Aus der Klinik und Poliklinik für Psychiatrie und Psychotherapie
der Ludwig-Maximilians-Universität München
Vorstand Prof. Dr. H.-J. Möller

Erfassung depressiver Störungen in der Primärversorgung: Validität des Brief Patient Health Questionnaire (B-PHQ) und des General Health Questionnaire (GHQ-12)

Dissertation
zum Erwerb des Doktorgrades der Humanbiologie an der Medizinischen Fakultät
der Ludwig-Maximilians-Universität zu München

vorgelegt von:

Simone Braun

aus Heilbronn

2002

Mit Genehmigung der Medizinischen Fakultät
der Universität München

Berichterstatter:	Prof. Dr. med. U. Hegerl
Mitberichterstatter:	Prof. Dr. J. Hasford Prof. Dr. R. L. Haberl
Mitbetreuung durch den promovierten Mitarbeiter:	Dr. med. V. Henkel
Dekan:	Prof. Dr. med. Dr. h. c. K. Peter
Tag der mündlichen Prüfung:	17.09.2002

Meinem Mann und meinen Eltern in Liebe und Dankbarkeit gewidmet.

Danksagung

Diese Arbeit hätte ohne die vielfältige Unterstützung der hier genannten Personen nicht verwirklicht werden können. Mein besonderer Dank gilt meinem Betreuer Herrn Prof. Dr. Ulrich Hegerl. Als Sprecher und Koordinator des Kompetenznetzes Depression und Leiter der Abteilung für Klinische Neurophysiologie der Psychiatrischen Klinik der Ludwig-Maximilians-Universität München war er der Hauptinitiator dieses Vorhabens und hat diese Arbeit erst möglich gemacht.

Weiterhin danke ich Herrn Prof. Dr. Hans-Jürgen Möller, der als Direktor der Psychiatrischen Klinik der Ludwig-Maximilians-Universität München die Rahmenbedingungen für diese Forschungsarbeit geschaffen hat. Mein Dank gilt insbesondere auch der Projektleiterin Frau Dr. Verena Henkel für ihre wertvolle Betreuung dieser Arbeit sowie meinen Kolleginnen, Kollegen und Freunden, die durch methodische und inhaltliche Anregungen mein Vorhaben immer wieder bereichert haben.

Dieses Vorhaben hätte ohne die engagierte Arbeit der Ärztinnen, Ärzte und der Praxismitarbeiterinnnen sowie die große Teilnahmebereitschaft und Offenheit der interviewten Patientinnen und Patienten nicht entstehen können. Ihnen gilt mein ausdrücklicher Dank.

Frau Prof. Dr. Ulla Mitzdorf und der Kommission zur Förderung des wissenschaftlichen Nachwuchses sowie der Fakultätskommission danke ich für die Bewilligung eines Stipendiums im Rahmen des Hochschul- und Wissenschaftsprogramms für die Abschlussphase meiner Promotion.

Inhaltsverzeichnis

1 Einleitung **1**

1.1 Depressive Störungen: Epidemiologie, Symptomatik, Diagnostik, Ätiologie und Therapie 1

1.2 Depressive Störungen in der Primärversorgung 13

2 Fragestellung **23**

3 Methoden **28**

3.1 Untersuchungsdesign und -durchführung 28

3.2 Statistische Zielgrößen 30

3.2.1 Sensitivität und Spezifität 30

3.2.2 Prädiktive Werte 37

3.2.3 Weitere Kenngrößen der Validität: Overall Accuracy, Kappa Index und Likelihood Quotient 41

3.2.4 Receiver Operating Characteristica (ROC) Kurven 49

3.3 Screeninginstrumente 54

3.3.1 General Health Questionnaire (GHQ-12) 54

3.3.2 Brief Patient Health Questionnaire (B-PHQ) 62

3.4 Goldstandard: DIA-X, eine rechnergestützte Version des CIDI 69

4 Ergebnisse **74**

4.1 Stichprobencharakteristika und Prävalenz 74

4.1.1 Alters- und Geschlechtsverteilung 74

4.1.2 Ergebnisse der Auswertung der Screeninginstrumente 77

4.1.3 Ergebnisse der Auswertung des DIA-X Interviews 82

4.2 Sensitivität und Spezifität 88

4.2.1 Sensitivität und Spezifität des GHQ-12 88

4.2.2 Sensitivität und Spezifität des B-PHQ Auswertungsalgorithmus A 92

4.2.3 Sensitivität und Spezifität des B-PHQ Auswertungsalgorithmus B 97

4.3 Prädiktive Werte 103

4.3.1 Prädiktive Werte des GHQ-12 103

4.3.2 Prädiktive Werte des B-PHQ Auswertungsalgorithmus A 108

4.3.3 Prädiktive Werte des B-PHQ Auswertungsalgorithmus B 113

4.4 Overall Accuracy, Kappa Index und Likelihood Quotienten 118
4.4.1 Overall Accuracy, Kappa Index und Likelihood Quotienten des GHQ-12 118
4.4.2 Overall Accuracy, Kappa Index und Likelihood Quotienten des B-PHQ Auswertungsalgorithmus A 124
4.4.3 Overall Accuracy, Kappa Index und Likelihood Quotienten des B-PHQ Auswertungsalgorithmus B 132
4.5 Receiver Operating Characteristica (ROC) Kurven 140
4.5.1 ROC-Kurven des GHQ-12, des B-PHQ Auswertungsalgorithmus A und des B-PHQ Auswertungsalgorithmus B für die gesamte Stichprobe 140
4.5.2 ROC-Kurven des GHQ-12, des B-PHQ Auswertungsalgorithmus A und des B-PHQ Auswertungsalgorithmus B getrennt nach Geschlecht 143
4.5.3 ROC-Kurven des GHQ-12, des B-PHQ Auswertungsalgorithmus A und des B-PHQ Auswertungsalgorithmus B getrennt nach Altersgruppen 149

5 Diskussion **156**
5.1 Stichprobencharakteristika und Prävalenz 157
5.2 Sensitivität und Spezifität 158
5.3 Prädiktive Werte 163
5.4 Overall Accuracy, Kappa Index und Likelihood Quotienten 165
5.5 ROC-Kurven 167
5.6 Empfehlungen für die Praxis 169

6 Zusammenfassung **171**

Literaturverzeichnis **174**

Anhang A Poster zur Information der Patienten im Wartezimmer **192**
Anhang B Vierfeldertafeln für die Cut-Off-Werte des GHQ-12, des B-PHQ Algorithmus A und des B-PHQ Algorithmus B **193**
Anhang C Weitere Kenngrößen der Validität: Rate der falsch Positiven, Rate der falsch Negativen, Misclassification Rate und Odds-Ratio **229**

1 Einleitung

Depressive Störungen zählen weltweit zu den häufigsten psychiatrischen Erkrankungen. Sie treten in jedem Lebensalter auf und beeinflussen, wie kaum eine andere Krankheit, in fundamentaler Weise die Lebensqualität der Betroffenen. In den letzten Jahren wurden weitreichende Fortschritte im Wissen um die Ursachen, Diagnostik und Therapiemöglichkeiten gemacht, so dass die Behandlungschancen heute als gut bis sehr gut bezeichnet werden können. Die Mehrheit der Patienten befindet sich nicht in fachärztlicher Behandlung, sondern sucht den Hausarzt auf. Hier werden die Weichen für die Diagnostik, die Therapie und den weiteren Verlauf der Erkrankung gestellt. Zahlreiche Studien belegen jedoch, dass bei über der Hälfte der Patienten die depressive Störung nicht erkannt wird. Verschiedene Studien haben gezeigt, dass durch eine routinemäßige Rückmeldung der Ergebnisse in Screeninginstrumenten und durch das Training von Ärzten die Erkennensrate und das Management depressiver Störungen verbessert werden können (Asvall, 2001; Rutz, 1999; Titus, van Os, Ormel, van den Brink, Jenner, Van der Meer, Tiemens, van der Doorn, Smith & van den Brink, 1999). Die Weltgesundheitsorganisation (WHO) schlägt vor, alle Patienten in der Primärversorgung mit Hilfe von Screeninginstrumenten zu untersuchen (WHO, 1998). Die vorliegende Arbeit ist die erste Studie in der die Validität von zwei Screeninginstrumenten, dem General Health Questionnaire (GHQ-12, Goldberg, 1972) und dem Brief Patient Health Questionnaire (B-PHQ, Spitzer, Kroenke & Williams, 1999) in Relation zu einem standardisierten klinisch-diagnostischen Interview, dem DIA-X, anhand einer einzigen Stichprobe in der Primärversorgung erhoben und verglichen wird.

1.1 Depressive Störungen: Epidemiologie, Symptomatik, Diagnostik, Ätiologie und Therapie

Epidemiologie

Depressionen gehören zu den affektiven Störungen und sind eine ernste Erkrankung des gesamten Menschen. Sie führen zu Veränderungen des Stoff-

wechsels, verschiedener Körperfunktionen und beeinflussen die Gedanken, Gefühle und Verhaltensweisen. Zentrale Beschwerden sind Niedergeschlagenheit, Freudlosigkeit, Interessenverlust und Verminderung des Antriebs. Depressionen zählen zu den häufigsten psychischen Erkrankungen und die Wahrscheinlichkeit an einer depressiven Störung zu erkranken, ist insgesamt als hoch einzuschätzen. Das Morbiditätsrisiko liegt für Frauen bei bis zu 26% und für Männer bei bis zu 12% (Hautzinger & de Jong-Meyer, 1998; Boyd & Weissman, 1981). Seit den 80er Jahren wurde eine Vielzahl von epidemiologischen Studien zur Schätzung der Prävalenz depressiver Störungen in der Allgemeinbevölkerung publiziert (vgl. Angst, 1992; Kessler, McGonagle, Zhao, Nelson, Hughes, Eshleman, Wittchen & Kendler, 1994; Lépine, Gastpar, Mendlewicz & Tylee, 1997; Rieger, Narrow, Rae, Manderscheid, Locke & Goodwin, 1993; Wittchen, Essau, von Zerssen, Krieg & Zaudig, 1992). Unterschiede in den Prävalenzraten ergeben sich vor allem durch die Verwendung verschiedener diagnostischer Kriterien und Instrumente sowie durch unterschiedliche Alterszusammensetzungen in den Stichproben (Wittchen, Knäuper & Kessler, 1994). Im Rahmen des 1998/1999 durchgeführten bundesweiten Gesundheitssurvey konnten Wittchen, Müller, Schmidtkunz, Winter und Pfister (2000) mittels des standardisierten Composite International Diagnostic Interviews (CIDI) zeigen, dass in den letzten vier Wochen vor der Untersuchung 6,3% der 18 bis 65 Jährigen in der Bundesrepublik Deutschland an einer Depression nach den Kriterien der ICD-10 und des DSM-IV litten. Die Rate der Frauen liegt mit 7,8% signifikant höher als die Rate der Männer (4,8%). Die Zwölfmonatsprävalenz für Depressionen liegt bei 11,5%, wobei Frauen wiederum deutlich häufiger (15%) betroffen sind als Männer (8,1%). Die häufigsten Erkrankungsformen sind depressive Episoden mit 8,3% und Dysthymie mit 2,5%. Es zeigte sich, dass für Frauen ein fast doppelt so hohes Risiko besteht, an einer depressiven Episode zu erkranken wie für Männer. Ebenfalls ist das Risiko bereits in jungen Jahren rezidivierende depressive Episoden zu entwickeln für Frauen höher. Für dysthyme und bipolare Störungen wurden keine signifikanten Unterschiede zwischen Frauen und Männern gefunden. Betrachtet man die lebenszeitbezogenen Daten des Gesundheitssurvey, so sind Frauen

mit 24,5% in allen Altersstufen mehr als doppelt so häufig von depressiven Störungen betroffen wie Männer (11,9%).

Symptomatik und diagnostische Kriterien

Depressionen sind eine Krankheit von allgemeinem Interesse, denn sie sind weit verbreitet, ernstzunehmend und nur wenige andere Erkrankungen beeinflussen so stark die Lebensqualität und Lebenszufriedenheit (Gray, 2001; Lépine et al., 1997; Palsson & Skoog, 1997). Depressive Störungen stellen eine enorme Belastung für den Einzelnen, seine Familie, das Gesundheitssystem und die Gesellschaft dar (Ballenger, 2000). Depressionen können zu Beeinträchtigungen des körperlichen Befindens, des Denkens, der Gefühle und der Gestimmtheit sowie zu Beeinträchtigungen der Person-Umweltbezüge und des Bezugs zur eigenen Person führen. Sie gehen mit einem hohen Maß an subjektiver Beeinträchtigung und Behinderung im Alltag einher. Wolfersdorf (2000, S.17) definiert Depression als „ein Zustandsbild allgemeiner und umfassender seelisch-körperlicher Herabgestimmtheit“. Der Begriff Depression stammt von dem lateinischen Begriff deprimere, was so viel bedeutet wie herunterdrücken, unterdrücken. Bei einer Depression sind drei wesentliche Funktionsbereiche gestört. Es treten psychische, psychomotorische und somatische Symptome auf. Die psychischen Symptome äußern sich etwa in einen Mangel an Antrieb, Gefühlen der Hoffnungslosigkeit, innerer Leere, Angst und gedrückter Stimmung. Zu den psychomotorischen Symptomen zählen psychomotorische Hemmung, Kommunikationshemmung aber auch psychomotorische Agitiertheit. Somatische Symptome sind beispielsweise Schlafstörungen, Appetitverlust, Schmerzen, Kraftlosigkeit, Druck- und Kältegefühl.

Die Abgrenzung affektiver Störungen vom gesunden Erleben ist häufig nicht einfach, da Depressionen durch eine Vielzahl heterogener Symptome gekennzeichnet sind. Aus diesem Grunde sind definierte Kriterien für die Diagnostik sehr wichtig (Ahrens, 2000). In der heutigen Zeit besteht die Tendenz den Begriff „depressiv“ für jede schlechte Gestimmtheit zu verwenden. Bei einer depressiven Störung handelt es sich jedoch um eine schwere und ernst zunehmende Erkrankung, die den ganzen Menschen umfasst und die lebensbedrohlich werden kann. So liegt die Lebenszeitsuizidalität bei Patienten mit schweren

depressiven Störungen bei circa 15%. Einen Suizidversuch unternehmen etwa 20 bis 60% der depressiv Erkrankten und 40 bis 80% leiden an Suizidideen während einer depressiven Erkrankung (Kasper, Möller & Müller-Spahn, 1997).
Die Diagnostik affektiver Störungen erfolgt heute meist anhand der beiden aktuell gültigen Klassifikationssysteme ICD-10 der Weltgesundheitsorganisation (WHO, 1992) und DSM-IV der American Psychiatric Association (APA, 1994). Beide Klassifikationssysteme beruhen auf einer genauen Deskription der diagnostischen Kategorien im Sinne einer operationalen Diagnostik. Die ICD-10 der WHO ist seit 1998 das offizielle Klassifikationssystem im ambulanten und stationären Bereich in der Bundesrepublik Deutschland (Schneider & Margraf, 1998). Ziel der WHO ist es ein für alle Länder akzeptables Klassifikationssystem zu schaffen. Nach Thangavelu und Martin (1995) hat die ICD-10 eher Richtliniencharakter während das DSM-IV eher Vorschriftencharakter besitzt. In den klinisch-diagnostischen Leitlinien der ICD-10 werden die affektiven Störungen unterteilt in F30 manische Episoden, F31 bipolare affektive Störung, F32 depressive Episode, F33 rezidivierende depressive Störungen, F34 anhaltende affektive Störungen, F38 sonstige affektive Störungen und F39 nicht näher bezeichnete affektive Störungen.
Da bei einem nicht unwesentlichen Prozentsatz der Patienten nur eine einzelne Episode auftritt, werden einzelne Episoden von wiederholt auftretenden episodischen Störungen und bipolaren Störungen unterschieden (Dilling, Mombour & Schmidt, 1991).
Bei den manischen Episoden (F30) werden drei Schweregrade unterschieden. Manische Episoden sind gekennzeichnet durch gehobene Stimmung sowie eine Steigerung der körperlichen und psychischen Aktivitäten in ihrer Geschwindigkeit und ihrem Ausmaß. Die Hypomanie (F30.0) ist definiert als eine leichtere Ausprägung von (F30.1), der Manie. Manische Störungen sind durch eine situationsinadäquat gehobene Stimmung gekennzeichnet, sowie durch vermehrten Antrieb, Überaktivität und ein vermindertes Schlafbedürfnis sowie starke Ablenkbarkeit und überhöhte Selbstüberschätzung. Die Episode dauert mindestens eine Woche und ist so schwer, dass sie die soziale und berufliche Leistungsfähigkeit unterbricht.

Bipolare affektive Störung (F31) sind durch mindestens zwei Episoden mit deutlich gestörter Stimmung und gestörtem Aktivitätsniveau definiert. Kennzeichen sind zum einen hypomanische oder manische zum anderen depressive Episoden. Meist tritt eine vollständige Symptombesserung zwischen den Episoden ein. Patienten mit ausschließlich manischen Symptomen sind sehr selten und werden in der ICD-10 als bipolar klassifiziert (F31.8). Die bipolaren affektiven Störungen werden nach der Art der gegenwärtigen Episode unterteilt.
Nach den klinisch-diagnostischen Leitlinien der ICD-10 werden die depressiven Episoden (F32) nach dem Schweregrad und dem Vorhandensein von psychotischen bzw. somatischen Symptomen unterteilt. Zu den Kernsymptomen einer depressiven Episode zählen gedrückte Stimmung, Interessenverlust und Freudlosigkeit sowie Verminderung des Antriebs. Durch die Verminderung der Energie kommt es zu erhöhter Ermüdbarkeit und zu Aktivitätseinschränkungen. Weitere häufige Symptome sind: Verminderung der Konzentration und Aufmerksamkeit, vermindertes Selbstwertgefühl und Selbstvertrauen, Schuldgefühle und Gefühle von Wertlosigkeit, negative und pessimistische Zukunftsperspektiven, Suizidgedanken, Selbstverletzung oder Suizidhandlungen, Schlafstörungen und verminderter Appetit.
Für die Diagnose einer depressiven Episode (leicht, mittelgradig, schwer) wird eine Dauer der Symptome von mindestens zwei Wochen gefordert. Kürzere Zeiträume können nach der ICD-10 dann berücksichtigt werden, wenn die Symptome ungewöhnlich schnell aufgetreten sind oder ungewöhnlich schwer sind. Die diagnostischen Leitlinien für die leichte depressive Episode (F32.0) fordern das Vorhandensein von mindestens zwei der drei Kernsymptome (depressive Stimmung, Verlust von Interesse oder Freude und Verminderung des Antriebs) und mindestens zwei der übrigen der oben genannten Symptome. Die diagnostische Leitlinien für mittelgradige depressive Episoden (F32.1) fordern das Vorhandensein von mindestens zwei Kernsymptomen und mindestens drei weiteren Symptomen. Für die Diagnose einer schweren depressiven Episode ohne psychotische Symptome (F32.2) wird das Vorhandensein aller drei Kernsymptome und mindestens vier weiterer Symptome gefordert. Von letzteren sollten einige besonders ausgeprägt sein. Bei besonders schweren Symptomen oder besonders raschem Auftreten der Symptome kann die Diagnose unter

Umständen auch in einem Zeitraum von weniger als zwei Wochen gestellt werden.

Mit F33 werden die rezidivierenden depressiven Störungen bezeichnet. Diese sind durch wiederholte depressive Episoden gekennzeichnet. Die rezidivierenden depressiven Störungen werden nach dem Typus der gegenwärtigen Episode und nach dem vorherrschenden Typus der vorangegangenen Episoden unterschieden.

Langanhaltende affektive Störungen werden mit F34 kodiert. Hierunter werden anhaltende, meist fluktuierende Stimmungsstörungen verstanden, wobei die einzelnen Episoden selten so schwer sind, dass sie die Kriterien für hypomanische bzw. leichte depressive Episoden erfüllen. Die diagnostischen Leitlinien für die Zyklothymie (F34.0) fordern eine anhaltende Stimmungslabilität mit zahlreichen Perioden leichter Depression und leicht gehobener Stimmung, wobei die einzelnen Episoden nicht die Kriterien für manische oder depressive Episoden erfüllen. Die Dysthymie (F34.1) ist definiert als eine langandauernde depressive Verstimmung, die nie oder höchstens selten so stark ausgeprägt ist, dass sie die diagnostischen Kriterien für eine rezidivierende leichte (F33.0) oder mittelgradige (F33.1) depressive Störung erfüllt.

Unter F38 fallen andere affektive Störungen. Mit F38.0 werden andere einzelne affektive Störungen bezeichnet und mit F38.00 gemischte affektive Episoden. Unter F38.1 fallen andere rezidivierende affektive Störungen. Mit F38.10 wird die rezidivierende kurze depressive Störung bezeichnet. Es handelt sich hierbei um kurze rezidivierende depressive Episoden, die im vergangenen Jahr etwa einmal im Monat auftraten. Die Episoden sind kürzer als zwei Wochen, meist dauern sie zwei bis drei Tage. Die Symptomkriterien erfüllen die Kriterien einer leichten, mittelgradigen oder schweren depressiven Episode.

Unterschwellige depressive Störungen

Unterschwellige psychiatrische Störungen erreichen die diagnostischen Schwellen der etablierten Kriterien nicht (Papassotiropoulos & Heun 1999). Die Diagnostik in der ICD-10 und dem DSM-IV basiert auf der Qualität, der Quantität und der zeitlichen Dauer der Symptome (Judd, Rapaport, Paulus & Brown,

1994). Unterschwellige depressive Störungen erfüllen entweder das Zeit- oder das Quantitätskriterium nicht und erreichen daher nicht den Schwellenwert für eine ICD-10 oder DSM-IV Diagnose einer Depression. Obwohl die vollen diagnostischen Kriterien nicht erfüllt werden, können unterschwellige Störungen zu Beeinträchtigung führen. Nach Ballenger, Davidson, Lecrubier, Nutt, Kirmayer, Lépine, Lin, Tajama & Ono (2001) ist es wichtig, den Schwellenwert einer Störung nicht mit der Behandlungsbedürftigkeit zu vermischen, da beispielsweise ein Patient zwar die vollen diagnostischen Kriterien erfüllen kann aber sich nicht beeinträchtigt fühlt, während ein anderer Patient, der zwar nicht die vollen diagnostischen Kriterien erfüllt, stark beeinträchtigt und behandlungsbedürftig sein kann.

Unterschwellige Formen depressiver Störungen sind ein wichtiges Ziel für die Prävention und Behandlung, da sie zu erheblicher Beeinträchtigung und Dysfunktion führen (Üstün & Chatterji, 2001) und in der Allgemeinbevölkerung (Angst, Merikangas & Preisig, 1997) sowie im nichtpsychiatrischen Bereich (Maier, Gänsicke & Weiffenbach, 1997; Pini, Perkonning, Tansella & Wittchen, 1999) häufig vorkommen. Nach Horwath, Johnson, Klerman & Weissman (1992) liegt die Lifetimeprävalenz für depressive Symptome bei 24%, nach Johnson, Weissman und Klerman (1992) bei 23,1% für subklinische Depression. Die Prävalenzraten variieren in verschiedenen Studien in Abhängigkeit von der verwendeten Definition der unterschwelligen Depression.

Unterschwellige Depressionen führen zu einer Reduktion der Lebensqualität (Helmchen & Linden, 2000; Spitzer, Kroenke, Linzer, Hahn, Williams, deGruy, Brody & Davies, 1995; Wells, Stewart, Hays, Burnam, Rogers, Daniels, Berry, Greenfield, & Ware, 1989) und zu starker Beeinträchtigung des Betroffenen (Lyness, King, Cox, Yoediono & Caine, 1999; Magruder & Calderone, 2000; Wagner, Burns, Broadhead, Yarnall, Sigmon & Gaynes, 2000) sowie zu verminderter Arbeitsfähigkeit (Skodol, Schwartz, Dohrenwend, Levav & Shrout, 1994).

Das Risiko, eine andere volle psychiatrische Störung zu entwickeln, ist erhöht (Broadhead, Blazer, George & Tse, 1990). So sind subklinische Depressionen ein signifikanter Risikofaktor für Major Depression (Kessler, Zhao, Blazer & Swartz, 1997). Broadhead et al. (1990) konnten zeigen, dass 10% der Patien-

ten, die an einer minoren Depression mit Stimmungsbeeinträchtigung litten, innerhalb von einem Jahr eine Major Depression entwickelten.

Ferner stehen unterschwellige Depressionen im Zusammenhang mit erhöhten Suizidraten und -versuchen (Judd, Akiskal & Paulus, 1997) sowie einer erhöhten Nutzung des Gesundheitswesen (Beck & Koening, 1996; Johnson et al., 1992). Zur Behandlung unterschwelliger Depressionen liegen bislang nur wenige Studien vor. Miranda und Munoz (1994) konnten zeigen, dass kognitive Verhaltenstherapie die depressive Symptomatik und die Somatisierung bei minoren Depressionen reduziert. Die Befunde verschiedener Therapiestudien deuten darauf hin, dass auch leichtere Depressionen durch Antidepressiva gebessert werden können (z.B. Rapaport & Judd, 1998; Szegedi, Wetzel, Angersbach, Philipp & Benkert, 1997). Insgesamt ist die Wirksamkeit von pharmako- und psychotherapeutischen Verfahren bei unterschwelligen Depressionen bislang jedoch erst unzureichend empirisch überprüft.

