

Ein Test auf kohärente Alternativen mit dem SAS-Makro POSET

Olaf Gefeller
Institut für Medizininformatik,
Biometrie und Epidemiologie
Universität Erlangen-Nürnberg
Waldstraße 6
91054 Erlangen
gefeller@imbe.med.uni-erlangen.de

Annette Pfahlberg
Institut für Medizininformatik,
Biometrie und Epidemiologie
Universität Erlangen-Nürnberg
Waldstraße 6
91054 Erlangen
pfahlberg@imbe.med.uni-erlangen.de

Zusammenfassung

Keine noch so ausgefeilte statistische Testprozedur zum Vergleich zweier nichtrandomisierter Behandlungsgruppen hinsichtlich der Verteilung von (einer) primären Zielvariablen kann das Fehlen der Randomisation als wichtiges Designmerkmal einer Studie ausgleichen. Nichtsdestotrotz gibt es einige Fortschritte bei der Entwicklung von statistischen Testprozeduren für nichtrandomisierte Studien, welche die Gleichheit zweier Gruppen gegen eine sogenannte kohärente Alternative testen. Das Konzept der Kohärenz kann man allgemein als Existenz eines konsistenten Musters in miteinander zusammenhängenden Zielvariablen beschreiben. Was eine kohärente Alternative konkret bedeutet, kann nur im Zusammenhang mit den Rahmenbedingungen einer Studie exakt angegeben werden. Ein einfacher nichtparametrischer statistischer Test gegen beliebig komplizierte kohärente Alternativen, der sogenannte Poset-Test, ist bei unterschiedlichen Skalenniveaus der Zielvariablen anwendbar und ist als Verallgemeinerung von üblichen Zweistichproben-Rangtests zu verstehen. Wir haben für seine Anwendung das SAS-Makro POSET entwickelt, um dieses methodische Vorgehen in das statistische Softwarepaket SAS zu integrieren. Die praktische Implementation und Anwendung des Makros wird in diesem Beitrag beschrieben. Das methodische Vorgehen wird dabei an einem Datensatz einer klinischen Therapiestudie illustriert.

Schlüsselworte: Poset, nichtrandomisierte Studien, nichtparametrische Testverfahren, SAS-Makro

1 Einleitung

Der Vergleich zwischen zwei (Therapie-)Gruppen in der klinischen Medizin sollte in randomisierten Studien anhand eines zuvor definierten primären Zielkriteriums erfolgen. In der Praxis sind allerdings sowohl bei der Art des Studiendesigns als auch bei der Fokussierung auf eine Zielvariable Fälle anzutreffen, die nicht der Idealvorstellung entsprechen. So finden Vergleiche therapeutischer und diagnostischer Vorgehensweisen auch in nichtrandomisierten Studien statt. Ebenso werden vielfach mehrere Zielgrößen zum Gruppenvergleich zugrunde gelegt, ohne dass eine explizite Festlegung auf ein primäres Zielkriterium erfolgt.

In dieser Arbeit wird ein auf den amerikanischen Statistiker Paul Rosenbaum zurückgehender Vorschlag vorgestellt, der sich dem Thema des statistischen Vergleichs von Therapiegruppen in nichtrandomisierten Studien widmet [1-3]. Er verwendet dazu das Konzept der Kohärenz, welches in zahlreichen älteren Publikationen [4-6] zur Beurteilung der Kausalität von Zusammenhängen diskutiert wird. Unterscheiden sich zwei Therapiegruppen kohärent, so spricht dies auch in nichtrandomisierten Studien mit all ihren Problemen für die Existenz einer auf die Therapie zurückzuführenden Differenz in der Wirkung der klinischen Vorgehensweise.

Das Konzept der Kohärenz kann man allgemein als Existenz eines konsistenten Musters in miteinander zusammenhängenden Zielvariablen beschreiben. Was eine kohärente Alternative konkret bedeutet, kann nur im Zusammenhang mit den inhaltlichen Rahmenbedingungen einer Studie exakt angegeben werden. Ein einfacher nichtparametrischer statistischer Test gegen beliebig komplizierte kohärente Alternativen, der sogenannte Poset-Test, ist bei unterschiedlichen Skalenniveaus der Zielvariablen anwendbar und ist als Verallgemeinerung von üblichen Zweistichproben-Rangtests zu verstehen.

