

Übereinstimmung zwischen konventioneller und Computer-Diagnostik bei ambulanten psychiatrischen Patienten

ROLF CASTELL, HELMUT BUSCH, JOSEF HOFFMANN, DIETER MITTELSTEN SCHEID, WERNER MOMBOUR und HELMUT WALDMANN*

Max-Planck-Institut für Psychiatrie, Kraepelinstr. 10, D-8000 München,
Bundesrepublik Deutschland

Eingegangen am 11. Februar 1976

AGREEMENT BETWEEN TRADITIONAL AND COMPUTER DIAGNOSIS IN PSYCHIATRIC OUTPATIENTS

SUMMARY. Psychiatric diagnosis based on personal interviews with 100 clinic outpatients was made and the data thus obtained were documented in Spitzer's and Endicott's CAPPS (Current and Past Psychopathology Scales). Two clinicians and the computer program DIAL - a modification of the DIAGNO II program of the same authors - diagnosed the same patients independently of each other basing their diagnoses solely on the entries in the CAPPS. The diagnostic agreement between the different diagnoses was evaluated and results compared with earlier findings by Schmid et al. These authors had, however, used CAPPS data gathered from case histories and not from direct interviews of patients. The modification in the procedure of gathering the data resulted in a slight, but statistically insignificant improvement in diagnostic agreement. The possibility of improving diagnostic agreement by the use of another scale is discussed.

KEY WORDS: Computer Diagnostics - Diagnostic Agreement - Psychiatric Rating Scales.

ZUSAMMENFASSUNG. Bei 100 psychiatrischen Patienten wurde in der poliklinischen Sprechstunde eine Diagnose gestellt und die erhobenen Daten in die anamnestischen und Befundbögen der CAPPS (Current and Past Psychopathology Scales) von Spitzer und Endicott eingetragen. Das Computerprogramm DIAL - eine Modifikation des DIAGNO II-Programms derselben Autoren - und zwei Kliniker stellten unabhängig davon und nur auf Grund der Angaben in den CAPPS-Bögen ebenfalls eine Diagnose für die gleichen Patienten. Die Übereinstimmung zwischen diesen verschiedenen Diagnosen wurde geprüft und die Ergebnisse mit früheren ähnlichen Untersuchungen von Schmid et al. verglichen, bei denen jedoch die CAPPS unter Verwendung

* Die Autoren danken Herrn Professor Dr. med. D. v. Zerssen und Herrn Dipl. Math. P. O. Degens für Ratschläge und Hilfe bei der Abfassung der Arbeit.

Sonderdruckanforderungen an Dr. W. Mombour.

der Krankengeschichte und nicht auf Grund direkter ambulanter Untersuchung ausgefüllt worden waren. Diese Modifikation der Datenerhebung erbrachte eine leichte, aber statistisch nicht signifikante Verbesserung in der Diagnostiker-Übereinstimmung. Die Möglichkeit, durch Verwendung einer anderen Skala die diagnostische Übereinstimmung zwischen ambulanter, Computer- und Fragebogendiagnostik zu erhöhen, wird diskutiert.

SCHLÜSSELWÖRTER: Computerunterstützte Diagnostik - diagnostische Übereinstimmung - psychiatrische Schätzskalen.

1. EINFÜHRUNG

Die mangelhafte Übereinstimmung in der psychiatrischen Diagnostik hat in den letzten Jahrzehnten immer wieder die Aufmerksamkeit auf die Details des diagnostischen Prozesses gelenkt (vgl. Cooper et al., 1972; Mombour, 1975). Da die Diagnosenstellung ein zusammengesetzter, aus mehreren Anteilen bestehender Vorgang ist (vgl. v. Zerssen, 1973), können Ungenauigkeiten in jedem dieser Anteile zur Gesamtdiskrepanz beitragen. So führen z. B. Ward et al. (1962) die Unterschiede in der psychiatrischen Diagnostik zu 62,5% auf ein ungenügendes Diagnoseschema zurück, zu 32,5% auf ein inkonsistentes Verhalten der Diagnostiker und nur zu 5% auf Widersprüche im psychopathologischen Befund. Die allgemeine Einführung des psychiatrischen Diagnosenschlüssels der Weltgesundheitsorganisation (ICD = International Classification of Diseases, 8. Revision, deutsche Ausgabe: Degkwitz et al., 1975) und des dazugehörigen Kommentars hat eine wichtige Voraussetzung für eine bessere Übereinstimmung in der psychiatrischen Diagnostik geschaffen. Es bleibt aber Aufgabe, andere Ursachen für eine diskrepante Diagnostik auszuschalten (vgl. z. B. Zubin, 1967; Shepherd et al., 1968). Hierzu gehören vor allem die Bestrebungen, mittels psychiatrischer Schätzskalen eine Standardisierung der Befunderfassung und -dokumentation (strukturiertes, evtl. vollstandardisiertes Interview) zu erreichen, Schweregrad und Häufigkeit der Symptome in graduierter Abstufung aufzuzeichnen und die verwendeten psychopathologischen Begriffe zu operationalisieren; auch eine Standardisierung des Weges von den Symptomen über Syndrome zur nosologischen Diagnose gemäß fester Regeln (Algorithmus) wird angestrebt (Lange, 1971). Die Anwendung dieser Methoden führte im "US-UK-diagnostic project" (Diagnosenvergleich zwischen London und New York an den gleichen Patienten, einmal ohne, das zweite Mal mit standardisiertem Verfahren; Cooper et al., 1972; Wing et al., 1974) zu einer überzeugenden Verbesserung bzgl. der diagnostischen Übereinstimmung. Auch die "International Pilot Study of Schizophrenia" (WHO, 1973) arbeitete mit der gleichen Methodik.