Zur Definition unterschwelliger Depressionen existieren verschiedene Konzepte mit verschiedenen Namen (z.B. minore oder subklinische Depression) und Definitionen, mit unterschiedlicher Zeitdauer, Symptomschwellen und Ausschlusskriterien. Gleiche Definitionen haben häufig verschiedene Namen und verschiedene Definitionen tragen zum Teil den gleichen Namen. Eine Literaturübersicht von Pincus, Davis und McQueen (1999) ergab, dass minore Depressionen auf neun verschiedene Arten definiert wurden, depressive Symptome auf drei und unterschwellige bzw. subsyndromale Depressionen auf fünf verschiedene Arten. Die Mindestzahl der geforderten Symptome lag zwischen einem und sechs. Meistens wurden mindestens zwei Symptome gefordert. In mehr als der Hälfte der Studien wurde vorausgesetzt, dass eines der Symptome depressive Stimmung ist. Häufig wurde auch der Verlust von Interesse oder Freude gefordert. In einigen Studien wurden Screeninginstrumente zur Bestimmung der minoren Depression verwendet. In insgesamt 13 Studien wurde keine Angabe über die geforderte Mindestdauer der Symptome gemacht, in 11 Studien wurde eine Mindestdauer von zwei Wochen gefordert. 25 der 36 untersuchten Studien enthielten kein Beeinträchtigungskriterium. In 14 Studien durfte für die Diagnose minore Depression aktuell keine andere affektive Störung vorliegen und in vier Studien durften keine Lebenszeitdiagnosen affektiver

Störungen vorhanden sein. Auch Feldman, Robbins und Jaffe (1998) fanden in ihrer Literaturübersicht verschiedene Definitionen für die Diagnose minore Depression. So setzten einige Autoren minore Depressionen mit Dysthymie gleich. Andere, wie beispielsweise Oxman, Barrett, Barrett und Gerber (1990) verwenden die Research Diagnostic Criteria (RDC) von Spitzer, Endicott und Robins (1978). Eine geringere Anzahl von Symptomen wie sie für die Diagnose einer Major Depression notwendig sind, forderten z.B. Jaffe, Froom und Galambos (1994). Zu der Kategorie andere Definitionen zählten Feldman et al. (1998) die Definition der sogenannten „subsyndromal symptomatic depression" (SSD) von Judd, Rapaport, Paulus und Brown (1994). Lyness et al. (1999) sowie Philipp, Delmo, Buller, Schwarze, Winter, Maier und Benkert (1992) verwendeten die Hamilton Depression Scale (HAMD, Hamilton, 1960) zur Bestimmung unterschwelliger Depressionen. In den Forschungskriterien (Anhang B) des DSM-IV (American Psychiatric Association, 1994) wird die leichte depressive Störung wie folgt definiert: Zwei bis vier der neun Symptome einer Major Depression lagen während derselben beiden Wochen vor, wobei eines davon depressive Stimmung oder Verlust von Interesse oder Freude ist. Es handelt sich dabei um eine Veränderung im normalen Funktionsniveau, die klinisch relevantes Leiden oder Beeinträchtigung hervorruft und nicht auf die direkte körperliche Wirkung einer Substanz oder eines medizinischen Krankheitsfaktors zurückgeht. Die Symptome dürfen nicht besser durch eine normale Trauerreaktion erklärbar sein. In der Vorgeschichte lag keine Major Depression, manische, gemischte oder hypomane Episode vor. Die Kriterien für eine dysthyme oder zyklothyme Störung sind nicht erfüllt. Das Störungsbild darf nicht ausschließlich während einer Schizophrenie, schizophrenieformen Störung, schizoaffektiven Störung, wahnhaften Störung oder nicht näher bezeichneten psychotischen Störung auftreten. Diese Definition findet beispielsweise bei Hance, Carney, Freedland und Skala (1996) Anwendung.

In der Einteilung von Papassotiropoulos und Heun (1999) auf Basis der Kriterien des DSM-III-R (American Psychiatric Association, 1987) zählen zu unterschwelligen Depressionen minore Depression, Recurrent Brief Depression und depressive Symptome. Minore Depressionen sind hier definiert durch das Vorliegen von weniger als fünf depressiven Symptomen für mindestens zwei

Wochen. Die Diagnose Recurrent Brief Depression erfordert fünf und mehr depressive Symptome, die weniger als zwei Wochen andauern und im letzten Jahr mindestens einmal im Monat auftraten. Die Diagnose depressives Syndrom wird gestellt, wenn eine Episode mit mindestens einem depressiven Symptom, ohne Spezifikation der Dauer, Häufigkeit und Beeinträchtigung vorliegt.

In der vorliegenden Arbeit wird in Anlehnung an die Kriterien der ICD-10 und des DSM-IV die aktuelle Diagnose „unterschwelliges depressives Syndrom" gestellt, wenn mindestens zwei verschiedene Symptome einer depressiven Episode in den letzten beiden Wochen vor der Durchführung des DIA-X Interviews erfüllt waren, aber noch nicht die Kriterien für eine volle affektive Störung erfüllt sind. In der Vorgeschichte können affektive Störungen nach den Kriterien der ICD-10 und des DSM-IV vorgelegen haben.

Pincus et al. (1999) kommen zusammenfassend zu dem Ergebnis, dass auf dem Gebiet der unterschwelligen Depressionen noch erheblicher Forschungsbedarf besteht, wie etwa die Untersuchung der Entstehung und des Verlaufs unterschwelliger Störungen anhand von Längsschnittstudien.

Verlauf

Depressive Störungen treten in allen Lebensaltern auf. Höheres Lebensalter stellt dabei keinen besonderen Risikofaktor dar (Hautzinger & de Jong-Meyer, 1998). Als Spätdepressionen werden Depressionen bezeichnet, die erstmals nach dem 45. Lebensjahr auftreten. Depressionen die erstmalig nach dem 60. Lebensjahr auftreten, werden als Altersdepressionen im engeren Sinne bezeichnet (Kasper, Möller & Müller-Spahn, 1997). Zusammenfassend kann festgehalten werden, dass es eine große individuelle Variabilität im Verlauf affektiver Erkrankungen gibt. 20-30% der affektiven Störungen verlaufen singulär, 70-80% sind rezidivierend. Bei unipolaren depressiven Störungen treten im Mittel vier bis sechs Episoden im Verlauf auf, bei bipolaren Erkrankungen sind es acht bis zehn Episoden. Die Phasendauer liegt bei rezidivierenden depressiven Störungen bei ca. fünf Monaten, bei bipolaren Störungen bei ungefähr vier Monaten. Die Zyklusdauer beträgt für unipolare Depressionen etwa fünf Jahre, für

bipolare Störungen zwei bis drei Jahre (Ahrens, 2000; Hautzinger & de Jong-Meyer, 1998).

Ätiologie

In den letzten zwanzig Jahren wurden weitreichende Fortschritte im Wissen um die Ursachen, Diagnostik und Therapiemöglichkeiten von depressiven Erkrankungen gemacht. Depressive Erkrankungen gelten heute als eine psycho-sozio-biologische Störung, da sie zum einen ätiopathologisch von psychologischen, biologischen und sozialen Faktoren ausgehen und zum anderen psychologische, biologische und soziale Folgen haben (Wolfersdorf, 2000). Es gibt allerdings bis heute kein allgemeingültiges ätiopathologisches Konzept. Insgesamt kann davon ausgegangen werden, dass multikausale Faktoren für depressive Störungen verantwortlich sind. Hierzu gehören genetische, biologische, psychologische, reaktive und soziale Faktoren.

Zu den bekanntesten psychologischen Theorien zählen die Theorie der erlernten Hilflosigkeit von Seligman (1974), die Verstärkerverlust-Theorie von Lewinsohn (1974) und die kognitive Theorie von Beck (1974). Nach Abramson, Seligman & Teasdale (1978) sind Depressionen das Ergebnis der erlebten Nichtkontrollierbarkeit von subjektiv bedeutsamen Ereignissen. Eine internale, globale und stabile Ursachenattribution führt auch in der Zukunft zu einer Misserfolgserwartung. Nach Lewinsohn (1974) befindet sich der depressive Patient in einer Art Löschungsbedingung. Auslösend für die Depression ist eine geringe Rate an verhaltenskontingenter Verstärkung. Nach Beck (1974) ist die Basis der depressiven Störung eine kognitive Störung. Die kognitive Störung führt zum sogenannten „kognitive Trias“ mit einer negativen Sicht der Welt, der eigenen Person und der Zukunft.

In der Neurobiologie ist seit mehr als 30 Jahren die führende Hypothese die sogenannte „Monoamin-Mangelhypothese“. Im neurobiochemischen Modell wird eine Dysregulation und/oder Beeinträchtigung der biogenen Amine, insbesondere der Neurotransmitter Serotonin, Noradrenalin und Dopamin angenommen. Als Ursache der Depression gilt entweder ein Mangel an Transmittern im synaptischen Spalt oder eine verringerte Sensibilität der postsynaptischen Rezeptoren (Stahl, 1999). Neueren Annahmen zufolge handelt es sich hierbei

jedoch nur um den ersten Schritt in einer „längeren pathopysiologisch bedeutsamen Kaskade“ (Kasper, Möller & Müller-Spahn, 1997, S.22). Neben den neuropharmakologisch abgeleiteten Hypothesen liegen verschiedene neuroendokrinologische Hypothesen wie beispielsweise die Glukokortikoid-Rezeptor-Hypothese, die eine Fehlfunktion der Glukokortikoidrezeptoren annimmt, vor (Holsboer-Trachsler & Vanoni, 2000).
Die traditionelle Einteilung in exogene, endogene und psychogene Störungen findet in den modernen Diagnosesystemen keine Verwendung mehr.

Therapie

Die Behandlung von Depressionen ist ein mehrdimensionaler Prozess mit guten bis sehr guten Behandlungschancen. Es können dabei biologische, psychotherapeutische, psychosozial-systemische Ansätze sowie eine Integration der verschiedenen Ansätze zum Einsatz kommen. Die am besten belegten Psychotherapieformen zur Behandlung von depressiven Störungen sind die kognitive Verhaltenstherapie, die tiefenpsychologische Psychotherapie und die interpersonelle Psychotherapie (Hautzinger, 1995; Wittchen, 1997). Zu den allgemeinen Wirkfaktoren der psychotherapeutischen Arbeit gehören eine vertrauensvolle, tragfähige Beziehung, das Angebot eines Erklärungsprinzips für die Erkrankung und die Behandlung, die Vermittlung von Hoffnung und Erfolgserlebnissen sowie die Analyse der Probleme und die Erarbeitung von Bewältigungsmöglichkeiten. Es hat sich dabei als sinnvoll erwiesen, tagesstrukturierende Maßnahmen einzusetzen (Wolfersdorf, 2000).
Im Bereich der Pharmakotherapie liegen heute zahlreiche wirksame antidepressive Medikamente zum gezielten Einsatz bei den verschiedenen depressiven Störungen vor. Antidepressiva lassen sich nach ihrer chemischen Struktur bzw. ihrem chemischen Wirkmechanismus unterteilen. Die wesentlichen Gruppen sind dabei die selektiven Serotonin-Wiederaufnahmehemmer (SSRI), die reversiblen selektiven Hemmer der Monoaminoxidase (RIMA), Medikamente mit dualem bzw. rezeptorspezifischem Wirkmechanismus, Trizyklische und Tetrazyklische Antidepressiva, Monoaminoxidase (MAO-) Hemmer und Phytopharmaka wie beispielsweise Johanniskraut. Nicht medikamentöse, biologisch fundierte Behandlungsmethoden sind z.B. Lichttherapie, Transkranielle

Magnetstimulation und Schlafentzug, (Kasper, Möller & Müller-Spahn, 1997; Wu, Buchsbaum, Gillin, Tang, Cadwell, Wiegand, Najafi, Klein, Hazen & Bunney, 1999).

Die heute verfügbaren Antidepressiva wirken auf das Noradrenalin- oder das Serotoninsystem bzw. auf beide Systeme gleichzeitig und gegebenenfalls auch auf weitere Transmittersysteme. Für die Gruppe der Antidepressiva konnte in klinischen Studien keine Abhängigkeit im Sinne der WHO-Definition beobachtet werden. Die Therapie sollte den unterschiedlichen Depressionsformen und dem Patienten jeweils individuell angepasst werden. Allgemein wird die Behandlungsstrategie mit Psychopharmaka in drei Phasen gegliedert. An die Akutbehandlung bis zur Remission schließt sich die Erhaltungstherapie mit einer durchschnittlichen Dauer von vier bis sechs Monaten an. Die dritte Phase, die Rezidivprophylaxe, kann sich über einen unbestimmten Zeitraum von mehreren Jahren erstrecken (Möller, Müller & Volz, 2000). Die Behandlung kann ambulant oder stationär erfolgen. Akute Suizidalität ist ein wichtiger Grund für die stationäre Behandlung. Die Erfolgsrate antidepressiver Medikamente liegt bei 60 bis 70% (vgl. Gastpar, 1986). Bei Patienten mit längerer Therapieresistenz können spezielle Maßnahmen ergriffen werden, beispielsweise die Erhöhung der Tagesdosis oder therapeutischer Schlafentzug (Kasper, Möller & Müller-Spahn, 1997).

Zusammenfassend kann festgehalten werden, dass trotz der bislang nicht endgültig geklärten Ätiologie depressiver Störungen die adäquate Anwendung medizinischer und psychologischer Behandlungsmethoden bei einem Großteil der Patienten erfolgreich ist.

1.2 Depressive Störungen in der Primärversorgung

Schlüsselrolle des Hausarztes

In der Allgemeinarztpraxis stellen die Erkennung und die Behandlung psychischer Störungen einen substantiellen Teil der Versorgungsaufgaben dar. Dies ergibt sich zum einen aus der hohen Prävalenz psychischer Störungen in der Primärversorgung und zum anderen aus der Schwere der Beeinträchtigung

dieser Patientengruppe. Die Mehrheit der Patienten mit einer depressiven Störung befindet sich nicht in fachärztlicher Behandlung sondern sucht den Hausarzt auf (Coyne, Klinkman, Gallo & Schwenk, 1997; Goldberg, 1999; Hegerl, 2000; Linden, Lecrubier, Bellantuono, Benkert, Kisley & Simon, 1999). Weltweit werden ca. 50 bis 90% der Patienten mit depressiven Störungen in der Primärversorgung betreut und weniger als 10% werden zum Spezialisten überwiesen (Gray, 2001; Linden, Maier, Achberger, Herr, Helmchen & Benkert, 1996; Tylee, 2001). So werden nach Paykel, Tylee, Wright, Priest, Rix & Hart (1997) in den USA ca. 50% der Patienten beim Hausarzt behandelt, in Großbritannien sind es sogar 90%. Damit kommt dem Hausarzt bei der Versorgung depressiver Patienten eine Schlüsselrolle zu (Von Korff, Shapiro, Burke, Teitlebaum, Skinner, German, Turner, Klein, & Burns, 1987). Er ist häufig die erste Anlaufstelle für die Betroffenen im Versorgungssystem und hat somit eine „Gatekeeperfunktion“. Hier werden die Weichen für die Diagnostik und Therapie und damit auch für den weiteren Verlauf der depressiven Erkrankungen gestellt (vgl. Wittchen, Höfler & Meister, 2000).

Schätzungen der Prävalenz depressiver Störungen ergeben, dass bis zu 30% der Patienten in der Primärversorgung unter depressiven Störungen leiden (Dawson & Tylee, 2001). Coyne, Fechner-Bates und Schwenk (1994) fanden mit dem SCID eine Prävalenz von 13,5% für Major Depression bei Patienten in der primärärztlichen Versorgung. Die Ergebnisse der WHO-Studie zu psychischen Erkrankungen in der Allgemeinarztpraxis ergaben mit Hilfe des CIDI eine Prävalenz von 11,2% in Mainz und 6,1% in Berlin für depressive Episoden. Die Prävalenzraten für depressive Episoden lagen hier zwischen 18,3% im Studienzentrum in Manchester und 4,7% in Verona (Maier, Linden & Sartorius, 1996). Sartorius, Üstün, Lecrubier und Wittchen (1996) ermittelten mit dem CIDI aktuelle (Ein-Monats-) Prävalenzen von 10,4% für depressive Episoden (F32/F33) und 2,1% für Dysthymie (F34) in der Primärversorgung.

In der Studie „Depression 2000“ erfüllten am Untersuchungsstichtag 10,9% der Patienten in der Primärversorgung die Kriterien für eine Depression anhand des Depressions-Screening-Questionnaires (DSQ, Winter, Wittchen, Höfler, Spiegel, Ormel, Müller & Pfister, 2000). Frauen waren mit 11,9% häufiger betroffen als Männer (9,4%). Die Schätzungen der Prävalenz variieren in Abhängigkeit

von der Falldefinition und den verwendeten Instrumenten (Üstün & Chatterji, 2001).

Diagnostisches und therapeutisches Defizit

Zahlreiche Studien belegen jedoch, dass bei einem großen Teil der Patienten in der Primärversorgung die Depressionen nicht richtig erkannt oder nicht konsequent behandelt werden (Ballenger et al., 2001; Hegerl & Möller, 2000; Rost, Zhang, Fortney, Smith, Coyne & Smith, 1998). So erkennen Allgemeinärzte weltweit über die Hälfte der betroffenen Patienten nicht (Coyne, Schwenk & Fechner-Bates, 1995; Coyne, Thompson, Palmer, Kagee & Maunsell, 2000; Ormel, Van Den Brink, Koeter, Giel, Van Der Meer, Van De Willige & Wilmink, 1990; Paykel et al., 1997). Weniger als 10% der Betroffenen erhalten eine angemessene Behandlung (Ballenger et al., 2001; Hirschfeldt, Keller, Panico, Arons et al., 1997; Kleinman & Cohnen, 2001). Häufig sind die verschriebenen Medikamente unangemessen oder nicht richtig dosiert (Coyne et al., 2000). Die größten Verbesserungsmöglichkeiten in der Versorgung depressiver Patienten liegen somit im Bereich der Allgemeinarztpraxen.

Die direkten und indirekten Kosten der Unterdiagnostizierung und Unterbehandlung depressiver Störungen sind sowohl für den Betroffenen als auch für das Gesundheitssystem und die Gesellschaft hoch. Depressionen führen zu besonderem Leid, da sie zentral und direkt die gesamte Lebensqualität des Patienten beeinträchtigen sowie die Sterblichkeitsrate an anderen Krankheiten erhöhen (Üstün & Chatterji, 2001). Circa 15% der Patienten mit einer schweren Depression sterben durch Suizid, die Rate der Suizidversuche wird als doppelt so hoch geschätzt (Chisholm, 2001). Depressionen beeinträchtigen gewöhnlich nicht nur den Betroffenen, sondern auch seine Familie und Freunde. Im Durchschnitt werden von einer depressiv erkrankten Person fünf weitere Personen ernsthaft mitbetroffen.

Durch unspezifische Behandlungen und nicht notwendige diagnostische Maßnahmen verursachen nicht erkannte depressive Störungen erheblichen Kosten für das Gesundheitssystem. Ferner ist die Inanspruchnahme ärztlicher Dienste gegenüber nicht depressiven Patienten drastisch erhöht. Winter et al. (2000) kamen zu dem Ergebnis, dass depressive Patienten im Mittel 11,9 Hausarztbe-

suche in den letzten 12 Monaten gemacht hatten, Nichtdepressive hingegen nur 6,6. Wird die Krankheit nicht in einem frühen Stadium erkannt und behandelt, können sich die Kosten verzehnfachen (Asvall, 2001).
Nach Untersuchungen der World Health Organisation (WHO) und der Weltbank nahmen 1990, gemessen an dem Indikator DALYs (Disability Adjusted Life Years), unipolare Major Depression den vierten Rang der 15 Hauptursachen für verlorenen Lebensjahre durch Tod oder schwerwiegende Behinderung hinter Infektionen der unteren Atemwege, Durchfallerkrankungen und perinatalen Erkrankungen (Murray & Lopez, 1997) ein. Nach Ballenger et al. (2001) wird die Major Depression bis zum Jahre 2020 einer der häufigsten Gründe für Beeinträchtigungen sein.

Ursachen

Die Ursachen für das diagnostische und therapeutische Defizit sind vielfältig. Nach Sackville (2001) sind sich Ärzte häufig der Ernsthaftigkeit der Krankheit und ihrer engen Verbindungen zu anderen Störungen nicht bewusst und haben es versäumt sich in der Diagnostik und Therapie weiterzubilden. Aus verschiedensten Gründen ist die Diagnose depressiver Störungen in der Primärversorgung nicht einfach, oft deshalb weil der Patient sich nicht mit den klassischen Symptomen depressiver Störungen vorstellt (Kruse, Heckrath, Schmitz, Alberti & Tress, 1999; Schüren, Kleinschmidt, Müßigbrodt & Schürmann, 1999). Fast 80% der Patienten in der Primärversorgung kommen mit somatischen Symptomen zum Arzt (Üstün & Chatterji, 2001). Es wird geschätzt, dass bei ca. 25% der Patienten mit Rückenschmerzen die zugrundeliegende Ursache eine Depression ist (Dawson & Tylee, 2001). Häufig liegen eine oder mehrere komorbide körperliche oder psychische Störungen vor (Wittchen, Lieb, Wunderlich & Schuster, 1999). Die Daten des National Comorbidity Survey (NCS) zeigen, dass 58% aller Patienten mit der Diagnose Major Depression auch eine komorbide psychiatrische Störung haben (Üstün & Chatterji, 2001).
Weitere Gründe liegen in negativen Einstellungen und der Zeitknappheit (Callahan, Hendrie, Dittus, Brater, Hui, & Tierney, 1994a). So haben Allgemeinärzte im Schnitt pro Patient nur 10 Minuten Zeit. Ferner konnten Botega, Mann, Blizard und Wilkinson (1992) mit dem Depression Attitude Questionnaire (DAQ),

der die Einstellung von Allgemeinärzten zu Depressionen misst, zeigen, dass es eine Gruppe von Ärzten gibt, die die Behandlung von Patienten mit Depressionen für undankbar hält und sich unwohl im Umgang mit depressiven Patienten fühlt. Auf der Seite der Patienten werden die depressiven Symptome häufig nicht erkannt und in ihrer Schwere unterschätzt bzw. aus Angst vor Stigmatisierung nicht erwähnt (Hirschfeld et al., 1997; Katon, Von Korff, Lin, Lipscomb, Russo, Wagner & Polk, 1990).

Verschiedene Studien haben gezeigt, dass besonders bei älteren Patienten und bei Männern Depressionen häufig unerkannt bleiben (Ames, 2001; Salzman, 1997). Die Symptome depressiver Erkrankungen werden gerade bei älteren Menschen häufig als Reaktionen auf den Alterungsprozess gewertet (Hegerl, 2001; Henkel & Hegerl, 2000; Skalad, 1995). Ältere Menschen haben möglicherweise andere Symptommuster als jüngere. Sie vermeiden es meist über die niedergeschlagene Stimmung zu sprechen, eher werden Angst und somatische Symptome eingestanden. Oft werden chronische Schmerzen in verschiedenen Körperteilen genannt (Gottfries, 1997; Katona, Livingston, Manela, Leek, Mullan, Orrell, D'Ath & Zeitlin, 1997). Ferner liegt häufig Multimorbidität vor (Zaudig, 2001). Nach Rutz (1999) hängen die Ursachen für die Unterdiagnostizierung und -behandlung bei Männern möglicherweise mit dem geringeren Hilfesuchverhalten von Männer sowie deren atypischen Symptomen mit aggressivem und Missbrauchsverhalten zusammen.

Ballenger et al. (2001) konnten zeigen, dass die Erkennensrate depressiver Störungen mit der Schwere der Symptome und dem spontanem Äußern von psychologischen Beschwerden ansteigt. Gerade in der Primärversorgung sind jedoch leichte und unterschwellige depressive Störungen häufig. Wittchen (2000) sowie Wittchen, Höfler und Meister (2001) weisen darauf hin, dass neben der Unterdiagnostizierung auch das Problem der hohen Rate an falsch positiven Diagnosen nicht übersehen werden darf. Ärzte in der Primärversorgung benötigen daher ein zeiteffizientes und effektives System für das Screening (Montano, 1994).

Verbesserungsmöglichkeiten

Verschiedene Studien haben gezeigt, dass durch eine routinemäßige Rückmeldung von Ergebnissen in Screeninginstrumenten und durch das Training von Ärzten die Erkennensrate und das Management depressiver Störungen verbessert werden können (Asvall, 2001; Titus, van Os, Ormel, van den Brink, Jenner, Van der Meer, Tiemens, van der Doorn, Smit & van den Brink, 1999; Rutz, 1999). So konnten Callahan, Hendrie, Dittus, Brater, Hui und Tierney (1994a) für die Primärversorgung zeigen, dass intensives Screening und die Rückmeldung von patientenspezifischen Behandlungsempfehlungen das Erkennen und die Behandlungsrate von Altersdepressionen erhöhten. Allerdings konnten keine signifikanten Verbesserungen der Depression oder der Schwere der Beeinträchtigung nachgewiesen werden. In einer weiteren Studie kommen dieselben Autoren (Callahan, Hendrie, Dittus, Brater, Hui & Tierney, 1994b) zu der Schlussfolgerung, dass ein formales Screenen aller älteren Patienten in der Primärversorgung die effektivste Methode zur Verbesserung der Erkennensrate darstellt. Simon, Goldberg, Tiemens und Ustun (1999) kamen zu dem Ergebnis, dass das Erkennen und die richtige Diagnosestellung einer depressiven Störung in der Primärversorgung zumindest kurzfristig zu einer signifikanten Verbesserung der Depression führen. Die Verbesserung des Erkennens ist ein wichtiger Schritt in Richtung eines effektiveren Managements der Depression (vgl. auch Rost et al., 1998).

So schlägt Montano (1994) vor jeden neuen Patient mit Hilfe von Screeninginstrumenten zu untersuchen, auch wenn die somatischen Symptome völlig unassoziiert mit Depressionen zu sein scheinen.

Die „Consensus Group on Depression and Anxiety“ um Ballenger (Ballenger et al., 2001) empfiehlt einen zweistufigen Prozess zur Verbesserung der Erkennensrate und Diagnose depressiver Störungen. In einem ersten einfachen Schritt findet die Fallidentifikation statt. Im zweiten Schritt werden dann weitere Details zur Erstellung einer Diagnose erhoben.

Die WHO (1998) fordert in ihrem Informationspaket „Mastering Depression in Primary Care“, dass das Screening des Wohlbefindens zu einem Standardverfahren werden sollte ebenso wie das Screening nach Bluthochdruck. Der Fra-

gebogen „WHO-5 Well-Being-Index“ sollte von allen Patienten ausgefüllt werden, denn fehlendes Wohlbefinden wird als ein Indikator für das mögliche Vorliegen einer Depression gesehen.

