

Dimensionale Diagnostik mit dem Fragebogen zur Erfassung von DSM-IV-Persönlichkeitsstörungen (ADP-IV)^{1,2}

Daniela Renn^{3,4}, Stefan Höfer⁴, Gerhard Schüßler⁴, Gerhard Rumpold⁴, Ulrike Smrekar⁴, Nicola Janecke⁴, Stephan Doering^{4,5,6}

Summary

A dimensional diagnostic approach using the „Assessment of DSM-IV Personality Disorders (ADP-IV)“ Questionnaire

Objectives: The ADP-IV is a self-assessment instrument of personality disorders providing categorical as well as dimensional diagnoses. The German translation of the Dutch original has already shown sufficient reliability as well as a rather weak convergent validity of categorical diagnoses. This study investigates the validity of the dimensional assessment of personality disorders.

Methods: Four hundred patients and a community sample of 385 subjects completed the ADP-IV. A SCID-II interview and an expert consensus rating were also performed in 210 patients. Norms for patients and the community sample were calculated by using z-transformation. Receiver operator curves (ROC) were employed for the determination of cut-offs for categorical diagnoses on the basis of the dimensional scores.

Results: The ROC analyses resulted in negative predictive values of 99 % for the five most frequent personality disorders in the sample; in contrast, positive predictive values were low.

Conclusions: The instrument is well suited for screening purposes, allowing for a reliable exclusion of personality disorders. For diagnosing personality disorders, however, clinical interviewing remains necessary. Norms and a dimensional personality disorder profile provide hints concerning the presence of pathological personality traits with respect to the reference group. The ADP-IV represents the only German-language questionnaire for the assessment of DSM-IV personality disorders which has been sufficiently validated. It is freely available on the internet at <http://zmkweb.uni-muenster.de/einrichtungen/proth/dienstleistungen/psycho/diag/index.html>.

Z Psychosom Med Psychother 54/2008, 214–226

¹ Gefördert vom Jubiläumsfonds der Österreichischen Nationalbank, Projekt Nr. 9141.

² Forschungsstipendium der Leopold-Franzens-Universität Nr. 18031/32–03.

³ www.psypraxis.org.

⁴ Klinik für Medizinische Psychologie und Psychotherapie, Medizinische Universität Innsbruck.

⁵ Poliklinik für Zahnärztliche Prothetik und Werkstoffkunde, Westfälische-Wilhelms-Universität Münster.

⁶ Klinik und Poliklinik für Psychosomatik und Psychotherapie, Westfälische-Wilhelms-Universität Münster.

Keywords

Personality Disorders – Dimensional Diagnoses – Questionnaire – Receiver Operator Curves – Norms

Zusammenfassung

Fragestellung: Der ADP-IV stellt ein Selbstbeurteilungsinstrument zur Erfassung von Persönlichkeitsstörungen dar, das sowohl eine kategoriale als auch eine dimensionale Diagnostik ermöglicht. Die deutsche Übersetzung des holländischsprachigen Originals zeigte bereits zuvor eine befriedigende Reliabilität und eine eher unzureichende Übereinstimmungsvalidität der kategorialen Diagnostik. In dieser Studie wird die Validität der dimensional Diagnostik untersucht.

Methode: 400 Patienten und eine Bevölkerungsstichprobe von 385 Personen wurden mit dem ADP-IV untersucht. Bei 210 Patienten wurden zusätzlich ein SKID-II-Interview und ein diagnostisches Expertenrating durchgeführt. Mit Hilfe der z-Transformation wurden Normwerte für die Patienten- und die Bevölkerungsstichprobe berechnet. ROC-Analysen dienten der Bestimmung von Cut-off-Werten für eine kategoriale Diagnostik auf der Basis der dimensional Summenscores.

Ergebnisse: Die ROC-Analysen lieferten negativ prädiktive Werte von 99 % für die fünf häufigsten Persönlichkeitsstörungen in der Stichprobe, die positiv prädiktiven Werte waren dagegen gering.

Diskussion: Das Instrument eignet sich in der Screening-Diagnostik zum relativ sicheren Ausschluss von Persönlichkeitsstörungen. Für die Diagnosestellung ist weiterhin eine klinische Interviewdiagnostik unumgänglich. Normwerte und Persönlichkeitsstörungsprofil geben Hinweise auf die Ausprägung pathologischer Persönlichkeitszüge in Relation zur jeweiligen Bezugsgruppe. Der ADP-IV stellt den einzigen umfassend validierten deutschsprachigen Fragebogen zur Erfassung von Persönlichkeitsstörungen nach DSM-IV dar; er ist im Internet frei verfügbar (download unter <http://zmkweb.uni-muenster.de/einrichtungen/proth/dienstleistungen/psycho/diag/index.html>).

1. Einleitung

Die kategoriale Diagnostik von Persönlichkeitsstörungen, wie sie derzeit in den Klassifikationssystemen von ICD-10 (Dilling et al. 2004) und DSM-IV (Saß et al. 1996) enthalten ist, steht stark in der Kritik. Diese gründet sich auf folgende Schwächen:

1. Die kategoriale Klassifikation konnte nicht empirisch (z. B. faktorenanalytisch) abgesichert werden (Westen u. Shedler 1999; Widiger 1993, Wöller u. Tress 2005).
2. Die Grenzziehung zwischen krank und gesund mithilfe eines Cut-offs – zum Beispiel fünf von neun Items erfüllt – ist willkürlich und entspricht nicht der klinischen Realität (Barnow et al. 2006; Leibing u. Doering 2006; Schüßler et al. 2006; Westen u. Shedler 1999; Widiger 1993).
3. Die einzelnen Kategorien sind sehr heterogen, so sind beispielsweise von der Borderline-Persönlichkeitsstörung 151 verschiedene Kombinationen der neun diagnostischen Merkmale des DSM-IV denkbar, was gegen die Existenz von „Krank-

heitsentitäten“ spricht (Barnow et al. 2006; Leibing u. Doering 2006; Trull u. Durrett 2005).