In diesem Beitrag wird das zugrunde liegende statistische Vorgehen des Poset-Tests beschrieben und anhand eines klinischen Beispiels illustriert. Zur Anwendung des Verfahrens haben wir das SAS-Makro POSET entwickelt, um die Integration der Methodik in das statistische Softwarepaket SAS zu erreichen. Die praktische Implementation und Anwendung des Makros wird in einem eigenen Abschnitt detailliert erläutert.

2 Herleitung des Poset-Tests

Ausgangssituation ist die typische Zweistichprobensituation mit N Beobachtungen, die von $i=1, \dots, n$ nummeriert sind, wobei die Beobachtungen aus einem K -dimensionalen Vektor $y_i \in \mathbb{R}^K$, $i=1, \dots, N$, bestehen. Die ersten N_1 (d.h. y_1, \dots, y_{N_1}) gehören zur ersten Gruppe und die verbleibenden $N_2=N-N_1$ (d.h. y_{N_1+1}, \dots, y_N) repräsentieren die zweite Gruppe. Um eine kohärente Hypothese formal aufzuschreiben, muss zunächst eine Relation $<_c$ im \mathbb{R}^K definiert werden. Eine solche Relation $<_c$ ist stets asymmetrisch, d.h. aus $a <_c b$ folgt, dass $b <_c a$ falsch ist. Dennoch unterscheidet sich $<_c$ fundamental von einer normalen Ungleichung, da es zulässig ist, dass für $a, b \in \mathbb{R}^K$ keine der drei Bedingungen ($a <_c b$, $b <_c a$, $a = b$) zutrifft. Demzufolge sind nicht alle Vektoren durch $<_c$ anzuordnen, da hiermit nur eine partielle Ordnung im \mathbb{R}^K definiert ist.

Betrachtet wird nun zunächst eine Zufallsvariable U_{ij} , $i, j=1, \dots, N$, die folgendermaßen definiert ist:

$$U_{ij} = \begin{cases} -1 & \text{für } y_i <_c y_j \\ +1 & \text{für } y_j <_c y_i \\ 0 & \text{sonst} \end{cases}$$

welche als Indikatorvariable interpretiert werden kann. Sie gibt an, ob y_i und y_j durch $<_c$ angeordnet werden können und, falls ja, in welcher Weise. Durch die Asymmetrie gilt: $U_{ij} = -U_{ji}$ und $U_{ii} = 0$. Für alle festen i wird durch die Summe

$$S_i := \sum_{j=1}^N U_{ij}$$

ein Score für y_i definiert. Dieser lässt sich analog zu einer Rangzahl interpretieren: Er gibt die genaue Position von y_i unter allen Beobachtungen im Hinblick auf die partielle Ordnungsrelation $<_c$ an. Genauer gesagt, gibt der Wert von S_i an, wie viele Beobachtungen es mit kleineren (gemäß $<_c$) Werten als y_i gibt minus der Anzahl der Beobachtung, die größere Werte haben. Wenn zum Beispiel $S_i = N-1$ ist, dann haben alle anderen Beobachtungen kleinere Werte als y_i und es ist die größtmögliche Rangzahl erreicht. Wenn $S_i < N-1$ kann allerdings nicht automatisch geschlossen werden, dass es Beobachtungen mit größeren Werten als y_i gibt, da dies auch dadurch verursacht sein kann, dass nicht alle Beobachtungen durch $<_c$ anzuordnen waren. Mit diesen so definierten Rangzahlen kann nun ein statistischer Test gegen eine kohärente Alternative durchgeführt werden. Voraussetzung hierfür ist zusätzlich die Teststatistik

$$T = \sum_{i=1}^{N_1} S_i$$

mit der alle Rangzahlen in der ersten Gruppe aufsummiert werden. Unter der Nullhypothese $P(y_i <_c y_j) = P(y_j <_c y_i)$ ist $E(T)=0$ und die Varianz kann mit den Argumenten aus der Theorie der linearen Rangstatistiken unter H_0 angegeben werden als

$$\text{Var}(T) = \frac{N_1 N_2}{N(N-1)} \sum_{i=1}^N S_i^2.$$

Der zentrale Grenzwertsatz für Rangstatistiken ist hier anwendbar für T , so dass die standardisierte Version von T asymptotisch normalverteilt ist. So können die Quantile der Standardnormalverteilung verwendet werden, um die kritischen Werte für den Test anzugeben bzw. bei gegebener Teststatistik für einen Datensatz den zugehörigen p -Wert zu berechnen.