Screeningmaßnahmen

Das Ziel von Screeningmaßnahmen ist es, mit Hilfe von schnell durchführbaren Untersuchungen oder Tests möglichst alle Personen zu erkennen, die mit hoher Wahrscheinlichkeit eine bestimmte Erkrankung oder einen bestimmten Risikofaktor aufweisen. Screeningverfahren sind jedoch nicht per se diagnostische Maßnahmen. Sie stellen vielmehr eine Filterfunktion dar zur Entdeckung derjenigen Personen, für die die größte Wahrscheinlichkeit besteht, die gesuchte Krankheit zu haben. Personen, die im Screeningtest auffällige Ergebnisse aufweisen, werden in einem nächsten Schritt eingehender diagnostisch untersucht. Das Konzept des Screenings beruht auf der Annahme, dass durch die frühe Entdeckung einer Erkrankung bei augenscheinlich gesunden Personen eine effektive Behandlung und Heilung erreicht werden kann (Maruish, 2000).
Screeninguntersuchungen sind Teil vieler Maßnahmen der primären Prävention und Teil aller Maßnahmen der sekundären Prävention. Maßnahmen der Primärprävention sollen das Auftreten einer Krankheit durch die Beseitigung der Krankheitsursachen verhindern. Maßnahmen der Sekundärprävention dienen der Früherkennung einer Krankheit mit dem Ziel, eine Progression rechtzeitig aufzuhalten (Fletcher, Fletcher & Wagner, 1999).
Wilson und Junger unterscheiden in ihrer von der World Health Organisation (WHO) 1968 publizierten Arbeit „Principles and practices of screening for diseases“ verschiedene Formen des Screenings. Unter *Massenscreening* wird das groß angelegte Screening ganzer Populationsgruppen verstanden. Beim sogenannten *selektiven Screening* werden ausgewählte Hochrisikogruppen mit Screeninginstrumenten untersucht. Das selektive Screening kann auch als eine Form des Populationsscreenings aufgefasst werden. Als *multiples Screening* wird die Anwendung einer Kombination von zwei oder mehr Screeningverfahren bezeichnet.
Ferner wird zwischen dem Einsatz von Screeninginstrumenten zum sogenannten „case-finding“ und dem Einsatz von Screeninginstrumenten in epidemiolo-

gischen Untersuchungen unterschieden. Ziel der Fallentdeckung ist es, Krankheiten zu erkennen, um anschließend den Patienten adäquat behandeln zu können. In epidemiologischen Studien werden ebenfalls Screeningtests eingesetzt, jedoch nicht mit dem primären Ziel der Behandlung des Patienten, sondern mit dem Ziel der Untersuchung von Prävalenz, Inzidenz und dem natürlichen Verlauf der Krankheiten. Die Fallidentifikation ist hier ein Nebenprodukt.
Unabhängig von dieser Unterscheidung können verschiedene Merkmale eines optimalen Screeningverfahrens formuliert werden. Ein idealer Screeningtest ist hoch empfindlich, d.h. er entdeckt alle Personen, die mit großer Wahrscheinlichkeit die gesuchte Erkrankung aufweisen. Ein Screeningverfahren sollte schnell durchführbar und kostengünstig sein, möglichst wenige Unannehmlichkeiten für den Patienten verursachen und sich durch hohe Akzeptanz auszeichnen.
Die Vorläufer der heutigen, modernen Screeninginstrumente entstanden im späten 19. und im frühen 20. Jahrhundert. Der erste Selbstbeurteilungsfragebogen war das „Personal Data Sheet“. Es wurde von Robert Woodworth im Jahre 1918 entwickelt. Mit diesem Verfahren sollten die amerikanischen Soldaten im ersten Weltkrieg auf psychiatrische Störungen untersucht werden. Ungefähr zur selben Zeit konstruierte der Psychiater A. Meyer die erste psychiatrische Beurteilungsskala, das sogenannte „Philipps Behaviour Chart“ (Maruish, 2000). Depressionsskalen werden seit ca. 40 Jahren entwickelt (Stieglitz, 1998).
Als Screeninginstrumente werden üblicherweise eher Selbstbeurteilungsinstrumente als klinische Fremdbeurteilungsskalen verwendet. Selbstbeurteilungsinstrumente sind meist kurz, kostengünstig in der Durchführung und haben eine hohe Akzeptanz bei den Probanden. Ihre Durchführung, Auswertung und Beurteilung kann häufig auch von Nicht-Klinikern vorgenommen werden. Heutzutage sind viele Selbstbeurteilungsskalen auch am Computer durchführ- und auswertbar.
Der größte Vorteil von Selbstbeurteilungsbögen besteht darin, dass der Test von der Person, die das Phänomen erlebt, selbst ausgefüllt wird. Nachteil aller Selbstbeurteilungsinstrumente ist, dass sie von der Bereitschaft und Kompetenz des Befragten zur Selbstbeurteilung abhängen. Bewusste und unbewusste

Verfälschungstendenzen des Patienten fallen somit stärker ins Gewicht. Ein weiterer Nachteil ist die geringe Flexibilität von Selbstbeurteilungsbögen, da sie meist keine adaptiven Testungen erlauben (vgl. Maruish, 2000; Stieglitz & Baumann, 1994).

Wilson und Junger (1968) diskutieren in Ihrer Arbeit zehn Prinzipien, die für das Screening von Krankheiten eminent wichtig sind:

1. Die untersuchte Krankheit ist ein bedeutendes Gesundheitsproblem.
2. Verfügbarkeit von effektive Behandlungsmöglichkeiten.
3. Verfügbarkeit von Einrichtungen zur Diagnose und Behandlung.
4. Frühes Erkennen der Krankheit senkt signifikant die Morbidität und Mortalität.
5. Reliabilität und Validität der eingesetzten Screeningverfahren, damit falsch positive und falsch negative Klassifikationen vermieden werden.
6. Hohe Akzeptanz des Tests in der Bevölkerung.
7. Bekanntheit des natürlichen Verlaufs der Krankheit.
8. Übereinstimmung darüber, welche Personen als behandlungsbedürftig gelten.
9. Kosteneffektivität der Screeningprogramme.
10. Kontinuierlicher Prozess der Fallerkennung.

Validität von Screeninginstrumenten

Durch das Hinzuziehen eines gut validierten Screeninginstrumentes ließen sich entscheidende Verbesserungen bei der Diagnostik depressiver Syndrome in der Primärversorgung erzielen. Die Validität eines Tests wird auch als Gültigkeit bezeichnet und ist das Maß an Genauigkeit, mit dem der Test dasjenige Merkmal misst oder vorhersagt, das er messen oder vorhersagen soll (Amelang & Zielinski, 1994; Lienert, 1994). Im Kontext des Screenings bedeutet die Validität wie gut ein Instrument zwischen Personen unterscheiden kann, die das gesuchte Merkmal besitzen und solchen Personen, die das gesuchte Merkmal nicht besitzen. Zur Beurteilung der Validität stehen verschiedene Validierungsmaße zu Verfügung. Die Validität eines Screeninginstruments wird üblicherweise durch seine Sensitivität und Spezifität ausgedrückt. Sensitivität und Spezifi-

tät sind Kenngrößen für den Anteil der mit Hilfe des Tests richtig beurteilten Patienten. Die Sensitivität eines Tests ist definiert als Anteil der Patienten mit einer Krankheit, bei dem der Test die Krankheit erkennt. Die Spezifität ist definiert als der Anteil der Patienten ohne die gesuchte Krankheit, den der Test als gesund einstuft (vgl. Dunn, 2000). Sensitivität und Spezifität sind fundamentale Validitätskennwerte eines Tests. Weitere häufig verwendete Maße sind neben den prädiktiven Werten die Overall Accuracy, der Kappa-Index und die Likelihood Quotienten.

Es kann nicht von „der“ Validität eines Tests gesprochen werden. Die Validität wird immer für eine bestimmte Gruppe von Personen erhoben und hängt von dem verwendeten Außenkriterium ab. Ein Test ist nur für einen bestimmten Zweck und eine bestimmte Population valide.

In der vorliegenden Arbeit wird die Validität von zwei für die Primärversorgung zur Verfügung stehenden Screeninginstrumenten in Relation zu einem standardisierten klinischen Interview erhoben und verglichen.

Gliederung der vorliegenden Arbeit

Im folgenden Kapitel 2 wird die Fragestellung der Arbeit vorgestellt. Kapitel 3 gibt einen Überblick über das Untersuchungsdesign und die Untersuchungsdurchführung sowie die statistischen Zielgrößen, die Validitätsmaße Sensitivität, Spezifität, prädiktive Werte, Overall Accuracy, Kappa Index, Likelihood Quotienten und Receiver Operating Characteristica (ROC) Kurven. Ferner werden die beiden untersuchten Screeninginstrumente, der Brief-Patient Health Questionnaire (B-PHQ, Spitzer, Kroenke, Williams, 1999) und der General Health Questionnaire (GHQ-12, Goldberg, 1972) sowie der Goldstandard, das DIA-X Interview (Wittchen & Pfister, 1997a) beschrieben. Die Ergebnisse der Validierung werden in Kapitel 4 dargestellt und in Kapitel 5 diskutiert. Kapitel 6 bietet eine Zusammenfassung der Arbeit.

2 Fragestellung

Ziel der vorliegenden Arbeit ist die Beurteilung und der Vergleich der Validität von zwei Screeninginstrumenten, validiert anhand eines standardisierten klinisch-diagnostischen Interviews, dem DIA-X (Wittchen & Pfister, 1997a).

Bei den Screeninginstrumenten handelt es sich um den *Brief Patient Health Questionnaire*, im Folgenden mit *B-PHQ* bezeichnet (Spitzer, Kroenke & Williams, 1999), und die 12-Item-Version des *General Health Questionnaire* (Goldberg, 1972; 1978), die im weiteren Text mit *GHQ-12* abgekürzt wird.

Die Validierung der beiden Fragebögen erfolgt an einer rechnergestützten Version des *Composite International Diagnostic Interviews* (*CIDI*, WHO, 1993), dem sogenannten *DIA-X* (Wittchen & Pfister, 1997a). Der CIDI gilt neben dem *SKID* (Strukturiertes Klinisches Interview, Wittchen, Zaudig, Schramm, Spengler, Mombour, Klug & Horn, 1990) als Goldstandard für die Diagnostik in klinischen psychopharmakologischen Wirksamkeitsprüfungen sowie in psychiatrischen epidemiologischen Erhebungen. Eine Auflistung der Depressionsdiagnosen (ICD-10) die das DIA-X standardmäßig ermöglicht, ist in Tabelle 7 (Kapitel 4.1.3) abgebildet. Zusätzlich zu diesen, im Folgenden als „Standarddepressionsdiagnosen des DIA-X“ bezeichneten, Diagnosen wurde die in der vorliegenden Untersuchung verwendete Version des DIA-X dahingehend modifiziert, dass auch die aktuelle Diagnose „unterschwelliges depressives Syndrom“ nach der in Kapitel 1.1 genannten Definition gestellt werden kann. Eine ausführliche Darstellung der beiden Screeninginstrumente sowie des DIA-X Interviews erfolgt in Kapitel 3.

Die Kernfrage der vorliegenden Arbeit lautet: Wie stellt sich die Validität des GHQ-12 im Vergleich zur Validität des B-PHQ validiert anhand der Standarddepressionsdiagnosen des DIA-X hinsichtlich der Erfassung depressiver Syndrome in der primären Gesundheitsversorgung dar?

Die Validität wird dabei anhand der folgenden Validitätskennwerte beurteilt: Sensitivität, Spezifität, prädiktive Werte, Accuracy, Kappa Index, Likelihood Quotient und ROC-Kurven[1].

[1] Eine ausführliche Darstellung der genannten Validitätskennwerte erfolgt in Kapitel 3.

Darüber hinaus werden in der vorliegenden Arbeit die beiden folgenden, rein explorativen Fragestellungen beantwortet:
Wie stellt sich die Validität des GHQ-12 im Vergleich zur Validität des B-PHQ validiert anhand der modifizierten Version[2] des DIA-X hinsichtlich der Erfassung depressiver Syndrome in der primären Gesundheitsversorgung dar?

Die zweite rein explorative Fragestellung lautet: Wie stellt sich die Validität des GHQ-12 im Vergleich zur Validität des B-PHQ für verschiedene Subgruppen der Gesamtstichprobe dar?
Bei der Auswahl der Subgruppen wurde der Schwerpunkt auf zwei wichtige demografische Variablen gelegt: Geschlecht und Alter. Die Berechnung der Validitätskennwerte erfolgt sowohl anhand der Standarddepressionsdiagnosen des DIA-X als auch anhand der modifizierten Version. Es ergeben sich daher die folgenden sechs Fragestellungen:

1. Wie stellt sich die Validität des GHQ-12 verglichen mit der Validität des B-PHQ validiert anhand der Standarddepressionsdiagnosen des DIA-X dar?
   Aus der Operationalisierung dieser Fragestellung ergeben sich vier nachgeordnete methodische Fragestellungen:
   1.1 Wie hoch sind die Sensitivität und Spezifität für verschiedene Cut-Off-Werte der beiden Fragebögen?
   1.2 Wie hoch sind die prädiktiven Werte für verschiedene Cut-Off-Werte?
   1.3 Wie hoch sind die Accuracy, der Kappa Index und die Likelihood Quotienten für verschiedene Cut-Off-Werte?
   1.4 Wie stellen sich die Receiver Operating Characteristica (ROC) Kurven für den GHQ-12 und den B-PHQ dar?

[2] Da diese modifizierte Version des DIA-X selbst noch der offiziellen Validierung bedarf, ist die Berechnung der Validierungsmaße rein explorativ.

2. Wie stellt sich die Validität des GHQ-12 verglichen mit der Validität des B-PHQ validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Berücksichtigung der Diagnose unterschwelliges depressives Syndrom dar?

   Aus der Operationalisierung dieser Fragestellung ergeben sich folgende vier methodische Fragestellungen:

   2.1 Wie hoch sind die Sensitivität und Spezifität für verschiedene Cut-Off-Werte der beiden Fragebögen?

   2.2 Wie hoch sind die prädiktiven Werte für verschiedene Cut-Off-Werte?

   2.3 Wie hoch sind die Accuracy, der Kappa Index und die Likelihood Quotienten für verschiedene Cut-Off-Werte?

   2.4 Wie stellen sich die Receiver Operating Characteristica (ROC) Kurven für den GHQ-12 und den B-PHQ dar?

3. Wie stellt sich die Validität des GHQ-12 verglichen mit der Validität des B-PHQ validiert anhand der Standarddepressionsdiagnosen des DIA-X in der Subgruppe der Frauen sowie in der Subgruppe der Männer dar?

   Aus der Operationalisierung dieser Fragestellung ergeben sich wiederum vier nachgeordnete methodische Fragestellungen:

   3.1 Wie hoch sind die Sensitivität und Spezifität für die verschiedenen Cut-Off-Werte in der Untergruppe der Frauen bzw. der Untergruppe der Männer?

   3.2 Wie hoch sind die prädiktiven Werte für die verschiedenen Cut-Off-Werte in der Subgruppe der Frauen bzw. der Subgruppe der Männer?

   3.3 Wie hoch sind die Accuracy, der Kappa Index und die Likelihood Quotienten in der Subgruppe der Frauen bzw. der Subgruppe der Männer für die verschiedenen Cut-Off-Werte?

   3.4 Wie stellen sich die Receiver Operating Characteristica (ROC) Kurven für den GHQ-12 und den B-PHQ in der Subgruppe der Frauen bzw. der Subgruppe der Männer dar?

4. Wie stellt sich die Validität des GHQ-12 verglichen mit der Validität des B-PHQ validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Berücksichtigung der Diagnose unterschwelliges depressives Syndrom in der Subgruppe der Frauen sowie in der Subgruppe der Männer dar?

   Aus der Operationalisierung dieser Fragestellung ergeben sich folgende vier methodische Fragestellungen:

   4.1 Wie hoch sind die Sensitivität und Spezifität für die verschiedenen Cut-Off-Werte in der Untergruppe der Frauen bzw. der Untergruppe der Männer?

   4.2 Wie hoch sind die prädiktiven Werte für die verschiedenen Cut-Off-Werte in der Subgruppe der Frauen bzw. der Subgruppe der Männer?

   4.3 Wie hoch sind die Accuracy, der Kappa Index und die Likelihood Quotienten in der Subgruppe der Frauen bzw. der Subgruppe der Männer für die verschiedenen Cut-Off-Werte?

   4.4 Wie stellen sich die Receiver Operating Characteristica (ROC) Kurven für den GHQ-12 und den B-PHQ in der Subgruppe der Frauen bzw. der Subgruppe der Männer dar?

5. Wie stellt sich die Validität des GHQ-12 verglichen mit der Validität des B-PHQ validiert anhand der Standarddepressionsdiagnosen des DIA-X für verschiedene Altersgruppen dar?

   Aus der Operationalisierung dieser Fragestellung ergeben sich die folgenden nachgeordneten methodische Fragestellungen:

   5.1 Wie hoch sind die Sensitivität und Spezifität für die verschiedenen Cut-Off-Werte für verschiedenen Altersgruppen?

   5.2 Wie hoch sind die prädiktiven Werte für die verschiedenen Cut-Off-Werte für verschiedenen Altersgruppen?

   5.3 Wie hoch sind die Accuracy, der Kappa Index und die Likelihood Quotienten für verschiedenen Altersgruppen für die verschiedenen Cut-Off-Werte?

5.4 Wie stellen sich die Receiver Operating Characteristica (ROC) Kurven für den GHQ-12 und den B-PHQ für verschiedenen Altersgruppen dar?

6. Wie stellt sich die Validität des GHQ-12 verglichen mit der Validität des B-PHQ validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Berücksichtigung der Diagnose unterschwelliges depressives Syndrom für verschiedene Altersgruppen dar?

   Aus der Operationalisierung dieser Fragestellung ergeben sich folgende vier methodische Fragestellungen:

   6.1 Wie hoch sind die Sensitivität und Spezifität für die verschiedenen Cut-Off-Werte für verschiedenen Altersgruppen?

   6.2 Wie hoch sind die prädiktiven Werte für die verschiedenen Cut-Off-Werte für verschiedenen Altersgruppen?

   6.3 Wie hoch sind die Accuracy, der Kappa Index und die Likelihood Quotienten für verschiedenen Altersgruppen für die verschiedenen Cut-Off-Werte?

   6.4 Wie stellen sich die Receiver Operating Characteristica (ROC) Kurven für den GHQ-12 und den B-PHQ für verschiedenen Altersgruppen dar?

Als Hypothese wird formuliert, dass beide Fragebögen ähnlich valide Instrumente zur Erfassung depressiver Störungen in der primären Gesundheitsversorgung sind, da der B-PHQ speziell für die Beurteilung psychischer Erkrankungen in der Allgemeinarztpraxis entwickelt wurde und der GHQ-12 sich bereits in zahlreichen Forschungsprojekten als Screeninginstrument bewährt hat (vgl. Goldberg, Gater, Sartorius, Ustun, Piccinelli, Gureje & Rutter, 1997).

3 Methoden

Die Validierung des General Health Questionnaire (GHQ-12) und des Brief Patient Health Questionnaire (B-PHQ) erfolgt an einem standardisierten klinisch-diagnostischen Interview, dem DIA-X. Der General Health Questionnaire wurde 1972 von Goldberg zur Erfassung psychischer Beeinträchtigung in der Allgemeinbevölkerung entwickelt und wurde seitdem in zahlreichen Studien eingesetzt. Mit dem B-PHQ liegt eine reine Selbstbeurteilungsversion des *Prime-MD* (*Primary Care Evaluation of Mental Disorders;* Spitzer, Williams, Kroenke, Linzer, deGruy, Hahn, Brody & Johnson, 1994) zur Diagnose von affektiven Störungen und Panikstörung anhand von DSM-IV Kriterien in der Primärversorgung vor.

In diesem Kapitel wird zunächst das Untersuchungsdesign und die Untersuchungsdurchführung dargestellt. Im zweiten Abschnitt erfolgt ein ausführlicher Überblick über die statistischen Zielgrößen: Die Validitätsmaße Sensitivität, Spezifität, prädiktive Werte, Overall Accuracy, Kappa Index, Likelihood Quotienten und Receiver Operating Characteristica (ROC) Kurven.

Abschließend werden die beiden Screeninginstrumente GHQ-12 und B-PHQ sowie der Goldstandard (DIA-X) eingehend beschrieben.

3.1 Untersuchungsdesign und -durchführung

Zur Validierung der beiden Screeninginstrumente GHQ-12 und B-PHQ am DIA-X Interview ist es erforderlich, dass die an der Untersuchung teilnehmenden Patienten in den beteiligten Hausarztpraxen die beiden Screeninginstrumente ausfüllen und am klinisch-diagnostischen Interview teilnehmen.

In der vorliegenden Arbeit wurde die Datenerhebung in Kooperation mit 18 hausärztlich tätigen Arztpraxen aus Nürnberg durchgeführt. Es handelt sich hierbei um Allgemeinärzte und Internisten aus dem sogenannten Praxisnetz

Nürnberg Nord und dem Praxisnetz Nürnberg Süd[3]. Um eine repräsentative Patientenstichprobe zu erhalten, wurden die kooperierenden Arztpraxen aufgefordert, an bestimmten, von ihnen selbst ausgewählten Stichtagen alle Patienten – unabhängig vom Anlass ihres Arztbesuches – zu bitten, die beiden Fragebögen auszufüllen und an einem diagnostischen Interview teilzunehmen. Dies konnte jedoch aus Gründen der Praktikabilität nicht immer von allen Arztpraxen eingehalten werden.

Das Ausfüllen der Screeningbögen fand in den Arztpraxen statt. Das DIA-X Interview als Validierungsinstrument wurde zu einem vom Patienten gewünschten Termin innerhalb der folgenden sechs Tage telefonisch von geschulten Interviewerinnen (zwei Ärztinnen und zwei Diplom-Psychologinnen) durchgeführt. Die Interviewerschulung erfolgte durch die Entwickler des DIA-X am Max-Planck-Institut für Psychiatrie in München.

Neben dem GHQ-12 und dem B-PHQ wurde im Rahmen dieser Studie noch ein weiteres Screeninginstrument, der sogenannte *WHO-5 Well Being Index* (WHO-5, WHO, 1998) durchgeführt. Die Validierung des WHO-5 ist Gegenstand der noch unveröffentlichten Dissertation von Frau Petra Ohlendorf. Um Positionseffekte zu vermeiden, wurde die Reihenfolge der drei Fragebögen von WHO-5, GHQ-12 und B-PHQ nach ungefähr der Hälfte der Studienzeit gewechselt. Die Reihenfolge war dann B-PHQ, GHQ-12 und WHO-5. Die Datenerhebung fand im Zeitraum von 10. Juli 2000 bis 28. Juli 2001 statt.

Die Fragebogen- und Interviewdaten wurden unabhängig und zudem verblindet erhoben. So waren den Interviewerinnen die Ergebnisse der Patienten in den Fragebögen nicht bekannt. Umgekehrt hatten die Mitarbeiterinnen in den Arztpraxen keinen Einblick in die anhand des Telefoninterviews gewonnenen Diagnosen der Patienten.

Zu Beginn der Untersuchung wurden zunächst umfangreiche Informationsmaterialien für die beteiligten Arztpraxen erstellt. Zu diesen Materialien gehörten beispielsweise ein Poster, das die Patienten im Wartezimmer über die Studie informierte, sowie Merkblätter für die Praxismitarbeiterinnen zur Studiendurch-

[3] Das Praxisnetz Nürnberg Süd ist ein Zusammenschluss von Ärzten aller Fachrichtungen aus dem Bereich Nürnberg Süd und Fürth. Das Praxisnetz Nürnberg Nord ist ein Zusammenschluss von Ärzten aus dem Norden Nürnbergs.

führung. Für die Patienten stand eine ausführliche Patienteninformation mit Einverständniserklärung zur Verfügung. Das DIA-X Interview wurde nur dann durchgeführt, wenn der Proband die Patienteninformation erhalten und schriftlich sein Einverständnis gegeben hatte. In Anhang A ist exemplarisch das Poster abgebildet.

Nach der Rekrutierungs- und Initiierungsphase der Arztpraxen erfolgte die Etablierung der organisatorischen Abläufe unter besonderer Berücksichtigung der Praktikabilität für die beteiligten Arztpraxen und einer Minimierung der Belastung für die Patienten. Die kooperierenden Arztpraxen wurden während der gesamten Studiendauer von Studienmitarbeitern betreut. Durch regelmäßige Treffen wurde ein kontinuierlicher Austausch zwischen den Arztpraxen und den Studienmitarbeitern gewährleistet.

Der Untersuchungsplan wurde der Ethik-Kommission der Medizinischen Fakultät der Ludwig-Maximilians-Universität München zur Begutachtung vorgelegt und von dieser genehmigt.

3.2 Statistische Zielgrößen

3.2.1 Sensitivität und Spezifität

Die Sensitivität und die Spezifität sind bedingte Wahrscheinlichkeiten, mit deren Hilfe die Brauchbarkeit eines diagnostischen Tests zur Erkennung einer bestimmten Krankheit beschrieben werden kann (vgl. Sachs, 1999). Die Sensitivität ist die Wahrscheinlichkeit, mit der ein Patient vom Test als krank (positiv) erkannt wird unter der Bedingung, dass er tatsächlich krank ist. Umgekehrt ist die Spezifität definiert als die Wahrscheinlichkeit, dass ein Patient vom Test als gesund (negativ) eingestuft wird unter der Bedingung, dass er tatsächlich auch gesund ist. Beide Validitätskennwerte sollten bei der Entscheidung, ob der Test durchgeführt werden sollte oder nicht, berücksichtigt werden, da es in der Regel wichtig ist, das Risiko beider Fehlklassifikationen zu minimieren.

Ziel ist es also allgemein, das Auftreten von Fehlern bei einer Klassenzuordnung zu verhindern. Ein Fehler liegt dann vor, wenn eine Diskrepanz zwischen der Zuordnung aufgrund eines Tests und der tatsächlichen Gruppenzugehörigkeit, die mit Hilfe des Goldstandards bestimmt wurde, besteht. Basis für die Berechnung der Validität ist die Anzahl der vom Test im Verhältnis zum Goldstandard richtig zugeordneten Personen. Im Fall von zwei Klassen ergeben sich somit zwei Möglichkeiten der Übereinstimmung und zwei Möglichkeiten der Nichtübereinstimmung. Die Beziehung zwischen den Ergebnissen eines Tests und der wahren Diagnose können in sogenannten Vier-Felder-Tafeln dargestellt werden. Mit Hilfe von Vier-Felder-Tafeln kann das Risiko von Fehlklassifikationen veranschaulicht werden. In Abbildung 1 sind die vier Arten richtiger und falscher Klassifikationen in Form einer Vier-Felder-Tafel dargestellt. Die Zeilen beziehen sich auf das Testergebnis, die Spalten auf den Goldstandard.