4. Es besteht eine hohe Komorbidität unter den verschiedenen Kategorien der Persönlichkeitsstörungen, was eine geringe Trennschärfe der Diagnosen nahe legt (Barnow et al. 2006; Leibing et al. 2008; Trull u. Durrett 2005; Westen u. Shedler 1999).

Daher werden bereits seit den neunziger Jahren des vergangenen Jahrhunderts dimensionale Modelle für die Diagnostik von Persönlichkeitsstörungen als Alternative zu ICD-10 und DSM-IV diskutiert. Diese basieren meist auf empirisch gewonnenen Dimensionen der Persönlichkeit, so zum Beispiel das Modell der „big five“ (Costa u. McCrea 1992; Costa u. Widiger 1993) oder Cloningers Sieben-Faktoren-Modell (Cloninger et al. 1993). Auch wenn weitgehend Einigkeit darüber besteht, dass eine dimensionale Diagnostik der Persönlichkeitsstörungen wünschenswert und der kategorialen in vielerlei Hinsicht überlegen ist, so konnte sich bislang jedoch keiner der dimensional Ansätze gegen die anderen Modelle durchsetzen. Eine Arbeitsgruppe der American Psychiatric Association (APA) und des National Institute of Mental Health (NIMH) formulierte daher im Hinblick auf die Entwicklung des DSM-V: „There is a clear need for dimensional models to be developed and for their utility to be compared with that of existing typologies.“ (Rounsaville et al. 2002, S. 13).

Bevor sich absehen lässt, ob und wenn ja, welches dimensionale Modell in das DSM-V und die ICD-11 aufgenommen werden wird, erscheint ein konservativer Ansatz Gewinn bringend, der von Widiger (1991, 1993) vorgeschlagen und inzwischen vielfach rezipiert (z. B. Farmer 2000; Oldham u. Skodol 2000; Trull u. Durrett 2005) und in empirischen Untersuchungen angewandt wurde (Skodol et al. 2005): Die dimensionale Diagnostik auf der Basis der vorhandenen diagnostischen Kriterien des DSM-IV. Widiger schlug vor, je nach Anzahl der erfüllten DSM-IV-Kriterien sechs Ausprägungsgrade jeder einzelnen Persönlichkeitsstörung zu definieren („absent“, „traits“, „subthreshold“, „threshold“, „moderate“, „prototype“), wobei die oberen drei Stufen über dem DSM-IV-Cut-off liegen und auch eine kategoriale Diagnose anzeigen.

Dieser Grundidee folgt der Fragebogen zur Erfassung von DSM-IV-Persönlichkeitsstörungen (ADP-IV), der von belgischen Autoren entworfen (Schotte u. De Doncker 1994, 1996; Schotte et al. 1998, 2004) und von Doering und Mitarbeitern ins Deutsche übersetzt und validiert wurde (Doering et al. 2007; Renn 2006) und bereits in einer klinischen Studie zum Einsatz kam (Renn et al. 2004). Das Selbstbeurteilungsinstrument enthält analog zu den diagnostischen Items der DSM-IV-Persönlichkeitsstörungen 94 Einzelitems, die auf einer siebenstufigen Skala eingeschätzt werden und für die darüber hinaus ein dreistufiges Distress-Rating vorgesehen ist. Die Auswertung erfolgt zum einen kategorial durch vorgegebene Algorithmen und zum anderen dimensional durch Summenscores über die Einzelitems die jeweils zu einer spezifischen Persönlichkeitsstörung gehören. Nachdem in einer früheren Arbeit die Reliabilität des Instrumentes sowie die Validität der kategorialen Diagnostik überprüft wurde (Doering et al. 2007; Renn 2006), wird in der hier vorgelegten Un-

tersuchung die dimensionale Diagnostik mit dem ADP-IV untersucht. Zu diesem Zweck werden zunächst Normen für die dimensional Summenscores ermittelt, anschließend werden mit Hilfe von Receiver Operator Curves (ROC) Cut-off-Werte für die Summenscores bestimmt, die auf der Basis der dimensional Scores wiederum eine kategoriale Diagnostik erlauben.