Das hier vorgestellte Testverfahren wurde als sogenannter „Poset-Test“ von Rosenbaum [1] vorgeschlagen und ist von ihm selbst in [2] und [3] diskutiert worden. Die Namensgebung geht auf die Abkürzung des Begriffs „partially ordered sets“ zurück und spielt damit auf die Eigenschaft der zur Definition des Kohärenzmusters festzulegenden Relation $<_c$ an, welche die Beobachtungen des Datensatzes partiell anordnet. In seinem ursprünglichen Vorschlag nutzt Rosenbaum eine andere Teststatistik, die aber algebraisch äquivalent zu der hier vorgestellten Version ist, welches sich einfach mit den Methoden von Mantel [7] nachrechnen lässt. Der Poset-Ansatz verallgemeinert verschiedene nichtparametrische Tests. Handelt es sich zum Beispiel um eine eindimensionale Zielvariable und wird $<_c$ als normale Ungleichung \leq definiert, dann ist die Poset-Teststatistik äquivalent zur Wilcoxon-Mann-Whitney-Statistik. Wenn die Zielvariablen hingegen aus einer Überlebenszeit und dem zugehörigen Zensierungsindikator besteht und $<_c$ passend definiert ist, dann entspricht der Poset-Test dem Gehan-Test [8].

3 Ein Kohärenzmaß

Im vorherigen Kapitel wurde eine Rangstatistik T vorgestellt, mit der die Hypothese $P(y_i <_c y_j) = P(y_j <_c y_i)$, für $i=1, \dots, N_1, j=N_1+1, \dots, N$, getestet werden kann. Der direkte Weg, um den Grad der Abweichung von dieser Nullhypothese zu messen, ist die Betrachtung der Größe

$$\zeta_c = P(y_i <_c y_j) - P(y_j <_c y_i).$$

ζ_c nimmt Werte zwischen -1 und $+1$ an. Die Extremwerte werden genau dann erreicht, wenn die Relation $<_c$ die zwei Gruppen völlig separiert. Im Allgemeinen ist $P(y_i <_c y_j) + P(y_j <_c y_i) < 1$, weil $<_c$ nur eine partielle Ordnung beschreibt. Falls ζ_c Werte in der Nähe 0 annimmt, liegt eine der beiden folgenden Situationen zugrunde:

1. ein großer Anteil von Beobachtungspaaren ist nicht anzuordnen (d.h. $U_{ij}=0$)
2. der Anteil von Beobachtungspaaren mit $U_{ij}=1$ und $U_{ij}=-1$ ist etwa gleich.

Beide Situationen reflektieren eine geringe Kohärenz der erwarteten Ergebnismuster und der beobachteten Daten im Hinblick auf die vorgegebene partielle Ordnungsrelation $<_c$.

Die Idee des Kohärenzkoeffizienten ist ähnlich zu der des Korrelationskoeffizienten, welcher den Grad der gemeinsamen Variation zweier Variablen beschreibt. Die Definition von ζ_c erinnert formal stark an das Kendall'sche τ , welches als Differenz der Wahrscheinlichkeiten von konkordanten und diskordanten Paaren von bivariaten Beobachtungen definiert ist [9].

