf eine binäre Zielvariable wie das Auftreten eines Ereignisses im Vergleich mit einer Kontrolltherapie bzw. unter Nicht-Exposition zu untersuchen, können die Anzahlen an Ereignissen verwendet werden, die in den beiden Gruppen beobachtet wurden. Diese seien im Folgenden zwei unabhängige Zufallsvariablen, die durch A im Falle der exponierten Gruppe und B im Falle der Kon-

trollgruppe gegeben sind. A sei binomialverteilt mit dem Stichprobenumfang n_1 und der Wahrscheinlichkeit bzw. dem Risiko π_1 und B sei binomialverteilt mit dem Stichprobenumfang n_2 und der Wahrscheinlichkeit bzw. dem Risiko π_2 . Diese Daten lassen sich in einer Vierfeldertafel der absoluten Häufigkeiten zusammenfassen (Vgl. Abbildung 1).

Das RR ist dann gegeben durch das Verhältnis der Risiken ($RR = \pi_1/\pi_2$) und nimmt Werte ≥ 0 an. Gilt $RR=1$, stimmt dies mit der Aussage überein, dass sich die Risiken in den beiden Gruppen nicht unterscheiden. Ein Wert kleiner als 1 bedeutet, dass sich das Risiko bei Exposition verringert, ein Wert größer als 1, dass es sich vergrößert. Hier ist zu beachten, dass der Parameterbereich des RR nach oben unendlich ist.

Zusätzlich gilt, dass die gemeinsame Wahrscheinlichkeit der Beobachtung $A=a$ und $B=b$ in diesem Falle durch das Produkt der binomialen Wahrscheinlichkeiten gegeben ist. Dies lässt sich reparametrisieren gemäß $\pi_2 = \pi_1/RR$, wodurch die gemeinsame Wahrscheinlichkeit für $A=a$ und $B=b$ gegeben das Risiko π_1 und das RR definiert ist durch:

$$P(A = a, B = b | RR, \pi_1) = \binom{n_1}{a} \pi_1^a (1 - \pi_1)^{n_1 - a} \binom{n_2}{b} \left(\frac{\pi_1}{RR} \right)^b \left(1 - \left(\frac{\pi_1}{RR} \right) \right)^{n_2 - b}.$$

2.2 Konfidenzintervalle

Konfidenzintervalle werden in der medizinischen Statistik standardmäßig zusätzlich zur Punktschätzung für einen Parameter verwendet. Punktschätzer liefern lediglich einen Wert um den Parameter zu schätzen, die Genauigkeit der Schätzung ist hierbei nicht ersichtlich. Bei der Konstruktion eines Konfidenzintervalls versucht man, den Schätzfehler miteinzubeziehen und ein Intervall zu konstruieren, welches den wahren Wert des Parameters mit einer vorher definierten Wahrscheinlichkeit von $(1-\alpha)$ überdeckt. Hierbei wird $(1-\alpha)$ als nominales Niveau des Konfidenzintervalls bezeichnet (Hartung et al., 2002). Man unterscheidet zwischen exakten und asymptotischen Konfidenzintervallen. Für ein exaktes Konfidenzintervall ist die minimale Überdeckungswahrscheinlichkeit gleich dem nominalen Niveau, für asymptotische Konfidenzintervalle gilt dies nur für unendlich große Stichprobenumfänge.

2.3 Exakte Konfidenzintervalle

Die Berechnung von exakten Konfidenzintervallen für das RR in der Prozedur PROC FREQ von SAS basiert auf der Invertierung zweier exakter unbedingter Tests jeweils zum Niveau $\alpha/2$. Dabei wird ausgenutzt, dass die Akzeptanzregion eines statistischen Tests zum Niveau α einem $(1-\alpha)$ -Konfidenzintervall entspricht. Verwendet man zwei einseitige Tests zum Niveau $\alpha/2$, so bildet die Schnittmenge der Akzeptanzregionen ein $(1-\alpha)$ -Konfidenzintervall. Die hier verwendete Version ist angelehnt an Chan und Zhang (1999).

Als Nullhypothesen für die exakten Tests definiert man $H_{\geq}(rr_0): RR \geq rr_0$ und $H_{\leq}(rr_0): RR \leq rr_0$. Für die entsprechenden Testentscheidungen werden die exakten p-Werte verwendet. Ein p-Wert ist definiert als die Wahrscheinlichkeit, unter der Nullhypothese das beobachtete oder ein extremeres Ergebnis zu erhalten. Die Einstufung aller möglichen Ergebnisse erfolgt hier mittels einer Ordnungsstatistik $D(a, b)$. In der Prozedur PROC FREQ stehen zwei alternative Ordnungsstatistiken zur Verfügung: das nicht standardisierte RR und die Mietinen-Nurminen-Score-Statistik. Das nicht standardisierte RR wird, falls a und b größer als 0 sind, berechnet durch $D(a, b) = an_2/bn_1$. Ist a oder b gleich 0, so wird in SAS automatisch zu jedem der Werte in der Formel 0,5 addiert. Die Mietinen-Nurminen-Score- oder auch kurz Score-Statistik wurde von Mietinen und Nurminen (1985) entwickelt, von Farrington und Manning (1990) aufgegriffen und für die Verwendung in exakten Tests von Chan und Zhang (1999) vorgeschlagen. Hierbei ist zu beachten, dass theoretisch auch diverse andere Statistiken denkbar sind. Die Score-Statistik lautet:

$$D(a, b) = \frac{\widehat{\pi}_1 - rr_0 \widehat{\pi}_2}{\sqrt{\frac{n_1+n_2}{n_1+n_2-1} \left(\frac{r_1(1-r_1)}{n_1} + rr_0^2 \frac{r_2(1-r_2)}{n_2} \right)}}$$

wobei $\widehat{\pi}_1$ und $\widehat{\pi}_2$ die Maximum-Likelihood-Schätzer und r_1 und r_2 die beschränkten Maximum-Likelihood-Schätzer gemäß der Beschränkung $\pi_2 = \pi_1/RR$ der Risiken sind. Laut Mietinen und Nurminen (1985) sind letztere unter der Nullhypothese gegeben durch

$$r_1 = \frac{(n_1+b)rr_0 - \sqrt{(n_1+b)^2 rr_0^2 - 4rr_0(a+b)}}{2(n_1+n_2)} \text{ und } r_2 = \frac{r_1}{rr_0}.$$