2. Methode

2.1. Stichprobe

643 ambulante Patienten der psychotherapeutischen Ambulanz an der Universitätsklinik für Medizinische Psychologie und Psychotherapie Innsbruck wurden um die Teilnahme an der Studie angefragt. 400 (62.2 %) gaben ihr Einverständnis und wurden in die Untersuchung eingeschlossen. Einschlusskriterien waren ein Alter von mindestens 18 Jahren und ausreichende Deutschkenntnisse. Ausschlusskriterien waren eine beeinträchtigte kognitive Leistungsfähigkeit sowie akute psychotische oder

Tabelle 1: Demographische Daten

| | Patienten (n = 400) | Bevölkerungstichprobe (n = 385) |
|-------------------------------------|---------------------|---------------------------------|
| Alter (Jahre) MW ± SD | 34.9 ± 12.0 | 35.2 ± 15.0 |
| Geschlecht | w: 279 (69.8 %) | w: 264 (68.6 %) |
| | m: 121 (30.3 %) | m: 121 (31.4 %) |
| Bildungsstand | | |
| Schule noch nicht abgeschlossen | 10 (2.5 %) | 0 (0 %) |
| Kein Schulabschluss | 18 (4.5 %) | 0 (0 %) |
| Hauptschulabschluss ohne Lehre | 107 (26.8 %) | 27 (7.0 %) |
| Hauptschulabschluss mit Lehre | 143 (35.8 %) | 112 (29.1 %) |
| Abitur/berufsbildende höhere Schule | 77 (19.3 %) | 155 (40.3 %) |
| Studium | 14 (7.3 %) | 90 (23.4 %) |
| fehlende Daten | 31 (7.8 %) | 1 (0.3 %) |
| Familienstand | | |
| ledig | 64 (30.8 %) | 88 (22.9 %) |
| unverheiratet mit Partner | 63 (30.3 %) | 131 (34.0 %) |
| verheiratet/Lebensgemeinschaft | 53 (25.5 %) | 137 (35.6 %) |
| geschieden | 27 (13.0 %) | 22 (5.7 %) |
| verwitwet | 1 (0.5 %) | 5 (1.3 %) |

schwere affektive Erkrankungen. Des Weiteren wurden 385 Personen aus der Tiroler Bevölkerung untersucht. Die Bevölkerungsstichprobe kann nicht als repräsentativ angesehen werden, da sie einen überproportional hohen Anteil an Frauen, jüngeren und höher gebildeten Probanden enthält (Tab. 1). Alle Patienten und Probanden füllten den ADP-IV aus, bei 210 Patienten wurde darüber hinaus ein SKID-II-Interview durchgeführt und in einem Expertenrating basierend auf der klinischen und der SKID-II-Diagnostik eine abschließende Einschätzung aller Persönlichkeitsstörungen vorgenommen.

2.2. Instrumente

Fragebogen zur Erfassung von DSM-IV-Persönlichkeitsstörungen (ADP-IV)

Das Instrument wurde bereits ausführlich dargestellt (siehe Doering et al. 2007) und kann unter <http://zmkweb.uni-muenster.de/einrichtungen/proth/dienstleistungen/psycho/diag/index.html> frei im Internet heruntergeladen werden. Der Fragebogen enthält 94 Items, die jeweils einem diagnostischen Kriterium für die spezifischen Persönlichkeitsstörungen des DSM-IV entsprechen. Der ADP-IV ergibt für jede Persönlichkeitsstörung einen dimensional Trait-Score und eine kategoriale (ja/nein) Diagnose. Jedes „Trait-Item“ wird zunächst auf einer siebenstufigen Skala eingeschätzt. Wird 5 oder mehr angekreuzt, ist ein zusätzliches „Distress-Rating“ auf einer dreistufigen Skala von 1 bis 3 auszufüllen. Der dimensionale Score wird durch das Addieren der einzelnen Trait-Items errechnet. Wird der Summenscore durch die Anzahl der Items geteilt, so ergibt sich ein Skalendurchschnittswert, der in einem Persönlichkeitsstörungsprofil abgebildet werden kann.

Für die kategoriale Diagnostik stehen zwei unterschiedlich sensible Auswertungs-Algorithmen zur Verfügung (siehe Doering et al. 2007; Renn 2006). Die deutsche Version des ADP-IV erwies sich als ausreichend reliabel ($r_{tt} = .37-.88$; Cronbachs $\alpha = .65-.91$). Die Übereinstimmungsvalidität mit SKID-II und Expertenrating (Sensitivität, Spezifität, Kappa-Maß) ist im Vergleich mit anderen Fragebögen zur Erfassung von Persönlichkeitsstörungen gut, für eine sichere klinische Diagnostik einzelner Persönlichkeitsstörungen jedoch nicht ausreichend (Doering et al. 2007).

Strukturiertes Klinisches Interview für DSM-IV, Achse II (SKID-II)

Das SKID-II (Fydrich et al. 1997) stellt das offizielle diagnostische Interview der American Psychiatric Association zur Erfassung von Persönlichkeitsstörungen nach den Kriterien des DSM-IV dar. Die zwölf Persönlichkeitsstörungen im DSM-IV werden durch 131 Fragen operationalisiert und auf einer dreistufigen Skala bewertet (1 = „Kriterium nicht erfüllt“, 2 = „Kriterium teilweise (oder unerschwerlich) erfüllt“, 3 = „Kriterium erfüllt“). Die Anzahl der positiven – mit „3“ bewerteten – Items wird addiert. Überschreitet die Summe die im DSM-IV vorgesehene Schwelle, wird die Diagnose der entsprechenden Persönlichkeitsstörung vergeben. Darüber hinaus können für jede Persönlichkeitsstörung dimensionale Summenscores aus den Ratings gebildet werden.