Für ζ_c kann ein unverzerrter Schätzer mit einer passenden Standardisierung von T angegeben werden. Wenn man die Scores innerhalb einer Gruppe summiert, dann egalisieren sie sich in ihren Effekten (da $U_{ik} = -U_{ki}$, so dass für $i, k \in \{1, \dots, N_1\}$ in der Summe $\sum_{i=1}^{N_1} S_i$ jeder Intragruppenvergleich zweimal mit unterschiedlichem Vorzeichen in der Summation auftritt). Daher kann T umgeschrieben werden in

$$T = \sum_{i=1}^{N_1} \sum_{j=N_1+1}^N U_{ij}$$

und ζ_c kann geschätzt werden, in dem T durch die Anzahl der Intergruppenvergleiche dividiert wird. Damit ist nun

$$\hat{\zeta}_c = \frac{T}{N_1 N_2}$$

ein unverzerrter Schätzer für ζ_c , da der Anteil der Beobachtungspaare aus den beiden Gruppen mit Indizes (i, j) , die zu $U_{ij}=1$ führen, ein unverzerrter Schätzer für den ersten Term der Differenz ist und der entsprechende Anteil, der zu $U_{ij} = -1$ führt, der unverzerrte Schätzer für den zweiten Term der Differenz ist.

Die Varianz von $\hat{\zeta}_c$ ist unter der Nullhypothese

$$\text{Var}(\hat{\zeta}_c) = (N_1 N_2)^{-2} \cdot \text{Var}(T) = (N(N-1)N_1 N_2)^{-1} \sum_{i=1}^N S_i^2.$$

Damit kann nun eine obere Grenze der Varianz unter der Alternative entsprechend der Form von Kendall's τ angegeben werden:

$$\text{Var}(\hat{\zeta}_c) \leq \frac{k}{N}(1 - \zeta_c^2),$$

wobei k eine positive Konstante ist. Mit diesem Resultat ist es möglich, konservative approximative Konfidenzintervalle für ζ_c anzugeben. Bessere (d.h. weniger konservative) Konfidenzintervalle erzielt man mit mittels der exakten Varianz von $\hat{\zeta}_c$. Diese kann unter einer gegebenen Alternative analog einer Idee beim Kendall'schen τ , wie bei Gibbons [9] formuliert, berechnet werden.

4 Praktisches Anwendungsbeispiel

Zur Illustration des Konzeptes der Kohärenz und der Anwendung des Poset-Tets dienen die Daten einer multizentrischen klinischen Studie zum Vergleich zweier Therapievorgehensweisen bei Frühgeborenen mit schweren Atemnotsyndrom [10]. Bei beiden Therapiearten wird ein natürlicher Schweinesurfactant als Mittel zur Lungenreifung der noch nicht voll entwickelten Lunge der Frühgeborenen gegeben, wobei in dem einen Arm eine hochdosierte Einfachgabe bei 176 teilnehmenden Patienten verwendet wurde, hingegen im anderen Arm 167 Patienten eine Mehrfachdosis (bis zu dreimal mit im Vergleich zum anderen Therapiearm geringeren Einzeldosen) des Surfactants zu fixierten Zeiten bekommen haben. Diese Mehrfachdosis wurde allerdings nur gegeben, wenn zu den definierten Zeitpunkten der Patient zusätzlichen Sauerstoff benötigte und dazu noch aktiv maschinell beatmet werden musste (Anteil des einzuatmenden Sauerstoffs (FIO_2) > 0.21). Die mit dieser Studie aufzuklärende Frage bestand darin, ob Frühgeborene mit schwerem Atemwegssyndrom von einer flexiblen einzusetzenden Mehrfachdosis profitieren. Das primäre Zielkriterium dieser Studie war das Überleben von 28 Tagen nach der Geburt. Weitere wichtige Zielgrößen umfassten die kumulative Zeit der zusätzlichen Sauerstoffgabe (O_2) in diesen ersten vier Wochen sowie der akute Effekt der Therapie auf die Beatmungssituation der Frühgeborenen (gemessen durch den FIO_2 -Wert 24 Stunden nach Therapiebeginn).

Der Überlebensstatus der Patienten wird als wichtiger angesehen als die Information über die anderen beiden Zielvariablen, d.h. nur unter den Überlebenden wird der Therapieeffekt durch Betrachtung dieser beiden Variablen analysiert. Betrachtet man den dreidimensionalen Zielvektor und definiert ein kohärentes Muster des Therapieeffekts, dann sollte ein positiver Effekt der Mehrfachdosis sich widerspiegeln in einer kleineren Anzahl von Toten, kombiniert mit kürzeren Zeiten unter O_2 und niedrigeren

FIO₂ Werten nach 24 Stunden für alle Überlebenden, jeweils verglichen mit der Einfachdosisgruppe.