Im Folgenden sei der Wert der Ordnungsstatistik, der aus den beobachteten Werten resultiert, durch $D(a_{obs}, b_{obs})$ gegeben. Wäre π_1 bekannt, so könnten mit Hilfe der in Abschnitt 2.1 genannten gemeinsamen Wahrscheinlichkeit $P(A = a, B = b | RR, \pi_1)$ und der Ordnungsstatistik $D(a, b)$ die exakten p-Werte für die Tests der oben genannten Nullhypothesen wie folgt bestimmt werden. Dabei ist zu beachten, dass im Falle der Hypothese $H_{\geq}(rr_0)$ gleiche oder kleinere Werte von $D(a, b)$ als der beobachtete $D(a_{obs}, b_{obs})$ und im Falle der Hypothese $H_{\leq}(rr_0)$ gleiche oder größere Werte von $D(a, b)$ gegen die Nullhypothese sprechen, wobei dies für beide in PROC FREQ möglichen Ordnungsstatistiken gilt. Somit sind die p-Werte in Abhängigkeit von π_1 definiert als

$$p_{\geq}(rr_0, \pi_1) = \sum_{D(a,b) \leq D(a_{obs}, b_{obs})} P(A = a, B = b | rr_0, \pi_1) \text{ und}$$

$$p_{\leq}(rr_0, \pi_1) = \sum_{D(a,b) \geq D(a_{obs}, b_{obs})} P(A = a, B = b | rr_0, \pi_1),$$

wobei $p_{\geq}(rr_0, \pi_1)$ der p-Wert zugehörig zum Test von $H_{\geq}(rr_0)$ und $p_{\leq}(rr_0, \pi_1)$ der p-Wert zugehörig zum Test von $H_{\leq}(rr_0)$ ist. Aufgrund der Tatsache, dass π_1 jedoch unbekannt ist, wird das Supremum der p-Werte über das Intervall $[0,1]$, also über alle möglichen Werte von π_1 , gebildet. Dadurch wird der Test gegebenenfalls konservativer,

die Einhaltung des Niveaus wird jedoch garantiert. Die p-Werte lauten, mit obiger Kennzeichnung,

$$p_{\geq}(rr_0) = \sup\{p_{\geq}(rr_0, \pi_1): \pi_1 \in [0,1]\} \text{ und}$$

$$p_{\leq}(rr_0) = \sup\{p_{\leq}(rr_0, \pi_1): \pi_1 \in [0,1]\}.$$

Die Nullhypothese wird genau dann abgelehnt, wenn der entsprechende p-Wert kleiner als das definierte Niveau $\alpha/2$ ist. Wie oben bereits genannt, führt die Invertierung dieser beiden einseitigen Tests zu einem exakten $(1 - \alpha)$ -Konfidenzintervall. Dieses enthält alle rr_0 , für die die entsprechenden Nullhypothesen von beiden Tests angenommen werden. Bei den p-Werten kann allerdings weder von Stetigkeit noch von Monotonie ausgegangen werden. Aus diesem Grund muss der gesamte Parameterraum des RR untersucht werden, folglich ganz \mathbb{R}^+ . Für die untere Grenze wird rr_0 von 0 ausgehend mit sehr kleinen Schrittweite genau so lange erhöht, bis die Bedingung $p_{\leq}(rr_0) \geq \alpha/2$ erfüllt ist. Der Wert von rr_0 entspricht dann der unteren Grenze. Die Schrittweite bestimmt die Genauigkeit der Grenze. Für die obere Grenze beginnt man bei einem sehr großen Wert von rr_0 und verringert diesen mit derselben Schrittweite solange, bis $p_{\geq}(rr_0) \geq \alpha/2$. Durch diese Vorgehensweise wird sichergestellt, dass das Niveau eingehalten wird. Allerdings ist hierdurch davon auszugehen, dass die Methode konservativer wird. Formal ist das $(1 - \alpha)$ -Konfidenzintervall $[L, U]$ definiert als

$$L := \sup\{rr_0: (p_{\leq}(rr_0) \geq \alpha/2) \wedge (\forall i < rr_0: p_{\leq}(i) < \alpha/2)\} \text{ und}$$

$$U := \sup\{rr_0: (p_{\geq}(rr_0) \geq \alpha/2) \wedge (\forall i > rr_0: p_{\geq}(i) < \alpha/2)\}.$$

Die Bestimmung des Konfidenzintervalls ist sehr aufwändig. Der Rechenaufwand steigt meist mit wachsenden Gesamtstichprobenumfang stark an. Aus diesem Grund ist dieses Intervall in der Praxis nur für kleine Stichproben empfohlen.

