

Konzepte psychiatrischer Diagnosen — eine Ergänzungsuntersuchung zur Reliabilität psychiatrischer Diagnosen*

J. Jakubaschk¹ und J. Werner^{2**}

¹Zentralinstitut für Seelische Gesundheit, I 5, D-6800 Mannheim, Bundesrepublik Deutschland

²Gesamthochschule Duisburg, Lotharstraße 65, D-4100 Duisburg, Bundesrepublik Deutschland

Concepts of Psychiatric Diagnoses: A Further Investigation of the Reliability of Psychiatric Diagnoses

Summary. A total of 128 physicians and psychologists, working in six psychiatric hospitals, were asked to allocate 10 psychiatric diagnoses on the basis of typical clinical criteria (symptoms, course of illness, etiological and anamnestic data), with which they were supplied. The correct diagnosis was made in 45.3%, while in 82% the broad diagnostic group was correctly ascribed.

The respondents were divided into two groups, according to the amount of diagnostic data they had been given (though in fact the information necessary for differential diagnosis was given equally to both groups). The group which made diagnosis on the grounds of more comprehensive data (ten characteristic diagnostic items) did significantly worse than the group which had few data (six characteristic items). The stereotypes for mania, endogenous depression, obsessional-compulsive neurosis, and antisocial personality disorder were correctly defined by 75% of more of the respondents. The profiles of the remaining six diagnostic groups were associated with the corresponding diagnoses by 50% or fewer of the respondents.

Key words: Psychiatric diagnoses – Concepts – Reliability.

Zusammenfassung. 128 Mediziner und Psychologen aus sechs Psychiatrischen Landeskrankenhäusern hatten anhand von Konzepten mit typischen Beurteilungskriterien (Symptome, Verlaufsangaben, ätiologische und anamnestische Daten) zehn psychiatrische Diagnosen zu stellen. In 45,3% der Fälle wurde die ursprünglich charakterisierte Diagnose aufgefunden, in 82% erfolgt zumindest eine treffende Zuordnung des Konzeptes zur entsprechenden Diagnosegruppe. — Die Beurteiler waren in zwei Gruppen aufgeteilt; die Konzepte beider

* Die Untersuchung wurde von der Stiftung Volkswagenwerk gefördert (Az: 11 2196)

** Herrn Prof. Häfner danken wir für die Mitwirkung bei der Planung und Anregungen bei der Durchführung und Darstellung des Forschungsprojektes.

Gruppen unterschieden sich zwar im Umfang — zehn gegenüber sechs charakteristischen Beurteilungskriterien — nicht jedoch im Informationsgehalt bezüglich der differentialdiagnostischen Abgrenzung. Die Gruppe, die anhand der ausführlicheren Konzepte diagnostizierte, erzielte signifikant schlechtere Ergebnisse. — Die Konzepte der Diagnose Manie, endogene Depression, Zwangsneurose und antisoziale Persönlichkeitsstörung wurden von 75% oder mehr der Beurteiler richtig zugeordnet. Die Charakteristika der übrigen sechs Krankheitsbilder wurden nur von 50% oder weniger der Beurteiler mit der entsprechenden Diagnose assoziiert.

Schlüsselwörter: Psychiatrische Diagnosen – Konzepte – Reliabilität.

Einleitung

Anamnese, Symptomatik und Verlauf kennzeichnen ein Krankheitsbild und sind damit die Grundpfeiler für jede Diagnose. Während des Studiums lernen wir, die charakteristischen Eigenheiten der verschiedenen Erkrankungen in Vorgeschichte, Befund und Krankheitsverlauf kennen und entwickeln eine Vorstellung vom „typischen Fall“; mit anderen Worten, wir prägen uns das Konzept einer Krankheit ein. Dadurch wird es möglich, auch dann ein Krankheitsbild zu diagnostizieren, wenn es in Einzelheiten vom „typischen Fall“ abweicht. Zwar wandeln sich meist mit zunehmender Erfahrung die Vorstellungen von den charakteristischen Eigenheiten einer Erkrankung, aber das Prinzip bleibt erhalten: Das Konzept erst ermöglicht, eine Diagnose zu stellen. Ausgehend von dieser Überlegung haben wir in den letzten Jahren die Konzepte von psychiatrischen Krankheitsbildern untersucht.

Fragestellung

Mediziner aller Fachbereiche, so auch Psychiater, stimmen in der diagnostischen Beurteilung von Patienten nicht voll überein. Unsere Forschungsarbeit sollte einerseits aufzeigen, welcher Anteil der Nicht-Übereinstimmung Unterschieden in den Konzepten anzulasten ist. Die entsprechenden Ergebnisse wurden bereits veröffentlicht.¹ Andererseits interessierte uns natürlich die Frage, ob sich überhaupt allgemein akzeptierte Konzepte für die untersuchten Krankheitsbilder ermitteln ließen. Das heißt, wir wollten nicht nur wissen, in welchem Prozentsatz Psychiater und Psychologen in den Konzepten übereinstimmen, sondern ob auch andere Psychiater wieder anhand dieser Konzepte die ursprünglich charakterisierten Diagnosen stellen können. Diese Fragestellung überprüfen wir in der vorliegenden Arbeit.

Es mag auf den ersten Blick scheinen, als handle es sich bei den beiden Fragestellungen im Kern um die Untersuchung des gleichen Problems. Aber der Weg von der Diagnose zu den charakteristischen Symptomen, zum Konzept, ist mit Sicherheit leichter zu finden als der

¹ „Die Abgrenzbarkeit psychiatrischer Diagnosen“, *Nervenarzt* 46 (1975) — „Übereinstimmung und Unterschiede in der Beurteilung psychiatrischer Diagnosen“, *Nervenarzt* 47 (1976).