Bei unseren eigenen Bemühungen um eine standardisierte Diagnostik verwendeten wir das - seinerzeit einzig erhältliche - von Spitzer und Endicott (1968 und 1969) entwickelte Computerprogramm DIAGNO II mit den Anamnese- und Befundbögen CAPPS = Current And Past Pychopathology Scales. In der von Gerster und Dirlich (1975) erarbeiteten Version DIAL ist die praktische Anwendbarkeit eines solchen Programms von Schmid et al. (1974) auf Grund von anamnestischen und psychopathologischen Befunden aus 98 Krankengeschichten untersucht worden. Dabei wird die Übereinstimmung von Klinikern, Computer und zwei sog. "Fragebogendiagnostikern"

miteinander verglichen. Die vorliegende Arbeit schließt sich in der Methode dem von Schmid et al. (1974) entwickelten Vorgehen an, jedoch mit einem wesentlichen Unterschied bezüglich der Datenerhebung. Bei Schmid et al. stellte der Kliniker seine Diagnose aufgrund der direkten Untersuchung des Patienten und trug diese in die Krankengeschichte ein. Die CAPPs-Fragebögen, die die Grundlage für die Diagnosenstellung des Computers und der beiden "Fragebogendiagnostiker" bildeten, waren nicht auf Grund einer direkten Patientenuntersuchung ausgefüllt worden, sondern von einem an der Untersuchung sonst nicht beteiligten Arzt aufgrund der Angaben in der Krankengeschichte. Bei der Untersuchung, über die hier berichtet wird, wurden von den Ärzten, die die klinische Diagnose stellten, im Anschluß an die Untersuchung gleichzeitig auch die CAPPs-Fragebögen mit 130 Items ausgefüllt. Diese auf Grund einer direkten Untersuchung ausgefüllten CAPPs-Bögen wurden sodann dem Computer und den beiden "Fragebogendiagnostikern" zur Diagnosenstellung vorgelegt. Eine zusätzliche Modifikation gegenüber Schmid et al. - vor allem arbeitsorganisatorisch bedingt - liegt darin, daß die klinische Diagnosenstellung und das Ausfüllen der CAPPs-Bögen bei ambulanten Patienten in der Poliklinik vorgenommen wurde, wo nur Ärzte mit längerer Berufserfahrung arbeiteten. Durch diese Modifikationen gegenüber Schmid et al. sollte folgende Hypothese überprüft werden: Die diagnostische Übereinstimmung läßt sich erhöhen, wenn die CAPPs-Fragebögen sofort im Anschluß an die Untersuchung des Patienten ausgefüllt werden und wenn hinsichtlich ihrer Berufserfahrung kein wesentlicher Unterschied zwischen den "Fragebogendiagnostikern" und den Ärzten, die die klinische Diagnose stellen und die Bögen ausfüllen, besteht. Die anderen Unterschiede in der Methodik gegenüber Schmid et al. (ambulante Patienten statt stationärer, andere Zusammensetzung in der Diagnosenhäufigkeit) schienen uns keine wesentlichen Unterschiede zu sein; die Angaben, die man zum Eintragen in die CAPPs-Bögen benötigt, können in entsprechend kurzer Untersuchung erhoben werden, so daß der unterschiedliche Zeitaufwand zwischen ambulanter und stationärer Patienten-Exploration nicht ins Gewicht fällt; auch bei unterschiedlicher Zusammensetzung in der Diagnosenhäufigkeit lassen sich die einzelnen Diagnosen-Gruppen direkt miteinander vergleichen. Auf Grund der Kenntnis der Patienten muß auch bezweifelt werden, ob sich die ambulanten Patienten im Schweregrad ihrer Erkrankung wesentlich von den stationären unterschieden. Es gibt sogar Anhaltspunkte dafür, daß der Unterschied zwischen den Ärzten in den beiden verglichenen Arbeiten nicht sehr wesentlich ist; in der vorliegenden Arbeit nahmen nur Fachärzte an der Untersuchung teil, in der früheren von Schmid et al. waren auch Anfänger an der Krankengeschichtsdiagnose beteiligt, die Diagnosen wurden jedoch erst nach Vorstellung der Patienten und Diskussion im Beisein von Fachärzten gestellt.