2.4 Score-Methode

Grundlage dieser Methode ist die Invertierung eines zweiseitigen Tests, welcher auf der in Abschnitt 2.3 beschriebenen Score-Statistik von Mietinen und Nurminen (1985) beruht. Die Nullhypothese ist dabei definiert durch $H_0: RR = rr_0$. Unter dieser Nullhypothese ist die Score-Statistik für $(n_1 + n_2) \rightarrow \infty$ ungefähr standardnormalverteilt. Daraus lässt sich ein asymptotischer Test zum Niveau α herleiten. Ist der Wert der Score-Statistik für die beobachteten Werte größer als das $(1 - \alpha/2)$ -Quantil $z_{1-\alpha/2}$ der Standardnormalverteilung, so wird H_0 abgelehnt. Das zugehörige Konfidenzintervall zum Niveau $1 - \alpha$ besteht aus allen rr_0 , welche die Ungleichung $D(a_{obs}, b_{obs}) \leq z_{1-\alpha/2}$ erfüllen, wobei $D(a_{obs}, b_{obs})$ der beobachtete Wert der Score-Statistik ist. Dabei ist die Existenz der Grenzen des Konfidenzintervalls für jede mögliche Beobachtung gegeben. So führt $a = 0$ zur unteren Grenze 0 und einer echt positiven oberen Grenze. Analog folgt aus $b = 0$, dass die obere Grenze gegen ∞ iteriert. Hier hat die untere Grenze einen reellen, nicht negativen Wert $\neq \infty$. Dies kann damit erklärt werden, dass man den Term $D(a_{obs}, b_{obs}) - z_{1-\alpha/2}$ als eine nach oben geöffnete Betragsfunktion in

Abhängigkeit von $D(a_{obs}, b_{obs})$ auffassen kann, welche die x-Achse genau zweimal schneidet. Die Schnittpunkte geben den Wert der Grenzen an. Ist $a = 0$ so liegt der kleinere Schnittpunkt im negativen Bereich. Mit der Begründung, dass das RR keine negativen Werte annehmen kann, wird die untere Grenze deshalb auf 0 gesetzt. Ist $b = 0$, gilt eine analoge Argumentation. Wurden in beiden Gruppen Ereignisse beobachtet, so liegen beide Schnittpunkte im positiven Bereich.

Für die Score-Methode existiert keine geschlossene Formel. Die obige Ungleichung muss numerisch gelöst werden. Laut Newcombe (2013) genügen allerdings 40 Iterationen eines Intervallhalbierungsverfahrens, um eine für die Praxis ausreichende Genauigkeit zu erreichen. Hierbei sollte allerdings beachtet werden, dass eine geeignete Transformation für das RR verwendet werden sollte, um das Verfahren durchzuführen. Eine solche ist gegeben durch $Q_{RR} = \frac{RR-1}{RR+1} \in [-1,1]$. Nur so lässt sich der Parameterraum gleichmäßig abschreiten.

2.5 MOVER-R-Methode

Zou und Donner (2008) haben einen generellen Ansatz entwickelt, um asymptotische Konfidenzintervalle für Differenzen zweier Parameter zu erstellen. Dieser Ansatz verwendet Konfidenzintervalle für einzelne Parameter, in welche die Varianz der Parameter eingegangen ist, um die Konfidenzintervalle für die Differenz der Parameter zu ermitteln. Dies ermöglicht die genauere Schätzung der Varianz der Parameter in der Nähe der zu erwartenden Grenzen für die Differenz. Aus diesem Grund ist die Methodik auch unter dem Namen „method of variance estimates recovery“, kurz MOVER-Methode, bekannt geworden. Dieser allgemeine Ansatz kann auch verwendet werden, um ein Konfidenzintervall für das RR basierend auf asymptotischen Konfidenzgrenzen für die einzelnen Risiken zu konstruieren (Newcombe, 2013). Dabei wird basierend auf der MOVER-Methode zunächst ein Konfidenzintervall für die Funktion $z_{rr_0} = \widehat{\pi}_1 - rr_0 \widehat{\pi}_2$ der Maximum-Likelihood-Schätzer der Risiken erstellt. Anschließend wird ausgenutzt, dass diese Funktion gleich 0 ist, wenn das wahre RR gleich rr_0 ist. Die vollständige Erläuterung kann Newcombe (2013) entnommen werden. Sind die Konfidenzintervalle für π_1 und π_2 gegeben durch $[l_1, u_1]$ und $[l_2, u_2]$, so ist das $(1 - \alpha)$ -Konfidenzintervall $[L, U]$ gegeben durch:

$$L = \frac{\widehat{\pi}_1 \widehat{\pi}_2 - \sqrt{(\widehat{\pi}_1 \widehat{\pi}_2)^2 - l_1 u_2 (2\widehat{\pi}_1 - l_1)(2\widehat{\pi}_2 - u_2)}}{u_2 (2\widehat{\pi}_2 - u_2)} \text{ und}$$

$$U = \frac{\widehat{\pi}_1 \widehat{\pi}_2 - \sqrt{(\widehat{\pi}_1 \widehat{\pi}_2)^2 - u_1 l_2 (2\widehat{\pi}_1 - u_1)(2\widehat{\pi}_2 - l_2)}}{l_2 (2\widehat{\pi}_2 - l_2)}.$$

Diese Methode wird mit MOVER-R-Methode bezeichnet, wobei das R für Ratio steht.

Generell gilt, dass die Güte des Verfahrens von der Wahl der Konfidenzintervalle für die Risiken abhängt. Umso besser diese Intervalle sind, umso besser ist die Schätzung der Stellen, in deren Nähe die Varianz geschätzt wird. Dadurch wird die Varianzschätzung genauer und somit auch die entsprechende Intervallschätzung. Von Newcombe

(2013) wird empfohlen, das Wilson-Score-Intervall zu verwenden, um die Konfidenzgrenzen für die Risiken zu bestimmen. Diese Methode gehört zu den besten Methoden, um Konfidenzintervalle für binomiale Wahrscheinlichkeiten zu erhalten (Newcombe, 2013). Das Wilson-Score-Intervall liefert immer sinnvolle reelle Grenzen. Dies führt dazu, dass unter Verwendung des Wilson-Score-Intervalls, die MOVER-R-Methode immer berechenbar ist. Ist $a = 0$ und somit $l_1 = 0$ so folgt $L = 0$ und $U \geq 0$. Ist $b = 0$, folgt $l_2 = 0$. Nach Newcombe (2013) gilt dann $U = \infty$.

