

## Probleme der Fallidentifikation mit einem halbstrukturierten psychiatrischen Interview am Beispiel einer epidemiologischen Untersuchung in Allgemeinpraxen\*

W. Bruder, S. Weyerer und H. Dilling

Psychiatrische Klinik der Universität München (Direktor: Prof. Dr. H. Hippius), München,  
Bundesrepublik Deutschland  
Klinik für Psychiatrie der Medizinischen Hochschule Lübeck (Direktor: Prof. Dr. H. Dilling),  
Lübeck, Bundesrepublik Deutschland

### Problems in Case Identification Using a Semi-Structured Psychiatric Interview as Exemplified by an Epidemiological Study in General Practices

**Summary.** From 1973 to 1974 research psychiatrists studied 1,274 patients in general practices in the county of Traunstein (FRG). On the basis of an interview developed by Goldberg et al. (1970), it could be shown that objective symptoms registered by the interviewers generally differentiate better between cases and non-cases than do subjective symptoms. According to the weighting suggested by the developers of the interview, singly weighting the subjective and doubly weighting the objective symptoms, a total score was compared with the psychiatric diagnosis made by the interviewer. At a cut-off of 20 points, the overall agreement reached an optimum of 83.1% and increased only slightly due to a-posterior weighting of the items. A cluster analysis of the symptoms showed that patients demonstrate significantly different combinations of symptoms according to the nature of their mental disorder.

**Key words:** Case identification – Psychiatric interview – Epidemiology – General practice

**Zusammenfassung.** In den Jahren 1973/74 haben psychiatrisch geschulte Interviewer 1274 Patienten in Allgemeinpraxen des Landkreises Traunstein untersucht. Aufgrund des dabei verwendeten Erhebungsinstruments nach

\* Diese Arbeit entstand mit Unterstützung der Deutschen Forschungsgemeinschaft im Rahmen des Sonderforschungsbereiches 116 „Psychiatrische Epidemiologie“ und war Thema einer Dissertation (Bruder 1978)

Wir danken Herrn Prof. Dr. Überla und Frau Dr. Kellhammer für die kritische Durchsicht des Manuskripts.

Sonderdruckanforderungen an: Dr. Siegfried Weyerer, Psychiatrische Universitätsklinik, Nußbaumstr. 7, D-8000 München 2, Bundesrepublik Deutschland

Goldberg et al. (1970) konnte gezeigt werden, daß vom Interviewer beurteilte Symptome in der Regel besser zwischen psychisch Auffälligen und Unauffälligen differenzieren als subjektive Beschwerden. Entsprechend der von den Entwicklern des Interviews vorgeschlagenen einfachen Gewichtung der subjektiven und der doppelten Gewichtung der objektiven Symptome wurde ein Gesamtscore gebildet und mit dem diagnostischen Gesamturteil des Interviewers verglichen. Die Übereinstimmung erreichte bei einem Cut-off von 20 Punkten ein Optimum von 83,1% und erhöhte sich mit Hilfe von a-posteriori-Gewichtungen der Items nur geringfügig. Eine Clusteranalyse der Itemscores des Interviews zeigte, daß Patienten je nach Art der psychischen Erkrankung statistisch signifikant unterscheidbare Symptomkombinationen aufweisen.

**Schlüsselwörter:** Fallidentifikation – psychiatrisches Interview – Epidemiologie – Allgemeinpraxis

## 1. Einleitung

Einem der größten Probleme in der psychiatrischen Epidemiologie, dem der Fallidentifikation, steht man vor allem bei Studien in der Allgemeinbevölkerung gegenüber. Bei einem in der Regel einmaligen und zeitlich sehr begrenzten Kontakt soll der Untersucher zu einem reproduzierbaren Urteil über das Vorliegen einer psychiatrischen Erkrankung beim Probanden kommen. Hier kann sich der Epidemiologe — im Gegensatz zu administrativen Untersuchungen — nicht auf medizinische Vorbefunde stützen, insbesondere aber fehlt die Selbstdefinition des Falles, die vom Patienten vorgenommen wird, indem er ärztliche Hilfe in Anspruch nimmt. Nach Häfner (1978) sollte eine Fallidentifikationsmethode objektiv, quantitativ gültig und trennscharf sein, eine Forderung, die für psychische Erkrankungen oft ungleich schwieriger zu erfüllen ist als für somatische. In diesem Zusammenhang sind die Bestrebungen zu sehen, psychische Störungen mit Hilfe weitgehend standardisierter Verfahren zu bestimmen, um subjektive Einflüsse beim "case finding" möglichst auszuschließen und dadurch die Vergleichbarkeit verschiedener Untersuchungen zu erhöhen. Besondere Bedeutung in der psychiatrisch-epidemiologischen Forschung hat — neben einem von Wing et al. (1974) entwickelten Erhebungsinstrument („Present State Examination“) — das klinisch-psychiatrische Interview nach Goldberg et al. (1970) gewonnen, wobei systematisch sowohl nach bestimmten Beschwerden gefragt wird, als auch bestimmte beobachtbare Symptome beurteilt werden. Bei diesen Verfahren kann die psychiatrische Fallidentifikation auch in der Weise versucht werden, daß ein Summenscore für die einzelnen entsprechend ihrem Schweregrad rangskalierten Items gebildet wird, und Probanden, die einen bestimmten „Cut-off“-Punktwert überschreiten, als psychisch auffällig klassifiziert werden. Als häufiges Außenkriterium für die Validierung dieser Meßmethode wird das Expertenurteil bzw. die im Interview ermittelte Diagnose verwendet.

Die Ziele der vorliegenden Arbeit, die Teil einer umfangreichen epidemiologischen Studie der Klientel in Allgemeinpraxen ist (Dilling et al. 1978; Dilling u. Weyerer 1978), lassen sich wie folgt angeben:

- Welche relative Bedeutung kommt einzelnen psychiatrischen Symptomen für die Fallidentifikation durch den Interviewer zu?
- Welcher Zusammenhang besteht zwischen der Fallidentifikation durch den Interviewer und derjenigen mit Hilfe eines Gesamtscores?
- Inwieweit lassen sich charakteristische Symptomkonstellationen für einzelne Diagnosengruppen eruieren?