„Rückweg“ vom Konzept zur Diagnose. Die Krankheitsbezeichnung gehört einem höheren Abstraktionsniveau an als Symptome, Verlaufskriterien und anamnestischen Daten, die das Konzept ausmachen. Wie wir aus psychologischen Untersuchungen wissen, ist es einfacher, von einem höheren Abstraktionsniveau auf ein niederes zu schließen als umgekehrt. Zum anderen bezeichnet die Diagnose ja nicht nur ein Krankheitsbild, das sich für uns mit charakteristischen Merkmalen verbindet, sondern beinhaltet gleichzeitig eine Aussage über die Stellung im hierarchisch gegliederten diagnostischen System. Damit ist gleichsam auch eine Vorauswahl in der Symptomatik verbunden. Wir finden es nicht verwunderlich, daß es bei organischen Psychosen typisch „endogene“ Zustandsbilder gibt; wir sind gewohnt, neurotisches Verhalten im Rahmen funktioneller Psychosen zu beobachten, wogegen psychotische Symptome bei einer Neurose nach landläufiger Auffassung nicht auftreten.

Weiterhin untersuchten wir, ob der Umfang der Konzepte, — die Anzahl der charakteristischen Merkmale, — Einfluß auf die Ergebnisse hat. Es bot sich an, auch dieser Frage nachzugehen, da wir aus den bisherigen Arbeiten für jede Diagnose zwei unterschiedlich ausführliche Konzepte zur Verfügung hatten, die sich jedoch nicht in ihrem Wert für die differentialdiagnostische Abgrenzung der Krankheitsbilder unterschieden.

Durchführung und Methodik der Untersuchung

Die Beantwortung der Fragestellung ist im Prinzip einfach. Zuerst gilt es, die Vorstellungen zu erfassen, die verschiedene Psychiater von den Charakteristika der Krankheitsbilder in Anamnese, Symptomatik und Verlauf haben. Aus diesen einzelnen Konzepten, die die ganze Variationsbreite der verschiedenen Vorstellungen repräsentieren, kann man ein Gesamtkonzept für jede Diagnose ableiten. Von diesen Konzepten ist dann im Idealfall zu erwarten, daß jeder Psychiater sie wieder der ursprünglich charakterisierten Diagnose zuzuordnen vermag.

In unseren Untersuchungen begrenzten wir die Zahl der Diagnosen auf zehn. Von diesen bezeichnen drei „Diagnosegruppen“ (Drei-Zahlen-Schlüssel der ICD-Klassifikation): 295 Schizophrenie, 300 Neurose und 301 Persönlichkeitsstörung (Psychopathie), die übrigen sieben benennen „diagnostische Untergruppen“ (Vier-Zahlen-Schlüssel der ICD-Klassifikation): 295.6 schizophrener Defekt, 296.1 Manie, 296.2 endogene Depression, 298.0 reaktive Depression, 300.3 Zwangneurose, 300.4 neurotische Depression und 301.7 antisoziale Persönlichkeitsstörung (antisoziale Psychopathie).

Als nächstes entwickelten wir eine sogenannte „Symptomliste“, die allerdings neben Symptomen auch Anamnesedaten, Hinweise zur Ätiologie und Verlaufskriterien enthält. Hier stellte sich ebenso wie bei den Diagnosen das Problem der Begrenzung. Es ist schwierig, eine Auswahl zu treffen, die einerseits noch eine differentialdiagnostische Abgrenzung ermöglicht, andererseits durch ein Überangebot die Beurteiler nicht überfordert. Wir entnahmen die Beurteilungskriterien verschiedenen Quellen und wählten 98 Begriffe aus, die unserer Meinung nach die wichtigsten Aspekte zur Charakterisierung von Diagnosen aus den Bereichen: endogene Psychosen, Neurosen und Persönlichkeitsstörungen beinhalten.

Auf der Symptomliste war hinter jedem Beurteilungskriterium eine Intervall-Skala vorgegeben, die mit „uncharakteristisch“, „mittelmäßig“ und „sehr charakteristisch“ bezeichnet war. Die Aufgabe der Beurteiler bestand darin, jede Einzelentscheidung, — für wie charakteristisch er das entsprechende Symptom bei der fragten Diagnose einschätzte, — auf der zugehörigen Skala zu markieren. Wenn wir einmal unterstellen, daß unsere Symptomliste alle wesentlichen Merkmale aus Vorgeschichte, Befund und Verlauf für die zehn untersuchten Krankheitsbilder enthält, haben wir damit die Möglichkeit, das Konzept des einzelnen Psychiaters von der vorliegenden Diagnose zu erfassen. Aus den Bewertungen aller Teilnehmer ist mühelos eine mittlere Beurteilung für alle 98 Begriffe bei der entsprechenden Diagnose zu berechnen. Die 98 Mittelwerte enthalten das „Gesamtkonzept“ der Diagnose, d. h. auch die acht oder 12 Symptome, anamnestischer Daten und Verlaufskriterien, die von den Beurteilern als sehr charakteristisch für die Diagnose angesehen wurden.

In einer Voruntersuchung wurde jede der zehn Diagnosen von über 30 psychiatrisch tätigen Medizinern und klinischen Psychologen anhand der 98 Begriffe unserer Liste charakterisiert. Damit konnten wir für jede Diagnose eine Gruppe von charakteristischen Merkmalen zusammenstellen, die zumindest das Konzept der Teilnehmer widerspiegeln. Ein anderes Ergebnis der Voruntersuchung war für uns erstaunlich: nur etwa die Hälfte der Symptome — 52 von 98 — trugen etwas zur differentialdiagnostischen Abgrenzung der zehn Krankheitsbilder bei.² Aus diesem Grunde verwendeten wir in unserer Hauptuntersuchung eine gekürzte Symptomliste mit nur mehr 52 Beurteilungskriterien. 181 Mediziner — die Hälfte Fachärzte — und zehn klinische Psychologen aus neun Psychiatrischen Landeskrankenhäusern charakterisierten die gleichen zehn Diagnosen wie in der Voruntersuchung anhand der gekürzten Liste. Da die Durchführung der Erhebungen und die Auswertung der Beurteilungen bei Vor- und Hauptuntersuchung identisch waren, hatten wir die Möglichkeit, für jede Diagnose eine zweite Gruppe von charakteristischen Merkmalen zusammenzustellen.