2. METHODE

Drei Ärzte der Ambulanz (AD = Ambulanz-Diagnostiker) des Max-Planck-Instituts für Psychiatrie in München füllten 1972 und 1973 nach Abschluß der ambulanten Diagnostik von 100 Patienten den Anamneseteil und den psychopathologischen Befundteil der CAPPs nach Spitzer & Endicott (1969) mit insgesamt 130 Items aus. AD_H (30 Patienten), AD_B (30 Patienten) und AD_W (40 Patienten) waren Fachärzte

Tabelle 1. Verzeichnis der DIAL-Diagnosen; in der ersten Spalte sind die Kategorien der Einteilung 1 (A, B, C, D), der Einteilung 2 (A1, A2, B, C1, C2, D) und der Einteilung 3 (A1a, A1b, B, C1a bis C1e, C2a, C2b, D) aufgeführt. Die zweite Spalte gibt die ICD-Nummern der in der dritten Spalte folgenden Diagnosen. Aus Gründen der nosologischen Ordnung wurde die Reihenfolge der ICD-Nummern an zwei Stellen und der Buchstabenummerierung an einer Stelle geändert

Ala	2910	Delirium tremens
Ala	2911	Alkoholisches Korsakowsyndrom
Ala	29152	Chronische Alkoholpsychose
Alb	2930	Psychose bei Hirnarteriosklerose
Alb	2934	Psychose bei degenerativen Erkrankungen des ZNS
Alb	29431	Akute Psychose bei Intoxikation durch Arzneimittel o. Gifte
Alb	29432	Chronische Psychose bei Intoxikation durch A. oder Gifte
Alb	2948	Psychose bei anderen körperlichen Erkrankungen
A2a	2950	Schizophrenia simplex
A2a	2951	Schizophrenie, hebephrene Form
A2a	2952	Schizophrenie, katatone Form
A2a	2953	Schizophrenie, paranoide Form
A2a	2954	Akute schizophrene Episode
A2a	2959	Nicht näher bezeichnete Schizophrenieform
A2a	2970	Paranoia
A2b	2957	Schizoaffektive Psychose
A2c	2961	Endogene Manie
B	2962	Endogene Depression
B	2968	Andere affektive Psychoosen
B	2960	Involutionsdepression
B	2980	Reaktive depressive Psychose
B	3004	Depressive Neurose
C1a	3000	Angstneurose
C1a	30013	Hysterische Neurose, Konversionssymptome
C1a	30014	Hysterische Neurose, Dämmerzustand
C1a	3002	Phobie
C1a	3003	Zwangsneurose
C1a	3072	Anpassungsstörung bei situativer Belastung in der Adoleszenz
C1a	3073	Anpassungsstörung bei situativer Belastung im Erwachsenenalter
C1a	3074	Anpassungsstörung bei situativer Belastung im höheren Alter
C1b	3010	Paranoide Persönlichkeit
C1b	3011	Cyclothyme Persönlichkeit
C1b	3012	Schizoide Persönlichkeit
C1b	3014	Anankastische Persönlichkeit
C1b	3015	Hysterische Persönlichkeit
C1b	3017	Antisoziale Persönlichkeit
C1b	30181	Passiv-aggressive Persönlichkeit
C1b	30182	Inadäquate Persönlichkeit
C1b	3163	Asoziales Verhalten
C1b	3018	Andere Persönlichkeitsstörungen
C1c	3029	Nicht näher bezeichnete sexuelle Verhaltensabweichungen
C2a	3039	Nicht näher bezeichnete Form des Alkoholismus
C2b	3049	Nicht näher bezeichnete Medikamentenabhängigkeit, Sucht u. mißbrauch
C1d	3055	Psychosomatische Störungen, Magen-Darm-Trakt
C1e	3091	Psychische Störungen bei Intoxikationen
C1e	3093	Psychische Störungen bei Kreislaufstörungen
C1e	3098	Psychische Störungen bei degenerativen Erkrankungen des ZNS
C1e	3158	Nicht näher bestimmbarer Schwachsinn bei Störungen des psychosozialen Milieus
D	3170	Psychiatrisch unklarer Fall
D	3180	Keine psychischen Erkrankungen