Expertenrating

Basierend auf dem SKID-II-Interview wurde ein diagnostisches Expertenrating für jeden der 210 interviewten Patienten durchgeführt. Das Vorhandensein einer Persönlichkeitsstörung nach den Kriterien des DSM-IV wurde auf einer fünfstufigen Skala geratet (1 = „Diagnosekriterien nicht passend“, 2 = „Patient zeigt einige Züge der Persönlichkeitsstörung“, 3 = „Patient zeigt deutliche Züge der Persönlichkeitsstörung“, 4 = „Patient hat die Persönlichkeitsstörung, Diagnosekriterien passend“, 5 = „Persönlichkeit des Patienten beispielhaft für die Persönlichkeitsstörung, prototypischer Fall“). Die Diagnose einer Persönlichkeitsstörung wurde für die Punkte 4 und 5 vergeben. Für die statistischen Analysen in dieser Arbeit wurden ausschließlich die kategorialen Expertenratings als Außenkriterium zugrunde gelegt.

2.3. Statistik

Normierung

Die Daten der Patienten- und der Bevölkerungsstichprobe wurden auf Normalverteilung geprüft und mit Hilfe der z-Transformation standardisiert. Aus der z-Transformation wurden die Werte in eine klassische T-Wert-Tabelle überführt ($MW \pm SD = 50 \pm 10$). Nicht normalverteilte Werte wurden mittels linearer Flächentransformation in einen Prozentrang umgerechnet. Werte, die mehr als eine Standardabweichung über oder unter dem Mittelwert liegen (68,3 % der erwarteten Fälle) werden als auffällig angesehen. Die auf diese Weise erstellten Normwerttabellen ermöglichen es, einen Patienten beziehungsweise Probanden in Relation zu seiner jeweiligen Bezugsgruppe zu setzen.

Receiver Operator Curves (ROC)

Um eine kategoriale Diagnostik auf der Basis der dimensional Summenscores der ADP-IV-Skalen zu ermöglichen, ist die Bestimmung eines Cut-off-Wertes nötig, bei dessen Überschreiten die Diagnose einer Persönlichkeitsstörung vergeben wird. Zu diesem Zweck wurden ROC als eine Methode der Grenzwertoptimierung angewandt. Dabei werden für jeden möglichen Grenzwert die resultierenden relativen Häufigkeitsverteilungen und die jeweils zugehörige Sensitivität und Spezifität errechnet. Das Ergebnis lässt sich in Form einer Kurve darstellen. Aus dieser Kurve wird grafisch der Grenzwert ermittelt, der mit der höchsten Sensitivität und Spezifität, und damit der höchsten diagnostischen Sicherheit einhergeht. Eine Steigerung der Sensitivität hat dabei eine Senkung der Spezifität zur Folge und umgekehrt.

Der Flächenanteil unterhalb der Kurve ist ein Maß für die Testgenauigkeit. Flächenanteilsgrößen über .90 gelten als ausgezeichnet, Werte über .80 als gut und solche zwischen .70 und .80 als noch akzeptabel (Janssen u. Laatz 2007, S. 697). Zur Berechnung der ROC wurden die Expertenratings herangezogen ($n = 210$).

Für die auf diese Weise ermittelten Cut-off-Werte wurden Sensitivität, Spezifität sowie positiv und negativ prädiktive Werte errechnet. Der prädiktive Wert gibt in Abhängigkeit von der Prävalenz einer Erkrankung an, wie hoch die Wahrscheinlichkeit ist, bei einem entsprechenden Testergebnis tatsächlich an dieser Erkrankung zu

leiden (positiv prädiktiver Wert) beziehungsweise nicht daran zu leiden (negativ prädiktiver Wert). Die Prävalenzen der einzelnen Persönlichkeitsstörungen in der Allgemeinbevölkerung wurden Herpertz und Saß (2003) entnommen.

3. Ergebnisse

3.1. Demographische Daten

Eine ausführliche Beschreibung der Stichproben findet sich bei Doering et al. (2007; siehe auch Renn 2006). Der Altersdurchschnitt lag in beiden Stichproben bei etwa 35 Jahren, es wurden jeweils knapp 70 % Frauen untersucht. Die Patienten zeigten im Vergleich zur Bevölkerungsstichprobe eine schlechtere Schulbildung und lebten seltener in einer Partnerschaft oder Ehe.

Bei den 210 Patienten, für die ein diagnostisches Expertenrating durchgeführt wurde, fand sich bei 82 (39.0 %) mindestens eine Persönlichkeitsstörung, zwei oder mehr Diagnosen wurden bei 32 (15.2 %) gestellt. Bei 79 % der Patienten lag eine Achse-I-Komorbidität vor, wobei Essstörungen, Anpassungsstörungen, affektive Störungen und Angsterkrankungen am häufigsten vertreten waren.

Tabelle 2: Normwerte für die Patientienstichprobe (Summenscore, Rohwerte)