Wie nicht anders zu erwarten, waren die Symptomgruppen für die gleiche Diagnose in beiden Untersuchungen sehr ähnlich. Zum anderen war es nicht verwunderlich, daß aus den 98 Begriffen der ersten Liste mehr charakteristische Beurteilungskriterien pro Krankheitsbild benannt wurden, als aus der auf 52 gekürzten Liste. Nach den Ergebnissen unserer Voruntersuchung konnten wir davon ausgehen, daß trotz der drastischen Kürzung kein zur differentialdiagnostischen Abgrenzung geeignetes Merkmal unter den Tisch gefallen war.

Die Hierarchie der diagnostisch relevanten (charakteristischen) Merkmale garantierte, daß in beiden Fällen die wichtigsten Kriterien erfaßt wurden. Wie hielten es aber für möglich und wahrscheinlich, daß das umfangreichere Konzept den Beurteilern das Auffinden der Diagnose erleichterte.

Aus diesem Grund prüften wir in der vorliegenden Arbeit, ob Psychiater anhand der wenigen Charakteristika, die wir in der Hauptuntersuchung pro Diagnose ermittelten, genau so sicher auf das ursprünglich charakterisierte Krankheitsbild rückschließen konnten wie anhand der Konzepte der Voruntersuchung.

Wir wählten sechs Psychiatrische Landeskrankenhäuser aus, die an den bisherigen Untersuchungen nicht beteiligt waren, und teilten nach Zufall auf, in welchem Haus welches Konzept beurteilt werden sollte. Nach dem Einverständnis der Klinikleitung informierten wir alle in der Diagnostik tätigen Mediziner und Klinischen Psychologen über die geplante Untersuchung und baten um Mitarbeit an unserem Projekt. Die Teilnehmer wurden vor Beginn der Erhebung in einer stets gleichlautenden mündlichen Information ausführlich mit der Untersuchung und der von ihnen erbetenen Beurteilung vertraut gemacht. Außerdem erhielt jeder Beurteiler zusammen mit den Erhebungsbögen noch eine schriftliche Information, in der die wichtigsten Punkte der mündlichen Unter-

² Berechnung und Bedeutung dieser Ergebnisse wurde in „Reliabilität psychiatrischer Diagnosen — eine Voruntersuchung“, Sozialpsychiatrie 8 (1973), dargestellt.

richtung nochmals dargelegt wurden. Wir erläuterten, wie wir die Charakteristika aus den Beurteilungen von insgesamt etwa 250 anderen Psychiatern gewonnen hatten und forderten die Teilnehmer auf, in unserem Beisein die zehn Gruppen von Symptomen, Verlaufskriterien, anamnestischen und ätiologischen Daten wieder zehn psychiatrischen Diagnosen zuzuordnen. Die Listen waren für beide Teilnehmergruppen gleich aufgebaut. Sie enthielten einmal die charakteristischen Merkmale, jeweils zehn von den 98 aus der Symptomliste der Voruntersuchung oder je sechs aus den 52 der Hauptuntersuchung. Darüberhinaus hatten wir je fünf, bzw. drei Symptome gesondert aufgeführt, die als besonders untypisch für die jeweilige Diagnose bewertet worden waren. Damit hatten die Teilnehmer die Möglichkeit, anhand der Charakteristika eine Diagnose zu erschließen, für die dann die uncharakteristischen Symptome nicht zutreffen durften. Die Diagnose wurde einmal ausformuliert niedergeschrieben und zum anderen als Zahl nach dem ICD-Schlüssel angegeben³. Zum Schluß erbaten wir auf einer Prozentskala eine Selbsteinschätzung der Sicherheit, mit der das Urteil gefällt worden war. Die Abbildungen 1 und 2 zeigen am Beispiel der Konzepte der Zwangsneurose wie die Fragebögen aussahen. Außerdem füllten die Teilnehmer einen Personalbogen aus, der unter anderem Angaben über Alter, Dauer der psychiatrischen Tätigkeit, Dienststellung, Facharztanerkennung, Lehrfähigkeit etc. erfragte.

In drei Psychiatrischen Landeskrankenhäusern erhielten die insgesamt 71 Mediziner und klinischen Psychologen die Konzepte — zehn charakteristische und fünf uncharakteristische Merkmale pro Diagnose —, die wir aus der Voruntersuchung gewonnen hatten. Die 57 Teilnehmer aus den anderen drei Landeskrankenhäusern beurteilten die Konzepte aus der Hauptuntersuchung mit sechs charakteristischen und drei uncharakteristischen Symptomen pro Krankheitsbild. Die Erhebungen wurden in der Zeit von September bis November 1974 durchgeführt.

Erhebungspopulation

Da wir die Konzepte derselben zehn Diagnosen in Vor- und Hauptuntersuchung auf die gleiche Weise ermittelten und die Teilnehmer an den beiden Projekten in Zusammensetzung und Ausbildungsstand nahezu identisch waren, konnten wir die Differenzen zwischen den beiden Gruppen der vorliegenden Untersuchung als Resultat aus den Unterschieden zwischen den Konzepten ansehen. Dabei unterstellen wir allerdings, daß auch die Beurteilergruppen dieses Projektes in Zusammensetzung und Psychiatrieerfahrung ähnlich und mit den Teilnehmern von Vor- und Hauptuntersuchung vergleichbar sind.