3 Umsetzung in SAS

3.1 Verschiedene verwendbare Dataset-Strukturen

Liegen die Daten in Form einer Vierfeldertafel mit absoluten Häufigkeiten so wie sie in Abbildung 1 gezeigt ist vor, so besteht das zugehörige Dataset aus drei Variablen und vier Beobachtungen. Eine der Variablen enthält die Kodierung der Gruppe und eine die Kodierung des Ereignisses. Im Folgenden seien diese beiden binären Variablen durch „gruppe“ und „ereignis“ bezeichnet, wobei „gruppe=1“ der exponierten Gruppe und „ereignis=1“ dem Fall, dass ein Ereignis eingetreten ist entspricht. Die dritte Variable enthält die entsprechenden beobachteten Häufigkeiten und sei mit „anz“ für Anzahl bezeichnet. Ein so aufgebautes Dataset sei mit „Vierfeldertafel-Dataset“ bezeichnet. Ein Beispiel findet sich in Abbildung 2. Hier besteht jeder Gruppe aus 50 Individuen. In der Gruppe der Exponierten wurde kein Ereignis, in der Gruppe der Nicht-Exponierten wurden 5 Ereignisse beobachtet.

	anz	gruppe	ereignis
1	0	1	1
2	50	1	2
3	5	2	1
4	45	2	2

Abbildung 2: Beispieldataset beruhend auf einer Vierfeldertafel

Sind die Daten noch in keiner Vierfeldertafel zusammengefasst sondern für jedes Individuum liegt ein Datensatz mit den beiden binären Variablen „gruppe“ und „ereignis“ vor, so wird das Dataset, das diese Datensätze enthält, mit „Individuum-Dataset“ bezeichnet. Eine weitere Möglichkeit ist, dass lediglich die relevanten Größen, die Realisierungen der Zufallsvariablen A und B, also die Anzahl an Ereignissen, in den Gruppen, in einem Dataset zusammengefasst sind. Hier besteht das Dataset aus zwei Variablen, die den Realisierungen a und b entsprechen, und einer einzigen Beobachtung. Ein solches Dataset wird hier „Ereignisanzahl-Dataset“ genannt. Zu beachten ist, dass mit Hilfe der Prozedur PROC FREQ und der Prozedur PROC TRANSPOSE aus einem Individuum-Dataset ein Ereignisanzahl-Dataset erstellt werden kann. Der zugehörige SAS-Code könnte beispielsweise wie folgt aussehen:

```
PROC FREQ DATA=Individuum-Dataset
          TABLES ereignis*gruppe / OUT=FreqCount (DROP=PERCENT);
RUN;
PROC TRANSPOSE DATA=FreqCount OUT=Ereignisanzahl-Dataset PREFIX=anz;
          VAR count;
RUN;
```

3.2 Exakte Konfidenzintervalle

3.2.1 Implementierung in PROC FREQ ab SAS Version 9.3

Seit der Version 9.3 von SAS ist es möglich, mit Hilfe der Prozedur PROC FREQ exakte Konfidenzintervalle für das RR basierend auf den in Abschnitt 2 dargestellten Grundlagen zu erstellen. Dabei kann sowohl ein Vierfeldertafel-Dataset als auch ein Individuum-Dataset als Datengrundlage verwendet werden. Es stehen in PROC FREQ zwei Ordnungsstatistiken zur Verfügung, und zwar das nicht standardisierte RR und die Miettinen-Nurminen-Score-Statistik. Soll das exakte 95%-Konfidenzintervall auf Basis eines Individuum-Dataset mit dem nicht standardisierten RR als Ordnungsstatistik berechnet werden, so lautet der SAS-Code:

```
PROC FREQ DATA=Individuum-Dataset
          TABLES ereignis*gruppe / ALPHA=0.05;
          EXACT RELRISK;
RUN;
```

Dabei wird die Berechnung des exakten Konfidenzintervalls für das RR durch das Statement „EXACT RELRISK“ bewirkt und mit Hilfe von „ALPHA=“ kann der Wert von α und somit das nominale Niveau des Konfidenzintervalls angepasst werden. Soll die Miettinen-Nurminen-Score-Statistik als Ordnungsstatistik dienen, so wird „EXACT RELRISK“ durch „EXAKT RELRISK(METHOD=FMSCORE)“ ersetzt. Dabei steht „FMSCORE“ für Farrington-Manning-Score-Statistik, die der Miettinen-Nurminen-Score-Statistik entspricht. Um Vierfeldertafel-Datasets verwenden zu können, wird der Prozeduraufruf um die Zeile „WEIGHT anz;“ erweitert. In Abbildung 3 findet sich ein Beispiel für den relevanten Teil des Output, das auf den Daten der Abbildung 2 und der Miettinen-Nurminen-Score-Statistik als Ordnungsstatistik beruht. Die exakten Konfidenzgrenzen werden von SAS als letztes ausgegeben, wobei das nominale Niveau des Konfidenzintervalls im Output erkennbar ist. Zudem ist im Output ersichtlich, welche Ordnungsstatistik verwendet wurde. Wurde die Miettinen-Nurminen-Score-Statistik verwendet, so wird, wie in Abbildung 3 zu sehen, an den Output im Englischen die Zeile „The exact confidence limits are based on the FM Score Statistik.“ angefügt. Bei Verwendung der nicht standardisierten RR-Statistik, die als Default in SAS eingestellt ist, entfällt dieser Zusatz. Anzumerken ist an dieser Stelle, dass SAS neben den exakten Konfidenzgrenzen auch asymptotische Grenzen für das RR angibt. Diese asymptotischen Grenzen werden auf Grundlage der gewöhnlichen Wald-Methode berechnet, die

auf der Normalapproximation des logarithmierten RR beruht. Diese Methode kann erfahrungsgemäß allerdings stark antikonservativ sein (Newcombe, 2013).