| Skalen | gesamte Patientenstichprobe (n = 400) | | | Männer (n = 121) | | | Frauen (n = 279) | | |
|---------------------------|--|-------|-------|---------------------|-------|-------|---------------------|-------|-------|
| | - | Ø | + | - | Ø | + | - | Ø | + |
| paranoid | 7–11 | 12–28 | 29–49 | 7–10 | 11–32 | 33–49 | 7–11 | 12–28 | 29–49 |
| schizoid | 7–10 | 11–25 | 26–49 | 7–11 | 12–28 | 29–49 | 7–9 | 10–23 | 24–49 |
| schizotypisch | 9–13 | 14–33 | 34–63 | 9–14 | 15–34 | 35–63 | 9–12 | 13–32 | 33–63 |
| antisozial | 8 | 9–21 | 22–56 | 8–9 | 10–21 | 22–56 | – | 8–21 | 22–56 |
| Borderline | 10–17 | 18–43 | 44–70 | 10–19 | 20–42 | 43–70 | 10–17 | 18–43 | 44–70 |
| histrionisch | 8–12 | 13–30 | 31–56 | 8–13 | 14–29 | 30–56 | 8–11 | 12–31 | 32–56 |
| narzisstisch | 9–13 | 14–28 | 29–63 | 9–15 | 16–28 | 29–63 | 9–12 | 13–27 | 28–63 |
| selbstunsicher-vermeidend | 7–11 | 12–32 | 33–49 | 7–11 | 12–32 | 33–49 | 7–10 | 11–32 | 33–49 |
| dependent | 8–12 | 13–30 | 31–56 | 8–13 | 14–28 | 29–56 | 8–11 | 12–32 | 33–56 |
| zwanghaft | 8–16 | 17–33 | 34–56 | 8–18 | 19–35 | 36–56 | 8–15 | 16–33 | 34–56 |
| depressiv | 7–12 | 13–32 | 33–49 | 7–12 | 13–31 | 32–49 | 7–13 | 14–33 | 34–49 |
| passiv-aggressiv | 7–10 | 11–24 | 25–49 | 7–10 | 11–24 | 25–49 | 7–9 | 10–19 | 20–49 |

– = unterdurchschnittlich, Ø = durchschnittlich, + = überdurchschnittlich

Tabelle 3: Normwerte für die Bevölkerungsstichprobe (Summenscore, Rohwerte)

| Skalen | gesamte Bevölkerungsstichprobe (n = 385) | | | Männer (n = 121) | | | Frauen (n = 264) | | |
|---------------------------|---|-------|-------|---------------------|-------|-------|---------------------|-------|-------|
| | - | Ø | + | - | Ø | + | - | Ø | + |
| paranoid | 7-9 | 10-22 | 23-49 | 7-11 | 12-24 | 25-49 | 7-9 | 10-21 | 22-49 |
| schizoid | 7-8 | 9-19 | 20-49 | 7-10 | 11-22 | 23-49 | 7-8 | 9-18 | 19-49 |
| schizotypisch | 9-10 | 11-23 | 24-63 | 9-11 | 12-23 | 24-63 | 9-10 | 11-23 | 24-63 |
| antisozial | - | 8-16 | 17-56 | 8-9 | 10-19 | 20-56 | - | 8-15 | 16-56 |
| Borderline | 10-13 | 14-30 | 31-70 | 10-14 | 15-31 | 32-70 | 10-13 | 14-30 | 31-70 |
| histrionisch | 8-9 | 10-24 | 25-56 | 8-9 | 10-24 | 25-56 | 8-9 | 10-24 | 25-56 |
| narzisstisch | 9-11 | 12-24 | 25-63 | 9-13 | 14-26 | 27-63 | 9-11 | 12-23 | 24-63 |
| selbstunsicher-vermeidend | 7-8 | 9-22 | 23-49 | 7-9 | 10-22 | 23-49 | 7 | 8-21 | 22-49 |
| dependent | 8-9 | 10-22 | 23-56 | 8-10 | 11-23 | 24-56 | 8-9 | 10-22 | 23-56 |
| zwanghaft | 8-13 | 14-28 | 29-56 | 8-14 | 15-30 | 31-56 | 8-13 | 14-27 | 28-56 |
| depressiv | 7 | 8-22 | 23-49 | 7 | 8-21 | 22-49 | 7-8 | 9-21 | 22-49 |
| passiv-aggressiv | 7-8 | 9-19 | 20-49 | 7-9 | 10-19 | 20-49 | 7-8 | 9-18 | 19-49 |

- = unterdurchschnittlich, Ø = durchschnittlich, + = überdurchschnittlich

3.2. Normen

Die Normwerte wurden für beide Stichproben getrennt sowie für Männer und Frauen berechnet. Basierend auf den transformierten Normwerten (T-Werte) finden sich die Rohwertbereiche für durchschnittliche, unter- und überdurchschnittliche Ausprägung in den Tabellen 1 und 2. Ein erhöhter Wert („+“) ist dabei nicht gleichbedeutend mit der Diagnose einer Persönlichkeitsstörung, sondern besagt lediglich, dass die untersuchte Person mehr als eine Standardabweichung über dem Durchschnittswert ihrer Bezugsgruppe liegt, das heißt die Person liegt außerhalb jenes Bereiches, in dem die Mehrheit aller Personen (68.3 %) zu erwarten ist.

3.3. Receiver Operator Curves

Es wurden nur für die Persönlichkeitsstörungen des ADP-IV, die in der Patientenstichprobe mindestens zwölf Mal vertreten waren, ROC berechnet. Die Flächen unter den Kurven lagen zwischen .83 und .89, was nach Janssen und Laatz (2007, S. 697) als gut angesehen werden kann (Tab. 4). In Abbildung 1 ist exemplarisch die ROC für die Skala „selbstunsicher-vermeidend“ dargestellt. Die Fläche unter der Kurve betrug .88. Bei einem errechneten Cut-off-Wert von 30 weist die Skala eine Sensitivität von .89 und eine Spezifität von .83 auf.