Formal betrachtet, bezeichne nun X_{1i} den Überlebensstatus (0=lebend, 1=tot), X_{2i} die Gesamtzeit unter O₂ (stetig, mögliche Spannweite: 0 – 672 Stunden) und X_{3i} den FIO₂-Wert nach 24 Stunden (stetig, mögliche Spannweite: 0,21 – 1,0) des i -ten Kindes. Dann ist der Zielvektor $y_i = (X_{1i}, X_{2i}, X_{3i}) \in \mathbb{R}^3, i = 1, \dots, 343$ gegeben durch alle Messungen dieser drei Variablen, wobei die ersten 176 Beobachtungen zum ersten Therapiearm (Einfachdosis) gehören und die verbleibenden 167 zur Mehrfachdosisgruppe.

Die Kohärenzalternative kann formuliert werden mit der passenden Definition von $<_c$ im \mathbb{R}^3 . Unter Berücksichtigung der Hierarchie der zuvor beschriebenen Variablen wird nun $<_c$ definiert als

$$y_i <_c y_j$$

$$\Leftrightarrow (X_{1i} > X_{1j}) \vee (X_{1i} = 0 \wedge X_{1j} = 0 \wedge X_{2i} \geq X_{2j} \wedge X_{3i} \geq X_{3j})$$

wobei mindestens eine der Ungleichungen \geq scharf sein muss. Mit dieser Definition haben Patienten mit kleinen Rangzahlen eine schlechtere Reaktion auf die Therapie als die mit großen Rangzahlen. Hierbei werden allerdings viele Vektoren durch $<_c$ nicht anzuordnen sein, da zum Beispiel innerhalb der Gruppe der Verstorbenen keine Anordnung möglich ist.

Die Anwendung des Poset-Tests auf die Daten dieser klinischen Studie führt zu einer Teststatistik $T=3648$ und $2931822,6$ für die geschätzte Varianz. Daraus resultiert ein standardisierter T-Wert von $2,131$ und damit ein asymptotischer p-Wert von $0,033$. Damit ergibt sich ein signifikanter Unterschied im Antwortmuster der Zielvariablen zwischen den beiden Therapiearten in Richtung einer kohärenten Alternative zugunsten der Mehrfachdosisgruppe.

Der geschätzte Kohärenzoeffizient, den man durch die Differenz der Anteile der negativen ($0,327$) und positiven Vergleiche ($0,451$) erhält, ist im Beispiel $\hat{\zeta}_c = 0,124$.

Die geschätzte Varianz von $\hat{\zeta}_c$ unter H_0 beträgt $0,00339$. Desweiteren sieht man, dass nur $77,8\%$ aller möglichen Vergleiche zwischen Beobachtungen aus den beiden Therapiegruppen mit der Relation $<_c$ anzuordnen sind.

Die Ergebnisse dieser Auswertung unter Verwendung der Poset-Methode werden nun im Folgenden mit der getrennten Analyse aller Zielvariablen verglichen. Obwohl der Anteil der Toten in der Einfachdosisgruppe ($37/176 = 21\%$) deutlich höher ist als in der Mehrfachdosisgruppe ($22/167 = 13\%$), ist der p-Wert des exakten Tests von Fisher nicht signifikant ($p = 0,063$). Bei der Analyse der Dauer unter O₂ werden die Verstorbenen als zensierte Beobachtungen betrachtet. Der Vergleich dieser Vertei-

lungen in beiden Gruppen mittels Logrank Test führt zu einem p-Wert von 0,472. Der Unterschied bei den FIO₂-Messungen nach 24 Stunden hingegen zeigt einen signifikanten Unterschied (p=0,006, Wilcoxon-Mann-Whitney-Test). Diese Analyse wurde auf die Gruppe Überlebenden beschränkt. Damit ergibt sich insgesamt, dass bei der univariaten Analyse nur für eine der drei Zielvariablen ein signifikantes Ergebnis erzielt wurde, das für einen Vorteil der Mehrfachdosistherapie spricht. Bei dem wichtigsten Zielkriterium, dem Überleben, ist der p-Wert nur grenzwertig signifikant.