		Goldstandard		
		Krankheit vorhanden (positiv)	Krankheit nicht vorhanden (negativ)	
Testergebnis	Krankheit vorhanden (positiv)	**a** richtig Positive	**b** falsch Positive Fehler 1. Art	**a + b** Testpositive
	Krankheit nicht vorhanden (negativ)	**c** falsch Negative Fehler 2. Art	**d** richtig Negative	**c + d** Testnegative
		a + c	**b + d**	**N = a + b +c + d**

Abbildung 1: Arten richtiger und falscher Klassifikationen in Form einer Vier-Felder-Tafel in Anlehnung an Sackett, Haynes, Guyatt und Tugwell (1991).

Wird im Folgenden von a, b, c oder d gesprochen, so ist damit die entsprechende Zelle in der Vier-Felder-Tafel gemeint, soweit nicht ausdrücklich auf eine andere Bedeutung hingewiesen wird. In die Zelle a werden die sogenannten richtig positiven Zuordnungen eingetragen. Unter richtig Positiven (RP) werden die vom Test richtig diagnostizierten Kranken verstanden. In die Zelle b werden die falsch Positiven (FP) geschrieben. Hierunter sind die fälschlicherweise als krank bezeichneten Gesunden zu verstehen. Als falsch Negative (FN) werden fälschlicherweise als gesund klassifizierte Kranke bezeichnet. Diese werden in die Zelle c eingetragen. In der Zelle d befinden sich die richtig Negativen (RN). Hierbei handelt es sich um die vom Test richtigerweise als gesund Identifizierten.

Es sind somit zwei Arten von Zuordnungsfehlern möglich. Unter dem Fehler erster Art werden falsch positive Zuordnungen des Tests verstanden, d.h. Personen werden als krank bezeichnet, obwohl sie eigentlich gesund sind. Bei einem Fehler zweiter Art werden falsch negative Zuordnungen vorgenommen, d.h. Personen werden als gesund diagnostiziert, obwohl sie tatsächlich der Klasse der Kranken angehören. Welche Bedeutung den beiden Fehlentscheidungen zugemessen wird, hängt jeweils von der diagnostischen Fragestellung ab (vgl. Amelang & Zielinski, 1994; Noack & Petermann, 1992).

Sensitivität und Spezifität sind Gütekriterien, die zur Beschreibung eines Tests berechnet werden können. Unter Sensitivität wird die Wahrscheinlichkeit verstanden, mit der ein Kranker von einem Test als krank diagnostiziert wird oder anders ausgedrückt, der Anteil der Patienten mit einer bestimmten Krankheit, die ein positives Testergebnis haben. Die Sensitivität wird auch als Empfindlichkeit eines Tests bezeichnet. Synonyme Begriffe, die sich vor allem in englischsprachigen Publikationen finden, sind nach Sackett, Haynes, Guyatt und Tugwell (1991) zum einen die sogenannte *Positive in Disease Rate (PiD rate)* und die *True positive Rate (TP rate).* Durch die Anwendung eines Tests mit einer hohen Sensitivität wird kaum eine kranke Person übersehen, d.h. es gibt nur wenige falsch negative Klassifikationen.

Unter Spezifität wird die Wahrscheinlichkeit verstanden, mit der ein Gesunder von einem Test als gesund erkannt wird oder anders ausgedrückt, der Anteil der Gesunden, die in dem Test ein negatives Ergebnis haben. Als synonyme

Begriffe für die Spezifität finden sich nach Sackett et al. (1991) *Negative in Health Rate (NiH rate)* und *die True Negative Rate (TN rate)* in der Literatur. Der Ausdruck „negative in health“ bedeutet, dass der Patient zwar die gesuchte Krankheit nicht hat, aber nicht notwendigerweise gesund ist. Ein Test mit einer hohen Spezifität wird Personen ohne die gesuchte Krankheit kaum als erkrankt fehlklassifizieren.

Ein Test mit einer hohen Sensitivität wird auch als sensitiver Test bezeichnet. Analog meint ein spezifischer Test, einen Test mit einer hohen Spezifität. Die Sensitivität und die Spezifität eines Tests werden für die Population der zu diagnostizierenden Personen über die relativen Häufigkeiten in einer untersuchten Stichprobe geschätzt (Noack & Petermann, 1992). In Abbildung 2 sind die mathematischen Formeln für die Berechnung der Sensitivität und der Spezifität dargestellt (vgl. Baldessarini, Finklestein, & Arana, 1983).

$$\text{Sensitivität} = \frac{a}{a+c}$$

$$\text{Spezifität} = \frac{d}{b+d}$$

Abbildung 2: Mathematische Berechnung von Sensitivität und Spezifität (vgl. Baldessarini, Finklestein, & Arana, 1983).

Bei der Auswahl eines diagnostischen Tests sollten die Sensitivität und die Spezifität in Abhängigkeit vom Ziel der diagnostischen Untersuchung berücksichtigt werden. Ein sensitiver Test sollte gewählt werden, wenn das Ziel darin besteht, falsch negative Klassifikationen zu vermeiden. Dies ist dann der Fall, wenn das Übersehen einer Krankheit zu schwerwiegenden Nachteilen für den Patienten führt, beispielsweise wenn eine schwere, aber behandelbare Krankheit wie Tuberkulose übersehen wird. Einen großen Nutzen bringen sensitive Tests auch zu Beginn eines diagnostischen Prozesses im Rahmen der Einzelfalldiagnostik. Bei einem negativen Testergebnis können anfänglich in Betracht

gezogene Diagnosen ausgeschlossen werden. Wenn das Testergebnis bei einem sensitiven Test negativ ist, dann ist der Patient auch mit hoher Wahrscheinlichkeit gesund, da bei einem sensitiven Test nur wenige falsch negative Zuordnungen erfolgen. Sensitive Tests können dazu verwendet werden, eine Diagnose auszuschließen. Im Rahmen von Screeningmaßnahmen werden sensitive Tests gewählt, um möglichst alle Personen mit der gesuchten Krankheit zu erfassen, da ein sensitiver Test nur sehr wenige Kranke übersehen wird. Ein spezifischer Test ist die Methode der Wahl, wenn ein falsch positives Testergebnis für den Patienten einen hohen materiellen oder immateriellen Schaden verursachen kann, z.B. durch eine fälschlicherweise gestellte Krebsdiagnose. Ist das Testergebnis eines hoch spezifischen Tests positiv, so kann die Diagnose mit hoher Wahrscheinlichkeit als bestätigt gelten, da ein spezifischer Test nur wenige falsch positive Zuordnungen vornimmt, d.h. bei Nichtvorliegen der Krankheit nur selten positiv ist.

In der diagnostischen Praxis ist es wünschenswert, Tests mit möglichst hoher Sensitivität und hoher Spezifität zu verwenden. Dies ist allerdings nur bedingt realisierbar, da die Sensitivität und die Spezifität wechselseitig voneinander abhängen, d.h. die eine Kenngröße läßt sich nur auf Kosten der anderen erhöhen. Zwar ist es möglich, eine Sensitivität von 100% zu erreichen, dafür sinkt aber die Spezifität drastisch ab und vice versa.

Eine Möglichkeit, die Beziehung zwischen der Sensitivität und der Spezifität eines Tests für eine bestimmte Stichprobe grafisch auszudrücken, ist die Erstellung einer ROC-Kurve. Das Akronym ROC steht für Receiver Operating Characteristica. Eine ausführliche Darstellung dieser Methode erfolgt unter Gliederungspunkt 3.2.4 dieses Kapitels.

Die Sensitivität und die Spezifität eines Tests hängen vom sogenannten Cut-Off-Wert ab und ändern sich mit diesem. Der Cut-Off-Wert wird auch als Cut-Off, Cutting-Score oder Trennpunkt bezeichnet (vgl. Dorsch, 1994). Es handelt sich dabei um den kritischen Wert in einem diagnostischen Test. Der Cut-Off wird in der Regel in Form eines Wertepaares $x / x+1$ angegeben. Personen mit einem Testwert oberhalb des Cut-Offs, d.h. mit einem Wert $\geq x+1$, erhalten eine andere diagnostische Zuweisung als Personen mit Testwerten unterhalb dieses Trennwertes ($\leq x$).

Ferner werden die Sensitivität und die Spezifität von vier weiteren Faktoren beeinflusst:

1. dem Spektrum der Patienten, an denen der Test validiert wird,
2. dem Validierungsstandard,
3. möglichen systematischen Fehlern (Bias),
4. eventuellen Zufallsschwankungen.

Das Spektrum der Patienten, an denen der Test valdiert wird, sollte für das Spektrum der Patienten, an denen der Test angewendet werden soll, repräsentativ sein, denn

> (...) theoretisch sollen Sensitivität und Spezifität eines Tests unabhängig von der Prävalenz der erkrankten Personen in der Stichprobe, in welcher der Test beurteilt wird, sein. In der Praxis aber können verschiedene Charakteristika der Patienten, z.B. Stadium und Schwere der Krankheit sowohl mit der Sensitivität und Spezifität eines Tests als auch mit der Prävalenz korrelieren, weil unterschiedliche Patienten in Situationen mit hoher und niedriger Prävalenz zu finden sind. (Fletcher, Fletcher & Wagner, 1999, S. 74)

Die Sensitivität und die Spezifität hängen immer vom gewählten Validierungsstandard ab, da sie den Grad der Übereinstimmung zwischen dem Test und dem Goldstandard widerspiegeln. Wenn ein ungeeigneter Validitätsstandard gewählt wurde, werden auch die Sensitivität und die Spezifität des Tests ungenau oder sogar falsch beschrieben.

Unter Bias versteht man systematische Fehler. Diese entstehen, wenn die Sensitivität und die Spezifität nicht unabhängig von den Methoden erhoben werden, mit denen auch der Validitätsstandard erhoben wird. Systematische Fehler können zu einer Verfälschung der Übereinstimmung zwischen den Ergebnissen des Tests und dem Goldstandard führen. Der Test erscheint dann valider als er tatsächlich ist. Selbst wenn keine systematischen Fehler vorliegen, kann es besonders bei kleinen Stichproben infolge von Zufallsschwankungen zu einer Fehleinschätzung der Sensitivität und Spezifität des Tests kommen. Daher kann es von Vorteil sein, das 95%-Konfidenzintervall für die Sensi-

tivität und die Spezifität anzugeben. Das 95%-Konfidenzintervall einer relativen Häufigkeit kann mit der nachfolgenden Formel, die auf der Binominalverteilung beruht, geschätzt werden.

$$p \pm 2\sqrt{\frac{p \times (1-p)}{n}}$$

p = die beobachtete relative Häufigkeit

n = die Anzahl der untersuchten Personen

Abbildung 3: Formel zur Schätzung des 95%-Konfidenzintervalls einer relativen Häufigkeit (Fletcher et al., 1999).

Die Sensitivität und die Spezifität einer diagnostischen Entscheidung lassen sich durch die Anwendung mehrerer diagnostischer Tests erhöhen. Wird für das Stellen einer positiven Diagnose verlangt, dass alle Tests positiv sind, so wird die Spezifität maximiert. Gleichzeitig wird die Anzahl der falsch positiven Klassifikationen minimiert. Jedoch muss dafür eine Zunahme der falsch negativen Zuordnungen in Kauf genommen werden. Fordert man umgekehrt für das Stellen einer negativen Diagnose, dass alle Testergebnisse negativ sind, so wird die Sensitivität maximiert und die Anzahl der falsch negativen Klassifikationen minimiert. Gleichzeitig steigt aber die Anzahl der falsch positiven Zuordnungen (vgl. Fletcher et al., 1999).

Nach Sackett et al. (1991) lassen sich verschiedene Regeln formulieren, die das Verständnis dieser beiden Validitätsmaßen erleichtern. So gilt, wenn die Rate der richtig positiven Zuordnungen größer bzw. kleiner wird, dann wird auch die Rate der falsch positiven Klassifikationen größer bzw. kleiner.

Wenn ein Test eine genügend hohe Sensitivität hat, dann schließt ein negatives Testergebnis das Vorhandensein der Zielkrankheit mit hoher Sicherheit aus. Umgekehrt gilt, wenn ein Test eine genügend hohe Spezifität hat, dann zeigt ein positives Testergebnis mit einer hohen Wahrscheinlichkeit das Vorliegen der Krankheit an. Je größer die Sensitivität und die Spezifität eines Tests sind, desto nützlicher ist der Test. Ein Test ist ungeeignet, wenn die Sensitivität

und die Spezifität kleiner oder gleich 50% sind, da diese Werte bestenfalls einer Zufallszuordnung entsprechen.
Sensitivität und Spezifität sind Eigenschaften des Tests, die bei der Entscheidung berücksichtigt werden sollten, ob der Test durchgeführt wird oder nicht bzw. welcher Cut-Off-Wert verwendet werden sollte. Nach der Testdurchführung sind für den Testanwender vor allem der positive und der negative prädiktive Wert interessant. Diese werden im folgenden Gliederungspunkt 3.2.2 ausführlich dargestellt.

3.2.2 Prädiktive Werte

Unter einem prädiktiven Wert versteht man die Wahrscheinlichkeit, dass die gesuchte Krankheit vorliegt bzw. nicht vorliegt, wenn das Testergebnis bekannt ist. Für einen Test können zwei prädiktive Werte, der positive prädiktive Wert und der negative prädiktive Wert bestimmt werden. Der positive prädiktive Wert wird in englischsprachigen Publikationen häufig mit *PPV (positive predictive value)* abgekürzt (vgl. Sackett et al., 1991). Als Synonym wird auch der Ausdruck „positiver Vorhersagewert“ (vgl. Noack & Petermann, 1992) oder die Formulierung „prädiktiver Wert einer positiven Zuordnung“ (Amelang & Zielinski, 1994) verwendet. Der positive prädiktive Wert ist definiert als die Wahrscheinlichkeit des Vorhandenseins der Krankheit bei einem Patienten mit einem positiven Testergebnis. Es handelt sich somit um die Wahrscheinlichkeit, mit der eine positive Testdiagnose tatsächlich zutreffend ist.
Der negative prädiktive Wert wird analog in englischsprachigen Publikationen mit *NPV (negative predictive value)* abgekürzt (Sackett et al., 1991). Als Synonym wird die Formulierung „negativer Vorhersagewert“ (vgl. Noack & Petermann, 1992) oder der Ausdruck „prädiktiver Wert einer negativen Zuordnung“ (Amelang & Zielinski, 1994) verwendet. Der negative prädiktive Wert ist definiert als die Wahrscheinlichkeit, die gesuchte Krankheit nicht zu haben, falls das Testergebnis negativ ist. Es handelt sich somit um die Wahrscheinlichkeit, mit der eine negative Testdiagnose richtig ist. Die Berechung der prädiktiven

Werte ist in Abbildung 4 dargestellt (vgl. Baldessarini, Finklestein, & Arana, 1983).

$$\text{positiver prädiktiver Wert} = \frac{a}{a+b} = \frac{\textit{richtig Positive}}{\textit{Testpositive}}$$

$$\text{negativer prädiktiver Wert} = \frac{d}{c+d} = \frac{\textit{richtig Negative}}{\textit{Testnegative}}$$

Abbildung 4: Berechnung der prädiktiven Werte (vgl. Baldessarini, Finklestein & Arana, 1983).

Die prädiktiven Werte geben dem Kliniker nach einer Testdurchführung Antwort auf folgende Frage: Wie hoch ist die Wahrscheinlichkeit bei Vorliegen eines positiven bzw. negativen Testergebnisses, dass der Patient die Erkrankung tatsächlich hat, respektive nicht hat?

Die prädiktiven Werte werden auch als A-posteriori- oder als Posttest-Wahrscheinlichkeit bezeichnet. Es handelt sich somit beim positiven Vorhersagewert um die Wahrscheinlichkeit der gesuchten Krankheit nach dem Bekanntsein eines positiven Testergebnisses oder, anders ausgedrückt, um den Anteil der richtig diagnostizierten Kranken an allen vom Test als krank diagnostizierten Personen. Umgekehrt handelt es sich beim negativen Vorhersagewert um die Wahrscheinlichkeit, dass der Patient gesund ist nach Bekanntwerden eines negativen Testergebnisses, oder anders formuliert, um den Anteil der vom Test richtig erkannten Gesunden an der Gesamtzahl der vom Test als gesund eingestuften Personen (vgl. Amelang & Zielinski, 1994; Noack & Petermann, 1992).

Unter A-priori- oder Pretest-Wahrscheinlichkeit wird die Prävalenz verstanden. Die Prävalenz ist definiert als der Anteil der Personen mit der gesuchten Krankheit, bezogen auf einen bestimmten Zeitpunkt in einer definierten Population oder Stichprobe.

Die prädiktiven Werte eines Tests werden von der Sensitivität und der Spezifität sowie von der Prävalenz der Krankheit in der untersuchten Stichprobe /Population bestimmt. Nach Sackett et al. (1999) gilt folgende Gesetzmäßigkeit: Wenn sich die Prävalenz ändert, können sich Sensitivität und Spezifität eines Tests ändern, die prädiktiven Werte hingegen müssen sich ändern.
Aus dem Satz von Bayes lässt sich die folgende mathematische Formel ableiten, die Sensitivität, Spezifität und Prävalenz zum positiven prädiktiven Wert in Beziehung setzt:

$$\text{Prävalenz} = \frac{a+c}{N}$$

$$\text{positiver prädiktiver Wert} = \frac{Sensitivität \times Prävalenz}{(Sensitivität \times Prävalenz) + ((1 - Spezifität) \times (1 - Prävalenz))}$$

Abbildung 5: Formel zur Berechnung der Prävalenz und des positiven prädiktiven Wertes (Fletcher et al., 1999).

Nach Sackett et al. (1991) gelten folgende Gesetzmäßigkeiten im Zusammenhang mit den prädiktiven Werten: Je sensitiver ein Test ist, desto besser ist der negative Prädiktive Wert, d.h. ein negatives Testergebnis schließt die Krankheit mit hoher Wahrscheinlichkeit aus. Umgekehrt gilt, je spezifischer ein Test ist, desto besser sein positiver prädiktiver Wert, d.h. ein positives Testergebnis bestätigt eine positive Diagnose mit einer Wahrscheinlichkeit in Höhe des positiven prädiktiven Wertes.
Der positive prädiktive Wert geht gegen Null, wenn die Prävalenz der Krankheit in der untersuchten Population gegen Null geht. Nähert sich die Prävalenz der gesuchten Krankheit allerdings 100 Prozent, so nähert sich der negative prädiktive Wert Null.
Wenn die Sensitivität und die Spezifität eines Tests gleich bleiben und die Prävalenz abnimmt, dann sinkt auch der positive prädiktive Wert und der negative prädiktive Wert nimmt zu. Das bedeutet somit Folgendes: Wenn die Pre-

testwahrscheinlichkeit einer bestimmten Krankheit abnimmt, dann sinkt auch die Posttestwahrscheinlichkeit der Krankheit.
Der positive prädiktive Wert kann auch als Prävalenz der Krankheit unter den Patienten mit einem positiven Testergebnis interpretiert werden.
Da sich die prädiktiven Werte mit der Prävalenz ändern, ergeben sich für verschiedene Untersuchungssituationen in Abhängigkeit von der Prävalenz unterschiedlich hohe prädiktive Werte. Allgemein gilt, dass bei einer sehr geringen Prävalenz auch bei Anwendung eines sehr spezifischen Tests positive Zuordnungen größtenteils falsch positive Klassifikationen sind. Umgekehrt gilt, dass bei einer sehr hohen Prävalenz die negativen Zuordnungen selbst bei einem sehr sensitiven Test größtenteils falsch negative Klassifikationen sind. Allgemein gilt, dass der Einsatz diagnostischer Tests einen besonders hohen Nutzen bringt, wenn die Wahrscheinlichkeit der Krankheit in der untersuchten Population weder sehr hoch noch sehr niedrig ist.
Grundsätzlich gilt, je geringer die Werte der Sensitivität und der Spezifität eines bestimmten Tests sind, desto größer ist der Einfluss der Prävalenz der gesuchten Erkrankung auf die Höhe der prädiktiven Werte. Auch hat die Prävalenz insofern einen größeren Einfluss auf die Höhe der prädiktiven Werte, als sie in Abhängigkeit von der untersuchten Population zwischen Bruchteilen von einem Prozent und 100 Prozent schwanken kann, während sich die Sensitivität und die Spezifität bei in der Praxis angewandten Tests meist nur zwischen 50 und 100 Prozent bewegt.
Zur Berechnung der prädiktiven Werte kann die Prävalenz, sofern keine eigene Validierungsstudie durchgeführt wird, auch aus der medizinischen Fachliteratur oder lokalen medizinischen Datenbanken geschätzt werden.
Ein Test ist unbrauchbar, wenn positiver und negativer prädiktiver Wert nahe Null sind.

3.2.3 Weitere Kenngrößen der Validität: Overall Accuracy, Kappa Index und Likelihood Quotient

Neben der Sensitivität, der Spezifität und den prädiktiven Werten existiert noch eine Reihe weiterer Kennwerte, die zur Beschreibung der Validität eines Tests berechnet werden können.

Gemeinsamer Ausgangspunkt für die Berechnung dieser Kennwerte ist die Vier-Felder-Tafel. Zur Veranschaulichung der Formeln in Tabelle 1 ist in Abbildung 6 nochmals das Prinzip der Vier-Felder-Tafel dargestellt. Voraussetzung für die Berechnung der oben angegebenen Kennziffern ist, dass es sich bei den angegebenen Werten in der Vier-Felder-Tafel um absolute Werte und nicht um Prozentangaben handelt.

		Goldstandard		
		Krankheit vorhanden (positiv)	Krankheit nicht vorhanden (negativ)	
Test-ergebnis	Krankheit vorhanden (positiv)	**a** richtig Positive	**b** falsch Positive	**a + b** Testpositive
	Krankheit nicht vorhanden (negativ)	**c** falsch Negative	**d** richtig Negative	**c + d** Testnegative
		a + c	**b + d**	**N = a + b +c + d**

Abbildung 6: Vier-Felder-Tafel (vgl. Sackett, Haynes, Guyatt und Tugwell, 1991).

Weitere wichtige Kenngrößen der Validität, ihre mathematischen Berechnungen und Beziehungen sind in Tabelle 1 abgebildet.

Tabelle 1: Kennwerte zur Beschreibung der Validität eines Tests.

Maß	**Berechnung**	**andere Bezeichnungen**	**Herleitung**
Sensitivität	$\frac{a}{a+c}$	Rate richtig Positiver	
Spezifität	$\frac{d}{b+d}$	Rate richtig Negativer	
positiver prädiktiver Wert	$\frac{a}{a+b}$		
negativer prädiktiver Wert	$\frac{d}{c+d}$		
Rate der falsch Positiven	$\frac{b}{b+d}$	Falsch Positiv Rate	$= 1 - Spezifität$
Rate der falsch Negativen	$\frac{c}{a+c}$	Falsch Negativ Rate	$= 1 - Sensitivität$
(Overall) Accuracy	$\frac{a+d}{N}$	Correct classification rate; Test efficiency	
Kappa (Index)	$\frac{2(ad-bc)}{((a+c)(c+d))+((a+b)(b+d))}$		
Likelihood Quotient	$\frac{Wahrscheinlichkeit\ eines\ Testergebnisses\ bei\ Kranken}{Wahrscheinlichkeit\ des\ Testergebnisses\ bei\ Gesunden}$		
Likelihood Quotient eines positiven Testergebnisses	$\frac{\frac{a}{a+c}}{\frac{b}{b+d}}$		$= \frac{Sensitivität}{1 - Spezifität}$
Likelihood Quotient eines negativen Testergebnisses	$\frac{\frac{c}{a+c}}{\frac{d}{b+d}}$		$= \frac{1 - Sensitivität}{Spezifität}$
Misclassification rate	$\frac{b+c}{N}$		$= 1 - Accuracy$
Odds-Ratio	$\frac{a \times d}{c \times b}$		$\frac{Sensitivität}{1 - Sensitivität} \times \frac{Spezifität}{1 - Spezifität}$

In Publikationen häufig angegebene Maße sind neben der Sensitivität, der Spezifität und den prädiktiven Werten, der Kappa Index und die Overall Accuracy. Ein alternativer Ansatz zur Berechnung von Sensitivität und Spezifität liegt mit den sogenannten Likelihood Quotienten vor. Kappa Index, Overall Accuracy und Likelihood Quotient werden im Folgenden genauer beschrieben.

Overall Accuracy

Die Overall Accuracy ist die Rate der Übereinstimmungen zwischen Test und Goldstandard. Sie wird berechnet aus der Summe der richtig positiven und der richtig negativen Klassifikationen, dividiert durch die Anzahl aller Klassifikationen und ist somit der Anteil aller korrekten Ergebnisse (vgl. Fletcher et al, 1999; Fielding & Bell, 1997; Grant, 1991). Sie hängt von der Prävalenz ab und kann nach Ruttiman (1994) wie folgt ausgedrückt werden:

$$Overall\ Accuracy = (Prävalenz \times Sensitivität) + [(1 - Prävalenz) \times Spezifität]$$

$$Overall\ Accuracy = \frac{a + d}{N}$$

Abbildung 7: Berechnung der Overall-Accuracy nach Ruttiman (1994).

Die Overall Accuracy ist eine Größe, die die Gesamtbedeutung eines Tests ausdrückt. Im Gegensatz zum Kappa Index berücksichtigt die Overall Accuracy nicht, wie groß der Anteil der zufälligen Übereinstimmung ist (vgl. Kramer & Feinstein, 1981). Synonyme Begriffe für die Overall Accuracy sind neben Accuracy auch *test efficiency* oder (*total) correct classification rate* (Ruttiman, 1994).

Kappa Index

Der Kappa Index kann einerseits als Kenngröße der Reliabilität und andererseits als ein Maß der Validität berechnet werden. Der Kappa Index geht auf Cohen (1960, 1968) zurück. Im Rahmen von Reliabilitätsstudien werden die

Urteile von zwei oder mehr Beobachtern hinsichtlich ihrer Übereinstimmung verglichen. Wenn der Kappa Index für den Vergleich zwischen einem diagnostischen Test und einem Goldstandard errechnet wird, dann ist er ein Maß der Genauigkeit, also der Validität (vgl. Sackett et al., 1991).