Tabelle 4: Cut-off-Werte, Sensitivität und Spezifität sowie positiv und negativ prädiktiver Werte basierend auf der Cut-off-Bestimmung mittels ROC-Analysen

| Skalen | ROC-Fläche | Cut-off | Sensitivität | Spezifität | positiv prädiktiver Wert | negativ prädiktiver Wert | n (%) (Experten-rating) |
|---------------------------|------------|---------|--------------|------------|--------------------------|--------------------------|-------------------------|
| paranoid | .83 | ≥ 23 | .91 | .69 | 4.3 % | 99.8 % | 12 (5.7 %) |
| Borderline | .88 | ≥ 44 | .82 | .90 | 11.1 % | 99.7 % | 23 (11.2 %) |
| selbstunsicher-vermeidend | .88 | ≥ 30 | .89 | .83 | 5.0 % | 99.9 % | 28 (13.7 %) |
| zwanghaft | .85 | ≥ 29 | .88 | .69 | 5.5 % | 99.7 % | 16 (7.9 %) |
| depressiv | .89 | ≥ 30 | .83 | .80 | 6.5 % | 99.8 % | 25 (11.9 %) |

Für die Berechnung der prädiktiven Werte wurden die Prävalenzen der einzelnen Persönlichkeitsstörungen in der Allgemeinbevölkerung zugrunde gelegt. Siehe hierzu Herpertz und Saß 2003.

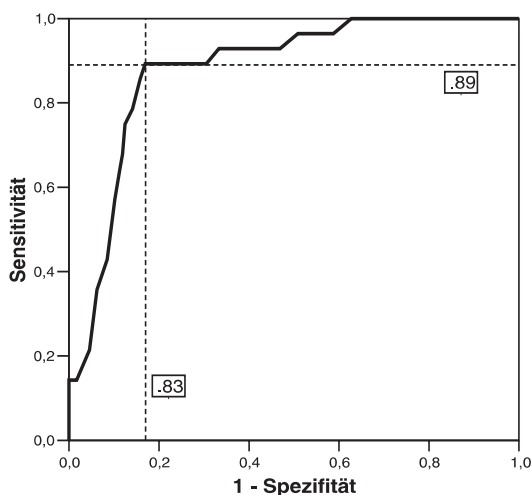


Abbildung 1: Receiver Operator Curve (ROC) für die ADP-IV-Skala „selbstunsicher-vermeidend“

Bei der Festlegung der Cut-off-Werte wurde das Ziel verfolgt, möglichst hohe negativ prädiktive Werte zu erreichen. Die Cut-off-Werte bewegten sich für die einzelnen Skalen zwischen 23 (Skala „paranoid“) und 44 (Skala „Borderline“) (Tab. 4). Die Sensitivität der Skalen lag zwischen .82 und .93, die Spezifität zwischen .69 und .90. Die positiv prädiktiven Werte waren durchwegs gering (< 12 %), die negativ prädiktiven Werte dagegen sehr hoch (> 99 %). Die Wahrscheinlichkeit, dass eine mit dem ADP-IV diagnostizierte Persönlichkeitsstörung tatsächlich vorliegt, ist aufgrund vieler falsch positiver Diagnosen relativ gering, die Wahrscheinlichkeit, dass bei negativer ADP-IV-Diagnostik tatsächlich keine Persönlichkeitsstörung vorliegt, dagegen sehr hoch.

4. Diskussion

Die valide Diagnostik einer Persönlichkeitsstörung setzt eine gründliche klinische Untersuchung des Patienten voraus, wobei oft mehr als ein diagnostisches Gespräch nötig ist, um zu einer abschließenden Einschätzung zu gelangen. Der Einsatz von strukturierten Interviews wie dem Strukturierten Klinischen Interview für DSM-IV (SKID-II; Fydrich et al. 1997) oder der International Personality Disorder Examination (IPDE; Loranger et al. 1994) ist in Ergänzung zur klinischen Untersuchung hilfreich. Spitzer (1983) prägte den Begriff des „Longitudinal Experts using All Data“ (LEAD) für eine diagnostische Prozedur, bei der mehrere Diagnostiker und Behandler („experts“) alle im Längsschnitt gewonnenen Informationen über den Patienten zusammentragen, um so zu einer validen Diagnose einer Persönlichkeitsstörung zu kommen. Dieses Vorgehen sah Spitzer als den „Gold Standard“ der Diagnostik an. Angesichts dieses Aufwandes, der offenbar für eine valide Diagnostik betrieben werden muss, erscheint es unrealistisch, mit einem Fragebogen allein zum selben Ziel zu gelangen. Tatsächlich ist dies auch bisher mit keinem der etablierten Instrumente gelungen: weder der „Personality Diagnostic Questionnaire“ (PDQ; Bronisch et al. 1993; Hyler 1994), noch der SKID-II-Screening-Fragebogen (Fydrich et al. 1997), das „Millon Multiaxial Inventory“ (MCMI; Millon et al. 1994), das „Wisconsin Personality Disorders Inventory“ (WISPI; Klein et al. 1993) oder die deutsche Version des ADP-IV (Doering et al. 2007; Renn 2006) zeigt eine annähernd befriedigende Übereinstimmungsvalidität mit kategorialen Interview-Diagnosen. Welchen Sinn haben dann überhaupt derartige Fragebögen?