für Psychiatrie und Neurologie. Sie stellten auf Grund der Untersuchung eine Haupt- und gegebenenfalls eine Zusatzdiagnose aus einem Verzeichnis von 50 leicht modifizierten ICD-Diagnosen (Tabelle 1), wie sie im DIAL-Programm Verwendung finden (vgl. Schmid et al., 1974). Die ausgefüllten CAPPS-Fragebögen wurden dann zwei Fachärzten für Psychiatrie und Neurologie (FD = "Fragebogendiagno-

stiker") zur Diagnosenstellung vorgelegt. Außerdem wurde die gleiche Information einem Computer (vgl. Gerster & Dirlich, 1975) eingegeben und die Diagnose aufgrund des DIAL-Programms gestellt (vgl. Schmid et al., 1974). Es wurden - wie bei den ambulanten Diagnostikern - aus den 50 im Programm vorgesehenen ICD-Diagnosen (Tabelle 1) eine Haupt- und gegebenenfalls eine Zusatzdiagnose ausgewählt.

Die Auswertung geschah in drei Stufen zunehmender diagnostischer Differenzierung wie bei Schmid et al. (1974). Für die Details dieser drei Einteilungen sei auf Tabelle 1 hingewiesen. Die Einteilung 1 umfaßt die Kategorien Psychosen ohne Depression (= A), alle Depressionsformen (= B), Nicht-Psychosen (wie Neurosen, Persönlichkeitsstörungen etc. = C) und psychiatrisch unklare Fälle (= D). Einteilung 2 enthält die Kategorien A1, A2, B, C1, C2 und D. Einteilung 3 setzt sich aus den Kategorien A1a, A1b, A2a bis A2c, B und C1a bis C1e, C2a, C2b und D zusammen. Verglichen wurden die Erstdiagnosen, die von jeweils einem AD, den beiden FD und auf Grund des Computer-Programms gestellt worden waren.

Die Übereinstimmung zweier Diagnostiker ist wie bei Schmid et al. der Prozentsatz der übereinstimmenden Diagnosenpaare, bezogen auf die Anzahl aller Diagnosenpaare der beiden Diagnostiker. Da für 100 Patienten Diagnosen erstellt wurden, ist bei der Prüfung der Gesamtübereinstimmung pro Einteilung (1, 2 oder 3) dieser Prozentsatz gerade die Anzahl der übereinstimmenden Diagnosenpaare, so weit nur die Gesamtübereinstimmung und nicht Übereinstimmung für einzelne Gruppen geprüft wird. Bei jeder Einteilung ergibt sich so für jeweils zwei Diagnostiker ein Wert für ihre Übereinstimmung (Tabelle 2). Da zwei Diagnostiker häufig zwar ähnliche Diagnosen stellen werden, die aber nicht genau übereinstimmen müssen (z. B. C1b und C2b), wird die Übereinstimmung bei der höheren Einteilung im allgemeinen höher sein, als bei der genauer differenzierenden Einteilung, d. h. die Übereinstimmung nimmt ab von Einteilung 1 zur Einteilung 3. Die Übereinstimmung zweier Diagnostiker in einer bestimmten diagnostischen Kategorie, z. B. C1a (Neurose) in Einteilung 3, ergibt sich als Prozentsatz der Diagnosenpaare, in denen beide Diagnosen der jeweiligen Kategorie angehören, bezogen auf die Diagnosenpaare, in denen mindestens eine der Diagnosen der jeweiligen Kategorie angehört (Tabelle 3).

Schließlich wurden zum Vergleich mit Schmid et al. auch die Zweitdiagnosen berücksichtigt. Die Berechnung erfolgte wie dort angegeben. Die entsprechenden Werte für die Einteilung 3 finden sich in Tabelle 2 (4. Zeile).

3. ERGEBNISSE

- A. Die Gesamtübereinstimmung der Erstdiagnosen geht aus Tabelle 2 hervor. Sie ist in Einteilung 1 mit vier diagnostischen Kategorien durchschnittlich 77%. In Einteilung 2 mit sechs diagnostischen Kategorien liegt sie im Durchschnitt bei 74% und in Einteilung 3 mit 14 diagnostischen Kategorien bei 67%. Die durchschnittliche diagnostische Übereinstimmung sinkt beim Übergang von 4 möglichen auf 14 mögliche Kategorien von 77% auf 67%.

