

Paarweises Testen und Korrelationsanalyse in SAS bei geringen Rücklaufquoten der Patientenbögen

R.-D. Recknagel und A. Spellig

Einleitung

Die Qualität einer medizinischen Leistung ist das komplexe Zusammenspiel von vielen verschiedenen Faktoren, die im einzelnen nicht exakt zu definieren und voneinander zu trennen sind. Unter Berücksichtigung der WHO-Qualitätsdefinitionen genießt neben den physiologischen Parametern und der Patientenzufriedenheit das Konzept der gesundheitsbezogenen Lebensqualität einen besonderen Stellenwert. Dabei versteht man unter „Lebensqualität“ ein psychologisches Konstrukt, das die körperlichen, psychischen, mentalen, sozialen und funktionalen Aspekte des Befindens und der Funktionsfähigkeit der Patienten aus ihrer Sicht beschreibt.

Das Interesse, das sich in der Entwicklungsphase der offenen Herzchirurgie ausschließlich auf die klinischen Belange und in der Folge auf die Überlebensstatistiken konzentrierte, verlagerte sich seit Mitte der 70iger Jahre zunehmend auf Fragen der veränderten und verbesserten Lebensqualität nach Herzoperationen für den Patienten in seiner Familie, am Arbeitsplatz, am Arbeitsmarkt, in der Gesellschaft usw.

Das Bindeglied zwischen Aufenthalt in der Herzchirurgischen Klinik mit der überwiegend apparate-technischen intensivmedizinischen Versorgung und dem Alltagsleben des Herzoperierten ist die Anschlußheilbehandlung (AHB) in einer Rehaklinik mit ihren spezifischen Zielstellungen. Die folgende Studie stammt aus der Medianklinik Bad Berka. Im Zeitraum eines Jahres von Juli 1993 bis Juli 1994 sind alle Patienten der Rehaeinrichtung, die sich vorher einer Herzoperation unterzogen hatten, nachuntersucht worden (n=809). Die gewonnenen Untersuchungsergebnisse wurden in Patientenbögen erfaßt und anschließend in SAS ausgewertet mit der Zielstellung, die Effizienz der Anschlußheilbehandlung zu bewerten. Ein Teilaspekt dieser Auswertung soll im folgenden beleuchtet werden, da er in derartigen Anwendungsbeobachtungen eine nicht unerhebliche Bedeutung haben kann für eine richtige Einschätzung einer Anschlußheilbehandlung.

1. Das STAI zur Erfassung von Angst als Zustand

Die *State-Angstskala* STAI (State-Trait-Angstinventar) besteht aus zwanzig Feststellungen, mit denen der Proband beschreiben soll, wie er sich **jetzt**, d.h. **in diesem Moment** fühlt.

Zehn Feststellungen sind in Richtung Angst formuliert, zehn andere in Richtung Angstfreiheit. Die Beantwortung erfolgt auf einer vierstufigen Skala mit Intensitätsangaben. Dazu stehen vier Antwortkategorien zur Verfügung, die jeweils durch die Ziffern 1 bis 4 hinter jeder Feststellung repräsentiert sind.

Die Kategorien stellen folgende Punkte auf einer Intensitätsdimension dar: „überhaupt nicht“ (1), „ein wenig“ (2), „ziemlich“ (3) und „sehr“ (4). Zur Ermittlung des Rohwertes eines Probanden werden die Summenwerte der jeweils 20 Feststellungen errechnet. Die Spannweite der möglichen Werte reicht von minimal 20 bis maximal 80. Vor Bildung eines Summenwertes, der das Ausmaß der Angst repräsentieren soll, muß eine *Inversion* derjenigen Feststellungen vorgenommen werden, die in Richtung *Angstfreiheit* formuliert sind (zu summierender Wert=5-angekreuzter Wert). Der Summenwert der State-Angstskala stellt ein Maß dar für die Intensität eines emotionalen Zustands, der gekennzeichnet ist durch Anspannung, Besorgtheit, Nervosität, innere Unruhe und Furcht vor zukünftigen Ereignissen. Dabei entspricht ein Wert von 20 dem Nichtvorhandensein und ein Wert von 80 der maximalen Intensität dieses Gefühls.




2. EDV und Statistik

Das erhobene Datenmaterial wurde mit Hilfe von SAS in einer entsprechenden Eingabemaske erfaßt und statistisch ausgewertet. Zur Anwendung kamen *deskriptive Statistiken und paarweiser t-Test (proc univariate)*, *t-Test (proc ttest)*, *Wilcoxon-Test (proc nlparway)* und *Blockdiagramme (proc chart)*. Der paarweise t-Test, der ja als SAS-Prozedur in Reinform fehlt, wurde in der üblichen Weise durch Differenzenbildung mit Test auf Null abgewickelt. Zu Vergleichszwecken wurde auch der *Wilcoxon paired rank* -Test herangezogen, da ein Kritikpunkt (unbegründet) bei der Auswertung angeblich die Nichtnormalverteilung der STAI-Werte ist. Doch ist die Datenlage so, daß zum einen die Effekte klar hervortreten und zum anderen die Verteilungen symmetrisch sind und gut approximierbar durch die Normalverteilung sind. Sowohl die parametrische als auch die nichtparametrische Testung liefert an all den Stellen Signifikanz, wo die Effekte auch fachlogisch zu erwarten sind.