Unseres Erachtens kann der ADP-IV unter gewissen Voraussetzungen bei der Diagnostik von Persönlichkeitsstörungen eine zeitökonomische Screeningfunktion erfüllen und so eine sinnvolle Ergänzung zur klinischen Diagnostik darstellen. Der Fragebogen kann in erster Linie feststellen, ob ein Patient *keine* Persönlichkeitsstörung hat. Dies gelingt mit der dimensionalen Auswertung allerdings äußerst zuverlässig, wenn die mit Hilfe der ROC bestimmten Cut-off-Werte für die Summenscores angewandt werden; hierfür sprechen die negativ prädiktiven Werte von durchweg über 99 %. Nicht-Erkrankte werden also fast immer richtig als solche erkannt. Indiziert dieser Auswertungsmodus dagegen das Vorliegen einer Persönlichkeitsstörung, so darf dies nicht als Diagnose verstanden werden. Wie bei den meisten anderen derartigen Fragebögen treten hier viele falsch positive Diagnosen auf. Da eine umfangreiche klinische Überprüfung aller positiven Screening-Ergebnisse sehr aufwändig wäre, empfiehlt es sich, erst dann klinisch vertieft zu explorieren, wenn ein positives Screening-Ergebnis mit einem entsprechenden klinischen Eindruck zusammen trifft.

Darüber hinaus können beim ADP-IV die andernorts beschriebenen algorithmischen Auswertungsmodi (Doering et al. 2007; Renn 2006) zusätzlichen Aufschluss geben, da diese weniger falsch positive und damit höhere positiv prädiktive Werte aufweisen und sich daher eher zur positiven Diagnostik von Persönlichkeitsstörungen eignen – allerdings um den Preis einer höheren Anzahl falsch negativer Diagnosen.

Neben der quasi kategorialen Diagnostik mit Cut-off-Werten kann das Persönlichkeitsstörungsprofil des ADP-IV ein Hilfsmittel für die weitere Diagnostik und Behandlungsplanung sein, da hier unterschwellige und komorbide pathologische Persönlichkeitszüge abgebildet werden. Darüber hinaus können die z-transformierten Normwerte Anhaltspunkte für auffällige Ausprägungen pathologischer Persönlichkeitsdimensionen liefern.

Die methodischen Grenzen dieser Studie liegen in erster Linie in der geringen Fallzahl bei einzelnen Untergruppen von Persönlichkeitsstörungen sowie in der Tatsache, dass die Bevölkerungsstichprobe nicht repräsentativ und für die Berechnung altersspezifischer Normen zu klein ist.

Der ADP-IV ist der einzige deutschsprachige Fragebogen zur Erfassung aller DSM-IV-Persönlichkeitsstörungen, der ausreichend validiert wurde und darüber hinaus frei verfügbar ist (download des Fragebogens, der Auswertungsbögen sowie wichtiger Publikationen unter <http://zmkweb.uni-muenster.de/einrichtungen/proth/dienstleistungen/psycho/diag/index.html>). Mit seiner kategorialen Auswertung kann der Fragebogen mit Einschränkungen (relativ hohe Zahl falsch negativer Diagnosen) zur Diagnostik von Persönlichkeitsstörungen eingesetzt werden, die dimensionale Auswertung kann Erfolg versprechend für eine Screening-Diagnostik zum Ausschluss von Persönlichkeitsstörungen verwandt werden. Beide Anwendungen sollten nicht allein zur Diagnosestellung genutzt werden, vielmehr sollten sie mit ihren unterschiedlichen Stärken und Schwächen für die Indikation weiterer spezifischer klinischer Diagnostik im Bereich einzelner Persönlichkeitsstörungen angewandt werden.

Literatur

- Barnow, S., Herpertz, S., Spitzer, C., Dudeck, M., Grabe, H. J., Freyberger, H. J. (2006): Kategoriale versus dimensionale Klassifikation von Persönlichkeitsstörungen: sind dimensionale Modelle die Zukunft? *Fortschr Neurol Psychiatr* 74, 706–713.
- Bronisch, T., Flett, S., Garcia-Borreguero, D., Wolf, R. (1993): Comparison of a self-rating questionnaire with a diagnostic checklist for the assessment of DSM-III-R personality disorders. *Psychopathology* 26, 102–107.
- Cloninger, C. R., Svrakic, D. M., Przybeck, T. R. (1993): A psychobiological model of temperament and character. *Arch Gen Psychiatry* 50, 975–990.
- Costa, P. T., McCrae, R. R. (1992): The five-factor model of personality and its relevance to personality disorders. *J Personal Disord* 6, 343–359.
- Costa, P. T., Widiger, T. A. (1993): *Personality disorders and the five-factor model of personality*. Washington DC: American Psychological Association.
- Dilling, H., Mombour, W., Schmidt, M. H. (2004): *Internationale Klassifikation Psychischer Störungen. ICD-10 Kapitel V(F). Klinisch-diagnostische Leitlinien*. 5. Auflage. Bern: Huber.
- Doering, S., Renn, D., Höfer, S., Rumpold, G., Smrekar, U., Janecke, N., Schatz, D., Schotte, C., DeDoncker, D., Schüßler, G. (2007): Validierung der deutschen Version des Fragebogens zur Erfassung von DSM-IV Persönlichkeitsstörungen (ADP-IV). *Z Psychosom Med Psychother* 53, 111–128.