Wie wir aus Tabelle 1 ersehen können, bestehen keine eklatanten Unterschiede in Ausbildungsstand (Prozentsatz der Fachärzte) und Berufserfahrung (Dauer der psychiatrischen Tätigkeit). Es ist jedoch der Anteil der Psychologen in Gruppe I mehr als doppelt so groß wie in den anderen Untersuchungen und in Gruppe II. Außerdem ist die Beteiligungsrate, d. h. der Teil von

³ Alle Teilnehmer erhielten einen Auszug aus der ICD-Klassifizierung für die endogenen Psychosen, Neurosen und Persönlichkeitsstörungen mit der Information, daß alle zehn zu erschließenden Diagnosen aus diesen drei Bereichen stammten.

Name.....

Schreiben Sie bitte die Diagnose auf, die Sie aufgrund der folgenden Symptome stellen.

Charakteristische Symptome:

Zwangserlebnisse und Zwangshandlungen
 Pedanterie
 anale Charakterzüge
 überstarkes Gewissen (Über-Ich)
 unbegründete Furcht vor bestimmten Situationen und Gegenständen
 Verdrängung
 Leidensdruck
 chronischer Verlauf mit und ohne Schwankungen
 äussere Schwierigkeiten infolge innerer Konflikte
 pathologisches Familienmilieu (Elternfamilie)

Uncharakteristische Symptome:

Orientierungsstörungen
 Halluzinationen
 Gedächtnisstörungen
 gehobene Stimmung
 formale Denkstörungen

Diagnose (frei formuliert).....

.....

.....

Diagnose Nr. (nach ICD-Klassifikation).....

Kreuzen Sie bitte auf der Skala an, mit welcher Sicherheit (0-100%) Sie die Diagnose gestellt haben.

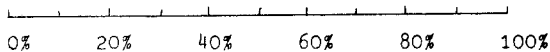


Abb. 1. Fragebogen mit dem Konzept der Zwangsneurose (ICD Nr.: 300.3) für die Gruppe I der Beurteiler

allen Medizinern und klinischen Psychologen an den ausgewählten Psychiatrischen Landeskrankenhäusern, der unserer Bitte um Mitarbeit bei den Erhebungen folgte, zwischen Gruppe I und II deutlich verschieden. Von den jeweils 99 Mitarbeitern nahmen in Gruppe I 71, in Gruppe II nur 57 an unserer Untersuchung teil. Wir wissen nicht, ob unsere Resultate durch die genannten Unterschiede beeinflusst sind, und wir können nur vermuten in welcher Hinsicht. Wenn auch die wünschenswerte Homogenität der Teilnehmergruppen nicht ganz erreicht ist, so

Name.....

Schreiben Sie bitte die Diagnose auf, die Sie aufgrund der folgenden Symptome stellen.

Charakteristische Symptome:

Zwangserlebnisse und Zwangshandlungen

Pedanterie

unbegründete Furcht vor bestimmten Situationen und Gegenständen

Verdrängung

Selbstunsicherheit

annale Charakterzüge

Uncharakteristische Symptome:

Halluzinationen

verminderte Intelligenz

gehobene Stimmung

Diagnose (frei formuliert)

.....

.....

Diagnose Nr. (nach ICD-Klassifikation).....

Kreuzen Sie bitte auf der Skala an, mit welcher Sicherheit (0-100%) Sie die Diagnose gestellt haben.

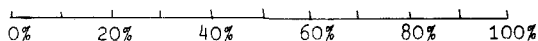


Abb. 2. Fragebogen mit dem Konzept der Zwangsneurose (ICD Nr.: 300.3) für die Gruppe II der Beurteiler

glauben wir doch die Ergebnisse nicht wegen dieser möglichen Selektionsfaktoren insgesamt verwerfen zu müssen.

Wir forderten alle Mediziner und klinischen Psychologen ohne Unterschied zur Teilnahme an unserer Untersuchung auf und vermieden damit weitgehend eine Auswahl, wie die Verteilung der Dienststellungen der 128 Beurteiler zeigt: 27 (= 21,1%) waren Direktoren, stellvertretende Direktoren oder Oberärzte, 29 (= 22,7%) fungierten als Abteilungs- oder Stationsärzte, 56 (= 43,8%) waren als Assistenzärzte, 2 (= 1,6%) als Medizinalassistenten und

Tabelle 1. Zusammensetzung und Ausbildungsstand der Teilnehmer in den verschiedenen Untersuchungen

	Vor- untersuchung	Haupt- untersuchung	Vorliegende Untersuchung	
			Gruppe I ^a	Gruppe II ^b
Zahl der Teilnehmer	52	191	71	57
Beteiligungsrate ^c	62,0%	66,4%	71,7%	57,8%
Mediziner	94,2%	94,8%	85,9%	93,0%
davon Fachärzte	53,0%	50,0%	57,4%	50,9%
Psychologen	5,8%	5,2%	14,1%	7,0%
Dauer der psychia- trischen Tätigkeit (Mittelwerte)	7,5 Jahre	8,4 Jahre	8,2 Jahre	7,1 Jahre
(Quartile) ^d	12 Monate— 12 Jahre	16 Monate— 15 Jahre	23 Monate— 11 Jahre	12 Monate— 10 Jahre

^a Gruppe 1: 71 Mediziner und klinische Psychologen, die die Konzepte aus unserer Voruntersuchung mit zehn charakteristischen und fünf uncharakteristischen Merkmalen beurteilten.