Der Kappa Index kann berechnet werden, wenn zwei Beobachter oder zwei Methoden unabhängig voneinander dieselben N Personen in k paarweise disjunkte Kategorien klassifizieren (Hanley, 1987). Die Beurteilung der Übereinstimmung zwischen zwei Tests/Beobachtern setzt voraus, dass die Daten denselben Skalentyp und dieselbe Anzahl von Kategorien besitzen (Kramer & Feinstein, 1981).

Es handelt sich bei Kappa um ein zufallskorrigiertes Maß der Übereinstimmung. Der Wertebereich reicht von –1 bis +1. Ein Kappa-Wert von 1 bedeutet, dass eine komplette Übereinstimmung zwischen den Beobachtern bzw. den diagnostischen Methoden vorliegt. Ein Wert nahe 0 bedeutet, dass über eine zufällige Übereinstimmung hinaus keine Übereinstimmung besteht. Ein Wert kleiner 1 zeigt an, dass weniger Übereinstimmung besteht, als dies zufällig möglich wäre (vgl. Hanley, 1987; Spitznagel & Helzer, 1985).

Der Kappa Index wird auch als Konkordanzindex bezeichnet. Konkordanz liegt dann vor, wenn beide Methoden zum selben Ergebnis führen. Die Methoden sind dann Surrogate füreinander. Nach Sachs et al. (1991) gelten für die Beurteilung des Kappa Index folgende Richtwerte:

Kappa Index	**Übereinstimmung**
< .10	keine
.10 – .40	schwache
.41 – .60	deutliche
.61 – .80	starke
.81 – 1.00	fast vollständige

Abbildung 8: Beurteilung des Kappa Index (Sachs et al., 1991).

Der Kappa Index ist eine Funktion von Sensitivität, Spezifität und Prävalenz. (Spitznagel & Helzer, 1985). Da Kappa von der Prävalenz der Krankheit in der untersuchten Stichprobe bzw. Population abhängt, empfehlen Grove, Andreasen, McDonald-Scott, Keller & Shapiro (1981) bei einer Prävalenz von <5% den

Kappa Index nicht zu berechnen. In Abbildung 9 ist die mathematische Berechnung des Kappa Index dargestellt.

$$\text{Kappa } K = \frac{2(ad - bc)}{((a+c)(c+d)) + ((a+b)(b+d))}$$

Abbildung 9: Mathematische Formeln zur Berechnung von Kappa (vgl. Dunn, 2000; Maxwell, 1977).

Likelihood Quotient

Mit dem Likelihood Quotient liegt ein weiteres Verfahren zur Beschreibung der Validität eines diagnostischen Tests vor. Synonym wird auch der Begriff Likelihood Ratio, abgekürzt LR verwendet. Mit Hilfe des Likelihood Quotienten kann analog zu den prädiktiven Werten, für den Kliniker die folgende Frage beantwortet werden: Wie groß ist die Wahrscheinlichkeit, dass ein Patient mit einem bestimmten Testergebnis die gesuchte Krankheit hat? Im Gegensatz zum Likelihood Quotienten hängen die prädiktiven Werte jedoch von der Prävalenz ab.
Der Likelihood Quotient für einen bestimmten Wert bzw. Wertebereich eines diagnostischen Tests ist definiert als Quotient der Wahrscheinlichkeit des Testergebnisses bei Personen mit der Krankheit und der Wahrscheinlichkeit des Testergebnisses bei gesunden Personen (vgl. Abbildung 10).

$$\text{Likelihood Ratio} = \frac{\textit{Wahrscheinlichkeit eines bestimmten Testergebnisses bei Kranken}}{\textit{Wahrscheinlichkeit des Testergebnisses bei Gesunden}}$$

Abbildung 10: Berechnung des Likelihood Quotienten (vgl. Schmitz, Kruse & Tress, 2000).

Der Likelihood Quotient bringt zum Ausdruck, um wieviel mehr bzw. weniger wahrscheinlich ein bestimmtes Testergebnis bei kranken Personen wie bei Gesunden gefunden wird (vgl. Fletcher et al., 1999). Für dichotome Tests, die zwischen krank und gesund unterscheiden und keine weitere Differenzierung in Subgruppen vornehmen, können nur zwei Likelihood Quotienten berechnet werden, der positive und der negative Likelihood Quotient. Der Likelihood Ratio

eines positiven Tests kann als das Verhältnis der Rate der richtig Positiven zur Rate der falsch Positiven oder auch als das Verhältnis von Sensitivität zu (1−Spezifität) ausgedrückt werden. Je größer der Wert, desto besser ist die diagnostische Fähigkeit des Tests. Der positive Likelihood Quotienten zeigt an, um wieviel mal wahrscheinlicher ein Kranker ein positives Testergebnis hat als ein Gesunder.

Der negative Likelihood Ratio eines Tests ist definiert als das Verhältnis der Rate der falsch Negativen zur Rate der richtig Negativen oder anders ausgedrückt als das Verhältnis von (1−Sensitivität) zur Spezifität. Je kleiner dieser Wert ist, desto besser ist die diagnostische Fähigkeit des Tests. Der negative Likelihood Quotient zeigt an, um wieviel mal wahrscheinlicher ein Kranker ein negatives Testergebnis hat als ein Gesunder.

Ein idealer Test sollte einen möglichst hohen positiven Likelihood Quotienten und einen möglichst niedrigen negativen Likelihood Quotienten haben (vgl. Biggerstaff, 2000). In Abbildung 11 sind die Möglichkeiten zur Berechnung der beiden Likelihood Quotienten für den dichotomen Fall abgebildet.

Likelihood Ratio eines positiven Tests p_+

$$= \frac{\textit{Rate der richtig Positiven}}{\textit{Rate der falsch Positiven}}$$

$$= \frac{\frac{a}{a+c}}{\frac{b}{b+d}} = \frac{\textit{Sensitivität}}{\textit{1 - Spezifität}}$$

Likelihood Ratio eines negativen Tests p_-

$$= \frac{\textit{Rate der falsch Negativen}}{\textit{Rate der richtig Negativen}}$$

$$= \frac{\frac{c}{a+c}}{\frac{d}{b+d}} = \frac{\textit{1 - Sensitivität}}{\textit{Spezifität}}$$

Abbildung 11: Berechnung des positiven und des negativen Likelihood Quotienten (vgl. Biggerstaff, 2000; Jaeschke, Guyatt & Sackett, 1994).

Der Likelihood Ratio kann auch für jedes einzelne Testergebnis berechnet werden. Die Sensitivität bezieht sich dann auf die Fähigkeit des einzelnen Testergebnisses, Personen mit der gesuchten Krankheit zu identifizieren. Dasselbe gilt auch für die Spezifität. Ferner besteht die Möglichkeit der Berechnung von *Stratum Specific* oder *Multilevel* Likelihood Quotienten. Hierbei werden die Likelihood Quotienten für einen Bereich von Testwerten berechnet (Schmitz, Kruse, Tress, 2000).

Allgemein formuliert, kann mit Hilfe von Likelihood Ratios die Wahrscheinlichkeit der Krankheit, nach Bekanntsein des Testergebnisses, berechnet werden. Die Anwendung des Likelihood Quotienten erfolgt mit Hilfe der sogenannten Odds. Odds werden auch als Chancen oder Risikoquotienten bezeichnet. Sie sind definiert als das Verhältnis von zwei Wahrscheinlichkeiten. Wahscheinlichkeiten und Odds enthalten dieselben Informationen, sie drücken diese lediglich unterschiedlich aus und können durch die folgenden beiden Formeln umgerechnet werden:

Formel A) $$\text{Odds} = \frac{\textit{Wahrscheinlichkeit des Ereignisses}}{\textit{1 - Wahrscheinlichkeit des Ereignisses}}$$

Formel B) $$\text{Wahrscheinlichkeit des Ereignisses} = \frac{\textit{Odds}}{\textit{1} + \textit{Odds}}$$

Abbildung 12: Berechnung von Odds und Wahrscheinlichkeiten (vgl. Fletcher et al., 1999).

Rechnerisch erfolgt die Anwendung des Likelihood Quotienten in folgenden Schritten:

1. Die A-priori-Wahrscheinlichkeit der Krankheit, die Prävalenz, wird über die Formel A aus Abbildung 12 in A-priori-Odds umgewandelt.
2. Durch Multiplikation mit dem Likelihood Quotienten werden aus den A-priori-Odds die sogenannten A-posteriori-Odds (vgl. Abbildung 13) gebildet.
3. Mit Hilfe der Formel B aus Abbildung 12 können die A-posteriori-Odds in Wahrscheinlichkeiten umgewandelt werden.

A-posteriori-Odds = *A - priori - Odds* × *Likelihood Quotient*

Abbildung 13: Berechnung der A-posteriori-Odds (vgl. Fletcher et al., 1999).

Eine Erleichterung für die Berechnung des Likelihood Quotienten bietet das Nomogramm von Fagan (1975), das in Abbildung 14 dargestellt ist.

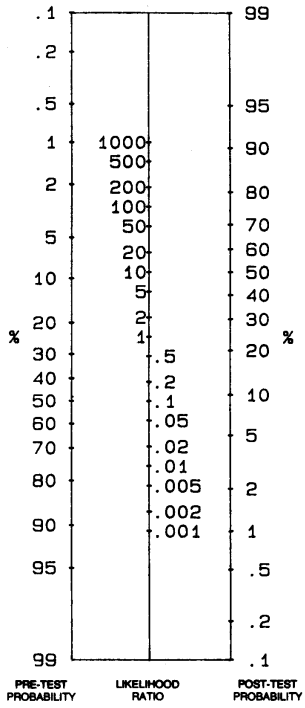


Abbildung 14: Nomogramm zur Berechnung der Posttestwahrscheinlichkeit (in Anlehnung an Fagan 1975).

In die erste Spalte des Nomogramms werden die Pretestwahrscheinlichkeiten eingetragen. In der zweiten Spalte stehen die Likelihood Quotienten und in der dritten Spalte die Posttestwahrscheinlichkeiten. Die Posttestwahrscheinlichkeit für einen bestimmten Patienten kann am Schnittpunkt der Verbindung der Pretestwahrscheinlichkeit mit dem Likelihood Quotienten des beobachteten Testergebnisses abgelesen werden.

Der Likelihood Quotient ist ein Indikator dafür, um wieviel ein diagnostischer Test die Pretestwahrscheinlichkeit der gesuchten Krankheit erhöht oder ernied-

rigt. Ein Likelihood Quotient von 1 bedeutet, dass die Posttestwahrscheinlichkeit gleich der Pretestwahrscheinlichkeit ist. Ist der Likelihood Quotient größer als 1, dann steigt die Posttestwahrscheinlichkeit, dass Patienten mit einem bestimmten Testergebnis die gesuchte Krankheit haben und zwar um so stärker, je größer der Likelihood Quotient ist. Umgekehrt senkt ein Likelihood Quotient kleiner 1 die Posttestwahrscheinlichkeit, dass die Krankheit bei Patienten mit einem bestimmten Testergebnis vorhanden ist.
Jaeschke et al. (1994) geben für die Bedeutung des LR folgende Richtwerte an:

Wert des Likelihood Ratios (LR)	Bedeutung
LR > 10 oder < 0.1	große Veränderung zwischen Pre- und Posttestwahrscheinlichkeit
LR = 5 bis 10 oder 0.1 bis 0.2	moderate Veränderung
LR = 2 bis 5 oder 0.5 bis 0.2	kleinere, aber manchmal wichtige Veränderungen
LR = 1 bis 2 oder 0.5 bis 1	kaum wichtige Veränderungen der Wahrscheinlichkeit

Abbildung 15: Bedeutung der Höhe des Likelihood Quotienten (Jaeschke et al., 1994).

3.2.4 Receiver Operating Characteristica (ROC) Kurven

Die Methode der *Receiver Operating Characteristica (ROC) Kurven* wurde im Kontext der elektronischen Signalentdeckung in den frühen 50er Jahren entwickelt. Ab Mitte der 60er Jahre wurden die ROC-Kurven in der experimentellen Psychologie und der Psychophysik genutzt. Der Radiologe Leo Lusted begann 1969 die Methode der ROC-Kurven in der Medizin für bildgebende Verfahren zu verwenden. Heute wird sie in vielen Gebieten der Medizin eingesetzt (Kraemer, 1988; Swets,1988; Zweig & Campbell, 1993).
Die ROC-Kurve liefert ein Bild des gesamten Spektrums der Sensitivität und Spezifität eines Tests. Jeder Punkt auf der ROC-Kurve repräsentiert das Paar

von Sensitivität und Spezifität an einem Schwellenwert. Eine ROC-Kurve ist somit eine einfache, aber vollständige Beschreibung der Genauigkeit eines Tests. Sie enthält die Sensitivität und die Spezifität für alle möglichen Cut-Off-Werte. Die ROC-Kurve beschreibt die Kompromisse, die zwischen der Rate der richtig Positiven und der Rate der falsch Positiven gemacht werden können (Mari & Williams, 1985; Metz, 1978). Die relative Position von zwei oder mehr ROC-Kurven stellt einen qualitativen Vergleich der Genauigkeit dar.
Für alle möglichen Cut-Off-Werte eines Tests wird die Rate der richtig Positiven (Sensitivität) als Funktion der Rate der falsch Positiven (1−Spezifität) aufgetragen. Auf der y-Achse wird somit die Sensitivität abgetragen und auf der x-Achse 1−Spezifität. In Abbildung 16 ist beispielhaft eine ROC-Kurve dargestellt. Unter der sogenannten ROC-Analyse wird der Vergleich von Tests anhand von ROC-Kurven verstanden.

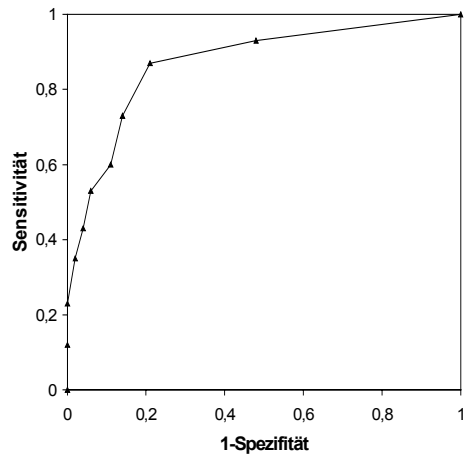


Abbildung 16: Beispiel einer ROC-Kurve.

Die Achsen umfassen einen Wertebereich für die Wahrscheinlichkeiten von 0 bis 1 oder anders ausgedrückt von 0 bis 100 %. Dies sind die Grenzen von richtig und falsch positiven Zuordnungen.
Nach Somoza, Steer, Beck & Clark (1993) besteht die Konstruktion einer ROC-Kurve im wesentlichen aus drei Schritten:

1. Auswahl der verschiedenen Cut-Off-Werte für einen Test,
2. Berechnung der Rate der richtig Positiven und der Rate der falsch Positiven für jeden Cut-Off-Wert und
3. Eintragung der Wertepaare von Sensitivität und 1−Spezifität in ein Koordinatensystem.

Tests, die gut zwischen Kranken und Gesunden diskriminieren, finden sich in der linken oberen Ecke der ROC-Kurve. Tests, die weniger gut zwischen Kranken und Gesunden unterscheiden, erzeugen Kurven, die näher an den diagonalen Verlauf von der unteren linken in die obere rechte Ecke heranrücken (Fletcher et al., 1999).
Die Diagnonale, die im 45° Winkel von der linken unteren Ecke in die rechte obere Ecke läuft, wird als „line of no information“ oder als „chance line“ bezeichnet und zeigt die Beziehung zwischen der Rate der richtig Positiven und der Rate der falsch Positiven an, die auftritt, wenn ein Test keine Informationen liefert. Je mehr die ROC-Kurve oberhalb dieser Linie liegt, desto besser ist die diskriminatorische Fähigkeit des Tests. Liegt die Kurve unterhalb dieser Linie, so nimmt der Test mehr falsch positive als richtig positive Zuordnungen vor und liefert damit schlechtere Ergebnisse, als dies durch eine zufällige Zuordnung möglich wäre (Fletcher et al., 1999; Erdreich & Lee, 1981). Tests mit perfekter Diskrimination laufen durch die obere linke Ecke. Die Rate der richtig Positiven ist hier 100%, die Rate der falsch Positiven ist 0%. Die meisten Tests liegen zwischen diesen beiden Extremen (vgl. Zweig & Campbell, 1993).
Die ROC-Kurve kann zur Festlegung des geeignetsten Cut-Off-Wertes genutzt werden. Im Allgemeinen ist der beste Schwellenwert jener Wert, der an oder in der Nähe der sogenannten „Schulter“ der ROC-Kurve liegt. Dieser Punkt liegt am Berührungspunkt der Tangente, die parallel zur Diagonalen vom Ursprung zum Punkt (1,1) verläuft und der ROC-Kurve (Ritchie & Fuhrer, 1992). Sprechen jedoch klinische Gründe dafür, die Rate der falsch positiven oder der falsch negativen Klassifikationen besonders niedrig zu halten, ist unter Umständen ein anderer Cut-Off-Wert besser geeignet. Die Festsetzung von Testtrennwerten ist nicht alleine statistisch begründbar, sondern erfordert zugleich

ein Werturteil, das sowohl soziale und ökonomische Werte als auch praktische Erwägungen berücksichtigt.

Neben der Auswahl von Cut-Off-Werten können ROC-Kurven zum Vergleich von alternativen Tests oder verschiedenen Auswertungsprozeduren eines Tests eingesetzt werden. Die ROC Analyse kann auch verwendet werden, um zu vergleichen, wie gut ein Test verschiedene Störungsbilder (z.B. Angst und Depression) misst. Mittels der ROC-Kurven besteht für den Kliniker die Möglichkeit, eine angemessene Testprozedur für eine bestimmte Population auszuwählen. ROC-Kurven sind heute somit eine Standardmethode in der psychiatrischen Forschung (Murphy, Berwick, Weinstein, Borus, Budman & Klerman, 1987).

Die Gesamtgenauigkeit eines Tests kann als Fläche unter der ROC-Kurve beschrieben werden. Die Fläche unter der ROC-Kurve wird als *Area under the curve (AUC)* oder *Area under the Receiver Operating Characteristica Curve (AUR)* bezeichnet (vgl. Murpy et al., 1987; Somoza et al., 1993). Per Konvention reicht der Wertebereich der Fläche unter der ROC-Kurve von 0.5 bis 1.0. Ist die Fläche unter der ROC-Kurve signifikant größer als 0.5, dann besteht ein klarer Zusammenhang zwischen Test und Goldstandard. Kreuzen sich zwei ROC-Kurven, dann kann der eine Test zwar eine größere Fläche unter der ROC-Kurve besitzen, aber in einem bestimmten Bereich der ROC-Kurve kann der andere Test dennoch bessere Werte der Sensitivität und Spezifität liefern. Daher sollte bei dem Vergleich der Flächen unter der ROC-Kurve immer die grafische Darstellung der ROC-Kurven berücksichtigt werden (Murphy et al., 1987).

Die Fläche unter der ROC-Kurve bringt zum Ausdruck, mit welcher Wahrscheinlichkeit ein zufällig ausgewählter Kranker einen höheren Testwert hat, als ein zufällig ausgewählter Gesunder. Beispielsweise bedeutet eine Fläche von .80, dass eine zufällig ausgewählte Person aus der Gruppe der Kranken in 80% der Fälle einen höheren Testwert hat, als eine zufällig ausgewählte Person aus der Gruppe der Gesunden. Die Fläche bedeutet nicht, dass 80% der Personen mit einem positiven Testergebnis krank sind bzw., dass ein positives Testergebnis mit einer Wahrscheinlichkeit von 80% auftritt (Zweig & Campbell, 1993).

Ruttimann (1994) sowie Zweig und Campbell (1993) listen folgende Vorteile von ROC-Kurven auf:

1. Einfache und gut verständliche grafische Darstellung.
2. Das gesamte Spektrum der möglichen Schwellenwerte wird abgebildet.
3. Unabhängigkeit von der Prävalenz in der Stichprobe.
4. Ein direkter visueller Vergleich verschiedener Tests ist auf einer gemeinsamen Skala möglich.
5. Es kann der statistische Vergleich zwischen verschiedenen Tests oder der statistische Vergleich eines Tests mit der zufälligen Übereinstimmung berechnet werden.
6. Die Daten müssen nicht gruppiert werden.
7. Die Sensitivität und die Spezifität können für jeden Cut-Off schnell abgelesen werden.

Nachteile der ROC-Kurven sind nach Ruttimann (1994) sowie Zweig und Campbell (1993), dass

1. anhand der Kurve nicht die Anzahl der Probanden abgelesen werden kann,
2. die Korrelationen zwischen Kurven, deren Werte aus derselben Stichprobe stammen, nicht sichtbar sind,
3. die statistischen Berechnungen ohne PC-Programm nur schwer durchführbar sind.

Zwischen den ROC-Kurven und dem Likelihood Ratio besteht eine enge Beziehung. Choi (1998) unterscheidet drei verschiedene Typen von Steigung bei einer ROC-Kurve:

1. die Tangente an einem Punkt der ROC-Kurve,
2. die Steigung zwischen einem Punkt auf der ROC-Kurve und dem Ursprung (0,0) bzw. dem Endpunkt (1,1),
3. die Steigung zwischen zwei Punkten auf der ROC-Kurve.

Analog unterscheidet er drei Typen von Likelihood Ratio:

1. Likelihood Ratio für eine bestimmten Testwert,
2. Likelihood Ratio für einen bestimmten Cut-Off-Wert und
3. Likelihood Ratio für ein Testergebnis innerhalb eines definierten Intervalls.

Die Steigung der Tangente an einem Punkt der ROC Kurve entspricht dem Likelihood Ratio für einen bestimmten Testwert. Die Steigung zwischen dem Ursprung (0,0) und einem Punkt der ROC Kurve korrespondiert zum positiven Likelihood Ratio für einen Cut-Off-Wert und die Steigung zwischen dem Endpunkt (1,1) und einem Punkt der ROC Kurve korrespondiert analog zum negativen Likelihood Ratio für einen Cut-Off-Wert. Die Steigung zwischen zwei Punkten der ROC Kurve ist gleich dem Likelihood Ratio für ein Testergebnis innerhalb eines definierten Intervalls (vgl. auch Peirce & Cornell, 1993)

3.3 Screeninginstrumente

3.3.1 General Health Questionnaire (GHQ-12)

Der *General Health Questionnaire (GHQ)* wurde 1972 von Goldberg entwickelt, um psychische Störungen in der Allgemeinbevölkerung und nichtpsychiatrischen Kliniken, wie in der Primärversorgung oder der allgemeinmedizinischen Ambulanz, zu untersuchen (Goldberg, 1972). Der GHQ ist ein Fragebogen zur Selbstbeurteilung des Befindens in den letzten Wochen. Goldbergs Ziel war die Konzeption eines Fragebogens, der kurz und leicht durchzuführen ist, hohe Akzeptanz bei den Befragten hat und objektiv ist, in dem Sinne, dass der Untersucher keine subjektiven Fremdbeurteilungen vornehmen muss (Goldberg & Williams, 1991).

Der General Health Questionnaire besteht aus einer Reihe von Fragen, die das kleinste gemeinsame Vielfache von Symptomen verschiedener psychiatrischer Syndrome darstellen (Goldberg & Hiller, 1979). Der GHQ macht keine theoretischen Annahmen über die Hierarchie innerhalb der Klasse der psychischen Krankheiten, sondern geht vielmehr davon aus, dass es auch einfachere und

weniger stark differenzierte Wege gibt, um Personen mit psychischen Krankheiten von Gesunden zu unterscheiden (Goldberg & Williams, 1991).

Nach Goldberg und Williams (1991) beruht die Entwicklung des GHQ u.a. auf den Arbeiten zum Cornell Medical Inventory (Brodman, Erdman, Lorge, Wolff & Broadbent, 1949) und anderen älteren Skalen wie dem Gurin Mental Status Index (Gurin, Veroff & Feld, 1969).

Heute stehen neben der Originalversion, dem aus 60 Fragen bestehenden GHQ-60, drei kürzere Versionen zur Verfügung. Der GHQ-30 setzt sich aus 30 Fragen zusammen, der GHQ-28 aus 28 Fragen und der GHQ-12 aus 12 Fragen. Der GHQ-28 bietet als einziger der vier Fragebögen die Möglichkeit neben einem Gesamttestergebnis auch Werte für einzelne Skalen zu berechnen. Er ist ein skaliertes Instrument mit den vier Subskalen A: somatische Symptome, B: Angst und Insomnie, C: soziale Dysfunktion und D: schwere Depression (Hobi, Gerhard & Gutzwiller, 1989; Power, 1988).

Ziel der vier Versionen des GHQ ist es, Unterbrechungen im normalen Funktionsniveau zu entdecken, nicht jedoch Lebenszeitdiagnosen zu stellen oder Persönlichkeitsstörungen zu erfassen. Daher werden mit dem GHQ Störungen ohne breites Symptomprofil, wie beispielsweise spezifische Psychosen, und chronische Krankheiten eher übersehen als aktuelle psychische Störungen mit einem breiten Symptomprofil (Hobbs, Ballinger, Greenwood, Martin & McClure 1984; Hobbs, Ballinger & Smith,1983).

Der GHQ kann zum einen als Screeninginstrument verwendet werden um wahrscheinlich Kranke zu erkennen, zum anderen als Maß für die Schwere der psychischen Beeinträchtigung. Bei der ersten Fragestellung ist die Unterscheidung zwischen krank und gesund interessant, bei der zweiten hingegen der Gesamttestwert.

In der vorliegenden Arbeit wird die 12-Item Version, der GHQ-12, verwendet. Abbildung 17 zeigt die in der vorliegenden Untersuchung verwendete, deutsche Übersetzung des GHQ-12 nach Linden, Maier, Achberger, Herr, Helmchen und Benkert (1996).

Der GHQ-12 enthält 12 Fragen über positive und negative Aspekte der psychischen Gesundheit in den letzten Wochen. Die Antwortskala ist in allen vier Versionen des GHQ jeweils vierstufig.