5 SAS-Makro Poset

Zur Berechnung der Teststatistik des Poset-Tests, des zugehörigen p-Werts sowie des Kohärenzmaßes steht ein SAS-Makro zur Verfügung. In dieses Makro können bis zu 20 Variablen zur Definition der Komponenten des Zielvektors mit einbezogen werden. Um Flexibilität in der Formulierung problemspezifischer kohärenter Alternativen zu gewährleisten, kann zu jeder einzelnen Komponente des Zielvektors einzeln ein Vergleichsoperator definiert werden, der die Art des für das entsprechende Kohärenzmuster notwendigen Vergleichs festlegt.

Besteht eine Hierarchie unter den Variablen, d.h. ist eine oder mehrere der Zielvariablen aus dem Vektor prioritär zu behandeln, müssen die Variablen mit höherer Priorität zuerst im Aufruf angegeben werden. Tritt dieser Fall auf, so muss zusätzlich beim Aufruf bestimmt werden, wieviel Variablen prioritär behandelt werden sollen. (`priori1` = Anzahl, default=0). Ist aufgrund der Ausprägungen der prioritären Variablen eine Anordnungsmöglichkeit gegeben, werden die anderen Variablen nicht mehr berücksichtigt. In jedem Fall ist eine die Gruppenzugehörigkeit definierende Variable in der Zuweisung der Makrovariablen `group` anzugeben. Falls diese Variable die Gruppenzugehörigkeit nicht mit den Ausprägungen 0 und 1 kodiert, so müssen mit den Makrovariablen `expo` und `unexpo` die tatsächlich vorkommenden Ausprägungen beim Aufruf angegeben werden. Mehr als zwei Ausprägungen sind nicht zulässig. Alle Beobachtungen mit fehlenden Werten in den Komponenten des Zielvariablenvektors oder der Gruppierungsvariable werden bei der Analyse nicht berücksichtigt.

Allgemeine Syntax des Aufrufs des Makros POSET:

```
%poset(dataset=SAS-Dateiname, <Name der auszuwertenden Datei>  
  var1=Name der 1. Variable des Zielvariablenvektors,  
  <var2=Name der 2. Variable des Zielvariablenvektors>  
  .  
  .  
  <, var20=Name der 20. Variable des Zielvariablenvektors>  
  Der Zielvariablenvektor kann höchstens 20 Komponenten umfassen  
  operan1=Vergleichsoperator für 1. Variable,  
  <operan2=Vergleichsoperator für 2. Variable>  
  .  
  .  
  <, operan20=Vergleichsoperator für 20. Variable>  
  <, prior1=Anzahl der Variablen mit Priorität,  
  Variablen mit Priorität müssen zuerst angegeben>  
  group=Name der die Gruppenzugehörigkeit definierenden Variable  
  <, expo=Ausprägung der Gruppierungsvariable für Exponierte (default=1)>  
  <, unexpo=Ausprägung der Gruppierungsvariable für Nichtexponierte  
  (default=0)>  
  );
```

Beispiele für den Aufruf des Makros POSET:

- Drei Zielvariablen definieren kohärente Alternative, keine Hierarchie:

```
%poset (    dataset    =    test,  
           var1       =    variab1,  
           operan1    =    <,  
           var2       =    variab2,  
           operan2    =    >=,  
           var3       =    variab3,  
           operan3    =    <=,  
           group      =    groupvar);
```

- Drei Zielvariablen definieren kohärente Alternative, erste Variable hat Priorität, die Gruppierungsvariable hat die Ausprägungen 2 und 5

```
%poset (    dataset =    test,
            var1    =    variab1,
            operan1 =    <,
            var2    =    variab2,
            operan2 =    >,
            var3    =    variab3,
            operan3 =    >=,
            prior1  =    1,
            group   =    groupvar,
            expo    =    2,
            unexpo  =    5);
```