## 2. Methodisches Vorgehen

### 2.1 Stichprobenziehung

In der kleinstädtisch-ländlichen Region des Landkreises Traunstein wurden von den dort vorhandenen 83 Allgemeinpraxen (Adressenliste der Kassenärztlichen Vereinigung: Stand vom April 1974) 18 entsprechend einer geschichteten Stichprobe in die Untersuchung einbezogen; als Schichtungskriterium diente die Überweisungshäufigkeit in psychiatrische Einrichtungen, die in einer Vorstudie bestimmt wurde. Eine ausführliche Darstellung der Stichprobenziehung ist von uns an anderer Stelle gegeben (Dilling und Weyerer 1978).

Insgesamt wurde eine Zufallsstichprobe aller Patienten (Auswahlsatz 46%) gezogen, die in einem Zeitraum von 14 Tagen den betreffenden Hausarzt während seiner regulären Sprechstunde aufsuchten. Durch den Vergleich mit den in der Hausarztkartei registrierten Sozialdaten und Diagnosen sämtlicher in diesem Untersuchungszeitraum die Praxis beanspruchenden Patienten konnte festgestellt werden, daß die Stichprobe der Interviewten bezogen auf alle Konsultationen repräsentativ war.

### 2.2 Psychiatrische Fallidentifikation

Bei der Bestimmung psychischer Störungen durch die drei Forschungspsychiater war sicher gestellt, daß diese unabhängig erfolgte und nicht durch Vorinformation des Hausarztes beeinflußt wurde. Gegenüber üblichen Explorationen in psychiatrischen Einrichtungen unterscheidet sich die Interviewsituation bei Studien in der Allgemeinpraxis ganz wesentlich. Das Interview muß auch für Personen akzeptabel sein, die sich nicht als psychiatrisch auffällig ansehen. Es soll möglichst zeitökonomisch sein, so daß zahlreiche Patienten untersucht werden können, die ausreichende Informationen über Symptome geben, um eine diagnostische Klassifikation zu ermöglichen. Schließlich sollten die klinischen Beurteilungen möglichst reliabel sein. Die genannten Kriterien werden von dem von Goldberg und Mitarbeitern (Goldberg et al. 1970) am Institute of Psychiatry in London entwickelten halbstrukturierten Interview in hohem Maße erfüllt. Es gliedert sich in folgende Bereiche:

In einem freien Interview wird die allgemeine soziale und die Krankheitsanamnese erhoben, wobei sich eine positive Beziehung zwischen den Patienten und dem Forschungspsychiater entwickeln soll.

Es werden die in den letzten 7 Tagen aufgetretenen Beschwerden aufgezeichnet und jedes Symptom anhand einer 5-stufigen Ratingskala beurteilt: Nach einer Gruppierung von Goldberg et al. werden 11 subjektive und 12 objektive Symptome unterschieden.

Bei den subjektiven Symptomen handelt es sich um vom Patienten angegebene Beschwerden und Störungen, bei den objektiven um solche, die vom Interviewer während der Untersuchung festgestellt werden. Hierbei ist es öfter schwierig, eine eindeutige Grenze zwischen beiden Bereichen zu ziehen.

Subjektive Symptome:	Objektive Symptome:
1. körperliche Symptome	12. langsam
2. Müdigkeit	13. abwehrend
3. Schlafstörungen	14. demonstrativ
4. Hypnotikagebrauch	15. deprimiert
5. Reizbarkeit	16. ängstlich

- |                         |                          |
|-------------------------|--------------------------|
| 6. Konzentrationsmangel | 17. gehoben              |
| 7. Depression           | 18. affektiv verflacht   |
| 8. Angst                | 19. Wahn                 |
| 9. Phobien              | 20. Halluzinationen      |
| 10. Zwangssymptome      | 21. gestörte Intelligenz |
| 11. Depersonalisation   | 22. Hypochondrie         |
|                         | 23. depressive Gedanken  |

Die Skala reicht von 0 (= Symptom nicht vorhanden) bis 4 (= Symptom häufig und stark vorhanden). Um die Beurteilung des Schweregrades zu vereinheitlichen, stand jedem Interviewer ein ausführlicher Leitfaden mit Fallbeispielen für jedes Symptom zur Verfügung. Eine Reliabilitätsstudie, die zu Beginn der Untersuchung durchgeführt wurde, ergab ein relativ günstiges Ergebnis. Diese wurde gemeinsam mit Kollegen einer Mannheimer Parallelstudie (Zintl et al. 1978) in Allgemeinpraxen durchgeführt. Bei den statistischen Auswertungen wurden die Merkmalsausprägungen der Symptome dichotomisiert (0,1 versus 2, 3, 4) und die Interviewerübereinstimmung mit Hilfe des Phi-Koeffizienten berechnet. Für 66% der Einschätzungen lagen die Übereinstimmungskoeffizienten über 0,80, für 29% zwischen 0,60 und 0,80 und nur für 5% — bei den Symptomen Müdigkeit, Zwang und ängstlich — unter 0,60. Bei der globalen Beurteilung wurde im Falle von „psychisch krank“ stets auch eine Diagnose (evtl. auch Nebendiagnose) nach dem Internationalen Diagnosenschlüssel psychiatrischer Krankheiten (8. Revision der ICD) gestellt. Aus Gründen der statistischen Vergleichbarkeit wurden die psychiatrischen Diagnosen zu 10 Gruppen zusammengefaßt.

1. (prä)senile Demenz (290; 293.0)
2. andere organische psychiatrische Erkrankungen (292, 293.1–9; 294 außer 294.3; 309)
3. Schizophrenie (295)
4. affektive und andere Psychosen (296–299)
5. depressive Neurose (300.4)
6. andere Neurosen (300 außer 300.4)
7. Persönlichkeitsstörungen (301; 302; 306 außer 306.8; 307; 308)
8. psychosomatische Erkrankungen (305; 306.8)
9. Alkoholismus oder Drogenabhängigkeit (291; 294.3; 303; 304)
10. Oligophrenie (310–315)

### *2.3 Statistische Auswertung der Daten*

Die Angaben in den Interviews wurden kodiert, auf Lochkarten übertragen und auf Fehler überprüft. Die mathematisch-statistische Auswertung erfolgte im Max-Planck-Institut für Plasmaphysik in Garching bei München. Dabei sind folgende Programme verwendet worden: das Datatext-Programmsystem zur Analyse sozialwissenschaftlicher Daten (Armor und Couch 1972), das SPSS-Programmsystem (Nie et al. 1975). Ein Programm zur hierarchischen Clusteranalyse (Ward 1963; Rollett und Bartram 1976) wurde weiterentwickelt und ergänzt. Einige Programme zur graphischen Darstellung mußten neu konzipiert werden.