Relative Risk (Column 1)	
Relative Risk	0.0000
Asymptotic Conf Limits	
95% Lower Conf Limit	.
95% Upper Conf Limit	.
Exact Conf Limits	
95% Lower Conf Limit	0.0000
95% Upper Conf Limit	0.9320
The exact confidence limits are based on the FM score statistic.	

Abbildung 3: Ausschnitt aus dem Beispieloutput der Prozedur PROC FREQ für die exakte Methode mit der Score-Statistik als Ordnungsstatistik

3.2.2 Anmerkung zur Laufzeit der exakten Methode mit PROC FREQ

Verwendet man das Statement „EXACT RELRISK“ bzw. „EXACT RELRISK(METHOD=FMSCORE)“ im Aufruf der Prozedur PROC FREQ um ein exaktes Konfidenzintervall für das RR zu berechnen, so ist die benötigte Rechenzeit vorab nur schwer einschätzen. Es existiert keine Formel, die die benötigte Rechenzeit angibt. Generell hängt die Rechenzeit sowohl von den Stichprobenumfängen als auch von den beobachteten Ereignisanzahlen und der verwendeten Ordnungsstatistik ab und kann, insbesondere für große Stichprobenumfänge, sehr hoch sein. Eine Veranschaulichung findet sich in Tabelle 1. In dieser Tabelle ist die von SAS gemessenen USER-CPU-Zeiten für verschiedene Wertekonstellationen eingetragen. Zu beachten ist, dass diese Zeiten sich für verschiedene Systeme, auf denen SAS ausgeführt wird, unterscheiden können. Die hier angegebenen Zeiten sollen lediglich einen Eindruck des Verhaltens der Prozedur vermitteln.

Mit Hilfe von Tabelle 1 wird deutlich, dass die Verwendung der exakten Methode in der Praxis gerade bei großen Stichprobenumfängen schwierig sein kann. Allerdings gibt es derzeit in SAS keine vorimplementierte Alternative, die exakte Konfidenzintervalle liefert und in Bezug auf die Laufzeit in jeder Situation vertretbar ist.

Tabelle 1: Gemessene USER CPU Zeiten von PROC FREQ

$n_1 + n_2$	n_1	n_2	a	b	USER CPU Zeit [mm:ss]	
					RR-Statistik	Score-Statistik
20	10	10	0	5	00:01	00:01
100	50	50	1	0	00:04	01:32
100	50	50	49	48	07:24	00:05
250	50	200	11	12	00:00	03:30
1000	500	500	0	1	01:47	18:31

3.3 Asymptotische Konfidenzintervalle

Wie in Abschnitt 2.3.2 gezeigt, kann die Laufzeit exakter Methoden insbesondere bei großen Stichprobenumfängen sehr hoch sein. Gerade in solchen Situationen bieten asymptotische Verfahren eine gute Alternative. Die Laufzeit asymptotischer Methoden ist im Allgemeinen sehr gering und zudem unabhängig von der Höhe der Stichprobenumfänge und den beobachteten Ereignisanzahlen. Allerdings ist in SAS in der Prozedur PROC FREQ lediglich die Wald-Methode zur Erstellung asymptotischer Konfidenzintervalle für das RR implementiert. Wie Newcombe (2013) herausstellt, gibt es jedoch bessere Methoden, die Konfidenzintervalle mit einer angemessenen Überdeckungswahrscheinlichkeit und einem hohen Informationsgehalt liefern. In diesem Zusammenhang sind sowohl die Miettinen-Nurminen-Score- als auch die MOVER-R-Methode, die sich in Abschnitt 2 erläutern lassen, zu nennen. Diese Methoden können mit Hilfe der Prozedur PROC FCMP vom Benutzer implementiert werden, sodass sie auf einen Ereignisanzahl-Dataset angewendet werden können.

3.2.1 Implementierung in PROC FCMP ab SAS Version 9.2

Die Prozedur PROC FREQ steht ab Version 9.2 in BASE SAS zu Verfügung. Mit ihrer Hilfe können eigene Funktionen und Routinen definiert werden, die in DATA-Steps, in der WHERE-Anweisung in verschiedenen Prozeduren wie PROC FCMP, PROC MCMC oder PROC SQL aufgerufen werden können. Im folgenden Codefragment findet sich ein Beispiel für PROC FCMP, indem die Deklaration einer Routine und einer Funktion im Pseudo-Code aufgezeigt wird.

```
PROC FCMP OUTLIB=SASUSER.tabellenname.versionsname;

/*Deklaration einer Subroutine mit mehreren Output-Argumenten */
SUBROUTINE routinename (input1,input2,...,output1,output2,...);
  OUTARGS output1,output2,...;
  output1 = funktion1(input1,input2,...);
  output2 = funktion2(input1,input2,...);
  .
  .
  .
ENDSUB;
```

```

/*Deklaration einer Funktion mit einem Output-Argument */
  FUNCTION funktionsname (input1,input2,...);
    output = funktion1(input1,input2,...);
    return(output);
  ENDSUB;

RUN;

```

Der SAS-Codes zur Deklaration der Routinen zur Bestimmung von Konfidenzintervallen nach der Score- oder der MOVER-R-Methode findet sich im Anhang. Zusätzlich kann der Code von der Internetseite <http://daniela-zoeller.de/programme.html> heruntergeladen werden. Die entsprechenden Konfidenzintervalle können, nachdem die zugehörigen Routinen mit Hilfe von PROC FCMP implementiert wurden, mit folgendem Code berechnet werden:

```

DATA ki;
  SET Ereignisanzahl-Dataset;
  L_mn_score=0;
  U_mn_score=0;
  L_mover_r=0;
  U_mover_r=0;
  CALL ki_rr_MN(a,b,&n1.,&n2.,0.05,20,L_mn_score,U_mn_score);
  CALL ki_rr_mover_r(a,b,&n1.,&n2.,0.05,L_mover_r,U_mover_r);

RUN;

```

Die Variable `L_mn_score` enthält dann den berechneten Wert für die untere Grenze des Score-Intervalls, `U_mn_score` die obere Grenze und `L_mover_r` und `U_mover_r` die entsprechenden Grenzen der MOVER-R-Methode.