^b Gruppe II: 57 Teilnehmer, die die Konzepte aus unserer Hauptuntersuchung mit sechs charakteristischen und drei uncharakteristischen Begriffen beurteilten.

^c Die *Beteiligungsrate* gibt an, welchen Prozentsatz die Teilnehmer an unseren Untersuchungen von der Gesamtheit der Mediziner und klinischen Psychologen der beteiligten Institutionen ausmachen. (Z. B. die 52 Beurteiler der Voruntersuchung sind 62% der insgesamt 84 Psychiater und Psychologen, die in den beteiligten psychiatrischen Landeskrankenhäusern und psychiatrischen Universitätskliniken arbeiteten.)

^d Die *Quartile* grenzen das obere und untere Viertel von der untersuchten Gesamtheit ab. (Z. B. in der Voruntersuchung haben 25% der Teilnehmer weniger als 12 Monate und 25% mehr als 12 Jahre Psychiatrie-Erfahrung.)

14 (= 10,9%) als klinische Psychologen tätig. 54,4% der Mediziner, d. h. 62 von 114 hatten ihre Facharztanerkennung erhalten. 41,4% beteiligten sich an einer Psychotherapieausbildung oder hatten eine solche abgeschlossen. Lehrverpflichtungen an Universitäten wurden von 11 (= 8,6%) der Teilnehmer wahrgenommen; 45 (= 35,2%) unterrichteten an Schwesternschulen oder ähnlichen Institutionen. Das Durchschnittsalter betrug 40,8 Jahre; die jüngsten Teilnehmer waren 25 Jahre, die ältesten 64 Jahre alt. An unseren Erhebungen beteiligten sich 64,6% (128 von 198) der Mediziner und Psychologen, die in den sechs Psychiatrischen Landeskrankenhäusern zum Zeitpunkt der Untersuchung arbeiteten.

Ergebnisse

Um zu verdeutlichen, was die Resultate dieser Arbeit aussagen, müssen wir noch einmal auf den bereits zitierten Idealfall zurückkommen.

Das Optimum, — zwar vorstellbar, aber wohl kaum zu erreichen, — sähe im Fall unserer Untersuchung folgendermaßen aus: Alle Psychiater verbinden mit der gleichen Diagnose identische Vorstellungen, d. h. das Konzept des einzelnen Teilnehmers für ein bestimmtes Krankheitsbild entspricht exakt dem jedes anderen Beurteilers und ist damit dem Gesamtkonzept gleichzusetzen. Träfe dies zu, wäre zu erwarten, daß alle Psychiater jedes Konzept nur einer einzigen Diagnose zuordnen. Wir wissen aber aus Ergebnissen einer unserer anderen

Untersuchungen, daß im Mittel nur 85% der Einzelkonzepte dem Gesamtkonzept einer Diagnose gleich oder ähnlich sind. Mit anderen Worten, wenn aus den unterschiedlichen Ansichten von hundert Psychiatern das Gesamtkonzept einer Diagnose extrahiert wird, weichen die Einzelkonzepte von durchschnittlich 15 Psychiatern soweit vom Gesamtkonzept der beurteilten Diagnose ab, daß sie mehr Ähnlichkeit mit anderen Krankheitsbildern haben.

Dies trifft zu, obgleich die Bewertungen dieser 15 Beurteiler in das Gesamtkonzept mit eingegangen sind⁴. Es ist notwendig, dieses Ergebnis bei der Betrachtung der vorliegenden Resultate in Erinnerung zu behalten. Ob die Konzepte der untersuchten psychiatrischen Krankheiten von verschiedenen Psychiatern mit der entsprechenden Diagnose assoziiert werden, ist im Überblick an den Tabellen 2 und 3 abzulesen.

Zwar ist kaum zu erwarten gewesen, daß die Konzepte so eindeutig und die Meinungen der Beurteiler so einhellig sind, daß alle Teilnehmer für jedes Konzept nur eine einzige Diagnose benennen würden. Andererseits ist erstaunlich, wieviele verschiedene Diagnosen einzelnen Konzepten zugeordnet wurden. Die charakteristischen Merkmale der „reaktiven Depression“ können — auch wenn sie Wissen und Erfahrung von ca. 250 Psychiatern repräsentieren — nicht allzu charakteristisch gewesen sein, da neben der richtigen Zuordnung einmal 15 und einmal zehn weitere Diagnosen gestellt wurden. — Im Mittel verteilen sich die Beurteilungen der Konzepte in Gruppe I auf mehr als 11 verschiedene Diagnosen, in Gruppe II wurden durchschnittlich neun unterschiedliche Diagnosen gestellt. Die Differenz der Mittelwerte zwischen den Gruppen I und II ist zwar nicht statistisch signifikant, aber es hat den Anschein, daß die ausführlicheren Konzepte der Gruppe I eher zur Verunsicherung der Beurteiler beitrugen als die Diagnosestellung erleichterten.

Dieser Trend bestätigt sich in Tabelle 3. Die 57 Mediziner und klinischen Psychologen der Gruppe II erreichten im Mittel höhere Prozentsätze im Auffinden der dem jeweiligen Konzept zugrundeliegenden Diagnose wie auch in der Zuordnung zur Diagnosegruppe. Die Differenz der Mittelwerte — 42,5% und 48,8% — ist statistisch auf dem 5%-Niveau signifikant (t-Test).

Wenn wir die Ergebnisse im einzelnen betrachten, fällt auf, daß die umfassenden Diagnosen 295 Schizophrenie, 300 Neurose und 301 Persönlichkeitsstörung offensichtlich schwerer aus den Konzepten abzuleiten waren als die übrigen Krankheitsbilder. Getrennte Mittelwerte für die spezifischen Diagnosen (Vier-Zahlen-Schlüssel) und die umfassenden Diagnosen (Drei-Zahlen-Schlüssel) zeigen diese Diskrepanz besonders augenfällig: 60% in Gruppe I und 66% in Gruppe II stehen nur 2% bzw. 9% richtiger diagnostischer Zuordnung gegenüber.