## **3. Ergebnisse**

### *3.1 Häufigkeit und Verteilung der Goldberg-Items bei psychisch Kranken und Nichtkranken*

Aufgrund der Analyse der Verteilung der Schweregrade der Goldberg-Symptome bei psychisch Kranken bzw. Nichtkranken können Aussagen über die relative Bedeutung der einzelnen Items für die Fallidentifikation gewonnen werden. Aus dem Aufbau der Beurteilungsskala ist für Items, die spezifisch für die Fallidentifi-

fikation geeignet sind, zu erwarten, daß Schweregrade  $\geq 2$  in der Gruppe der Nichtkranken kaum vorkommen, in der Gruppe der Kranken jedoch zu einem erheblichen Anteil.

Nicht alle Items haben für die beiden Gruppen einen deutlich voneinander verschiedenen Verlauf der Verteilungskurven (Abb. 1). Dies trifft vor allem für die subjektiven Symptome Müdigkeit, Hypnotika, Phobien und Zwangssymptome zu. Bei den objektiven Symptomen finden wir nur für das Item „gehoben“ bei Kranken und Nichtkranken eine sehr ähnliche Häufigkeitsverteilung, bei den übrigen Symptomen sind die Unterschiede zwischen beiden Gruppen sehr viel stärker ausgeprägt. Dies weist darauf hin, daß objektive Symptome eine wesentlich größere Bedeutung für die Fallidentifikation durch den Interviewer haben als subjektive. — Die Symptome Depersonalisation, Wahn und Halluzinationen sind für die Gruppentrennung irrelevant, da sie in der gesamten Stichprobe nur bei weniger als  $n = 5$  Probanden auftraten (siehe auch Diskriminanzanalyse).

### *3.2 Die Beziehung zwischen Gesamtscore und Fallidentifikation durch den Psychiater*

*3.2.1 Zuordnung mit a-priori-Gewichtung der Items.* Als „Außenkriterium“ für die Validierung der Fallidentifikation mittels der 23 Interviewitems verwendeten wir das Urteil „psychisch krank — psychisch nicht krank“ des Forschungspsychiaters. Die Beurteilung der Hausärzte als Außenkriterium zu verwenden, wäre sehr fragwürdig, da deren diagnostische Gewohnheiten sehr stark voneinander und von denen der Psychiater abweichen, wie sich in dem entsprechenden Teil unserer Studie bestätigte. Hierbei sind wir uns darüber im klaren, daß bei diesem Vergleich zwischen Globalurteil und der GBI-Gesamtscore das diagnostische Urteil des Forschungspsychiaters nur mit Einschränkungen als Außenkriterium zu betrachten ist, da in sein Globalurteil ja auch die Goldbergbeurteilungen miteinfließen. Die klinische Diagnose wurde von den Forschungspsychiatern der Psychiatrischen Universitätsklinik München nach den dort üblichen Kriterien gestellt. Die Reliabilitätsstudie bezog sich ausschließlich auf die Beurteilung von Symptomen und nicht auf Diagnosen.

Da bei jedem Patienten die Schweregrade der Symptome als 23 Zahlen zwischen 0 und 4 erfaßt wurden, liegt es nahe, aus einer Linearkombination dieser Punktzahlen einen Gesamtscore zu bilden. Aus der Untersuchung des Zusammenhangs zwischen der erreichten Gesamtpunktzahl und der klinischen Diagnose wird dann versucht, ein Kriterium abzuleiten, um auch bei Probanden ohne bekannte psychiatrische Diagnose allein aus dem ermittelten Gesamtscore eine Aussage über das mögliche Vorliegen einer psychiatrischen Erkrankung zu machen. Die Trennformel, die man hierzu benötigt, lautet in ihrer allgemeinsten Form:

$$\text{Formel (1): } t = \sum_{i=1}^n s(i) w(i)$$

Dabei ist  $t$  die Gesamtpunktzahl,  $s(i)$  der Schweregrad und  $w(i)$  eine eventuelle Gewichtung des Items  $i$ .