Tabelle 2. Gesamtübereinstimmung zwischen den verschiedenen Diagnostikern in den diagnostischen Einteilungen 1, 2 und 3 (Einteilung 3 mit Zusatzdiagnosen; in Klammern steht die Gesamtzahl der Diagnosenpaare)

	AD-FD _M	AD-FD _C	AD-DIAL	FD _M -DIAL	FD _C -DIAL	FD _M -FD _C	TOTAL
Einteilung 1 Erstdiagnose	82 % (100)	72 % (100)	76 % (100)	80 % (100)	73 % (100)	80 % (100)	77 % (100)
Einteilung 2 Erstdiagnose	75 % (100)	72 % (100)	70 % (100)	77 % (100)	70 % (100)	78 % (100)	74 % (100)
Einteilung 3 Erstdiagnose	65 % (100)	66 % (100)	62 % (100)	71 % (100)	65 % (100)	71 % (100)	67 % (100)
Einteilung 3 Erst- u. Zu- satzdiagnose	52 % (130)	56 % (140)	45 % (154)	51 % (153)	48 % (161)	55 % (141)	51 % (147)

Die Abnahme der Übereinstimmung bei der Zunahme der diagnostischen Kategorien wird von einer Reihe von Autoren beschrieben (Ellendorf, 1975; Krauß, 1974; Kreitman et al., 1961; Schmid et al., 1974; Zubin, 1967).

Unter Berücksichtigung der Erst- und Zusatzdiagnosen liegt die durchschnittliche Übereinstimmung in Einteilung 3 um durchschnittlich 16% unter der der Erstdiagnosen. Die Anzahl der Erst- und Zusatzdiagnosenpaare beträgt durchschnittlich 147.

Die Gesamtübereinstimmung zwischen den einzelnen Diagnostikern in den drei Einteilungen geht ebenfalls aus Tabelle 2 hervor. Die Werte lie-

Tabelle 3. Übereinstimmung in den 14 diagnostischen Kategorien der Einteilung 3 (Erstdiagnose). Die erste Spalte gibt die Diagnosen; in der Spalte 2 bis 7 sind die Zahl der von jedem Diagnostiker gestellten Diagnosen und in Klammern die Zahl der übereinstimmenden Diagnosen angegeben. Die letzte Spalte gibt die dazugehörige Gesamtzahl der Diagnosen aus 6 Paarvergleichen; in Klammern steht die Zahl der Übereinstimmungen in 6 Paarvergleichen. Die Übereinstimmung in Prozent berechnet sich aus der Zahl der übereinstimmenden Diagnosen, dividiert durch die Gesamtzahl minus Zahl der übereinstimmenden Diagnosen mal 100. Bei Schmid et al. ist in der letzten Spalte der entsprechenden Tabelle von der Gesamtzahl der Diagnosen die Zahl der übereinstimmenden Diagnosen bereits abgezogen. Zur Rubrik "keine psychischen Erkrankungen" wurden keine Diagnosen gestellt

Diagnosen	AD - FD _M	AD - FD _C	AD - DIAL	FD _M - DIAL	FD _C - DIAL	FD _M - FD _C	Total
A1a Alkoholpsychosen	4 (4)	4 (3)	4 (3)	4 (3)	3 (3)	4 (3)	42 (19)
A1b Organische Psychosen	7 (2)	3 (3)	7 (3)	3 (2)	3 (2)	3 (2)	60 (30 %)
A2a Schizophrenie	10 (10)	12 (10)	10 (8)	12 (8)	14 (8)	12 (10)	10 (14)
A2b Schizoaaffektive Psychosen	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
A2c Manie	0 (0)	0 (0)	1 (0)	0 (0)	0 (0)	1 (0)	1 (0)
B Depression	27 (23)	33 (23)	27 (19)	33 (23)	27 (22)	33 (28)	33 (38)
C1a Neurosen	22 (10)	20 (12)	21 (12)	20 (15)	20 (11)	20 (13)	21 (24)
C1b Psychopathien	10 (4)	7 (3)	5 (4)	8 (4)	7 (4)	8 (3)	5 (0)
C1c Sexuelle Abweichungen	3 (3)	3 (3)	3 (2)	2 (2)	3 (2)	2 (3)	0 (0 %)
C2a Alkoholabusus	4 (4)	6 (4)	5 (4)	4 (4)	6 (4)	4 (4)	5 (33 %)
C2b Sonstige Sichter	5 (5)	11 (5)	7 (3)	10 (5)	11 (10)	10 (7)	11 (6) 7 (36)
C1d Psychosomatische Krankheiten	7 (0)	0 (0)	7 (2)	2 (2)	0 (0)	2 (1)	2 (0) 2 (5)
C1e Psychische Störung bei organischen Krankheiten	1 (0)	0 (0)	1 (0)	0 (0)	0 (0)	0 (0)	3 (0) 0 (0 %)
D Unklar	0 (0)	1 (0)	0 (0)	4 (0)	7 (0)	4 (2)	7 (1) 4 (2) 4 (0) 4 (2) 6 (2) 6 (2 %)