Zur Erklärung dieses Faktums lassen sich mehrere Gründe anführen; uns erscheinen die folgenden am naheliegendsten. Vermutlich sind die meisten Psychiater gewohnt, eine spezifische Diagnose (Vier-Zahlen-Schlüssel) zu stellen und nicht nur die Diagnosegruppe (Drei-Zahlen-Schlüssel) zu benennen. Obgleich wir die Teilnehmer in unseren Informationen vor der Erhebung ausdrücklich darauf hingewiesen haben, daß es sowohl spezifische als auch globale Diagnosen aufzufinden gelte, — wir erläuterten das an Beispielen, — mag die

⁴ Die hier erwähnten Ergebnisse sind ausführlich in „Die Abgrenzbarkeit psychiatrischer Diagnosen“, Nervenarzt 46 (1975) dargestellt.

Tabelle 2. Tabelle 2 zeigt, wieviele verschiedene Diagnosen — einschließlic der dem jeweiligen Konzept zugrundeliegenden Diagnose — anhand der zehn Konzepte in den beiden Teilnehmergruppen der vorliegenden Untersuchung gestellt wurden. Die Differenz der Mittelwerte ist nicht statistisch signifikant.

ICD Nr.	Konzept der Diagnose	Zahl der Diagnosen	
		Gruppe I ^a	Gruppe b ^b
296.1	Manie	3	3
296.2	endogene Depression	4	6
300.3	Zwangsneurose	7	5
295.6	schizophrener Defekt	11	11
301.7	antisoziale Persönlichkeitsstörung	14	10
295	Schizophrenie	15	9
300.4	neurotische Depression	14	11
298.0	reaktive Depression	16	11
301	Persönlichkeitsstörung	15	12
300	Neurose	15	13
Mittelwert		11,4	9,1

^a und ^b siehe Legende zu Tabelle 1.

diagnostische Gewohnheit dennoch häufig gesiegt haben. - Eine andere Erklärungsmöglichkeit kann in den Konzepten liegen. Sicherlich ist eine Zwangsneurose mit charakteristischen Merkmalen treffender zu umschreiben als die Neurose, die so unterschiedliche Krankheitsbilder wie die Hysterie und die neurotische Depression umfaßt.

Der Vollständigkeit halber ist zu erwähnen, daß sich selbstverständlich dieser Unterschied nicht mehr nachweisen läßt, wenn man die Zuordnung der Konzepte zu einer Diagnosegruppe betrachtet. Denn in diesem Fall werden ja zum Prozentsatz der umfassenden Diagnose, (z. B. 295 Schizophrenie), alle Fehlklassifizierungen sofern sie zur gleichen Diagnosegruppe gehören (z. B. 295.0, 295,1 bis 295,9) addiert.

Die unter der Überschrift „Diagnosegruppe“ in der zweiten Hälfte der Tabelle 3 aufgeführten Prozentsätze zeigen, daß im Durchschnitt vier von fünf Psychiatern das Konzept einer Diagnose, wenn nicht dem eigentlich charakterisierten Krankheitsbild, so doch zumindest der richtigen diagnostischen Gruppe zuordnen. — Die Prozentsätze für die reaktive Depression (298,0) fallen in diesem Teil der Tabelle aus dem Rahmen. Die einfachste und vermutlich zutreffende Erklärung dafür ist wohl in der Diagnosegruppe 298 zu suchen. Da diese Kategorie kaum Krankheitsbilder umfaßt, die mit der reaktiven Depression zu verwechseln wären, war der „Zugewinn“ für diese Diagnose minimal. Eine Fehlklassifizierung des Konzeptes der reaktiven Depression psychotischen Ausmaßes (298,0) kommt vermutlich am ehesten der neurotischen Depression (300.4) oder der endogenen Depression (296.2) zugute.

Ein Ergebnis, das wir eigentlich erwartet hatten, ließ sich nicht nachweisen. Wir vermuteten, daß die Konzepte der Diagnosen, die zur Gruppe der endogenen Psychosen gehörten, von den Psychiatern sicherer erkannt und damit treffender

Tabelle 3. Prozentsätze der Teilnehmer, denen es gelang, anhand des jeweiligen Konzeptes die zugrundeliegende Diagnose aufzufinden oder der entsprechenden diagnostischen Gruppe zuzuordnen.

ICD Nr.	Diagnose	Diagnose ^c				Diagnosegruppe ^d			
		Gruppe I ^a	Gruppe II ^b	Mittelwert (gewichtet)		Gruppe I ^a	Gruppe II ^b	Mittelwert (gewichtet)	
296.1	Manie	97,2%	96,5%	96,9%		98,6%	98,3%	98,4%	
296.2	endogene Depression	91,5%	86,0%	89,1%		98,5%	94,8%	96,9%	
300.3	Zwangsneurose	80,3%	73,7%	77,3%		87,3%	79,0%	83,6%	
301.7	antisoziale Persönlichkeitsstörung	69,0%	82,5%	75,0%		95,6%	91,5%	93,8%	
300.4	neurotische Depression	36,6%	68,4%	50,8%		69,0%	72,0%	70,3%	
295.6	schizophrener Defekt	33,8%	17,5%	26,6%		87,3%	84,2%	85,9%	
298.0	reaktive Depression	11,3%	35,1%	21,9%		11,3%	38,7%	23,4%	
301	Persönlichkeitsstörung	0,0%	14,0%	6,3%		88,7%	94,7%	91,4%	
300	Neurose	4,2%	8,8%	6,3%		78,8%	91,4%	84,4%	
295	Schizophrenie	1,4%	5,3%	3,1%		90,2%	94,8%	92,2%	
Mittelwert:		42,5%	48,8%	45,3%		80,5%	83,9%	82,0%	

a und b siehe Legende zu Tabelle 1.