1. Haben Sie in den letzten Wochen wegen Sorgen weniger geschlafen?
□ nein, gar nicht □ nicht schlechter als üblich □ schlechter als üblich □ viel schlechter als üblich

2. Haben Sie das Gefühl gehabt, dauernd unter Druck zu stehen?
□ nein, gar nicht □ nicht mehr als üblich □ mehr als üblich □ viel mehr als üblich

3. Haben Sie sich in den letzten Wochen auf das, was Sie gemacht haben konzentrieren können?
□ besser als üblich □ so wie üblich □ schlechter als üblich □ viel schlechter als üblich

4. Haben Sie in den letzten Wochen das Gefühl gehabt, für etwas nützlich zu sein?
□ mehr als üblich □ so wie üblich □ weniger als üblich □ viel weniger als üblich

5. Haben Sie in den letzten Wochen das Gefühl gehabt, sich mit Ihren Problemen auseinander zu setzen?
□ besser als üblich □ so wie üblich □ weniger als üblich □ viel weniger als üblich

6. Ist es Ihnen in den letzten Wochen schwer gefallen, Entscheidungen zu treffen?
□ nein, gar nicht □ so wie als üblich □ schwerer als üblich □ viel schwerer als üblich

7. Haben Sie in den letzten Wochen den Eindruck gehabt, dass Sie mit Ihren Schwierigkeiten nicht zu Rande gekommen sind?
□ nein, gar nicht □ nicht schlechter als üblich □ schlechter als üblich □ viel schlechter als üblich

8. Alles in allem, haben Sie sich in den letzten Wochen einigermaßen zufrieden gefühlt?
□ mehr als üblich □ so wie üblich □ weniger als üblich □ viel weniger als üblich

9. Konnten Sie in den letzten Wochen Ihren Alltagsverpflichtungen mit Freude nachgehen?
□ mehr als üblich □ so wie üblich □ weniger als üblich □ viel weniger als üblich

10. Haben Sie sich in den letzten Wochen unglücklich und deprimiert gefühlt?
□ nein, gar nicht □ nicht mehr als üblich □ mehr als üblich □ viel mehr als üblich

11. Haben Sie in den letzten Wochen einen Mangel an Selbstvertrauen gespürt?
□ nein, gar nicht □ nicht mehr als üblich □ mehr als üblich □ viel mehr als üblich

12. Haben Sie sich in den letzten Wochen wertlos gefühlt?
□ nein, gar nicht □ nicht mehr als üblich □ mehr als üblich □ viel mehr als üblich

Abbildung 17: GHQ-12 (deutsche Übersetzung: Linden, Maier, Achberger, Herr, Helmchen & Benkert, 1996).

Es handelt sich beim GHQ-12 um einen häufig untersuchten Fragebogen, der bereits in zahlreichen Studien zur Fallidentifikation eingesetzt wurde. Wenn ein angemessener Cut-Off-Wert festgelegt wurde, kann der GHQ als Screeninginstrument zur Erfassung von psychischem Distress verwendet werden (Schmitz, Kruse, Heckrath, Alberti & Tress, 1999).

Da der GHQ nicht mit dem Ziel entwickelt wurde, zwischen verschiedenen psychiatrischen Störungen zu unterscheiden, kann mit dem GHQ keine psychiatri-

sche Diagnose gestellt werden. Vielmehr ist der GHQ ein Screeninginstrument für generelle psychiatrische Beeinträchtigung (Schmitz, Kruse & Tress, 1999).
Im GHQ wird zum einen gefragt, ob es dem Probanden in den letzten Wochen möglich war, seine normalen, gesunden Aktivitäten weiter auszuführen und zum anderen, ob neue distressende Phänomene aufgetreten sind (Goldberg & Williams, 1991).
Der GHQ-30 und der GHQ-12 wurden aus der 60-Item Version des GHQ entwickelt, indem zunächst diejenigen Fragen gestrichen wurden, die von körperlich kranken Probanden positiv beantwortet wurden. In einem zweiten Schritt wurden die verbleibenden Items in zwei Gruppen eingeteilt: Solche bei denen Zustimmung Gesundheit anzeigt und solche, bei denen Zustimmung Krankheit anzeigt. Innerhalb dieser Gruppen wurden dann diejenigen Items ausgewählt, welche die höchsten Werte in der ursprünglichen Itemanalyse des GHQ-60 hatten (Goldberg & Williams, 1991).
Der GHQ-28 wurde nicht auf diesem Weg gewonnen, sondern mit Hilfe einer Hauptkomponentenanalyse erzeugt.
Die Vorteile des GHQ-12 gegenüber den längeren Versionen sind die schnelle Durchführbarkeit und Auswertung. Trotz der geringen Anzahl von Fragen ist der GHQ-12 hinsichtlich Reliabilität und Validität mit den längeren Versionen durchaus vergleichbar (Goldberg, Gater, Sartorius, Ustun, Piccinelli, Gureje & Rutter, 1997). Aufgrund des geringen Zeitaufwandes wird der GHQ-12 häufig in Forschungsstudien verwendet.
Seit den 70er Jahren wurde der GHQ-12 international in zahlreichen Studien sowohl im klinischen Bereich als auch in der Primärversorgung eingesetzt. Der GHQ-12 wird heute in mehr als 36 Sprachen in der ganzen Welt in den verschiedensten Populationen durchgeführt (Goldberg & Williams, 1991). Sartorius, Üstün, Costa e Silva et al. (1993) sehen für den häufigen Einsatz des GHQ-12 folgende Gründe: Der GHQ-12 ist kurz und verfügt über akzeptable Werte der Sensitivität und Spezifität. Ferner wurde er in verschiedenen kulturellen Umgebungen erfolgreich eingesetzt.
In Tabelle 2 ist eine Auswahl internationaler Studien abgebildet, die den GHQ-12 im Primary Care Bereich eingesetzt haben.

Tabelle 2: Internationale Studien, die den GHQ-12 im Primary Care Bereich einsetzen.

Autor	Publikationsjahr	Land
Araya, Wynn & Lewis	1992	Chile
Gureje, Üstün & Simon	1997	International, in 14 Ländern wie China, Griechenland, Indien, USA
Gureje	1991	Nigeria
Hotopf, Sharp & Lewis	1998	England
Kihç et al.	1997	Türkei
Kruse, Heckrath, Schmitz, Alberti & Tress	1999	Deutschland
Mari & Williams	1985	Brasilien
Simon & Von Korff	1995	USA
Titus et al.	1999	Niederlande

Auch die Antwortskala des GHQ-12 ist vierstufig. Die Antwortmöglichkeiten der positiv formulierten Fragen lauten: *Besser als üblich, so wie üblich, schlechter als üblich* und *viel schlechter als üblich*. Bei den negativ formulierten Items lauten die Antworten: *Gar nicht, nicht mehr als üblich, mehr als üblich* und *viel mehr als üblich*. Positiv formulierte Items tragen zu einem höheren GHQ Gesamtwert bei, wenn sie mit *schlechter* bzw. *viel schlechter als üblich* beantwortet werden. Negativ formulierte Fragen führen zu einer Erhöhung des Gesamtscores, wenn sie mit *mehr* bzw. *viel mehr als üblich* beantwortet werden.

Es gibt vier Möglichkeiten den GHQ-12 auszuwerten. Goldberg (Goldberg et al., 1997) empfiehlt eine dichotome Auswertung nach dem Muster 0-0-1-1: *Gar nicht* = 0, *nicht mehr als üblich* = 0, *mehr als üblich* = 1 und *viel mehr als üblich* = 1 bzw. bei positiv formulierten Items *besser als üblich* = 0, *so wie üblich* = 0, *schlechter als üblich* = 1, *viel schlechter als üblich* = 1 (vgl. auch Newmann, Bland & Orn, 1988). Die Werte der 12 Items werden zu einem Summenwert addiert. Der Gesamttestwert kann daher zwischen 0 und 12 liegen. Diese Auswertungsmethode wird auch als GHQ-Scoring bezeichnet. Sie wird in der vor-

liegenden Arbeit zur Auswertung des General Health Questionnaires verwendet.

Beim sogenannten Likert-Scoring erfolgt die Auswertung nach dem Schema 0-1-2-3 und beim modifizierten Likert-Scoring nach dem Muster 0-0-1-2. In das Likert-Scoring und in das modifizierte Likert-Scoring geht somit die Schwere der Symptome ein.

Die vierte Möglichkeit, das C-GHQ-Scoring, bewertet positive Items, bei denen Zustimmung Gesundheit anzeigt, nach dem Schema 0-0-1-1 und negative Items, bei denen Zustimmung Krankheit anzeigt, nach den Muster 0-1-1-1. Mit dieser von Goodchild und Duncan-Jones (1985) vorgeschlagene Auswertungsmethode sollen auch chronische Erkrankungen erfasst werden. Goldberg & Williams (1991) empfehlen diese Methode nicht als Alternative zum GHQ-Scoring einzusetzen, sondern sie eher als eine zusätzliche Auswertungsmethode zu verstehen.

Durch die Methode der Likert-Bewertung erhält man eine weniger schiefe und breitere Verteilung als beim GHQ-Scoring. Die C-GHQ Bewertung erzeugt eher eine normalverteilte Kurve als die dichotome Auswertung (Goldberg & Williams, 1991; Goldberg et al., 1997). Durch die Methode des GHQ-Scoring werden Fehler, die durch zentrale Antworttendenzen und die Neigung zu extremen Antworten entstehen können, vermieden.

Im World Wide Web wird auf der Homepage des General Health Questionnaires vom Testherausgeber NFER Nelson (http://www. nfer-nelson.co.uk/ghq/) für den GHQ-12 in Übereinstimmung mit Goldberg (1972) und Goldberg et al. (1997) der Cut-Off-Wert 1/2 für die Methode des GHQ-Scorings vorgeschlagen. Der Schwellenwert wird im GHQ mit zwei Zahlen angegeben. Patienten, die einen Wert kleiner oder gleich der ersten Zahl haben, liegen unterhalb der Schwelle. Patienten, die ein Ergebnis gleich oder größer der zweiten Zahl haben, liegen über der Schwelle. Nach Furukawa und Goldberg (1999), Goldberg et al. (1997) sowie Goldberg, Oldehinkel und Ormel (1998) variiert der beste Cut-Off für den GHQ-12 bei Verwendung des GHQ-Scorings zwischen 0/1 und 6/7. Wenn der GHQ bei Patienten mit starken körperlichen Krankheiten eingesetzt wird, ist es manchmal nötig, den Schwellenwert zu erhöhen (Goldberg & Williams, 1991). Um jedoch ein optimales Verhältnis von Sensitivität und

Spezifität zu erreichen besteht, nach Goldberg et al. (1998) die sicherste Methode weiterhin darin, eine eigene Validitätsstudie durchzuführen, wie dies auch im Benutzerhandbuch empfohlen wird (Goldberg & Williams, 1991).
Nach Papassotiropoulos, Heun und Maier (1997) ist der GHQ-12 auch bei Probanden über 65 Jahren indiziert und wird nicht durch milde kognitive Beeinträchtigung beeinflusst. Papassotiropoulos et al. (1997) empfehlen für die Methode des GHQ-Scorings einen Cut-Off-Wert von 3/4 für Probanden ab 65 Jahren und einen Cut-Off von 1/2 für Personen unter 65 Jahren zu verwenden.
Die Validierung des GHQ-12 wurde in verschiedenen internationalen Studien sowohl für klinische Populationen als auch für die Allgemeinbevölkerung durchgeführt. Eine exemplarische Darstellung neuerer Validierungsstudien ist in Tabelle 3 wiedergegeben.

Tabelle 3: Validierungsstudien zum GHQ-12.

Autor	Publikationsjahr	Vergleichsinstrumente
Araya, Wynn & Lewis	1992	Self Report Questionnaire (SRQ-20), Revised Clinical Interview Schedule (CIS-R)
Caraveo-Anduaga, Martinez, Saldivar, Lopez & Saltijeral	1998	Composite International Diagnostic Interview (CIDI)
Goldberg, Gater, Sartorius, Ustun, Piccinelli, Gureje & Rutter	1997	GHQ-28, CIDI
Gureje & Obikoya	1990	CIDI
Hardy, Shaipiro, Haynes & Rick	1999	CIS-R
Hotopf, Sharp & Lewis	1998	Beck Depression Inventory (BDI), Hamilton Rating Scale for Depression (HRSD), CIS-R
Le Fevre, Devereux, Smith, Lawrie & Cornbleet	1999	Hospital Anxiety Depression Scale (HADS), CIS-R
Mari & Williams	1985	SRQ-20, Clinical Interview Schedule (CIS)
Politi, Piccinelli & Wilkinson	1994	Psychiatrisches Urteil
Schmitz, Kruse, Heckrath, Alberti & Tress	1999	Symptom Check-List (SCL-90-R), Structured Clinical Interview (SCID)
Schmitz, Kruse & Tress	2000	SCID, SCL-90

Goldberg et al. (1997) kamen zu dem Ergebnis, dass Geschlecht, Alter und Bildungsstand keinen signifikanten Einfluss auf die Validität des GHQ-12 haben und der GHQ-12 über ebenso gute Validitätskennwerte verfügt wie der GHQ-28.

Ein Maß für die Reliabilität ist die Interne Konsistenz, gemessen anhand des Alpha Koeffizienten von Cronbach. Banks, Clegg, Jackson, Kemp, Stafford und Wall (1980) fanden für den GHQ-12 Alpha Koeffizienten zwischen .82 und .90. Schmitz, Kruse und Tress (1999) ermittelten einen Alpha Koeffizienten von .91 im Bereich der Primärversorgung in der Bundesrepublik Deutschland.

Faktorenanalytische Untersuchungen haben zu zwei und drei Faktorenlösungen geführt. In einer WHO-Studie zu psychischen Störungen in der Allgemeinarztpraxis wurde der GHQ-12 in fünfzehn Zentren rund um die Welt eingesetzt. Werneke, Goldberg, Yalcin und Üstün (2000) konnten in dieser Studie zeigen, dass es substantielle Varianzen in den Faktoren zwischen den Zentren gibt. Zwischen den Faktoren scheinen hohe Interkorrelationen zu bestehen. Das Vorhandensein von Interkorrelationen und die Variabilität der Stichproben ist nach Wernecke et al. (2000) für die Instabilität der Faktoren verantwortlich. Durch Rotation der Faktoren ergab sich jedoch eine Zweifaktorenlösung mit den beiden Faktoren A: Angst und Depression und B: soziale Dysfunktion. Auch Gureje (1991) fand zwei Faktoren A: soziales Funktionsniveau und B: Dysphorie. Kihç et al. (1997) fanden in der Primärversorgung in der Türkei ebenfalls die beiden Faktoren A: Angst und Depression sowie B: soziale Dysfunktion.

GHQ-Scores können allgemein auf drei verschiedene Arten interpretiert werden. Erstens als Maß für die Schwere einer psychischen Störung, zweitens zur Schätzung der Prävalenz psychischer Störungen und drittens als Indikator für das Vorliegen einer psychischen Krankheit (Goldberg & Williams, 1991). Die letztgenannte Funktion wird genutzt, wenn der GHQ als Screeninginstrument eingesetzt wird.

3.3.2 Brief Patient Health Questionnaire (B-PHQ)

Der *Brief Patient Health Questionnaire (B-PHQ)* ist eine Kurzversion des *Prime-MD Patient Health Questionnaire (PHQ)* von Spitzer, Kroenke und Williams (1999). Vorläufer des Prime-MD Patient Health Questionnaire ist der sogenannte *Prime-MD* (*Primary Care Evaluation of Mental Disorders;* Spitzer, Williams, Kroenke, Linzer, deGruy, Hahn, Brody & Johnson, 1994). Der Prime-MD war das erste Instrument, das zur Diagnose von spezifischen Störungen anhand von DSM-III-R bzw. DSM-IV Kriterien für den Gebrauch in der Primärversorgung entwickelt wurde.

Ziel des Prime-MD ist es, die speziellen Bedürfnisse von Ärzten in der Primärversorgung zu erfüllen. Aus diesem Grunde wurde bei der Entwicklung des Prime-MD neben der Genauigkeit besonderer Wert auf Kürze und Effektivität des Verfahrens gelegt. Der Prime-MD wurde von Spitzer und Williams in Zusammenarbeit mit fünf Ärzten aus dem Bereich der Primärversorgung entwickelt.

Der Prime-MD besteht aus zwei Teilen, zum einen dem sogenannten Patientenfragebogen, der mit *PQ (Patient Questionnaire)* abgekürzt wird, und zum anderen dem *Clinician Evaluation Guide (CEG),* einem strukturierten Interview.

Der Prime-MD zeichnet sich durch seine flexible Handhabbarkeit aus, d.h. er kann routinemäßig bei allen Patienten oder nur bei denjenigen Patienten, bei denen eine psychische Störung vermutet wird, eingesetzt werden. Ferner ist der CEG als Ganzes verwendbar oder es werden nur diejenigen Module eingesetzt, die dem Arzt aufgrund der Auswertung des Patientenfragebogens am wichtigsten erscheinen (vgl. Spitzer et al., 1994).

Der Patientenfragebogen ist ein Selbstbeurteilungsfragebogen, der aus einer einzige Seite besteht. Diese kann vom Patienten bereits im Wartezimmer ausgefüllt werden. Es handelt sich um 26 Fragen über das Vorliegen von körperlichen und psychischen Symptomen während der letzten vier Wochen. Die Fragen können mit *ja* oder *nein* beantwortet werden. Die 26 Fragen des Patientenfragebogens und die Module des CEG beziehen sich auf die fünf häufigsten Störungsgruppen im Primary Care Bereich: Depressive Störungen, Angststörungen, Alkoholprobleme, somatoforme Störungen und Essstörungen (vgl.

Johnson, Spitzer, Williams, Kroenke, Linzer, Brody, deGruy & Hahn, 1995; Van Hook, Berkman, Dunkle, 1996).
In der zweiten Stufe des Prime-MD, dem CEG, findet eine klinische Bewertung des Patienten durch den Arzt statt. Beim Clinical Evaluation Guide handelt es sich um standardisierte Fragen für ein kurzes und zielgerichtetes Interview. Die Fragen sind nach einem algorithmischen Modell geordnet, das die Differentialdiagnose erleichtert.
Im Patientenfragebogen werden depressive Störungen mit den beiden folgenden Fragen untersucht: „Fühlten Sie sich im Verlauf der letzten vier Wochen deutlich beeinträchtigt durch wenig Interesse oder Freude an Ihren Tätigkeiten?“ und „Fühlten Sie sich im Verlauf der letzten vier Wochen deutlich beeinträchtigt durch Niedergeschlagenheit, depressive Gefühle oder Hoffnungslosigkeit?“
Wenn eine der beiden Fragen mit ja beantwortet wird, sollte der Arzt das Modul affektive Störungen im CEG durchführen.

Folgende affektive Störungen können mit dem Prime-MD erfasst werden:

1. Major Depression,
2. teilweise remittierte Major Depression,
3. Dysthymie,
4. Geringgradige depressive Störung,
5. Verdacht auf affektive Störung aufgrund medizinischer Krankheitsfaktoren, Medikamente oder Drogenkonsum und
6. bipolare Störung.

Der Prime-MD Clinical Evaluation Guide enthält jedoch nur eine Frage zum Verdacht auf eine bipolare Störung. Eine volle Evaluation der bipolaren Störung ist nach Angabe der Autoren nicht enthalten, da die meisten Patienten mit bipolaren Störungen im Primary Care Bereich bereits diagnostiziert sind und von Fachärzten behandelt werden (Williams, Spitzer, Linzer, Kroenke, Hahn, deGruy & Lazev, 1995).
Die Validierung des Prime-MD erfolgte 1994 durch Spitzer et al. in der sogenannten „Prime MD 1000 Studie“ und ergab eine gute Übereinstimmung mit

den Diagnosen von Klinikern. Der Kappa-Index lag bei .71 und die Overall-Accuracy bei 88%. Die Mehrheit der Patienten beurteilten den Prime-MD als hilfreich für den Arzt zum Verständnis und zur Behandlung ihrer Beschwerden.
Loerch, Szegedi, Kohnen und Benkert (1999) validierten die deutsche Version des Prime-MD am Münchner Composite International Diagnostic Interview (M-CIDI). Die Ergebnisse der Studie ergaben gute Validitätskennwerte für alle erfassten psychiatrischen Störungen mit einer Sensitivität von .73 und einer Spezifität von .67 sowie einer Overall Accuracy von .70.
Mit Kobak, Taylor, Dottl, Greist, Jefferson, Burroughs, Katzelnick und Mandell (1997) liegen zwei rechnergestützte Versionen des Prime-MD im Primary Care Setting vor. Bei der einen Version handelt es sich um eine Desktop-Version des Prime-MD. Bei der anderen Version wird das Interview mittels eines Touchtone-Telefons durchgeführt, bei dem der Computer die Fragen über das Telefon stellt und der Patient die Antwort via Tastendruck am Telefon eingibt.
Die Bewertung des Prime-MD dauert im Durchschnitt 8,4 Minuten, 5 bis 6 Minuten bei Patienten ohne Prime-MD Diagnose und 11,4 bei Patienten mit einer Diagnose. Dies stellt für die klinische Primärversorgung einen relativ hohen Zeitaufwand dar, da pro Patient im Durchschnitt nur jeweils 10 bis 15 Minuten Zeit zur Verfügung stehen. Der Prime-MD wird zwar in der klinischen Forschung relativ häufig eingesetzt, sein Gebrauch in der Primärversorgung ist aber aufgrund des Zeitaufwandes limitiert.
Der Prime-MD Patient Health Questionnaire (PHQ) ist eine reine Selbstbeurteilungsversion des Prime-MD. Beim PHQ wurden die beiden Komponenten des Prime-MD, der Patientenfragebogen und die Arztbeurteilung zu einem einzigen Fragebogen kombiniert. Der PHQ hat eine vergleichbare diagnostische Validität wie der Prime-MD, er ist aber kürzer und damit effektiver. Der PHQ kann vom Patienten selbständig ausgefüllt werden oder falls nötig, dem Patienten auch vorgelesen werden (Spitzer et al., 1999).
Zur Stellung einer Diagnose wendet der Arzt die im Fragebogen angegebenen Auswertungsalgorithmen an. Lediglich für die Auswertung der somatoformen Störungen ist zusätzlich das Urteil des Arztes hinsichtlich der Adäquatheit der Erklärungen des Patienten für die bei ihm vorhandenen körperlichen Symptome notwendig.

In seiner ausführlichesten Version umfasst der PHQ drei Seiten. Eine vierte Seite, die Fragen zur Menstruation, Schwangerschaft und Geburt sowie nach in der letzten Zeit vorhandenen psychosozialen Stressoren stellt, ist optional. Die zweiseitige Version wird als *Brief Patient Health Questionnaire* (B-PHQ) bezeichnet und enthält affektive Störungen und die Diagnose Paniksyndrom sowie nichtdiagnostische Informationen. Ferner kann auch nur die erste Seite dieser zweiseitigen Version verwendet werden, die affektive Störungen und die Diagnose Paniksyndrom enthält (Spitzer et al., 1999).

Der PHQ bewertet acht Diagnosen, die in unterschwellige und volle DSM-IV Diagnosen eingeteilt sind. Unterschwellige Diagnosen haben weniger Symptome als für eine DSM-IV Diagnose nötig sind. Bei den Diagnosen des PHQ handelt es sich um somatische Störung, Major Depression, geringgradige depressive Störung, Panikstörung, andere Angststörungen, Bulimia Nervosa, Binge Eating Disorder und Alkoholmissbrauch. Die in der vorliegenden Studie verwendete einseitige Version des B-PHQ ermöglicht die Erfassung der drei folgenden Diagnosen: Major Depression, geringgradige depressive Störung und Paniksyndrom.

Eine wichtige Modifikation wurde bei den Antwortkategorien vorgenommen. In der Originalversion des Prime-MD sind die Antwortmöglichkeiten bei den affektiven und bei den somatoformen Störungen lediglich dichotom. Im PHQ ist die Abstufung der Antwortkategorien bei den affektiven Störungen vierstufig. Die Frage lautet: „Wie oft fühlten Sie sich im Verlauf der letzten zwei Wochen durch die folgenden Beschwerden beeinträchtigt?“

Die Antwortmöglichkeiten sind: *Überhaupt nicht, an einzelnen Tagen, an mehr als der Hälfte der Tage* und *beinahe jeden Tag.* Diese Neuerung macht es möglich, dass der PHQ nicht nur als diagnostisches Instrument dient, sondern auch der Schweregrad der depressiven Störung bestimmt werden kann (Spitzer et al., 1999).

Bei den somatoformen Störungen wird der Patient gebeten, für jede der dreizehn Beschwerden anzugeben, ob er im letzten Monat davon *nicht*, *wenig* oder *stark* beeinträchtigt wurde. Im Prime-MD sind die Antwortkategorien dieser Sektion lediglich dichotom. Dies hat aufgrund der großen Häufigkeit von kör-

perlichen Beschwerden zur Folge, dass viele Patienten Beschwerden vermerken, die klinisch nicht signifikant sind.

Am Ende des Fragebogens wurde zusätzlich die folgende Frage eingefügt: „Wenn eines oder mehrere der bisher in diesem Fragebogen beschriebenen Probleme bei Ihnen vorliegen, geben Sie bitte an, wie sehr diese Probleme es Ihnen erschwert haben, Ihre Arbeit zu tun, Ihren Haushalt zu regeln oder mit anderen Menschen zurecht zu kommen“. Die Antwortkategorie ist vierstufig: *Überhaupt nicht erschwert, etwas erschwert, relativ stark erschwert* und *sehr stark erschwert*. Zusammen mit den anderen Symptomen des Fragebogens stellt sie einen guten Indikator für die Wahrscheinlichkeit einer psychiatrischen Diagnose und für die funktionelle Beeinträchtigung dar.

Wie auch beim Prime-MD sollte der Kliniker, bevor er eine abschließende Diagnose stellt, körperliche Ursachen für Angststörungen und depressive Störungen ausschließen. Zusätzlich sollte bei Depressiven Störungen überprüft werden, ob es sich um eine normale Trauerreaktion handelt und ob in der Vorgeschichte eine oder mehrere manische Episoden vorhanden waren.