gen bei den Erstdiagnosen zwischen 72% und 82% (Einteilung 1), 70% und 78% (Einteilung 2) und 62% und 71% (Einteilung 3) bei den Erstdiagnosen.

B. Betrachtet man das Ausmaß der durchschnittlichen Übereinstimmung in den einzelnen diagnostischen Kategorien der Einteilung 3 (Tabelle 3), so ergeben sich für die Erstdiagnosen die besten Übereinstimmungen bei Alkoholismus (C2a) und Alkoholpsychosen (A1a) und außerdem bei sexuellen Verhaltensabweichungen (C1c). Eine mittlere Gruppe hinsichtlich diagnostischer Übereinstimmung bilden die schizophrenen Psychosen (A2a), die Depressionen (B), die Neurosen (C1a) und die Süchte (C2b), während die Psychopathien (C1b), die psychosomatischen Erkrankungen (C1d) und die organischen Psychosen (A1b) nur wenig Übereinstimmung zeigen. Für A1a, C2a, C2b, C1b, C1c, C1d und A1b liegt allerdings die Häufigkeit, mit der ein Diagnostiker jeweils eine dieser Diagnosen wählte, bei maximal 10, so daß wohl nur die Werte für schizophrene Psychosen, Depressionen und Neurosen interpretierbar sind.

4. DISKUSSION

Die diagnostische Übereinstimmung bei der Erstdiagnose (77% bei Einteilung 1 mit 4 diagnostischen Kategorien, 67% bei 14 diagnostischen Kategorien) entspricht der Übereinstimmung, wie sie z. B. bei Zubin (1967) und bei Kreitman (1961) zu finden sind. Das Ausmaß der diagnostischen Übereinstimmung in unserer Arbeit liegt höher als die von Schmid et al. (1974) gefundenen. Die Tendenz ist eindeutig, da in jeder Einteilung sowohl bei Erstdiagnosen als auch bei Erst- und Zusatzdiagnosen die Übereinstimmungen höher ausfallen. Schmid et al. (1974) haben zum Ausfüllen der CAPPS Informationen verwendet, die aus Krankengeschichten ausgewählt wurden, während in der vorliegenden Untersuchung die ambulant diagnostizierenden Ärzte (AD) selbst unmittelbar nach der Diagnostik die CAPPS ausfüllten. Die Unterschiede zwischen Schmid et al. und uns bezüglich der Übereinstimmung wurde im Chi²-Test (einseitig, bei einem Freiheitsgrad) geprüft. Obwohl sich bei den Prozentwerten in allen Einteilungen eine Verbesserung gegenüber den Werten von Schmid et al. (1974) ergab, sind die Unterschiede aufgrund des Tests nicht signifikant. Die Ergebnisse sollen kurz aufgeführt werden.

Die Verbesserung in der Übereinstimmung, wenn nur die Erstdiagnosen verglichen werden, liegt in Einteilung 1 bei 4% ($\text{Chi}^2 = 0,43$), in Einteilung 2 bei 9% ($\text{Chi}^2 = 1,91$) und bei Einteilung 3 bei 11% ($\text{Chi}^2 = 2,48$). Dasselbe gilt für die Beschreibung eines Krankheitsbildes durch Erst- und Zusatzdiagnose. Die Ergebnisse dieser Untersuchung sind in Einteilung 3 um 9% ($\text{Chi}^2 = 2,32$) höher als bei Schmid et al. (1974), wobei zu berücksichtigen ist, daß bei Erst- und Zusatzdiagnosen 147 Diagnosenpaare in dieser Untersuchung 160 in der von Schmid et al. (1974) gegenüberstehen.