- Farmer, R. F. (2000): Issues in the assessment and conceptualization of personality disorders. *Clin Psychol Rev* 20, 823–851.
- Fydrich, T., Renneberg, B., Schmitz, B., Wittchen, H. (1997): Strukturiertes Klinisches Interview für DSM-IV Achse II: Persönlichkeitsstörungen. Interviewheft. Göttingen: Hogrefe.
- Herpertz, S. C., Saß, H. (2003): Persönlichkeitsstörungen. Stuttgart: Thieme.
- Hyer, S. E. (1994): Personality questionnaire, PDQ-4+. New York: New York Psychiatric Institute.
- Janssen, J., Laatz, W. (2007): Statistische Datenanalyse mit SPSS für Windows. 6. Auflage. Berlin: Springer-Verlag.
- Klein, M. H., Benjamin, L. S., Rosenfeld, R., Treece, C., Husted, J., Greist, J. H. (1993): The Wisconsin personality disorders inventory: Development, reliability, and validity. *J Personal Disord* 7, 285–303.
- Leibing, E., Doering, S. (2006): Diagnostik von Persönlichkeitsstörungen. *Psychotherapeut* 51, 229–244.
- Leibing, E., Jamrozinski, K., Vormfelde, S. V., Stahl, J., Doering, S. (2008): Dimensions of personality – relationship between DSM-IV personality disorder symptoms, the five-factor model, and the biosocial model of personality. *J Personal Disord* 22, 101–108.
- Loranger, A. W., Sartorius, N., Andreoli, A., Berner, W., Buchheim, P., Channabasavanna, S. M., Coid, B., Dahl, A., Diekstra, R. F., Ferguson, B., Jacobsberg, L. B., Mombour, W., Pull, C., Ono, Y., Regier, D. A. (1994): The international personality disorder examination. The World Health Organization/Alcohol, Drug Abuse, and Mental Health Administration international pilot study of personality disorders. *Arch Gen Psychiatry* 51, 215–24.
- Millon, T., Millon, C., Davis, R. D. (1994): Millon Clinical Multiaxial Inventory – III. Minneapolis: National Computer Systems.
- Oldham, J. M., Skodol, A. E. (2000): Charting the future of axis II. *J Personal Disord* 14, 17–29.
- Renn, D., Emshoff, R., Höfer, S., Doering, S. (2004): Psychopathologie bei PatientInnen mit Temporomandibulären Störungen. *Psychologische Medizin* 15, 27–32.
- Renn, D. (2006): Fragebogen zur Erfassung von Persönlichkeitsstörungen: ADP-IV – psychometrische Prüfung der deutschen Version. Dissertation an der Leopold-Franzens-Universität Innsbruck, Österreich.
- Rounsaville, B. J., Alarcon, R. D., Andrews, G., Jackson, J. S., Kendell, R. E., Kendler, K. (2002): Basic nomenclature issues for DSM-V. In: Kupfer, D. J., First, M. B., Regier, D. E. (Hg.): A research agenda for DSM-V, S. 1–29. Washington DC: American Psychiatric Association.
- Saß, H., Wittchen, H.-U., Zaudig, M. (1996): Diagnostisches und statistisches Manual psychischer Störungen DSM-IV. Göttingen: Hogrefe.
- Schotte, C., De Doncker, D. (1994): ADP-IV Questionnaire. Antwerpen: University Hospital Antwerp.
- Schotte, C., De Doncker, D. (1996): ADP-IV Questionnaire: Manual and Norms. University Hospital Antwerp.
- Schotte, C. K. W., De Doncker, D., Vankerckhoven, C., Vertommen, H., Cosyns, P. (1998): Self-report assessment of the DSM-IV personality disorders. Measurement of trait and distress characteristics: the ADP-IV. *Psychol Med* 28, 1179–1188.
- Schotte, C. K. W., De Doncker, D., Dmitruk, D., Van Mulders, I., D’Haenen, H., Cosyns, P. (2004): The ADP-IV Questionnaire: Differential validity and concordance with semistructured interview. *J Personal Disord* 18, 405–419.
- Schüßler, G., Burgmer, M., Mans, E., Schneider, G., Dahlbender, R., Heuft, G. (2006): Neurotische Konfliktsystematik – Grundlagen der Konfliktdiagnostik (OPD-2). *Z Psychosom Med Psychother* 52, 23–38.
- Skodol, A. E., Oldham, J. M., Bender, D. S., Dyck, I. R., Stout, R. L., Morey, L. C., Shea, M. T.,

- Zanarini, M. C., Sanislow, C. A., Grilo, C. M., McGlashan, T. H., Gunderson, J. G. (2005): Dimensional representations of DSM-IV personality disorders: Relationships to functional impairment. *Am J Psychiatry* 162, 1919–1925.
- Spitzer, R. L. (1983): Psychiatric diagnosis: are clinicians still necessary? *Compr Psychiatry* 24, 399–411.
- Trull, T. J., Durrett, C. A. (2005): Categorical and dimensional models of personality disorder. *Annu Rev Clin Psychol* 1, 355–380.
- Westen, D., Shedler, J. (1999): Revising and assessing axis II, part I: Developing a clinically and empirically valid assessment method. *Am J Psychiatry* 156, 258–272.
- Widiger, T. A. (1991): Personality disorder dimensional models proposed from DSM-IV. *J Personal Disord* 5, 386–398.
- Widiger, T. A. (1993): The DSM-III-R-R categorical personality disorder diagnoses: A critique and an alternative. *Psychol Inq* 4, 75–90.
- Wöller, W., Tress, W. (2005): Die psychotherapeutische Behandlung von Persönlichkeitsstörungen. *Z Psychosom Med Psychother* 51, 110–127.

Korrespondenzadresse: Univ.-Prof. Dr. med. Stephan Doering, Bereich Psychosomatik in der Zahnheilkunde, Poliklinik für Zahnärztliche Prothetik und Werkstoffkunde, Universitätsklinikum Münster, Waldeyerstraße 30, D-48149 Münster
E-Mail: stephan.doering@ukmuenster.de