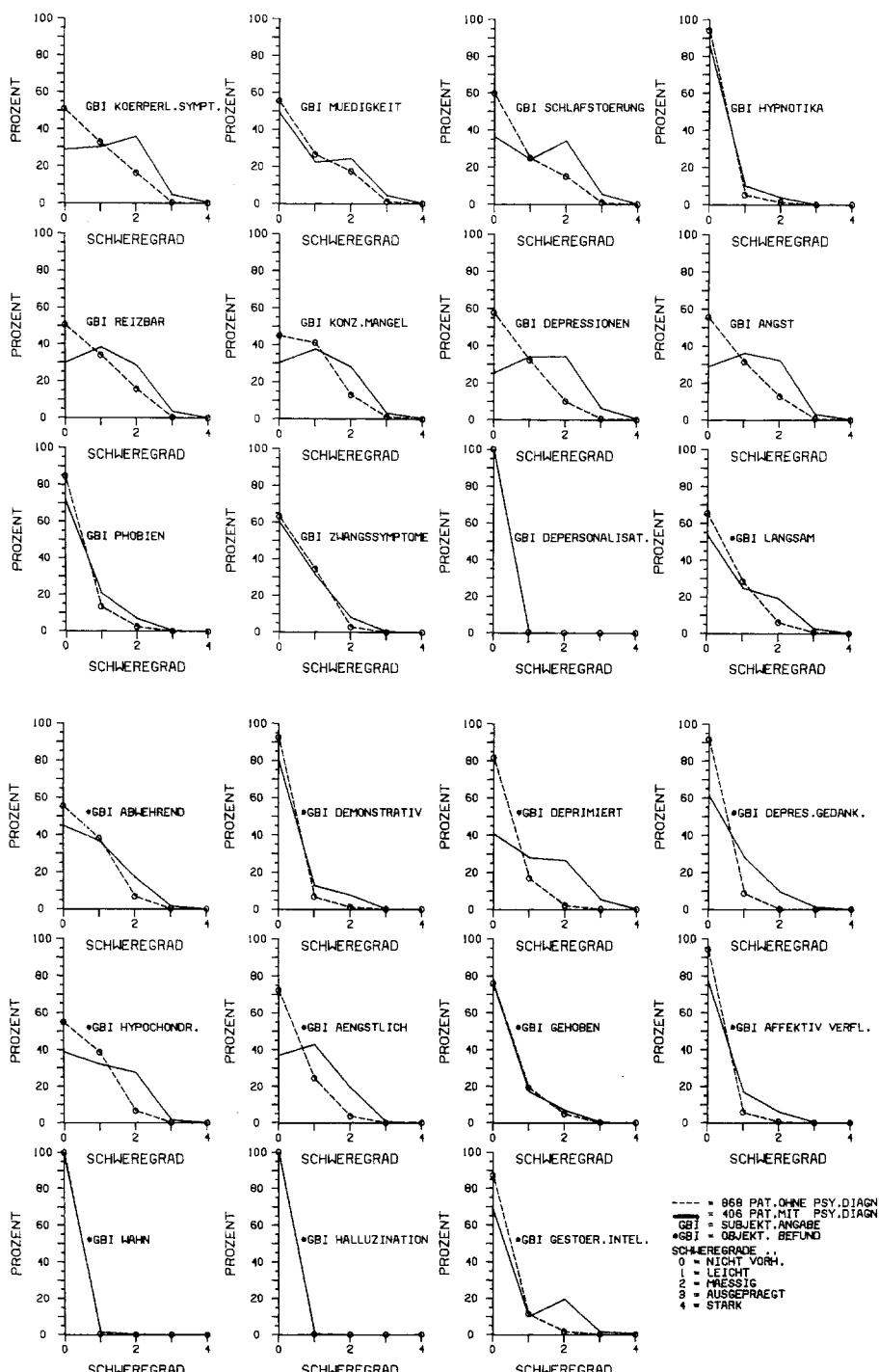


Abb. 1. Prozentuale Verteilung der Schweregrade der psychiatrischen Symptome von Patienten ohne und mit psychiatrischer Diagnose

Bezüglich der Gewichtung gibt es grundsätzlich zwei verschiedene Möglichkeiten: Die Gewichte können, ausgehend von Überlegungen über die inhaltliche Bedeutung der verwendeten Items, der statistischen Auswertung vorgegeben werden (*a-priori*-Gewichtung). Sie können aber auch aus den erhobenen Daten selbst in einem Optimierungsverfahren gewonnen werden (*a-posteriori*-Gewichtung). Im letzten Fall dient die erstmalige Erhebung des Interviews gewissermaßen als Eichung der Auswertung für eine spätere Wiederverwendung des Untersuchungsinstruments.

Von den Entwicklern des Interviews wurde bisher für die Trennformel eine *a-priori*-Gewichtung der subjektiven Items mit  $w = 1$  und der objektiven Items mit  $w = 2$  vorgeschlagen. Abbildung 2 enthält die damit errechnete Verteilung der Gesamtpunktzahl für 406 psychisch Kranke und 868 Nichtkranke.

Beide Kurven der unteren Graphik verlaufen näherungsweise wie Gauss-Verteilungen. Da sie sich überlappen, ist eine exakte Zuordnung jedes Patienten zu einer der beiden Kurven aus der erreichten Punktzahl allein nicht möglich. Innerhalb des Überlappungsbereiches kann aber eine gewisse Punktzahl  $T$  gefunden werden, die eine bestmögliche Zuordnung erlaubt: Alle Fälle mit einer Punktzahl  $t \geq T$  werden als psychisch krank eingestuft. Der kritische Wert  $T$  ist aus der oberen Graphik der Abb. 2 zu entnehmen. Die dort eingezeichnete Trefferquote gibt den prozentualen Anteil der richtig zugeordneten Fälle in Abhängigkeit von der Punktzahl  $t$  an. In unserem Fall liegt der kritische Wert  $T$  bei einer Punktzahl von 20, dem Maximum der Trefferquote, wobei 83,1% der Patienten

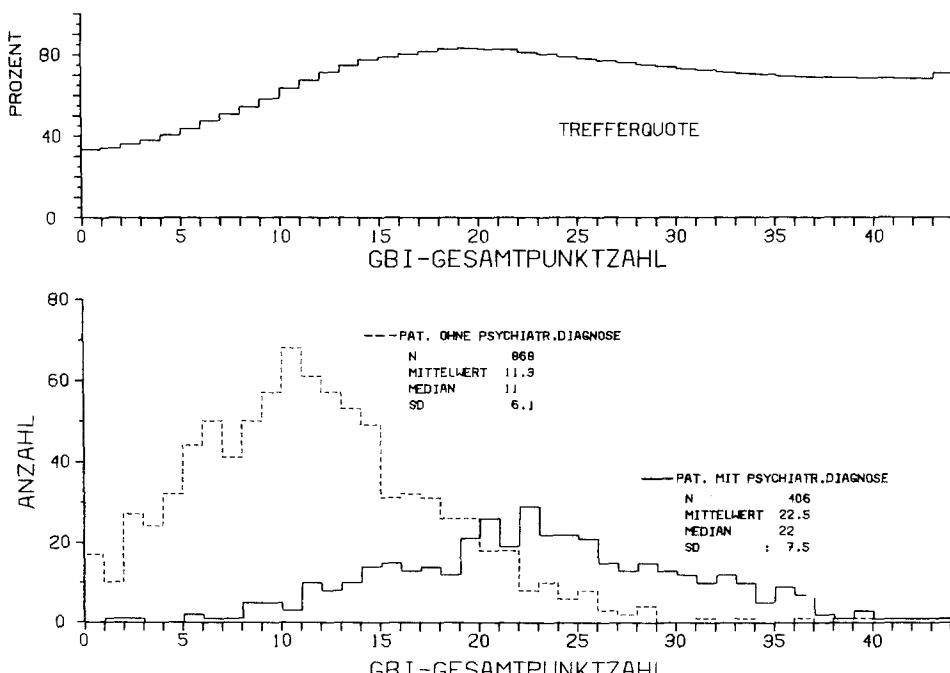


Abb. 2. Häufigkeitsverteilung von Gesamtpunktzahl und Trefferquote des Goldberg-Interviews (GBI) bei Gewichtung der subjektiven Symptome mit 1 und der objektiven Symptome mit 2

Goldberg-Score	Psychiatrische Diagnose	Keine psychiatrische Diagnose	Gesamt
≥ 20	240 (a)	49 (c)	289
0–19	166 (b)	819 (d)	985
Gesamt	406	868	1274

**Tabelle 1.** Die Beziehung zwischen Goldberg-Score und Urteil des psychiatrischen Interviewers

richtig zugeordnet sind. Dieser Wert entspricht der sogenannten Gesamteffizienz  $\left(\frac{a+d}{a+b+c+d}\right)$ . Die Spezifität  $\left(\frac{d}{c+d}\right)$  und die Sensitivität  $\left(\frac{a}{a+b}\right)$ , die sich aus Tabelle 1 berechnen lassen, liegen bei 94,9% bzw. 59,1%. Von den insgesamt 406 psychisch Kranken sind bei einem Cutt-off von 20 Punkten 40,9% (166 Patienten) nicht richtig zugeordnet.