Eine Verbesserung der Diagnostik ist in den Kategorien hoher diagnostischer Übereinstimmung wie Alkoholpsychosen (83%) und Alkoholismus (73%) zu finden, doch liegt die durchschnittliche Anzahl der Diagnosen pro Diagnostiker nur in jeder Kategorie bei 4 und damit ebenso niedrig wie bei Schmid et al. (1974). Diese beiden diagnostischen Kategorien lassen sich offenbar leichter durch anamnestische sowie durch eindeutige Leitsymptome erfassen als z. B. die Kategorie "Sonstige Süchte" (57%; durchschnitt-

liche Diagnosenhäufigkeit pro Diagnostiker 8), bei der die Übereinstimmung um 30% gegenüber den oben genannten Kategorien geringer ist. Wegen der geringen Anzahl haben wir auch hier auf eine Berechnung der Signifikanz verzichtet.

Die Diagnostik der Schizophrenie (64% Übereinstimmung; durchschnittliche Diagnosenhäufigkeit 11) und Depressionen (Übereinstimmung 62%; Diagnosenhäufigkeit 30) und Neurosen (Übereinstimmung 41%; Diagnosenhäufigkeit 21) ist vergleichsweise nicht durch eine so eindeutige Vorgeschichte und Symptomatik gekennzeichnet. Die Ergebnisse liegen für die Schizophrenie um 13% und die Depression um 18% über den Werten von Schmid et al. (1974). Für die Neurosen (3% niedriger) liegen sie praktisch gleich. Auch hier ergab sich im Chi-Quadrat-Test kein signifikanter Unterschied (Chi²-Werte für Schizophrenen 0,27; für Depressionen 1,75; für Neurosen 0,18). Das Ausmaß der Übereinstimmung ist für die diagnostische Kategorie Psychopathie (Übereinstimmung 29%; Diagnosenhäufigkeit 8) niedriger als z. B. bei Beck et al. (1962; Übereinstimmung 38%). Der Prozentwert ist zwar deutlich besser als bei Schmid et al. (1974), wo für Psychopathie 9% (Diagnosenhäufigkeit 3) Übereinstimmung erzielt wurde; doch muß man sich fragen, ob die Charakterisierung der diagnostischen Kategorie Psychopathie (etwa durch das Glossar der ICD der psychiatrischen Krankheiten) zu wenig scharf ist oder ob die geringe Häufigkeit, mit der diese Diagnose gestellt wurde, die Ursache für die niedrige Übereinstimmung darstellt.

Die Ergebnisse der vorliegenden Untersuchung zeigen, daß sich durch die oben beschriebene Modifikation in der Art der Datenerhebung keine signifikante Verbesserung in der diagnostischen Übereinstimmung zwischen verschiedenen Beurteilern erreichen ließ; denn obwohl die Prozentwerte auf eine Tendenz zu höherer Übereinstimmung hinweisen, zeigt die statistische Prüfung keine über den Zufall hinausgehende Unterschiede.

Grundlage der Diagnosenstellung für die "Fragebogendiagnostiker" und den Computer waren auch bei dieser Untersuchung die CAPPS-Fragebögen, in die die "Ambulanz-Diagnostiker" ihre Befunde eintrugen, während sie selber noch die direkte Beobachtung des jeweiligen Patienten mit zur Grundlage ihrer Diagnosenstellung machen konnte. Es erhebt sich nach dieser Untersuchung natürlich die Frage, ob sich nicht durch Verwendung einer anderen Skala die diagnostische Übereinstimmung zwischen Ambulanz-Diagnostikern, Fragebogen-Diagnostikern und Computer erhöhen ließe. Bei der praktischen Arbeit mit den CAPPS-Fragebögen fielen uns zahlreiche Mängel auf, auf die z. B. bereits Schmid (1973) hingewiesen hat, die den Wert dieser Skala für differentialdiagnostische Zwecke etwas in Frage stellen. So werden z. B. formale Denkstörungen registriert, nicht aber ob diese in Zerfahrenheit, Denkhemmung, Ideenflucht oder anderen Störungen bestehen; erst dadurch würden diese Items jedoch differentialdiagnostisch bedeutsam. Als Alternative kämen eventuell die im oben erwähnten "US-UK diagnostic project" verwendeten Skalen in Frage oder eine der zahlreichen in einer Übersicht von Mombour (1972) zusammengestellten Skalen, die sich jedoch z. T. nur auf den psychopathologischen Befund ohne Anamnese beziehen. Auch auf ein Instrument, das sich z. Zt. in unserem Institut in Entwicklung befindet, sei hingewiesen, das System DiSiKa (Diagnostische Sichtloch-Kartei, v. Zerssen, i. Vorb.; Ellendorf, 1975). Dieses System bezieht sich auf Anamnese und Befund, es umfaßt etwa 140 Items; gegenwärtig können 45 ICD-Diagnosen mit ihm gestellt werden. Bei einer Zwi-

schenauswertung in Einteilung 3 mit 9 diagnostischen Kategorien ließ sich eine Übereinstimmung zwischen Diagnosen der Krankengeschichten und der DiSiKa von 63% für die Depressionen und 76% für die Neurosen erzielen.