^c *Diagnose:* Prozentsätze, in denen die Beurteiler auf die den Konzepten zugrundeliegenden Diagnosen schlossen. (Z. B. 17,5% der Gruppe II erkannten aus dem Konzept die ursprünglich charakterisierte Diagnose: 295.6, schizophrener Defekt.) — Die Unterschiede zwischen Gruppe I und II sind auf dem 5%-Niveau statistisch signifikant (t-Test).

^d *Diagnosegruppe:* Prozentsätze für die Zuordnung der Konzepte zur zugrundeliegenden Diagnose *und* zu anderen Diagnosen aus der gleichen diagnostischen Gruppe (z. B. 84,2% der Gruppe II ordneten das Konzept des schizophrenen Defekts einer der folgenden Diagnosen zu: 295.6, 295, 295.0, 295.1, 295.5, 295.7 und 295.9).

zugeordnet würden als die Konzepte aus den diagnostischen Bereichen Neurose und Persönlichkeitsstörung. Es zeigt sich jedoch, daß sich keine Bevorzugung einer Diagnosekategorie finden läßt. — Wie oben bereits dargestellt, ist lediglich ein deutlicher Unterschied zwischen speziellen Krankheitsbildern (Vier-Zahlen-Schlüssel) und umfassenden Diagnosen (Drei-Zahlen-Schlüssel) nachweisbar.

Bleibt die Frage zu beantworten, wie weit Konzepte psychiatrischer Diagnosen übereinstimmend verwendet werden. Bei vier der zehn von uns untersuchten Krankheitsbilder werden die charakteristischen Merkmale von 75% oder mehr der Beurteiler mit der entsprechenden Diagnose (296.1 Manie, 296.2 endogene Depression, 300.3 Zwangsneurose und 301.7 antisoziale Persönlichkeitsstörung) assoziiert. Wenn wir unterstellen, daß die Teilnehmer an unseren Erhebungen in etwa die in der psychiatrischen Diagnostik tätigen Mediziner und Psychologen repräsentieren, können wir sicher von diesen Krankheitsbildern sagen, daß ihre Konzepte allgemein anerkannt sind. Wie sieht es aber mit den anderen sechs Diagnosen aus? Die Charakteristika der neurotischen Depression (300.4) werden noch von jedem zweiten Beurteiler als zu dieser Diagnose gehörig erkannt, beim schizophrenen Defekt (295.6) ist es nur noch jeder vierte Beurteiler, bei der reaktiven Depression (298.0) jeder fünfte, die Konzepte von Persönlichkeitsstörung (301) und Neurose (300) ordnet nur noch einer von 20 Psychiatern diesen Diagnosen zu und die Schizophrenie (295) wird schließlich anhand — vorgeblich — charakteristischer Merkmale nur mehr von 3%, d. h. nicht einmal von jedem dreißigsten Beurteiler erkannt. Für die Konzepte der zuletzt genannten sechs Krankheitsbilder kann man wohl kaum Allgemeingültigkeit postulieren. Trotz dieser ernüchternden Erkenntnis sind die Ergebnisse nicht allzu deprimierend. Immerhin stimmen die Teilnehmer in mehr als 80% überein, wenn wir die richtige Zuordnung der Konzepte zur Diagnosegruppe betrachten. Das bedeutet schließlich, daß zumindest darüber weitgehend Einigkeit besteht, welche charakteristische Eigenheiten in Symptomatik, Verlauf und Anamnese eine grobe — in der Praxis häufig ausreichende — diagnostische Klassifizierung erlauben.

Diskussion

Als wir diese Untersuchung planten und vorbereiteten, diskutierten wir mit mehreren erfahrenen Psychiatern Sinn und Zweck dieses Forschungsvorhabens. Dabei wurden wir häufig mit der Auffassung konfrontiert, daß die Durchführung dieses Projektes eigentlich überflüssig sei. Man hielt uns vor, wir wüßten ja aus unseren bisherigen Untersuchungen, daß im Durchschnitt etwa 85% der Teilnehmer die zehn Krankheitsbilder gleich oder ähnlich charakterisierten. Daher sei mit Sicherheit zu erwarten, daß annähernd der gleiche Prozentsatz den Charakteristika die richtigen Diagnosen zuordnen würde, wenn wir das Verfahren lediglich umkehrten. Diese Argumentation klingt logisch, doch unsere Ergebnisse widerlegen sie. Es bleibt zu erörtern, welche Gründe dafür anzuführen sind. Das bedeutsamste Gegenargument wurde in der Fragestellung bereits diskutiert; es handelt sich um zwei qualitativ unterschiedliche Entscheidungsprozesse. Darüberhinaus kommen prinzipiell drei Möglichkeiten in Frage: Die Ursache kann bei den Beurteilern, den Diagnosen oder den Konzepten (und den diesen

zugrundeliegenden Symptomlisten) zu suchen sein. Wir konnten zeigen, — vergleiche Tabelle 1 —, daß die Beurteiler aus Vor- und Hauptuntersuchung wie auch dieses Projektes in Zusammensetzung und Ausbildungsstand nicht eklatant verschieden waren. Mit Sicherheit erklären die relativ geringfügigen Unterschiede nicht die Differenz zwischen durchschnittlich 85% übereinstimmender Charakterisierung in unserer Hauptuntersuchung und im Mittel nur 45% treffender Zuordnung der Konzepte zu den ursprünglich charakterisierten Krankheitsbildern in der vorliegenden Arbeit.