**3.2.2 Zuordnung mit a-posteriori-Gewichtung der Items.** Um die Zuordnung möglicherweise noch zu verbessern, wurden a-posteriori-Gewichtungen analog den in dem BAYES-Theorem benutzten Größen (Logarithmen der Likelihood-Quotienten) getestet.

$$\text{Formel (2): } w(i) = \frac{10 \log p(K, i)}{p(NK, i)}$$

$p(K, i)$  ist die prozentuale Häufigkeit des Symptoms  $i$  mit Schweregraden  $\geq 2$  bei den Kranken bezogen auf die Gesamtstichprobe (analog  $p(NK, i)$ ). Dabei brauchen keine weiteren mathematischen Forderungen wie etwa statistische Signifikanz der Unterschiede oder Unabhängigkeit der Items an die Rohdaten gestellt werden. Die Formel wurde in Analogie zu einer Arbeit von Koller et al. (1972) gewählt.

Für unsere empirischen Daten ergeben sich nach Formel (2) folgende a-posteriori-Gewichte:

Symptom-Nr.:	1	2	3	4	5	6	7
Gewicht:	0,6	-1,4	0,7	1,6	-0,3	0,2	2,6
Symptom-Nr.:	8	9	10	11	12	13	14
Gewicht:	1,0	1,5	1,5	—	2,0	1,1	5,0
Symptom-Nr.:	15	16	17	18	19	20	21
Gewicht:	8,6	4,3	-1,5	8,2	—	—	7,3
Symptom-Nr.:	22	23					
Gewicht:	3,1	16,3					

Ebenso wie bei der schon beschriebenen a-priori-Gewichtung erhalten auch hier die objektiven Befunde (Symptom 12–23) durchschnittlich höhere Gewichte als die subjektiven Beschwerden (Symptome 1–11). Die belegt ihre größere Bedeutung für die Fallidentifikation.

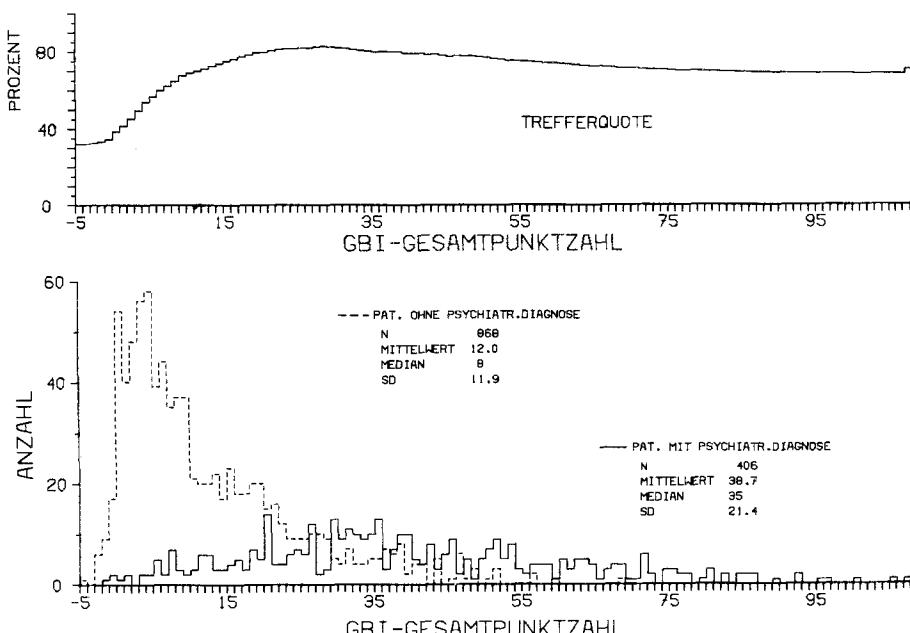


Abb. 3. Häufigkeitsverteilung der Gesamtpunktzahl und Trefferquote bei Gewichtung der Symptome mit den Logarithmen der Likelihood-Quotienten

Mit den obigen Gewichten und der Formel (1) wurde nun für jeden Patienten eine neue Gesamtpunktzahl berechnet und ihre Häufigkeitsverteilung in den beiden Gruppen „krank“ und „nicht krank“ untersucht. Die Trennschärfe dieses Ansatzes geht aus Abb. 3 hervor.

Analog wie in Abb. 2 lässt sich das Maximum der Trefferquote bestimmen; der kritische Wert liegt bei  $T = 28$  (Abb. 3). Betrachtet man diesen Wert als untere Grenze für die Zuordnung „krank“, so sind 82,7% der untersuchten Personen richtig zugeordnet, ein Wert, der in derselben Größenordnung liegt wie bei der a-priori-Gewichtung und somit keine Verbesserung der Trennschärfe erbrachte.

**3.2.3 Zuordnung mit linearer Diskriminanzanalyse.** Ein weiterer Versuch wurde unternommen, um mittels einer linearen Diskriminanzanalyse eine Zuordnung zu treffen. Es handelt sich dabei um ein aufwendiges mathematisches Verfahren, aus einem Satz quantitativer Variablen, die an Patienten mit bekannter Gruppenzugehörigkeit gemessen wurden, eine Linearkombination dieser Variablen zu finden, die eine möglichst fehlerfreie Zuordnung von Personen mit unbekannter Gruppenzugehörigkeit ermöglicht (Nie et al. 1975). Dafür werden quantitative, normalverteilte Variablen benötigt. Die lineare Diskriminanzanalyse ist aber in ihren Ergebnissen stabil gegen die Verletzung dieser Voraussetzungen durch die vorliegenden rangskalierten Daten (Koller 1967; Lachenbruch 1977).