LITERATUR

- Beck, A. T., Ward, C. H., Mendelson, M., Mork, J. E., Erbaugh, J. K.: Reliability of psychiatric diagnosis: A study of consistency of clinical judgements and ratings. Amer. J. Psychiat. 119, 351-357 (1962)
- Cooper, J. E., Kendell, R. E., Gurland, B. J., Sharpe, L., Copeland, J. R. M., Simon, R.: Psychiatric diagnosis in New York and London. London-New-York-Toronto: Oxford University Press 1972
- Degkwitz, R., Helmchen, H., Kockott, G., Mombour, W. (Hrsg.): Diagnosenschlüssel und Glossar psychiatrischer Krankheiten. Deutsche Übersetzung der internationalen Klassifikation der WHO (ICD), 8. Revision, und des internationalen Glossars. 4. Aufl. Berlin-Heidelberg-New York: Springer 1975
- Ellendorf, Ch.: Studie zur Validierung der diagnostischen Sichtlochkartei DiSiKa. Med. Diss., Univ. München (1975)
- Gerster, F., Dirlich, G.: DIAL - Ein Programm zur Entscheidungshilfe in der Psychiatrie für das System IBM 1130. IBM-Nachrichten, 25 (Heft 224) (1975)
- Krauß, B.: Empirische Untersuchung zur Verlässlichkeit psychiatrischer Diagnosen. Med. Diss., Univ. München (1974)
- Kreitman, N.: The Reliability of Psychiatric Diagnosis. J. of ment. Science 107, 876-886 (1961)
- Kreitman, N., Sainsbury, P., Morrissey, J., Towers, J., Scrivener, J.: The Reliability of Psychiatric Assessment: An Analysis. J. of ment. Science 107, 887-908 (1961)
- Lange, H. J.: Algorithmische Diagnostik. Münch. med. Wschr. 113, 577-580 (1971)
- Mombour, W.: Verfahren zur Standardisierung des psychopathologischen Befundes. Teil 1. Psychiat. clin. 5, 73-120 (1972); Teil 2. Psychiat. clin. 5, 137-157 (1972)
- Mombour, W.: Klassifikation, Patientenstatistik, Register. In: K. P. Kissner, J. E. Meyer, M. Müller, E. Strömgren (Hrsg.): Psychiatrie der Gegenwart, Bd. 3, 2. Aufl. Berlin-Heidelberg-New York: Springer 1975
- Schmid, W.: Untersuchung über die diagnostische Übereinstimmung zwischen Klinikern und dem System DIAL zur computerunterstützten psychiatrischen Diagnostik. Med. Diss., Univ. München (1973)
- Schmid, W., Castell, R., Mombour, W., Mittelsten Scheid, D., v. Zerssen, D.: Die diagnostische Übereinstimmung zwischen Klinikern und dem DIAL-Programm. Arch. Psychiat. Nervenkr. 218, 339-351 (1974)
- Shepherd, M., Brooke, E. M., Cooper, J. E., Lin, T.: An experimental approach to psychiatric diagnosis. Acta Psychiat. scand., suppl. 201 ad vol. 44 (1968)
- Spitzer, R. L., Endicott, J.: DIAGNO: A computer program for psychiatric diagnosis utilizing the differential diagnostic procedure. Arch. gen. Psychiat. 18, 746-756 (1968)

- Spitzer, R. L., Endicott, J.: DIAGNO II: Further developments in a computer program for psychiatric diagnosis. Amer. J. Psychiat. 125 Suppl., 12-21 (1969)
- Ward, C. H., Beck, A. T., Mendelson, M., Mock, J. E. Erbaugh, J. K.: The psychiatric nomenclature: Reasons for diagnostic disagreement. Arch. gen. Psychiat. 7, 198-205 (1962)
- WHO: The International Pilot Study of Schizophrenia. World Health Organization, Geneva 1973
- Wing, J. K., Cooper, J. E., Sartorius, N.: Measurement and classification of psychiatric symptoms. An instruction manual for the PSE and Catego program. Cambridge: University Press 1974
- Zerssen, D. v.: Diagnose. In: C. Müller (Hrsg.): Lexikon der Psychiatrie. Berlin-Heidelberg-New York: Springer 1973
- Zerssen, D. v.: Die Diagnostische Sichtloch-Kartei (DiSiKa). (In Vorber.)
- Zubin, J.: Classification of behavior disorders. Ann. Rev. Psychol. 18, 373-406 (1967)