Da es in unseren Untersuchungen stets um die gleichen zehn Diagnosen ging, ist auszuschließen, daß hier ein Grund für die Differenz zwischen der hohen Übereinstimmung bei der Charakterisierung und der geringen Sicherheit in der Zuordnung der Konzepte zu suchen ist. Wir können also weder mit der Zusammensetzung der Beurteiler noch mit Unterschieden in den Diagnosen die Diskrepanz hinreichend erklären; bleibt zu prüfen, ob in den Konzepten, bzw. in den Symptomlisten die Ursache zu finden ist.

Zwar haben wir einerseits die Beurteilungskriterien unserer Liste sehr sorgfältig ausgesucht, und zum anderen hat keiner der Erhebungsteilnehmer von der Möglichkeit Gebrauch gemacht, fehlende Merkmale am Ende der Liste einzufügen, und dennoch kann hier beim Entstehen der Konzepte aus der Symptomliste eine Erklärung für die aufgezeigte Diskrepanz liegen. Es ist denkbar, daß unser Merkmalskatalog nicht umfassend genug war oder die Beurteilungskriterien nicht ausreichend spezifisch. Wenn wir dies als richtig unterstellen, ist der Effekt auf die Übereinstimmung in der Charakterisierung der Krankheitsbilder sicherlich deutlich geringer als die Auswirkung auf die Zuordnung der Konzepte zu den ursprünglich charakterisierten Diagnosen. Aber es gibt zumindestens noch zwei weitere Faktoren, die in diesem Zusammenhang eine Rolle spielen. Zum einen wissen wir, daß die durchschnittliche Übereinstimmung von 85% eine Maximalschätzung darstellt⁵. Andererseits ist es nach aller Erfahrung leichter, ein Krankheitsbild charakterisierend zu beschreiben als aus wenigen typischen Merkmalen eine treffende Diagnose zu stellen. Da wir den jeweiligen Einfluß der genannten Erklärungsmöglichkeiten nicht quantifizieren können, müssen wir auf eine präzise Antwort verzichten. Wir können lediglich festhalten, daß etwa 85% der Psychiater und klinischen Psychologen in der charakterisierenden Beschreibung von zehn psychiatrischen Diagnosen, d. h. in der Bewertung der typischen Merkmale dieser Krankheitsbilder, weitgehend übereinstimmen. Wenn allerdings eine vergleichbare Gruppe von Beurteilern aufgefordert wird anhand eben dieser charakteristischen Symptome, Verlaufsangaben etc. wieder die ursprünglich beschriebenen Diagnosen aufzufinden, ist nur eine Erfolgsrate von 45% zu ermitteln. Läßt man Fehlklassifizierungen innerhalb einer diagnostischen Gruppe außer acht, erhöht sich der Prozentsatz treffender Zuordnungen auf 82.

Wir haben sicher zu Recht darauf hingewiesen, daß die relativ geringfügigen Unterschiede zwischen den Teilnehmern der verschiedenen Untersuchungen nicht die erhebliche, oben besprochene Diskrepanz der Ergebnisse erklären. Aber es ist zu diskutieren, wieweit sich unterschiedliche Beteiligungsrate und Zusammensetzung auf die beobachteten Gruppendifferenzen ausgewirkt haben.

⁵ „Die Abgrenzbarkeit psychiatrischer Diagnosen“, Nervenarzt 46 (1975).

Wenn wir einmal annehmen, daß diejenigen am bereitwilligsten unserer Aufforderung Folge leisteten, die sich in der psychiatrischen Diagnostik sicher fühlten und mit den Charakteristika der Krankheitsbilder sehr vertraut waren, dann wird eine hohe Beteiligungsrate zu schlechteren Ergebnissen führen. Allerdings ist auch das Gegenteil denkbar: Die erfahrenen Diagnostiker beteiligen sich nicht, da sie derartige Untersuchungen für überflüssig erachten, während diejenigen, die die Probleme psychiatrischer Diagnostik aus eigenem Erleben kennen, eher zur Mitarbeit bereit sind. In diesem Fall wird eine hohe Beteiligung die Resultate verbessern.

Der unterschiedliche Prozentanteil der Psychologen in den beiden Gruppen dieser Untersuchung kann die Resultate ebenfalls beeinflußt haben. Es ist möglich, daß in der Regel Mediziner mit den Charakteristika von Krankheitsbildern besser vertraut sind als Psychologen, auch wenn diese in der klinischen Diagnostik arbeiten. Eine nach Medizinern und Psychologen getrennte Auswertung unserer Untersuchungsergebnisse steht noch aus. Die Gruppe I hat im Vergleich mit der zweiten Beurteilergruppe einerseits eine hohe Beteiligungsrate, zum anderen relativ viele Psychologen unter den Teilnehmern. Da nicht auszuschließen ist, daß beide Faktoren in der gleichen Richtung wirken, — mit anderen Worten die Ergebnisse verschlechtern, — bleibt zu fragen, ob die beobachtete Differenz in den Resultaten zwischen den Gruppen I und II den verschiedenen Konzepten anzulasten ist oder lediglich die unterschiedliche Zusammensetzung der Teilnehmergruppen widerspiegelt. Wenn wir diese Frage auch nicht zweifelsfrei beantworten können, so halten wir doch die genannten Faktoren nicht für derart bedeutsam, daß sie alleine zu den signifikant unterschiedlichen Ergebnissen führen.

Literatur

Den Autoren sind keine wissenschaftlichen Veröffentlichungen bekannt, die sich mit dem speziellen Thema der vorliegenden Arbeit befaßt haben. Ein Literaturverzeichnis zum diagnostischen Entscheidungsprozeß und zur Reliabilität psychiatrischer Diagnosen kann bei den Verfassern angefordert werden.

Eingegangen am 2. Juli 1976