3.3.2 Anmerkung zur Laufzeit der asymptotischen Methoden

Die Laufzeit der asymptotischen Methoden ist weitgehend unabhängig vom Stichprobenumfang und den beobachteten Ereigniszeiten. Dies zeigt die Tabelle 2, in der die USER CPU Zeiten für die asymptotischen Methoden aufgerufen in einem Data-Step für dieselben Situationen wie in Tabelle 1 aufgezeigt sind. Hier zeigt sich, dass die asymptotischen Methoden nahezu keinen Rechenaufwand haben und sofort ein Ergebnis liefern.

Tabelle 2: Gemessene USER CPU Zeiten der asymptotischen Methoden

$n_1 + n_2$	n_1	n_2	a	b	USER CPU Zeit [s]	
					MN-Score	MOVER-R
20	10	10	0	5	0.00	0.00
100	50	50	1	0	0.01	0.03
100	50	50	49	48	0.03	0.01
250	50	200	11	12	0.01	0.00
1000	500	500	0	1	0.00	0.01

4 Zusammenfassung

In SAS bieten sich mehrere Möglichkeiten um Konfidenzintervalle für das RR zu berechnen, wobei mit der Prozedur PROC FREQ seit Version 9.3 auch exakte Konfidenzintervalle bestimmt werden können. Diese bieten im Vergleich mit asymptotischen Methoden den Vorteil, dass sie die Einhaltung des Niveaus garantieren. Dabei stehen zwei Ordnungsstatistiken, die nicht standardisierte RR-Statistik und die Miettinen-Nurminen-Score-Statistik, für den zugrundeliegenden exakten Test zur Verfügung. Allerdings ist die Laufzeit abhängig von den Stichprobenumfängen und der Effektstärke bzw. den beobachteten Häufigkeiten. Zudem sind insbesondere für große Stichprobenumfänge hohe Laufzeiten erwartet werden, wobei die Laufzeit vorab nur schwer einschätzbar ist. Alternativ können asymptotische Verfahren mit geringer, nahezu konstanter Laufzeit verwendet werden. Als Beispiele sind die Miettinen-Nurminen-Score- und die MOVER-R-Methode zu nennen. Diese liefern oftmals adäquate Konfidenzintervalle und lassen sich einfach mit Hilfe der Prozedur PROC FCMP, die seit Version 9.2 in BASE SAS enthalten sind, umsetzen. Die Konfidenzintervalle lassen sich dann einfach in einem Data-Step berechnen. Die zugehörigen SAS-Codes können von der Internetseite <http://daniela-zoeller.de/programme.html> heruntergeladen werden.

Literatur

- [1] A. Agresti: Categorical data analysis (2nd). New York: Wiley, 2002
- [2] I. S. Chan, Z. Zhang: Test-based exact confidence intervals for the difference of two binomial proportions. *Biometrics* 1999, 55, 1202-1209
- [3] C. P. Farrington, G. Manning: Test statistics and sample size formulae for comparative binomial trials with null hypothesis of non-zero risk difference or non-unity relative risk. *Statistics in Medicine* 1990, 9, 1447-1454
- [4] J. Hartung, B. Elpelt, K. Klösener: Statistik Lehr und Handbuch der angewandten Statistik; mit zahlreichen, vollständigen druckgerechneten Beispielen. 13. München; Oldenburg, 2002
- [5] O. Miettinen, M. Nurminen: Comparative analysis of two rates. *Statistics in Medicine* 1985, 4, 513-517
- [6] R. G. Newcombe: Confidence intervals for proportions and related measures of effect size. Boca Raton, FL: CRC Press 2013 (Chapman and Hall: CRC biostatistics series)
- [7] G.Y. Zou, A. Donner: Construction of confidence limits about effect measures: A general approach. *Statistics in Medicine* 2008, 27, 1693-1702

Anhang A

PROC FCMP-Code zur Deklaration einer Subroutine zur Berechnung von Konfidenzintervallen nach der Miettinen-Nurminen-Score-Methode

```

*****
*   Zweiseitiges Konfidenzintervall für das Relative Risiko   *
*   Miettinen-Nurminen Score Intervall (1985)               *
*   Version Newcombe 2013, S. 209                           *
*   Iteratives Verfahren                                     *
*   Autor: D. Zoeller                                       *
*   E-Mail: daniela.zoeller@uni-mainz.de                    *
*   Datum: 31.08.2013                                       *
*****;
*-----*
| Input: (Im Aufruf feste Werte einsetzen)                  |
| a:      Anzahl Ereignisse in Gruppe 1                     |
| b:      Anzahl Ereignisse in Gruppe 2                     |
| n1:     Stichprobenumfang der Gruppe 1                   |
| n2:     Stichprobenumfang der Gruppe 2                   |
| alpha:  Niveau der Untersuchung                           |
| maxiter: Maximale Anzahl an Iterationen (iteratives Verfahren) |
|
| Output: (Diese Var. werden im Aufruf benannt, dann verändert) |
| l:      Untere Grenze des (1-alpha)-KIs                   |
| u:      Obere Grenze des (1-alpha)-KIs                   |
|
| Speicherort der Routinen an diesen Stellen anpassen:     |
| Festlegung: PROC FCMP OUTLIB=LIBNAME.TABNAME.PAKET      |
| Aufruf:     OPTIONS CMPLIB=LIBNAME.TABNAME              |
| In dieser Datei: 1XFestlegung                            |
|
| Aufruf der Subroutine:                                    |
| OPTIONS CMPLIB=Funktion.KIrr                             |
| CALL ki_rr_MN(a,b,n1,n2,alpha,maxiter,l,u)               |
*-----*