Nach Spitzer et al. (1999) kann bei der Auswertung des PHQ die im DSM-IV geforderte Anzahl von fünf Symptomen für die Diagnose Major Depression übernommen werden. Da jedoch die Antwortkategorien von zwei auf vier ausgedehnt wurden, hat es sich als sinnvoll erwiesen, die Schwelle der Beeinträchtigung in den letzten zwei Wochen von *an fast allen Tagen* wie sie in DSM-IV gefordert wird auf *an mehr als der Hälfte der Tage* zu senken. Dies führt zu einer beträchtlichen Erhöhung der Sensitivität von 37% auf 73% bei einer gleichbleibend hohen Spezifität von 94%. Die Sensitivität des PHQ ist für affektive Störungen vergleichbar mit der Sensitivität des Prime-MD, 61% versus 67%. Für die Diagnose Major Depression ist die Sensitivität des PHQ höher als die Sensitivität des Prime-MD, 73% versus 57%. Die Spezifität ist ebenfalls vergleichbar. Sie beträgt 94% beim PHQ versus 92% beim Prime-MD für affektive Störungen und 98% beim PHQ versus 94% beim Prime-MD für Major Depression.

In der vorliegenden Arbeit wurde ausschließlich die kürzeste Version des PHQ, die aus der ersten Seite des Brief Patient Health Questionnaire (B-PHQ) besteht, verwendet. Abbildung 18 zeigt diese Version des B-PHQ.

1. Wie oft fühlten Sie sich im Verlauf der letzten 2 Wochen durch die folgenden Beschwerden beeinträchtigt?	**Überhaupt nicht**	**An einzelnen Tagen**	**An mehr als der Hälfte der Tage**	**Beinahe jeden Tag**
a Wenig Interesse oder Freude an Ihren Tätigkeiten	☐	☐	☐	☐
b Niedergeschlagenheit, Schwermut, Hoffnungslosigkeit	☐	☐	☐	☐
c Schwierigkeiten ein- oder durchzuschlafen oder vermehrter Schlaf	☐	☐	☐	☐
d Müdigkeit oder das Gefühl, keine Energie zu haben	☐	☐	☐	☐
e Verminderter Appetit oder übermäßiges Bedürfnis zu essen	☐	☐	☐	☐
f Schlechte Meinung von sich selbst; das Gefühl ein Versager zu sein oder die Familie enttäuscht zu haben	☐	☐	☐	☐
g Schwierigkeiten, sich auf etwas zu konzentrieren, z.B. beim Zeitung lesen oder Fernsehen	☐	☐	☐	☐
h Waren Ihre Bewegungen oder Ihre Sprache so verlangsamt, dass es auch anderen auffallen würde? Oder waren Sie im Gegenteil „zappelig" oder ruhelos und hatten dadurch einen stärkeren Bewegungsdrang als sonst?	☐	☐	☐	☐
i Gedanken, dass Sie lieber tot wären oder sich Leid zufügen möchten	☐	☐	☐	☐

2. Fragen zum Thema „Angst"	**NEIN**	**JA**
a Hatten Sie in den letzten 4 Wochen eine Angstattacke (plötzliches Gefühl der Furcht oder Panik?)	☐	☐
Wenn „NEIN", gehen Sie bitte weiter zu Frage 3		
b Ist dies bereits früher einmal vorgekommen?	☐	☐
c Treten manche dieser Anfälle völlig unerwartet auf – d.h. in Situationen, in denen Sie nicht damit rechnen, dass Sie angespannt oder beunruhigt reagieren?	☐	☐
d Empfinden Sie diese Anfälle als stark beeinträchtigend, und/oder haben Sie Angst vor erneuten Anfällen?	☐	☐
e Hatten Sie während Ihrem letzten schlimmen Angstanfall Gefühle wie Kurzatmigkeit, Schwitzen, Herzrasen oder -klopfen, Schwindel oder Ohnmachtsgefühle, Kribbeln oder Taubheitsgefühle, Übelkeit oder Magenbeschwerden?	☐	☐

3. Wenn eines oder mehrere der bisher in diesem Fragebogen beschriebenen Probleme bei Ihnen vorliegen, geben Sie bitte an, wie sehr diese Probleme es Ihnen erschwert haben, Ihre Arbeit zu tun, Ihren Haushalt zu regeln oder mit anderen Menschen zurecht zu kommen?

Überhaupt nicht erschwert	Etwas erschwert	Relativ stark erschwert	Sehr stark erschwert
☐	☐	☐	☐

Abbildung 18: Erste Seite des B-PHQ von Spitzer et al. (1999; © Pfizer, Karlsruhe).

Nach Spitzer et al. (1999) können folgende drei Diagnosen mit dieser Version des B-PHQ gestellt werden: Major Depression, geringgradige depressive Störung und Paniksyndrom.

Die Diagnose Major Depression wird gestellt, wenn fünf oder mehr Items von 1a bis 1i positiv, also mit *an mehr als der Hälfte der Tage* oder *beinahe jeden Tag* beantwortet werden und darunter die Fragen 1a oder 1b sind. Die Diagnose geringgradige depressive Störung wird gestellt, wenn zwei bis vier der Items von 1a bis 1i positiv, d.h. mit *an mehr als der Hälfte der Tage* oder *beinahe jeden Tag* beantwortet werden und darunter die Items 1a oder 1b sind. Für beide Diagnosen zählt die Frage 1i bereits als positiv beantwortet, wenn sie mit *an einzelnen Tagen* beantwortet ist. Die Diagnose Paniksyndrom wird gestellt, wenn alle Fragen von 2a bis e mit *ja* beantwortet sind.

In der vorliegenden Arbeit wird für die Validierung des B-PHQ anhand der aktuellen Depressionsdiagnosen des DIA-X die Depressionssektion des B-PHQ, also die Fragen 1a bis 1i, betrachtet. Die Diagnose Paniksyndrom spielt für die weiteren Berechnungen somit keine Rolle.

Für die Festlegung der Cut-Off-Werte werden zwei Auswertungsalgorithmen verwendet. Der Auswertungsalgorihmus A folgt dem Auswertungsschema von Spitzer et al. (1999), das entsprechend den Kriterien des DSM-IV für die Verdachtsdiagnose einer Depression in jedem Fall die positive Beantwortung der Items 1a oder 1b erfordert. Nach diesem Schema sind somit neun verschiedene Schwellenwerte denkbar. Die niedrigste Schwelle lautet: Es ist mindestens die Frage 1a oder die Frage 1b positiv beantwortet. Der nächst höhere Schwellenwert ist erreicht wenn mindestens zwei der Items 1a bis 1i positiv beantwortet werden und darunter sind die Fragen 1a oder 1b sind. Dieses Muster wird so lange fortgesetzt, bis alle neun Items positiv beantwortet sind und somit der letzte Schwellenwert erreicht ist.

Der zweite, einfachere Auswertungsalgorithmus B orientiert sich an der in dieser Arbeit verwendeten Definition der Diagnose „unterschwelliges depressives Syndrom“, die nicht das Vorhandensein von mindestens einem der beiden Symptome „depressive Stimmung“ oder „Verlust von Interesse oder Freude“ fordert (vgl. Kapitel 1, Gliederungspunkt 1.1). Der Auswertungsalgorithmus B verzichtet daher auf die Forderung dass die Items 1a oder 1b zwingend positiv

beantwortet werden müssen. Der erste Schwellenwert lautet somit: Es ist mindestens eine der neun Fragen von 1a bis 1i positiv beantwortet. Der nächste Cut-Off ist erreicht, wenn mindesten zwei Fragen von 1a bis 1i positiv beantwortet werden. Der Gesamtscore kann somit zwischen 0 und maximal 9 positiv beantworteten Items liegen. Der Bereich der Cut-Off-Werte reicht dann von 0/1 bis 8/9.

Nach Spitzer et al. (1999) sollte der PHQ an alle neuen Patienten einer Praxis, an alle Patienten, bei denen eine psychiatrische Diagnose vermutet wird sowie an alle bereits bekannten Patienten in regelmäßigen Zeitintervallen (ca. jedes Jahr) ausgehändigt werden. Somit könnte der PHQ zu einem nützlichen klinischen Werkzeug werden, das der Verbesserung der Diagnostik und Behandlung psychischer Störungen dient.

3.4 Goldstandard: DIA-X, eine rechnergestützte Version des CIDI

Beim *Diagnostischen Expertensystem für Psychische Störungen* (*DIA-X*, Wittchen & Pfister, 1997a) handelt es sich um ein standardisiertes klinisches Interview, das die Diagnosenstellung entsprechend den Kriterien der aktuellen Klassifikationssysteme ICD-10 und DSM-IV ermöglicht. Das DIA-X (...) „beruht auf dem WHO-CIDI 1.2, das entsprechend der DSM-IV-Forschungskriterien und Fragen modifiziert und erweitert wurde, aber auch den Forschungskriterien und Fragen nach ICD-10 entspricht“ (Wittchen & Pfister, 1997a, S. 74). Das Composite International Diagnostic Interview (CIDI) ist ein voll standardisiertes und umfassendes klinisches Interview zur Bewertung von lebenszeitbezogenen und Querschnittsdiagnosen nach ICD-10 und DSM-IV (vgl. Andrews & Peters, 1997; Brickenkamp, 1997). Die Entwicklung des CIDIs begann Anfang der 80er Jahre im Rahmen des „Joint Project on Diagnosis and Classification of Mental Disorders, Alcohol- and Drug-Related Problems“, eines Projektes der Weltgesundheitsbehörde (WHO) und der US Alcohol Drug Abuse and Mental Health Administration (ADAMHA) (Robins, Wing et al., 1988; Wittchen & Pfister, 1997b; Wittchen, Beloch et al., 1995). Der CIDI basiert auf dem *Present State Examination* (PSE, Wing, Cooper & Sartorius, 1974) und dem *Diagnostic Interview*

Schedule (DIS, Robins, Helzer, Croughan, Williams & Spitzer, 1981) (vgl. Robins & Sartorius, 1993; Stieglitz, Dittmann & Mombour, 1992).

Das Composite International Diagnostic Interview wurde ursprünglich für den Einsatz in epidemiologischen Untersuchungen entwickelt. Heute liegt der CIDI in 16 Sprachen vor und wird in zahlreichen internationalen Forschungsprojekten eingesetzt (Essau & Wittchen, 1993). Die Durchführung des Interviews erfordert ein einwöchiges Training und kann auch durch Nicht-Kliniker erfolgen. Neben den Fragen sind auch die Kodierung, die Dateneingabe und die computergestützte Auswertung standardisiert (Wittchen, 1993; Wittchen, 1994).

Das DIA-X selbst besteht aus drei Komponenten: Den DIA-X Fragebögen, dem DIA-X Interview und dem PC-Auswertungsprogramm des DIA-X Interviews. Das Interview und die Fragebögen können sowohl gemeinsam als auch unabhängig voneinander eingesetzt werden. Im sogenannten *Stamm-Screening-Questionnaire (SSQ)* werden die 16 „Stammfragen" des DIA-X Interviews gestellt. Bei den Stammfragen handelt es sich um die Eingangskriterien jeder Sektion, die zur Diagnosestellung erforderlich sind. Wird eine Stammfrage vom Probanden verneint, so kann mit großer Sicherheit davon ausgegangen werden, dass die entsprechene Störung nicht vorhanden ist. Der *Anxiety Questionnaire (ASQ)* erfasst mit Hilfe von 15 Fragen das Vorhandensein von Angsterkrankungen. Anhand von 15 Fragen werden im *Depressions-Screening-Questionnaire (DSQ)* depressive Störungen sowie deren Schweregrad erfasst. Ein vierter Fragebogen, der *PMS (Prämenstruelles Syndrom)* befindet sich in Vorbereitung.

In der vorliegenden Arbeit wurde das rechnergestützte DIA-X Interview und das PC-Auswertungsprogramm in der Version 1.2 (deutsch) von 2000 (© Max Planck Institut für Psychiatrie, München) verwendet. Ein großer Vorteil des DIA-X Interviews liegt in seinem Aufbau aus verschiedenen diagnostischen Modulen. Diese können in ihrer Gesamtheit oder entsprechend der Fragestellung flexibel eingesetzt werden. Die Module des DIA-X sind in Tabelle 4 aufgelistet. „Mit dem DIA-X Interview können über 100 ICD-10- und DSM-IV-kompatible Diagnoseklassen über Symptom-, psychosoziale Einschränkungs- und zeitbezogene Fragen beurteilt werden" (Wittchen & Pfister, 1997b, S. XII). Es handelt sich dabei um die häufigsten Diagnosen des Jugend- und Erwach-

senenalters. Nicht beinhaltet sind Persönlichkeitsstörungen, psychische und Verhaltensstörungen des Kindes- und Jugendalters, sowie schwerste akute psychotische Erkrankungen und schwerste dementielle Erkrankungen. Das DIA-X kann ab einem Alter von 17 Jahren durchgeführt werden. Nicht sinnvoll ist der Einsatz des Interviews bei Probanden mit akuten Psychosen oder schwerer Demenz.

Das DIA-X liegt in zwei Versionen vor: Einer lebenszeit bezogenen Version, der sogenannten Lifetime-Version, und einer jeweils auf die letzten 12 Monate bezogene Querschnittsversion. Die Durchführung des Interviews dauert für die Lifetime-Version ungefähr 75 Minuten und für die 12-Monats-Querschnittsversion ungefähr 55 Minuten. In der vorliegenden Arbeit wurde die lebenszeitbezogene Version des DIA-X Interviews verwendet, da diese sowohl Lebenszeitdiagnosen als auch aktuelle Querschnittsdiagnosen sowie Informationen bezüglich des ersten und des letzten Auftretens der Symptomatik erfasst und über die computergestützte Auswertung ausgibt.

Tabelle 4: Module des DIA-X Interviews (vgl. Wittchen & Pfister, 1997b).

Sektion	Inhalt
A	Demographischer Teil
B	Störungen durch Tabak
C	Somatoforme und Dissoziative Störungen
D	Phobien und andere Angststörungen
E	Depressive Störungen und Dysthymie
F	Manie und Bipolare Affektive Störungen
G	Schizophrene und andere psychotische Störungen
H	Essstörungen
I	Störungen durch Alkohol
K	Zwangsstörungen
L	Drogenmissbrauch und –abhängigkeit
M	Organisch bedingte psychische Störungen
N	Posttraumatische Belastungsstörungen
P	Interviewbeobachtungen
Q	Abschlussfragen und Fragebögen
X	Interviewbeurteilungen

Wittchen (Wittchen & Pfister, 1997b, S. 2) nennt folgende Einsatzgebiete des DIA-X Interviews:

- In der Praxis zur routinemäßigen klinischen Eingangsdiagnostik sowie
- zur Ausschlussdiagnostik,
- zur Feststellung von Lebenszeit- und Querschnittskomorbiditätsmustern,
- zur Feststellung von zeitlichen Beziehungen (Symptomprogression),
- für den Einsatz in epidemiologischen Forschungsprojekten,
- für Standarddokumentationszwecke in Kliniken und anderen Einrichtungen,
- für den Einsatz bei klinischen Projekten,
- und in der Lehre sowie der Fort- und Weiterbildung.

Das DIA-X kann von Klinikern und Nicht-Klinikern gleichermaßen durchgeführt werden, da die Fragestellungen voll standardisiert und wörtlich vorzulesen sind. Lediglich die Sektion *X: Interviewbeurteilungen* erfordert das Urteil des Untersuchers. Eine zweitägige Interviewerschulung wird von Wittchen & Pfister (1997b) jedoch nachdrücklich empfohlen. Die Fragen sind einfach formuliert, damit sie in unterschiedlichen Kulturen und unabhängig vom Bildungsstand der Probanden beantwortet werden können.

Die Reliabilität des DIA-X ist nach Wittchen (Wittchen & Pfister, 1997b) als sehr gut zu bezeichnen. So liegt die Inter-Rater-Reliabilität für Symptominformationen bei Kappa-Werten von .81 bis 1.0 und für diagnostische Entscheidungen bei Kappa-Werten von .82 bis .98. Auch die Test-Retest-Reliabilität erreicht für Wiederholungen in ein- bis vierzehntägigem Abstand befriedigende Kappa-Werte. Für die Dysthymie (F34.1) ergibt sich beispielsweise ein Kappa-Wert von .51 und für rezidivierende depressive Störungen (F33) ein Kappa-Wert von .62. Der Kappa-Wert für depressive Episoden (F32) liegt bei .69 und für depressive Störungen (F32/F33) bei .77. Auch die Validität ist nach Wittchen (Wittchen & Pfister, 1997b) befriedigend hoch. Der Vergleich mit Konsensus-Diagnosen erfahrener Psychiater ergab Kappa-Werte von .39 für psychotische Störungen bis maximal .82 für Panikstörungen.

In Tabelle 5 sind die, in der vorliegenden Arbeit verwendeten Sektionen des DIA-X Interviews aufgelistet.

Tabelle 5: In der vorliegenden Arbeit verwendete Module des DIA-X Interviews.

Sektion	Inhalt
A	Demographischer Teil
C	Somatoforme und Dissoziative Störungen
D	Phobien und andere Angststörungen
E	Depressive Störungen und Dysthymie
F	Manie und Bipolare Affektive Störungen
G	Schizophrene und andere psychotische Störungen
H	Eßstörungen
K	Zwangsstörungen
M	Organisch bedingte psychische Störungen
P	Interviewbeobachtungen
Q	Abschlußfragen und Fragebögen
X	Interviewbeurteilungen

Besondere Aufmerksamkeit galt dabei der Diagnostik von depressiven Störungen. Die Sektion *E (Depressive Störungen und Dysthymie)* wurde ergänzt durch das DIA-X Zusatzmodul *Recurrent Brief Depression* und um die Diagnose unterschwelliges depressives Syndrom erweitert. Wie bereits in Kapitel 1 (siehe Gliederungspunkt 1.1) beschrieben wurde, wird in Anlehnung an die Kriterien des ICD-10 und des DSM-IV die aktuelle Diagnose „unterschwelliges depressives Syndrom“ dann gestellt, wenn mindestens zwei verschiedene Symptome einer depressiven Episode in den letzten beiden Wochen vor der Durchführung des DIA-X Interviews erfüllt waren, aber noch nicht die Kriterien für eine volle affektive Störung erfüllt sind. In der Vorgeschichte können affektive Störungen nach den Kriterien der ICD-10 und des DSM-IV vorgelegen haben.

4 Ergebnisse

Während der Datenerhebungsphase wurden insgesamt 566 DIA-X Interviews durchgeführt. In die hier vorliegende Datenauswertung gingen die vollständigen Datensätze von N = 448 Patienten ein. Ein Datensatz gilt als vollständig, wenn alle Fragen in den Screeninginstrumenten beantwortet wurden und das klinisch-diagnostische Interview (DIA-X) innerhalb von 6 Tagen nach dem Ausfüllen der Fragebögen durchgeführt wurde.

4.1 Stichprobencharakteristika und Prävalenz

4.1.1 Alters- und Geschlechtsverteilung

283 (63%) Patienten sind weiblich und 165 (37%) männlich. Das Alter der Patienten liegt zwischen 18 und 88 Jahren. Der Mittelwert liegt bei 53,01 Jahren, die Standardabweichung beträgt 16,60. Der Median liegt bei 56,21 Jahren. Die Altersverteilung der gesamten Stichprobe ist in Abbildung 19 dargestellt.

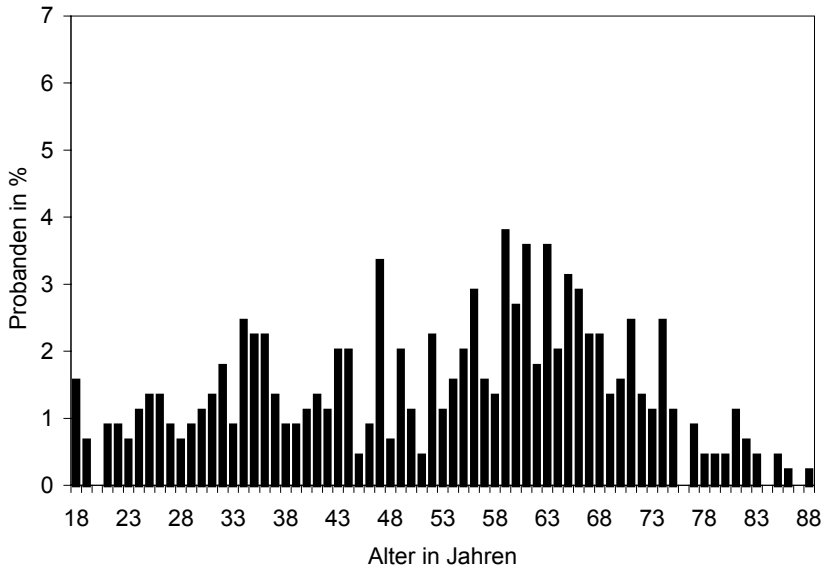


Abbildung 19: Altersverteilung der gesamten Stichprobe (N = 448).

Die Datenauswertung wurde zum einen für die gesamte Stichprobe und zum anderen getrennt nach den Variablen Geschlecht und Alter durchgeführt. In Abbildung 20 ist die Altersverteilung der 283 Frauen wiedergegeben. Die Altersverteilung in der Subgruppe der Frauen reicht von 18 bis 88 Jahre. Der Mittelwert liegt bei 53,05 Jahren, die Standardabweichung beträgt 16,28.

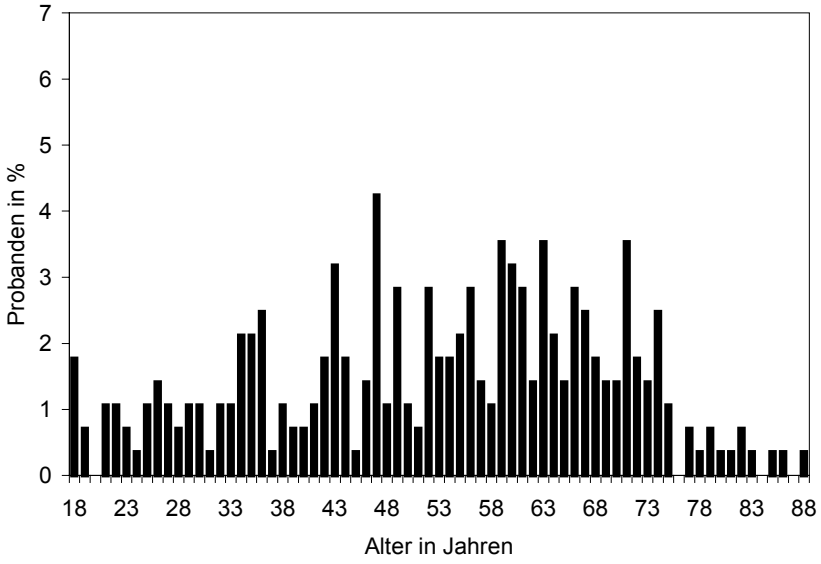


Abbildung 20: Altersverteilung in der Subgruppe der Frauen (N = 283).

Abbildung 21 zeigt die Altersverteilung der 165 Männer. In der Subgruppe der Männer liegt das Alter zwischen 18 und 85 Jahren. Der Mittelwert liegt bei 52,95 Jahren, die Standardabweichung beträgt 17,14.

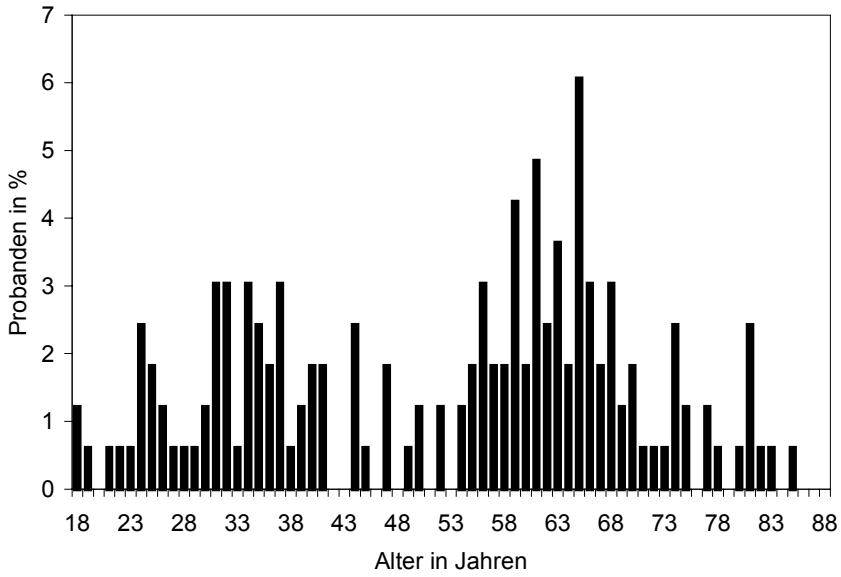


Abbildung 21: Altersverteilung in der Subgruppe der Männer (N = 165).

Tabelle 6 zeigt die Verteilung der verschiedenen Altersgruppen in der gesamten Stichprobe.

Tabelle 6: Altersverteilung der gesamten Stichprobe (N = 448).

Alter (Jahre)	Absolute Anzahl	Anzahl in Prozent (%)
18 bis 24 Jahre	26	5,80%
25 bis 34 Jahre	56	12,50%
35 bis 44 Jahre	68	15,18%
45 bis 54 Jahre	63	14,06%
55 bis 64 Jahre	113	25,22%
65 bis 74 Jahre	93	20,76%
75 bis 88 Jahre	29	6,47%

Um zu überprüfen ob die beiden Fragebögen depressive Störungen in verschiedene Altersgruppen valide erfassen, wurde ein Mediansplit durchgeführt. Der Median liegt bei 56,21 Jahren, d.h. die Variable Alter ist bei 224 Probanden kleiner als 56,21 Jahre. Davon sind 149 (67%) Frauen und 75 (33%) Männer. Bei 224 Probanden ist die Variable Alter größer als 56,20 Jahre. In dieser Subgruppe sind 134 (60%) Frauen und 90 (40%) Männer.

4.1.2 Ergebnisse der Auswertung der Screeninginstrumente

Ergebnisse der Auswertung des GHQ-12

In Abbildung 22 ist die Verteilung der GHQ-12 Scores (%) in der gesamten Stichprobe (N= 448) wiedergegeben.