Ziel der Auswertung war eine einfache deskriptive Statistik und die statistische Absicherung der Effizienz der stationären AHB-Maßnahmen mit Hilfe der biometrischen Bewertung verschiedener diagnostischer Verfahren und der Summe aller therapeutischen Maßnahmen. Die Grundgesamtheit bestand wie oben bereits erwähnt aus allen kardiologische Patienten des Zeitraume Juli 1993 bis Juli 1994. Die erfaßten Patientendaten umfaßten 3 Bereiche und wurden auch durch dementsprechend eigenständige SAS-Dateien strukturiert. Der erste Bereich umfaßte die psychologischen Daten, bestehend aus der weiter oben beschriebenen Komponente des STAI und des FPI (Freiburger Persönlichkeitsinventar), der aber hier nicht betrachtet wird. Der zweite Bereich umfaßt alle kardiologische Daten und der dritte schließlich die sozialen Daten der Patienten. Die Patientendaten wurden jeweils nach 1 Woche (Visite 1), nach 4 Wochen (Visite 2) und nach einem halben Jahr (Visite 3) zum Zeitpunkt der Entlassung aus der Klinik erhoben.

Die folgende Tabelle gibt eine Übersicht über missing values und Mittelwerte der Rohwerte zum STAI. Ein Vergleich von sowohl der Anzahl der fehlenden Werte als auch deren Mittelwerte und mittleren Differenzen läßt den Verdacht aufkommen, daß hier Abhängigkeiten existieren, die unter Umständen Verzerrungen und damit Fehler in der Bewertung erzeugen.

Tabelle: Erfassung der Rohwerte jedes Patienten

Visite 1 (zum Ende der Behandlung)	Visite 2 (zwischenzeitlich)	Visite 3 (Abschlußuntersuchung)
Fragebogen mit Rohwert 1	Fragebogen mit Rohwert2	Fragebogen mit Rohwert 3
Gesamtzahl n=809	Gesamtzahl n=809	Gesamtzahl n=809
vorhanden: 549 missing: 260 38..52	vorhanden: 599 missing:210 37.180	vorhanden: 425 missing: 384 39.25
 vorhanden: 422 missing: 387 d=-1,57		
 vorhanden: 296 missing: 513 d=1,08		
	 vorhanden: 317 missing: 492 d=2,71	

3. Missing values und paarweises Testen beim STAI

Die Rücklaufzeiten der STAI-Fragebögen hängen vom Gegenstand der Untersuchung selbst ab: Patienten mit hohen Summenwerten auf der Angstskala tendieren zu geringeren Rücklaufzeiten der Fragebögen. Es ist aufgrund der obigen Tabelle ersichtlich, daß bei größeren Mittelwertsdifferenzen auch die Anzahl der missing values zunimmt. Diese Erscheinung wird noch maskiert und ist in Wirklichkeit noch krasser, als sich durch die Mittelwertsdifferenzen offenbart: Der wahre Indikator für eine vorliegende Verzerrung sind die Mittelwerte der Differenzen, weil in diese Mittelwerte nur solche Paardifferenzen eingehen, bei denen beide Paarlinge vorhanden sind. Damit erhält man realistische Differenzen, die nicht durch den unterschiedlichen Stützungsgrad beider Verteilungen verzerrt sind. Eine Verzerrung der mittleren Rohwerte ist somit bei der ersten Vorgehensweise unvermeidlich. Für die SAS-Auswertungsroutine folgt daher die Empfehlung der paarweisen Testung den Vorrang zu geben und eine Vergleichsrechnung der unverbunden Stichprobe mitzuführen.

Wenn man eine zuverlässige (reliable) Variable in bestimmten zeitlichen Abstand an denselben Personen mehrmals erhebt, dann ist damit zu rechnen, daß die Korrelation der Wertereihen beider Zeitpunkte hoch ist (Retest-Reliabilität). Man hat also das doppelte Dilemma, daß die intraindividuellen Beobachtungswerte hochkorrelieren und außerdem mit wachsenden STAI-Werten auch die Wahrscheinlichkeit für das Auftreten von missing values zunimmt.

4. Regeln für die Korrelationsanalyse

1. Es zeigen sich hohe positive Korrelationen der Rohwerte aus Visite 1, Visite 2 und Visite 3. Um die Verzerrung des Korrelationskoeffizienten gering zu halten, wird Abstand genommen von der automatischen Routine, Wertepaare aus der Auswertung zu eliminieren, bei denen ein Paarling fehlt.
2. Die Standardabweichungen s_1 und s_2 werden aus den ungekürzten Stichproben und die Kovarianz aus den vollständigen Paaren und Verwendung der Maximum-Likelihood-Schätzung von \mathbf{r} aus dem bereinigten Datensatz berechnet.
3. Die Testgröße zum paarweisen Test in der Korrelationsform zur Bewertung der Varianzreduzierung durch Korrelation wird in folgender Form verwendet:

$$|t| = \frac{|\bar{d}|}{\sqrt{(s_1^2 + s_2^2 - 2 \cdot r_{12} \cdot s_1 \cdot s_2) / n}}$$

4. Überlegungen zur Verzerrung durch Vergleich von mittleren Differenzen und Differenzen von Mittelwerten im Zusammenhang mit Rücklaufzeiten sind durchzuführen:

- Ermittlung der Rücklaufzeiten R_1 und R_2
- Berechnung der Mittelwertdifferenzen $D = \bar{x}_1 - \bar{x}_2$ und der mittleren Differenz d
- Schätzung des Bias bezogen auf 100 fehlende Patienten:

$$b = \frac{D - \bar{d}}{R_1 - R_2} \cdot 100$$

falls $R_1 \neq R_2$

- Schätzung der wahren Mittelwerte der unverzerrten Stichprobe durch Elimination des Bias:

$$m_1 = x_1 + \frac{R_1}{100} \cdot b$$

falls $R_1 \neq R_2$

- Vergleich mit den verzerrten Werten und fachliche Beurteilung der medizinischen Relevanz.