Die Reihenfolge, in der die Variablen in die Analyse eingingen, wurde bestimmt nach ihrem Beitrag zum Anstieg der generalisierten Distanz (Methode Wilks Lambda). Die Gewichte der Variablen werden bei diesem Verfahren so berechnet, daß die Varianz der Items zwischen den Gruppen zur Varianz innerhalb der

Gruppen maximiert wurde. Die Items Müdigkeit, Konzentrationsmangel, Depersonalisation, Wahn und Halluzinationen schieden in der Diskriminanzanalyse aus, da sie keine statistisch signifikante Verbesserung der Zuordnung erbrachten. Folgende standardisierte Koeffizienten der Diskriminanzfunktion wurden damit berechnet:

Symptom-Nr.:	1	2	3	4	5	6	7
Gewicht:	-0,149	—	-0,067	+0,039	-0,096	—	-0,115
Rang:	5	—	15	17	12	—	9
Symptom-Nr.:	8	9	10	11	12	13	14
Gewicht:	-0,092	-0,087	-0,099	—	+0,108	-0,118	-0,136
Rang:	13	14	11	—	18	8	6
Symptom-Nr.:	15	16	17	18	19	20	21
Gewicht:	-0,451	-0,183	-0,120	-0,189	—	—	-0,327
Rang:	1	4	7	3	—	—	2
Symptom-Nr.:	22	23					
Gewicht:	-0,101	-0,066					
Rang:	10	16					

Schwerpunkt der Gruppe „krank“: -0,985

Schwerpunkt der Gruppe „nicht krank“: 0,460

Zuordnung zur Gruppe „krank“: Diskriminanzfunktion 0,225

Zuordnung zur Gruppe „nicht krank“: Diskriminanzfunktion 0,225

Die größere Bedeutung der objektiven Symptome für den Zuordnungsprozeß geht daraus hervor, daß diese Items überwiegend die Rangliste anführen. Die größte Bedeutung für die Zuordnung mit Hilfe der Diskriminanzanalyse hat das Item 15 (deprimiert). Es ergab sich nach Reklassifizierung im Vergleich zu den beiden ersten Verfahren eine nur geringfügig höhere Trefferquote von 84,3% („Gesunde“: 76,1%; „Kranke“: 88,1%).

### 3.3 Der Zusammenhang zwischen Symptomen und Diagnosengruppen

Die bisher verwendeten Zuordnungsverfahren sind einfache kumulierende Verfahren. Die Zuordnung wird hauptsächlich nach der Anzahl der beobachteten Symptome und ihrem Schweregrad getroffen. Dabei bleibt die Möglichkeit, daß eine bestimmte Symptomkombination gehäuft bei einer bestimmten Diagnose auftritt, unberücksichtigt. Für das Problem der Fallidentifikation ist das auch nicht wesentlich. Möchte man eine breite gefächerte Zuordnung treffen, kann jedoch auf eine Symptomkombinationsanalyse nicht mehr verzichtet werden. Dazu sind aufwendige statistische Verfahren notwendig und verschiedene Voraussetzungen bezüglich der Datenstruktur zu erfüllen (Jesdinsky 1972).

Je subtiler die Zuordnung sein soll, um so mehr Informationen müssen über jeden Patienten vorliegen. Für die Erfassung des gesamten psychopathologischen Befundes verschiedener Patientengruppen darf die Anzahl der zu erhebenden Symptome in der Regel ein gewisses Minimum (ca. 100) nicht unterschreiten

Tabelle 2. Mediane der Schweregrade der psychiatrischen Symptome in den clusteranalytischen Gruppen ( $n = 406$ )

Goldberg-Symptome (Mediane)	Cluster									Signifikanz 394 df/11 df <i>F</i> -Test
	I	II	III	IV	V	VI	VII	VIII	IX	
<b>Subjektiver Befund</b>										
1. Körperliche Symptome	1	2	2	0	2	1	0	2	1	8,7
2. Müdigkeit	0	0	0	2	0	0	0	2	1	12,7
3. Schlafstörungen	1	0	2	2	1	1	0	2	1	9,4
4. Schlaftabletten	0	0	2	0	0	0	0	0	0	12,6
5. Reizbarkeit	1	1	2	2	1	1	0	1	2	7,2
6. Mangel an Konzentration	1	1	2	2	1	1	0	1	1	7,0
7. Depression	1	1	2	2	2	1	0	2	1	8,2
8. Angst	1	1	2	1	2	0	0	2	2	10,9
9. Phobien	0	0	0,5	0	1	0	0	0	0	14,1
10. Zwangssymptome	0	0	1	0	1	0	0	0	1	11,9
11. Depersonalisation	0	0	0	0	0	0	0	0	0	—
<b>Beurteilter Befund</b>										
1. Langsam	0	1	0	2	0	0	2	1	0	35,2
2. Abwehrend	1	1	0	1	0	1	1	0	0	8,7
3. Demonstrativ	0	0	0,5	0	0	2	0	0	0	24,0
4. Deprimiert	1	1	2	2	1	0	0	2	0	14,5
5. Ängstlich	1	1	1	1	1	0	1	1	0	6,9
6. Gehoben	0	0	0	0	0	1	0	0	1	27,2
7. Affektiv verflacht	0	1	0	1	0	0	1	0	0	18,2
8. Wahn	0	0	0	0	0	0	0	0	0	—
9. Halluzinationen	0	0	0	0	0	0	0	0	0	—
10. Gestörte Intelligenz	0	2	0	1	0	0	2	0	0	59,4
11. Hypochondrie	0	2	2	0	2	0	0	1	1	19,5
12. Depressive Gedanken	0	1	1	1	0	0	0	0	0	5,6
Anzahl	108	47	14	30	59	17	45	45	31	

(Mombour 1972). In unserem Fall (23 Items) ist es daher nur sinnvoll, eine Zuordnung zu relativ groben Diagnosengruppen zu untersuchen. Die Fallzahlen bei den insgesamt 10 diagnostischen Untergruppen schwanken zwischen 5 (Schizophrenie) und 140 (depressive Neurose). Eine Untersuchung der Symptomkombinationen von nur 23 Items könnte immerhin modellhaft gewisse neue Hinweise für die klinische Diagnostik ergeben. Ein wesentlicher Aspekt der Analyse von Symptomkombinationen ist das Identifizieren von Krankheitseinheiten im Sinne von Syndromklassifikationen. Es wurde deshalb mit den Schweregraden der Items der Patienten mit psychiatrischer Diagnose ein hierarchisches Klassifikationsverfahren (Clusteranalyse) nach Ward (1963) durchgeführt. Dabei handelt es sich um ein parameterfreies Verfahren, das keine besonders einschränkenden Bedingungen an die Datenstruktur stellt. Die Symptommuster aller Patienten werden nach dem Kriterium der größten Ähnlichkeit (kleinste Varianz) iterativ in Gruppen zusammengefaßt.