```

```
PROC FCMP OUTLIB=Funktion.KIrr.paket;
```

```

*-----*
|           Berechnung der Teststatistik                   |
|           Wird automatisch von ki_rr_MN aufgerufen       |
*-----*
| Input: (Im Aufruf feste Werte einsetzen)                |
| a:      Anzahl Ereignisse in Gruppe 1                     |
| b:      Anzahl Ereignisse in Gruppe 2                     |
| n1:     Stichprobenumfang der Gruppe 1                   |
| n2:     Stichprobenumfang der Gruppe 2                   |
| rr:     Auswertung an der Stelle rr (theta)              |
| it:     Iterationsnr. (Für log bei Wert<0 unter Wurzel) |
| grenze: 0(Untere), 1(Obere) (Für log bei Wert<0 unter Wurzel) |

```

```

|
| Output: (Diese Var. werden im Aufruf benannt, dann verändert)
|   z_rr:      Wert der Teststatistik an der Stelle rr
|   zwischen:  Wert unter der Wurzel (0 falls dieser negativ war)
|
|
| Aufruf:
|   OPTIONS CMPLIB=Funktion.KIrr
|   CALL t_rr_MN_sub (a,b,n1,n2,rr,it,grenze,z_rr,zwischen)
*-----*
SUBROUTINE t_rr_MN_sub (a,b,n1,n2,rr,it,grenze,z_rr,zwischen);
  OUTARGS z_rr, zwischen;

  B_stat = (n1 + b) * rr + a + n2;
  zwischen = B_stat*B_stat - (4 * rr * (n1 + n2) * (a + b));
  IF zwischen < 0 THEN zwischen = 0;
  r_1 = (B_stat - sqrt(zwischen))/(2*(n1 + n2));
  r_2 = r_1 / rr;
  V = ((r_1 * (1 - r_1) /n1) + (rr * rr * r_2 * (1 - r_2) /
n2)) * ((n1 + n2)/(n1 + n2 -1));
  z_rr = abs((a / n1) - (rr * (b / n2))) / sqrt(V);
ENDSUB;
*-----*
|
|          Berechnung der Grenzen
|          Eigentliche Routine
|
*-----*
SUBROUTINE ki_rr_MN (a,b,n1,n2,alpha,maxiter,l,u);
  OUTARGS l, u;
  /*Berechnung der unteren Grenze des KIs*/
  IF a=0 THEN DO;
    l = 0;
  END; ELSE DO;
  /*Berechnung für q_rr = (rr-1)/(rr+1) (in [-1,1])*/
  /*Numerische Lösung über Intervallschachtelung (binäre Suche)*/
  *Startintervall nach Newcombe 2013, oddsratioandrr.xls;
  l_1 = -0.9999999999;
  l_2 = 0.9999999999;
  IF a=n1 and b=n2 THEN DO;
    IF n1>n2 THEN DO;
      l_2 = 0.0000000001;
    END; ELSE DO;
      l_2 = -0.0000000001;
    END;
  END; ELSE IF b=0 THEN DO;
    l_2 = 0.9999999999;
  END; ELSE DO;
    rr_dat = (a/n1)/(b/n2);
    l_2 = (rr_dat - 1)/(rr_dat + 1);
  END;
  diff=1; *Abstand z_rr und (1-alpha/2)-Quantil von N(0,1);
  q_rr_it = (l_1 + l_2)/2;
  theta_it = (1+q_rr_it)/(1-q_rr_it);

```

```

DO i = 1 to maxiter;
  q_rr_it = (l_1 + l_2)/2;
  theta_it = (1+q_rr_it)/(1-q_rr_it);
  CALL t_rr_MN_sub(a,b,n1,n2,theta_it,i,0,z_rr_it,zwischen);
  diff = z_rr_it - QUANTILE('normal',1-(alpha/2),0,1);
  IF (diff>0) THEN DO; *Linke Grenze ersetzen;
    l_1 = q_rr_it;
  END; ELSE DO; *Rechte Grenze ersetzen;
    l_2 = q_rr_it;
  END;
  i=i+1;
END;
l = theta_it; *Linke Grenze des KIs;
END;
/*Berechnung der oberen Grenze des KIs */
IF b=0 THEN DO;
  u = .I; *.I=inf;
END; ELSE DO;
  u_1 = -0.9999999999;
  u_2 = 0.9999999999;
  IF a=n1 and b=n2 THEN DO;
    IF n1>n2 THEN DO;
      u_1 = 0.0000000001;
    END; ELSE DO;
      u_1 = -0.0000000001;
    END;
  END; ELSE IF b=0 THEN DO;
    u_1 = -0.9999999999;
  END; ELSE DO;
    rr_dat = (a/n1)/(b/n2);
    u_1 = (rr_dat - 1)/(rr_dat + 1);
  END;
  diff=1;
  q_rr_it = (u_1 + u_2)/2;
  theta_it = (1+q_rr_it)/(1-q_rr_it);
  DO i = 1 to maxiter;
    q_rr_it = (u_1 + u_2)/2;
    theta_it = (1+q_rr_it)/(1-q_rr_it);
    CALL
t_rr_MN_sub(a,b,n1,n2,theta_it,i,1,z_rr_it,zwischen);
    diff = z_rr_it - QUANTILE('normal',1-(alpha/2),0,1);
    IF (diff>0) THEN DO;
      u_2 = q_rr_it;
    END; ELSE DO;
      u_1 = q_rr_it;
    END;
    i=i+1;
  END;
  u = theta_it;
END;
ENDSUB;
RUN;