Das in Abschnitt 4 vorgestellte Beispiel wurde mit einem im ersten Moment überraschenden Aufruf des Makros POSET gerechnet, da zur Implementation des Kohärenzmusters vorab zwei Veränderungen im ursprünglichen Datensatz vorgenommen werden mussten: (i) Die beiden Variablen `time_o2` und `fio2_24` wurden vor der Analyse für alle Verstorbenen auf Null gesetzt, um diese bei der Beurteilung des Kohärenzmusters unter den Verstorbenen auszublenden, (ii) es musste eine zusätzliche Variable `sum_time_fio2` als Summe aus den beiden beobachteten Variablen `time_o2` und `fio2_24` gebildet werden, um die Forderung, dass mindestens eine der beiden Ungleichungen \geq in der Definition der partiellen Ordnungsrelation \prec_c strikt sein muss, umsetzen zu können. Der konkrete Aufruf lautete dann:

```
%poset (    dataset =    surfactant,
            var1    =    life_status,
            operan1 =    >,
            var2    =    time_o2,
            operan2 =    >=,
            var3    =    fio2_24,
            operan3 =    >=,
            var4    =    sum_time_fio2,
            operan3 =    >,
            prior1  =    1,
            group   =    therapie);
```

Der zugehörige Output des Makros POSET in diesem Beispiel hatte folgende Gestalt:

PROC POSET - Results of a two-sample test for coherence
in multidimensional data

 Data set: surfactant
 Number of observations: 343
 Group variable: therapie
Poset analysis for the following variables:
 life_status time_o2 fio2_24 sum_time_fio2

Definition of the coherence pattern:

variable	operand
life_status	>
time_o2	>=
fio2_24	>=
sum_time_fio2	>

 Hierarchy among variables: yes
First variable of the list above has priority

Poset test statistic T:	3648.000
Variance of T under H0:	2931822
Standardized T:	2.131
Prob > T under H0:	0.033
Measure of coherence C:	0.124
Variance of C under H0:	0.003
Prop. of decidable comparisons:	0.778
Proportion of "pos." comparisons:	0.451
Proportion of "neg." comparisons:	0.327

Das Makro POSET ist Bestandteil des am Institut für Medizininformatik, Biometrie und Epidemiologie (IMBE) der Universität Erlangen-Nürnberg aufgebauten Systems ISSAN (IMBE Statistical Software Archive Network). Der SAS-Code des Makros kann - für nichtkommerzielle Anwendungen kostenlos - mit einer Beschreibung unter folgender Web-Adresse heruntergeladen werden: <http://www.imbe.med.uni-erlangen.de/issan/issan.htm>. Diese Web-Seite wird kontinuierlich gepflegt, so dass dort auch Updates des Makros zu finden sein werden.

Literatur

- [1] Rosenbaum, P.R. (1991). Some poset statistics. *Annals of Statistics* 19, 1091-1097.
- [2] Rosenbaum, P.R. (1994). Coherence in observational studies. *Biometrics* 50, 368-374.
- [3] Rosenbaum, P.R. (2002). *Observational Studies* (2nd edition). New York: Springer-Verlag.
- [4] Susser, M. (1973). Causal Thinking in the Health Sciences: Concepts and Strategies in Epidemiology. New York: Oxford University Press.
- [5] Hill, A.B. (1965). The environment and disease: association or causation? *Proceedings of the Royal Society of Medicine* 58, 295-300.
- [6] Evans, A.S. (1978). Causation and disease: a chronological journey. *American Journal of Epidemiology* 108, 249-258.
- [7] Mantel, N. (1967). Ranking procedures for arbitrarily restricted observations. *Biometrics* 23, 65-78.
- [8] Gehan, E. (1965). A generalized Wilcoxon test for comparing arbitrarily singly-censored samples. *Biometrika* 52, 203-223.
- [9] Gibbons, J.D. (1971). *Nonparametric Statistical Inference*. New York: McGraw-Hill.
- [10] Speer, C.P., Robertson, B., Curstedt, T., Halliday, H.L., Compagnone, D., Gefeller, O. et al. (1992). Randomized European multicenter trial of surfactant replacement therapy for severe neonatal respiratory distress syndrome: single versus multiple doses of Curosurf. *Pediatrics* 89, 13-20.