Es wurde ein Optimum von 9 Clustern gefunden. Zur Charakterisierung der Symptommuster der Cluster wurden die Mediane der Schweregrade verwendet (Tabelle 2).

Die Unterschiede der Symptommuster zwischen den Clustern sind hochsignifikant. Die typischen Symptommuster der Cluster können wie folgt beschrieben werden:

Cluster	Beschreibung
I	die Gruppe mit der größten Fallzahl; die Gruppe mit unspezifischen, leichten Beschwerden.
II	wie I, jedoch zusätzlich ausgeprägte objektive Befunde.
III	die Gruppe mit einer Vielzahl ausgeprägter subjektiver Beschwerden; wenige objektive Befunde.
IV	ähnlich wie III, jedoch zusätzlich ausgeprägte objektive Befunde.
V	die Gruppe mit vielen subjektiven Beschwerden; wenige objektive Befunde.
VI	die Gruppe mit wenigen subjektiven Beschwerden; wenige objektive Befunde.
VII	die Gruppe mit keinen subjektiven Beschwerden; aber ausgeprägte objektive Befunde.
VIII	wie I, jedoch ausgeprägtere Schweregrade.
IX	die Gruppe mit vielen subjektiven Beschwerden; aber weniger ausgeprägt als bei Gruppe III; einziger objektiver Befund: Hypochondrie.

Um mit diesen Symptommustern zu einem diagnostischen Ansatz zu kommen, ist es notwendig zu untersuchen, in welchen anderen Variablen die gefundenen Cluster statistisch signifikante Unterschiede besitzen. Die mit der Clusteranalyse gefundenen Krankheitseinheiten decken sich weitgehend mit den aus dem ICD-Schema gebildeten Diagnosengruppen, denn jeder Cluster enthält hauptsächlich eine bestimmte Diagnosengruppe oder zwei symptomatisch nahe verwandte (Tabelle 3).

Tabelle 3. Prozentuale Verteilung der Diagnosen in den clusteranalytischen Gruppen

Psychiatrische Diagnose (Interviewer)	Cluster									Gesamt
	I	II	III	IV	V	VI	VII	VIII	IX	
(Prä)senile Demenz	3,7	10,6	7,1	26,7	—	5,9	44,4	15,6	—	9,9
Andere organische psychiatrische Erkrankungen	—	19,2	—	—	3,4	5,9	11,1	—	—	4,2
Schizophrenie	0,9	—	—	6,7	—	—	2,2	—	—	1,2
Affektive und andere Psychosen	7,4	8,5	7,1	20,0	8,5	—	2,2	—	9,7	9,9
Depressive Neurose	44,4	21,3	42,9	20,0	35,6	11,8	2,2	66,7	38,7	34,0
Andere Neurosen	9,3	8,5	35,7	6,7	20,3	17,6	—	11,1	38,7	13,1
Persönlichkeitssstörungen	4,6	12,8	7,1	6,7	10,2	58,8	4,4	2,2	6,5	8,9
Psychosomatische Erkrankungen	24,1	4,3	—	—	22,0	—	—	4,4	6,5	11,1
Alkoholismus oder Drogenabhängigkeit	3,7	4,3	—	6,7	—	—	6,7	—	—	2,7
Oligophrenie	1,9	10,6	—	6,7	—	—	26,7	—	—	5,2
Gesamt ( $n = 100\%$ )	108	47	14	30	59	17	45	45	31	406

**Tabelle 4.** Demographische und klinische Kennwerte der clusteranalytischen Gruppen

	Cluster									Gesamt
	I	II	III	IV	V	VI	VII	VIII	IX	
Alter (Jahre)	45,2	61,3	57,2	56,9	43,3	48,4	60,5	48,0	48,6	50,8
Geschlecht (Anteil der Männer)	32,4	51,1	28,6	53,3	22,1	52,9	55,6	24,4	25,8	35,7
Überweisung zum Psychiater erforderl.	30,6	25,5	64,3	46,7	55,9	17,7	17,8	51,1	35,5	37,8
Einnahme von Psychopharmaka	12,0	14,9	14,3	16,7	25,4	—	—	20,0	32,3	16,0
% bezogen auf n =	108	47	14	30	59	17	45	45	31	406

Das diagnostische Vorgehen könnte jetzt nach einem „wenn — dann“-Schema ablaufen. Zum Beispiel: Wenn ein sonst unbekannter Patient in den Items des Goldberg-Interviews eine Symptomverteilung ähnlich des Clusters VIII besitzt, dann ist die Wahrscheinlichkeit, daß er an einer depressiven Neurose leidet, gleich 66,7%. Beziehen wir noch weitere soziodemographische und klinische Parameter ein (Tabelle 4), so können wir die Gruppen wie folgt charakterisieren: Zum Beispiel sind Patienten der Gruppe VIII im Mittel 48 Jahre alt und mit einer Wahrscheinlichkeit von 24,4% männlich, die Wahrscheinlichkeit für eine psychiatrische Überweisungsbedürftigkeit liegt bei 51,1%, die einer psychopharmakologischen Medikation bei 20,0%.

#### 4. Diskussion

In der vorliegenden Untersuchung sollte die Eignung des Goldberg-Interviews zur psychiatrischen Fallidentifikation von Patienten in Allgemeinpraxen analysiert werden. Da ein objektives Außenkriterium für die Gültigkeit einer Fallidentifikation nicht existiert, wurde als Maßstab das diagnostische Urteil der psychiatrischen Interviewer verwendet.

Die gewonnenen Ergebnisse bedürfen also noch der Bestätigung an unabhängigem Material, da — wie die Erfahrung zeigt — Trefferquoten, die an dem gleichen Material wie die Trennregeln gewonnen sind, die Güte des Algorithmus überschätzen.

Für die Einzelitems konnte gezeigt werden, daß die erhobenen subjektiven Beschwerden im Gegensatz zu den meisten objektiven Symptomen nur mit relativ geringem Gewicht an der Fallidentifikation beteiligt sind, da auch Nichtkranke relativ häufig subjektive Störungen mit krankheitswertigen Schweregraden angeben.