```

Anhang B

PROC FCMP-Code zur Deklaration einer Subroutine zur Berechnung von Konfidenzintervallen nach der MOVER-R-Methode inkl. Wilson-Score-Methode zur Bestimmung von asymptotischen Konfidenzintervallen für binomiale Wahrscheinlichkeiten.

```

*****
*   Zweiseitiges Konfidenzintervall für eine binomiale W'keit   *
*   Wilson Score Interval 1927 (Newcombe 2013 S. 62)           *
*   Nicht-iteratives Verfahren                                 *
*   Autor: D. Zoeller                                         *
*   E-Mail: daniela.zoeller@uni-mainz.de                       *
*   Datum: 31.08.2013                                         *
*****;
*-----*
| Input: (Im Aufruf feste Werte einsetzen)                    |
| a:      Anzahl Ereignisse                                    |
| n:      Stichprobenumfang                                  |
| alpha:  Niveau der Untersuchung                            |
|                                                |
| Output: (Diese Var. werden im Aufruf benannt, dann verändert)|
| l:      Untere Grenze des (1-alpha)-KIs                     |
| u:      Obere Grenze des (1-alpha)-KIs                     |
|                                                |
| Speicherort der Routinen an diesen Stellen anpassen:      |
| Festlegung: PROC FCMP OUTLIB=LIBNAME.TABNAME.PAKET        |
| Aufruf:     OPTIONS CMPLIB=LIBNAME.TABNAME                 |
| In dieser Datei: 1XFestlegung                              |
|                                                |
| Aufruf der Subroutine:                                     |
| OPTIONS CMPLIB=Funktion.KIrr                               |
| CALL ki_p_wilson(a,n,alpha,l,u)                            |
*-----*
PROC FCMP OUTLIB=Funktion.KIrr.paket;
  SUBROUTINE ki_p_wilson (a,n,alpha,l,u);
    OUTARGS l,u;
    z = QUANTILE('normal',1-(alpha/2),0,1);
    /*Berechnung der unteren Grenze des KIs*/
    l = (2*a + z*z - z*sqrt(z*z + 4*n*(a/n)*(1-a/n)))/(2*(n +
z*z));
    /*Berechnung der oberen Grenze des KIs*/
    u = (2*a + z*z + z*sqrt(z*z + 4*n*(a/n)*(1-a/n)))/(2*(n +
z*z));
  ENDSUB;
RUN;

```



```

*****
*   Zweiseitiges Konfidenzintervall für das Relative Risiko   *
*   MOVER-R nach Zou, Donner & Taleban 2010                 *
*   Vgl. Newcombe 2013 S. 137                               *
*   Nicht-iteratives Verfahren                             *
*
*   !!!ACHTUNG!!!                                           *
*!!Dieses Verfahren benötigt ki_p_wilson(a,n,alpha,l,u)-Routine!! *
* Autor: D. Zoeller                                         *
* E-Mail: daniela.zoeller@uni-mainz.de                     *
* Datum: 31.08.2013                                        *
*****;
*-----*
| Input: (Im Aufruf feste Werte einsetzen)                 |
| a:      Anzahl Ereignisse in Gruppe 1                    |
| b:      Anzahl Ereignisse in Gruppe 2                    |
| n1:     Stichprobenumfang der Gruppe 1                  |
| n2:     Stichprobenumfang der Gruppe 2                  |
| alpha:  Niveau der Untersuchung                         |
|
| Output: (Diese Var. werden im Aufruf benannt, dann verändert) |
| l:      Untere Grenze des (1-alpha)-KIs                  |
| u:      Obere Grenze des (1-alpha)-KIs                  |
|
| Speicherort der Routinen an diesen Stellen anpassen:    |
| Festlegung: PROC FCMP OUTLIB=LIBNAME.TABNAME.PAKET     |
| Aufruf:     OPTIONS CMPLIB=LIBNAME.TABNAME             |
| In dieser Datei: 1XFestlegung, 1XAufruf (optional)     |
|
| Aufruf der Subroutine:                                  |
| OPTIONS CMPLIB=Funktion.KIrr                           |
| CALL ki_rr_mover_r(a,b,n1,n2,alpha,l,u)                 |
*-----*

OPTIONS    CMPLIB=Funktion.KIrr;    /*Gegebenenfalls ändern bzw.
entfernen*/

PROC FCMP OUTLIB=Funktion.KIrr.paket;
  SUBROUTINE ki_rr_mover_r(a,b,n1,n2,alpha,l,u);
    OUTARGS l,u;

    /*Berechnung der KI-Grenzen der einzelnen W'keiten, nach Wunsch
ersetzen*/
    CALL ki_p_wilson(a,n1,alpha,l_1,u_1);
    CALL ki_p_wilson(b,n2,alpha,l_2,u_2);

    /*MLE der binomialen Risiken pi_1 und pi_2*/
    p1 = a/n1;
    p2 = b/n2;

    /*Berechnung der unteren Grenze des KIs*/

```

```
IF a = 0 THEN l = 0;
ELSE l = ((p1*p2) - sqrt((p1*p1*p2*p2) - (l_1*u_2*((2*p1)-
l_1)*((2*p2)-u_2)))) / (u_2*((2*p2)-u_2));

/*Berechnung der oberen Grenze des KIs*/
IF b = 0 THEN u = .I; *.I=inf;
ELSE u = ((p1*p2) + sqrt((p1*p1*p2*p2) - (u_1*l_2*((2*p1)-
u_1)*((2*p2)-l_2)))) / (l_2*((2*p2)-l_2));

ENDSUB;
RUN;
```