Unsere Ergebnisse der Optimierung des Verfahrens der psychiatrischen Fallidentifikation bestätigen Feststellungen, die Koller et al. (1972) in einem diagnostischen Simulationsmodell an einer synthetischen Krankheit trafen: Die Treffer-

quoten, die mit Hilfe eines Gesamtscores aus einer einfachen Gewichtung und Summierung der Punktewerte der Items zu erreichen sind (83,1%), liegen in etwa der gleichen Größenordnung wie diejenigen, die mit einer a-posteriori-Gewichtung analog den in dem Bayes-Theorem benutzten Größen (82,7%) bzw. eines diskriminanzanalytischen Ansatzes (84,2%) gewonnen wurden. Für künftige epidemiologische Untersuchungen besteht somit keine Veranlassung, die von Goldberg et al. vorgeschlagene einfache Gewichtung der subjektiven und doppelte Gewichtung der objektiven Symptome aufzugeben.

In einer analogen Untersuchung in Mannheimer Allgemeinpraxen (Zintl-Wiegand et al. 1978) lag die Gesamteffizienz bei einem Cut-off von 20 Punkten mit 92,0% sogar noch höher. In einer Feldstudie von Macmillan (1959) betrug die Übereinstimmung zwischen psychiatrischer Beurteilung und „Health Opinion Survey“ 86%. Eine Gegenüberstellung der Punktewerte des „General Health Questionnaire“ und der Fallidentifikation durch den Psychiater ergab Übereinstimmungsraten von 76% (Finlay-Jones und Murphy 1979), 80% (Goldberg et al. 1976) und 92% (Tennant 1977). Die übereinstimmende Beurteilung aufgrund der Gesamtpunktzahl und der Fallidentifikation durch den Interviewer liegt deutlich über derjenigen zwischen Hausarzt und Interviewer, wo bezüglich der unabhängig voneinander beurteilten Patienten eine Gesamteffizienz von 73,5% erzielt wurde (Dilling und Weyerer 1978). Zu einem ähnlichen Ergebnis kam auch die Mannheimer Forschungsgruppe; hier stimmten Hausarzt und Interviewer zu 71,1% in der Fallidentifikation überein.

Eine wesentliche Schwierigkeit, die Gesamteffizienz zu erhöhen, besteht darin, daß in dem Spektrum psychischer Störungen monosymptomatische Krankheitsbilder nur einen niedrigen Gesamtscore erreichen und als falsch negativ klassifiziert werden. Darüber hinaus kann ein bestimmtes Krankheitsverhalten wie etwa die Abwehr psychischer Konflikte oder ein Mangel an Krankheitseinsicht zu Fehlklassifikationen beitragen.

Das Goldberg-Interview hat sich als ein nützliches Instrument für die epidemiologische Forschung in Allgemeinpraxen erwiesen. Möglicherweise könnte aber die Aufnahme weiterer relevanter Symptome wie etwa Alkoholabhängigkeit die Validität des Interviews noch erhöhen.

## Literatur

- Armor DJ, Couch AS (1972) Data-text primer. The Free Press, New York  
Bruder W (1978) Psychiatrische Fallidentifikation mit einem standardisierten psychiatrischen Interview bei epidemiologischer Forschung in Allgemeinpraxen. Dissertation, München  
Dilling H, Weyerer S (1978) Patienten mit psychischen Störungen in der Allgemeinpraxis und ihre psychiatrische Überweisungsbedürftigkeit. Bericht an die Deutsche Forschungsgemeinschaft über das Projekt A 2b im Sonderforschungsbereich 116, München  
Dilling H, Weyerer S, Enders I (1978) Patienten mit psychischen Störungen in der Allgemeinpraxis und ihre psychiatrische Überweisungsbedürftigkeit. In: Häfner H (Hrsg) Psychiatrische Epidemiologie. Springer, Berlin Heidelberg New York  
Finlay-Jones RA, Murphy E (1979) Severity of psychiatric disorders and the 30-item general health questionnaire. Br J Psychiat 134:609-616  
Goldberg DP, Cooper B, Eastwood MR, Kedward HB, Shepherd M (1970) A standardized psychiatric interview for use in community surveys. Br J Prev Soc Med 24:18-23

- Goldberg DP, Rickels K, Downing R, Hesbacher P (1976) A comparison of two psychiatric screening tests. *Br J Psychiat* 129:61–67
- Häfner H (1978) Einführung in die psychiatrische Epidemiologie. Geschichte, Suchfeld, Problemlage. In: Häfner H (Hrsg) *Psychiatrische Epidemiologie*. Springer, Berlin Heidelberg New York
- Jesdinsky HJ (1972) Diagnose — Modelle in der Medizin. *Meth Inform Med* 11:48–59
- Koller S (1967) Mathematisch-statistische Grundlagen der Diagnostik. *Klin Wochenschr* 45: 1065–1072
- Koller S, Michaelis J, Scheidt E (1972) Untersuchungen an einem diagnostischen Simulationsmodell. *Meth Inform Med* 11:213–227
- Lachenbruch PA (1979) Some misuses of discriminant analysis. *Meth Inform Med* 16:255–258
- Macmillan AM (1959) A survey technique for estimating the prevalence of psychoneurotic and related types of disorders in communities. In: Pasamanick B (ed) *Epidemiology of mental disorders*. American Association for the Advancement of Science, New York
- Mombour W (1972) Verfahren zur Standardisierung des psychopathologischen Befundes. *Psychiat Clin* 5:73–120, 137–157
- Nie NH, Hull CH, Jenkins JG, Steinbrunner K, Bent DH (1975) *SPSS — Statistical package for the social sciences*. McGraw Hill, New York
- Rollett B, Bartram M (1976) Einführung in die hierarchische Clusteranalyse. Klett, Stuttgart
- Tennant C (1977) The general health questionnaire: a valid index of psychological impairment in Australian populations. *Med J Austr* 2:392–394
- Ward JH (1963) Hierarchical grouping to optimize an objective function. *Am Stat Assoc J* 58:236–244
- Wing JK, Cooper JE, Sartorius N (1974) Description and classification of psychiatric symptoms. Cambridge University Press, Cambridge
- Zintl-Wiegand A, Schmidt-Maushart C, Leisner R, Cooper B (1978) Psychische Erkrankungen in Mannheimer Allgemeinpraxen. In: Häfner H (Hrsg) *Psychiatrische Epidemiologie*. Springer, Berlin Heidelberg New York

Eingegangen am 3. September 1980