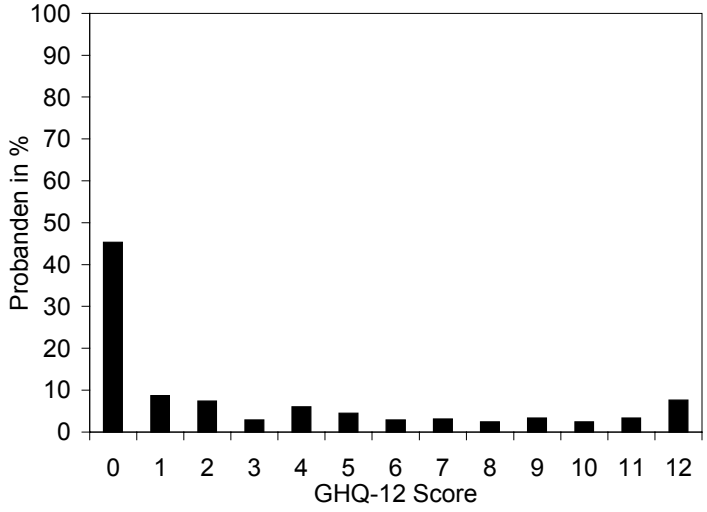


Abbildung 22: Verteilung der GHQ-12 Scores (%) in der gesamten Stichprobe (N = 448).

Abbildung 23 zeigt die Verteilung der GHQ-12 Scores (%) getrennt nach der Variable Geschlecht (N= 448).

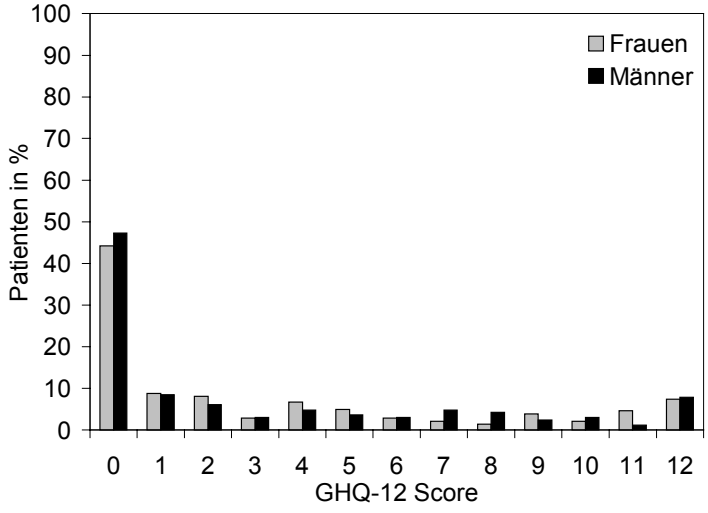


Abbildung 23: Verteilung der GHQ-12 Scores (%) getrennt nach der Variable Geschlecht (N = 448).

In Abbildung 24 ist die Verteilung der GHQ-12 Scores (%) getrennt nach Altersgruppen (N= 448) dargestellt.

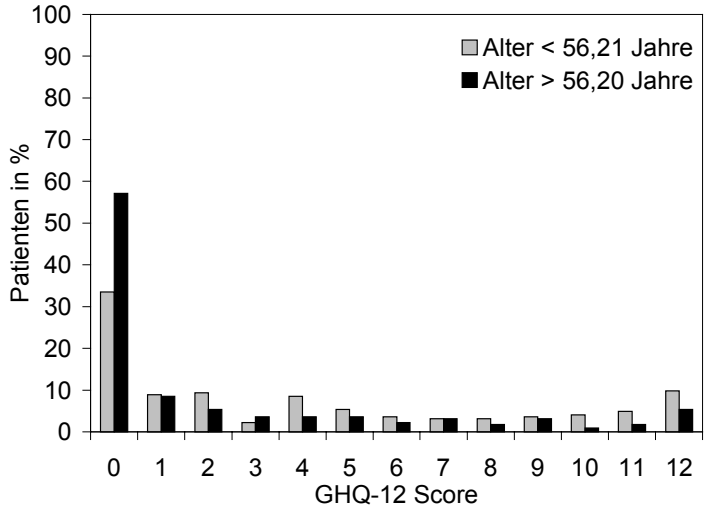


Abbildung 24: Verteilung der GHQ-12 Scores (%) getrennt nach Altersgruppen (N = 448).

Ergebnisse der Auswertung des B-PHQ Auswertungsalgorithmus A

In Abbildung 25 ist die Verteilung der B-PHQ Auswertungsalgorithmus A Scores (%) in der gesamten Stichprobe (N= 448) wiedergegeben. 0 bedeutet, dass die Fragen 1a und 1b nicht positiv beantwortet wurden. 1 heißt, dass die Frage 1a oder 1b positiv beantwortet wurde. 2 bedeutet, dass zwei Fragen positiv beantwortet wurden und darunter die Fragen 1a oder 1b sind. Der letzte Wert 9 bedeutet, dass alle neun Fragen von 1a bis 1i positiv beantwortet wurden.

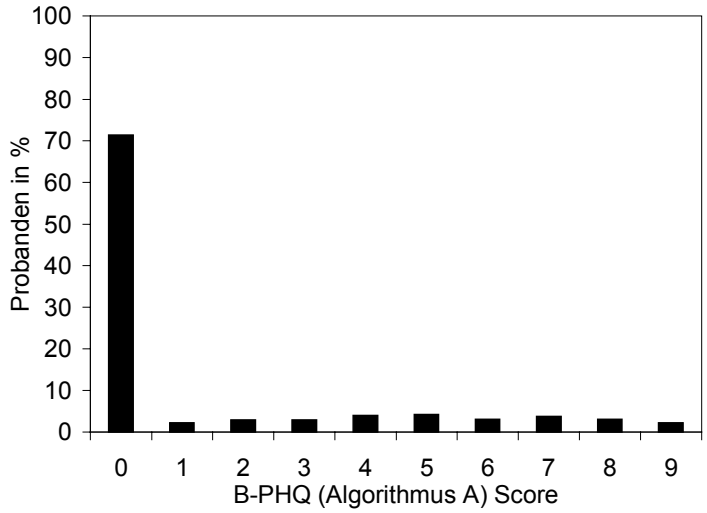


Abbildung 25: Verteilung der B-PHQ Auswertungsalgorithmus A Scores (%) in der gesamten Stichprobe (N = 448).

Abbildung 26 zeigt die Verteilung der B-PHQ Auswertungsalgorithmus A Scores (%) getrennt nach der Variable Geschlecht (N= 448).

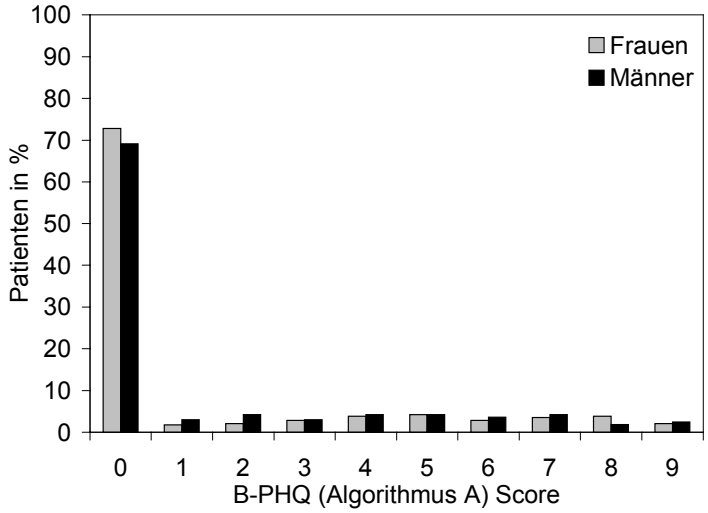


Abbildung 26: Verteilung der B-PHQ Auswertungsalgorithmus A Scores (%) getrennt nach der Variable Geschlecht (N = 448).

In Abbildung 27 ist die Verteilung der B-PHQ Auswertungsalgorithmus A Scores (%) getrennt nach Altersgruppen dargestellt.

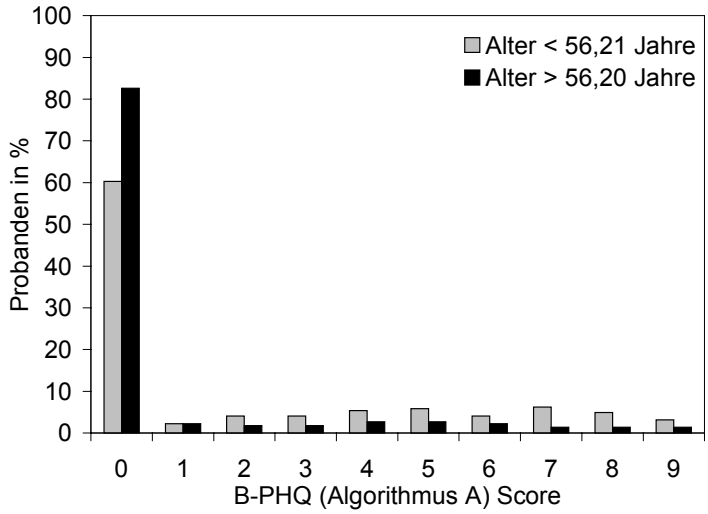


Abbildung 27: Verteilung der B-PHQ Auswertungsalgorithmus A Scores (%) getrennt nach Altersgruppen (N = 448).

Ergebnisse der Auswertung des B-PHQ Auswertungsalgorithmus B

Abbildung 28 zeigt die Verteilung der B-PHQ Auswertungsalgorithmus B Scores (%) in der gesamten Stichprobe (N= 448). 0 bedeutet, dass keine Frage positiv beantwortet wurde. 1 heißt, dass eine der Fragen von 1a bis 1i positiv beantwortet wurde. 2 bedeutet, dass zwei Fragen von 1a bis 1i positiv beantwortet wurden. Der letzte Wert ist 9, d.h. alle neun Items wurden positiv beantwortet.

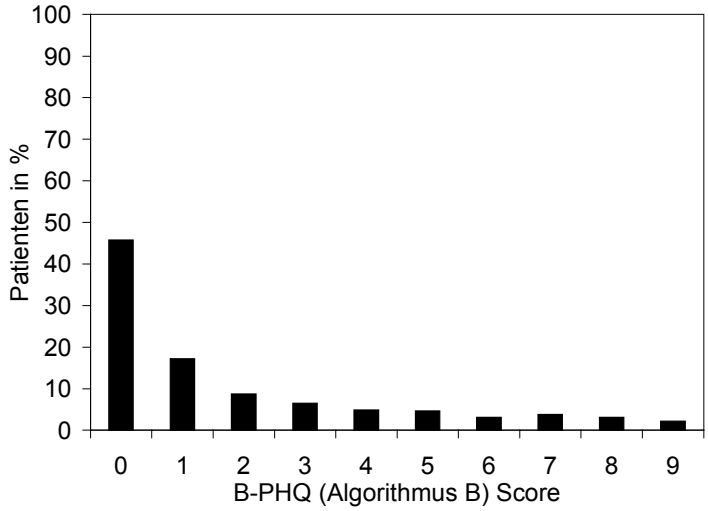


Abbildung 28: Verteilung der B-PHQ Auswertungsalgorithmus B Scores (%) in der gesamten Stichprobe (N = 448).

Abbildung 29 zeigt die Verteilung der B-PHQ Auswertungsalgorithmus B Scores (%) getrennt nach der Variable Geschlecht (N= 448).

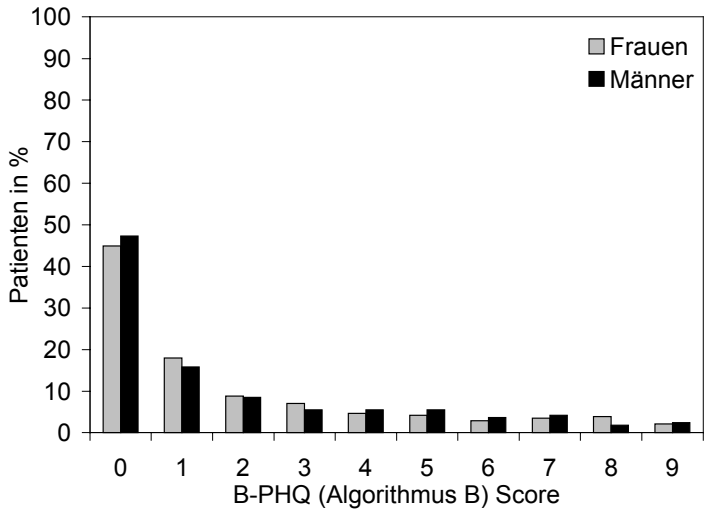


Abbildung 29: Verteilung der B-PHQ Auswertungsalgorithmus B Scores (%) getrennt nach der Variable Geschlecht (N = 448).

In Abbildung 30 ist die Verteilung der B-PHQ Auswertungsalgorithmus B Scores (%) getrennt nach Altersgruppen (N= 448) dargestellt.

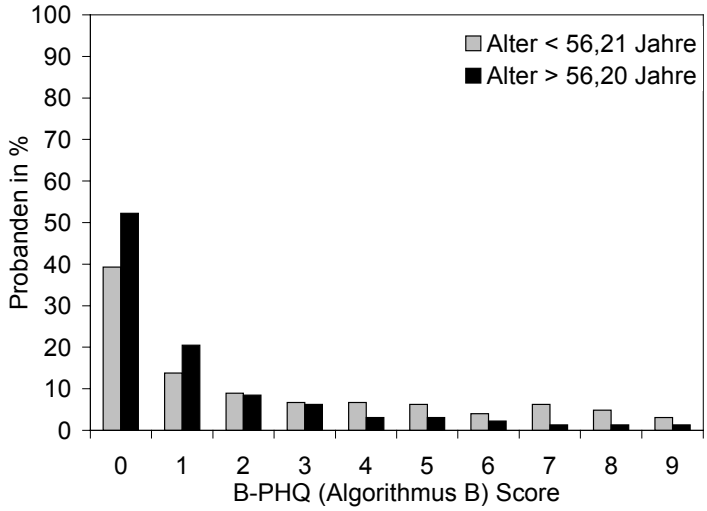


Abbildung 30: Verteilung der B-PHQ Auswertungsalgorithmus B Scores (%) getrennt nach Altersgruppen (N = 448).

4.1.3 Ergebnisse der Auswertung des DIA-X Interviews

In Tabelle 7 wird ein Überblick über die Depressionsdiagnosen und ihre korrespondierenden ICD-10 Kodenummern, die das DIA-X Modul *E Depressive Störungen und Dysthymie* und das DIA-X Modul *Recurrent Brief Depression* standardmäßig ermöglichen, gegeben.

Tabelle 7: Standarddepressionsdiagnosen des DIA-X Moduls E *Depressive Störungen und Dysthymie* sowie des Moduls *Recurrent Brief Depression* und deren korrespondierenden ICD-10 Kodierung (vgl. auch Wittchen & Pfister, 1997b).

F32.0	Major Depression, Einzelne Episode, Leicht
F32.1	Major Depression, Einzelne Episode, Mittelschwer
F32.2	Major Depression, Einzelne Episode, Schwer, ohne Psychotische Merkmale
F32.3	Major Depression, Einzelne Episode, Schwer, mit Psychotischen Merkmalen
F33.0	Major Depression, Rezidivierend, Leicht
F33.1	Major Depression, Rezidivierend, Mittelschwer
F33.2	Major Depression, Rezidivierend, Schwer, ohne Psychotische Merkmale
F33.3	Major Depression, Rezidivierend, Schwer, mit Psychotischen Merkmalen
F34.1	Dysthyme Störung
F31.31	Bipolar I Störung, Letzte Episode Depressiv, Leicht
F31.32	Bipolar I Störung, Letzte Episode Depressiv, Mittelschwer
F31.4	Bipolar I Störung, Letzte Episode Depressiv, Schwer ohne Psychotische Merkmale
F31.5	Bipolar I Störung, Letzte Episode Depressiv, Schwer mit Psychotischen Merkmalen
F31.61	Bipolar I Störung, Letzte Episode Gemischt, Leicht
F31.62	Bipolar I Störung, Letzte Episode Gemischt, Mittelschwer
F31.63	Bipolar I Störung, Letzte Episode Gemischt, Schwer ohne Psychotische Merkmale
F31.64	Bipolar I Störung, Letzte Episode Gemischt, Schwer mit Psychotischen Merkmalen
F31.81	Bipolar II Störung, Letzte Episode Depressiv, Leicht
F31.82	Bipolar II Störung, Letzte Episode Depressiv, Mittelschwer
F31.83	Bipolar II Störung, Letzte Episode Depressiv, Schwer ohne Psychotische Merkmale
F31.84	Bipolar II Störung, Letzte Episode Depressiv, Schwer mit Psychotischen Merkmalen
F06.32	Affektive Störung Aufgrund eines Medizinischen Krankheitsfaktors mit Depressiven Merkmalen
F06.33	Affektive Störung Aufgrund eines Medizinischen Krankheitsfaktors mit Gemischten Merkmalen
F1x.8	Substanzinduzierte Affektive Störungen
F38.10	Rezidivierende kurze depressive Störung

Wird in der vorliegenden Arbeit von *Standarddepressionsdiagnosen des DIA-X* gesprochen, so ist damit die Gesamtheit der in Tabelle 7 aufgelisteten Diagnosen gemeint.

Das in der vorliegenden Untersuchung verwendete DIA-X Interview wurde in Zusammenarbeit mit den Entwicklern des DIA-X dahingehend modifiziert, dass zusätzlich die Diagnose *unterschwelliges depressives Syndrom,* nach der in Kapitel 1 (siehe Gliederungspunkt 1.1) wiedergegebenen Definition, gestellt werden kann. Diese Diagnose hat keine korrespondierende ICD-10 Kodenummer, da sie keine volle Diagnose nach ICD-10 ist.

Bei den hier angegebenen Diagnosen handelt es sich um aktuelle Querschnittsdiagnosen, die sich auf den Zeitraum der letzten beiden Wochen vor der Durchführung des diagnostischen Interviews beziehen.

Von den insgesamt 448 Probanen haben 82 Patienten eine Standarddepressionsdiagnose. Die Prävalenz für die gesamte Stichprobe beträgt somit 18,30%. Für die Subgruppe der Frauen liegt die Prävalenz bei 20,14%, für die Subgruppe der Männer bei 15,16%. Für die Subgruppe der Probanden mit der Variable Alter < 56,21 liegt die Prävalenz bei 24,55% und für die Subgruppe der Patienten mit der Variable Alter > 56,20 bei 12,06%.

Zusätzlich haben insgesamt 54 Patienten die Diagnose *unterschwelliges depressives Syndrom* erhalten. Die Prävalenz für die Diagnose unterschwelliges depressives Syndrom liegt für die Gesamtstichprobe somit bei 12,05%. Für die Subgruppe der Frauen liegt die Prävalenz für die Diagnose unterschwelliges depressives Syndrom bei 10,60%, für die Subgruppe der Männer bei 14,55%. Für die Subgruppe der Probanden mit der Variable Alter < 56,21 liegt die Prävalenz bei 13,39% und für die Subgruppe der Patienten mit der Variable Alter > 56,20 bei 10,71%.

Die Gesamtprävalenz für depressive Störungen nach den DIA-X Depressionsmodulen einschließlich der Diagnose unterschwelliges depressives Syndrom beträgt 30,35%. Für die Subgruppe der Frauen liegt die Gesamtprävalenz bei 30,74% und für die Subgruppe der Männer bei 29,71%. Für die Subgruppe der Probanden mit der Variable Alter < 56,21 liegt die Prävalenz bei 37,94% und für die Subgruppe der Patienten mit der Variable Alter > 56,20 bei 22,77%.

Abbildung 31 zeigt die prozentuale Verteilung der aktuellen Depressionsdiagnosen für die Gesamtstichprobe. Unter der Gruppe *Major Depression* wurden folgende Diagnosen zusammengefasst: F32.0, F32.1, F32.2, F32.3, F33.0, F33.1, F33.2, F33.3. In die Kategorie *Bipolare Affektive Störung* wurden die Diagnosen F31.31, F31.32, F31.4, F31.5, F31.61, F31.62, F31.63, F31.64, F31.81, F31.82, F31.83, F31.84 eingeordnet. Zu *Organische Affektive Störung* wurden F06.32 und F 06.33 subsumiert. Die Gesamtprävalenz beträgt 30,35%.

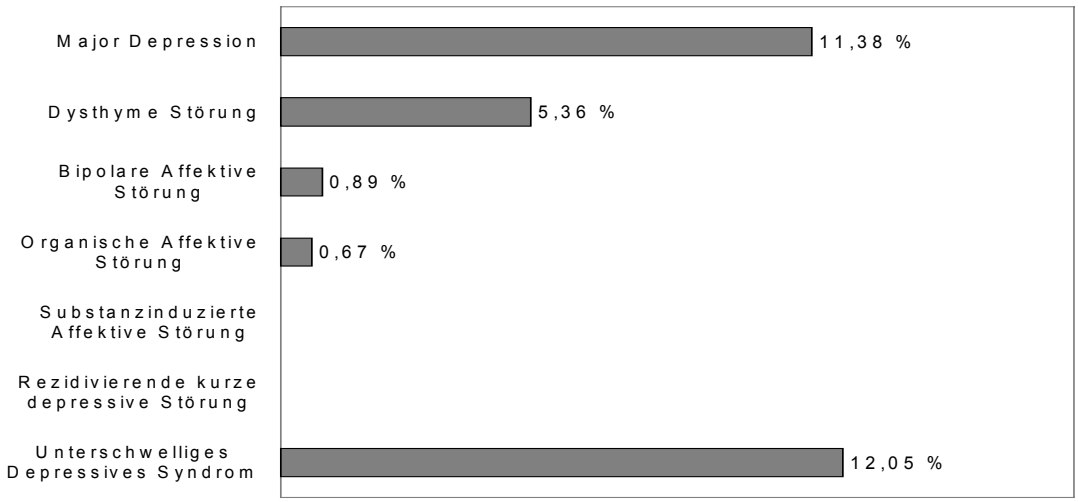


Abbildung 31: Prozentuale Verteilung der Depressionsdiagnosen in der gesamten Stichprobe (N = 448).

In Abbildung 32 ist die prozentuale Verteilung der Depressionsdiagnosen für die Subgruppe der Frauen dargestellt. Die Gesamtprävalenz beträgt 30,74%.

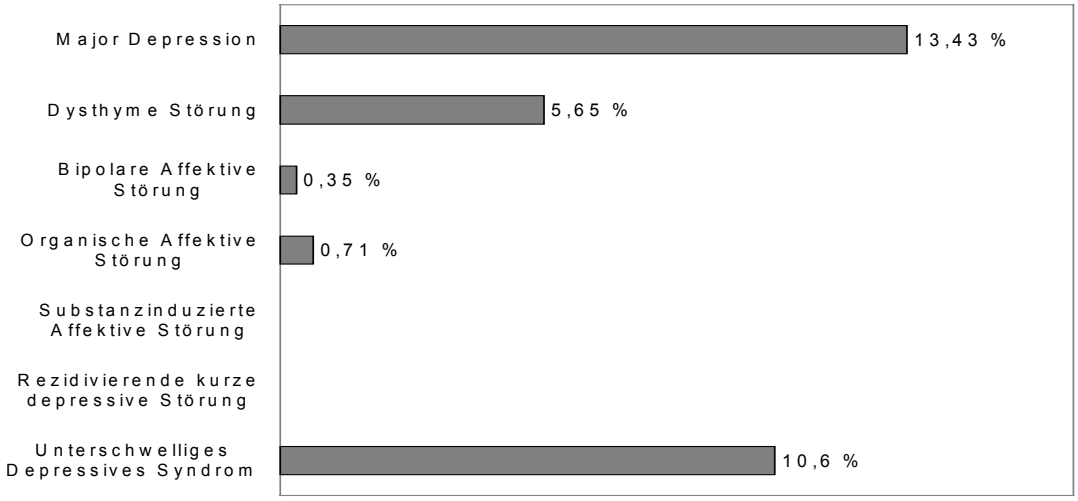


Abbildung 32: Prozentuale Verteilung der Depressionsdiagnosen für die Subgruppe der Frauen (N = 283).

In Abbildung 33 ist die prozentuale Verteilung der Depressionsdiagnosen für die Subgruppe der Männer abgebildet. Die Gesamtprävalenz beträgt 29,71%.

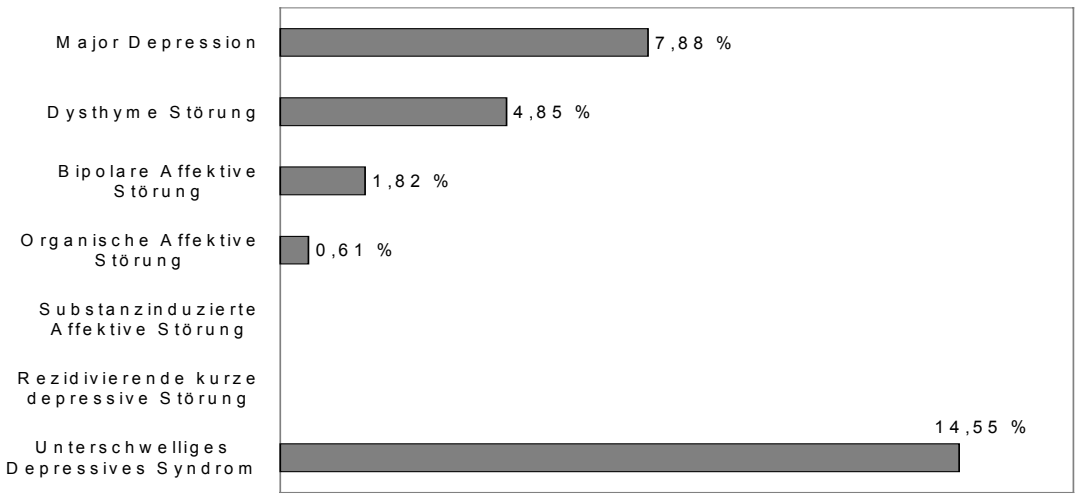


Abbildung 33: Prozentuale Verteilung der Depressionsdiagnosen für die Subgruppe der Männer (N = 165).

Abbildung 34 zeigt die prozentuale Verteilung der Depressionsdiagnosen für Probanden mit der Variable Alter < 56,21. Die Prävalenz ist 37,94%.

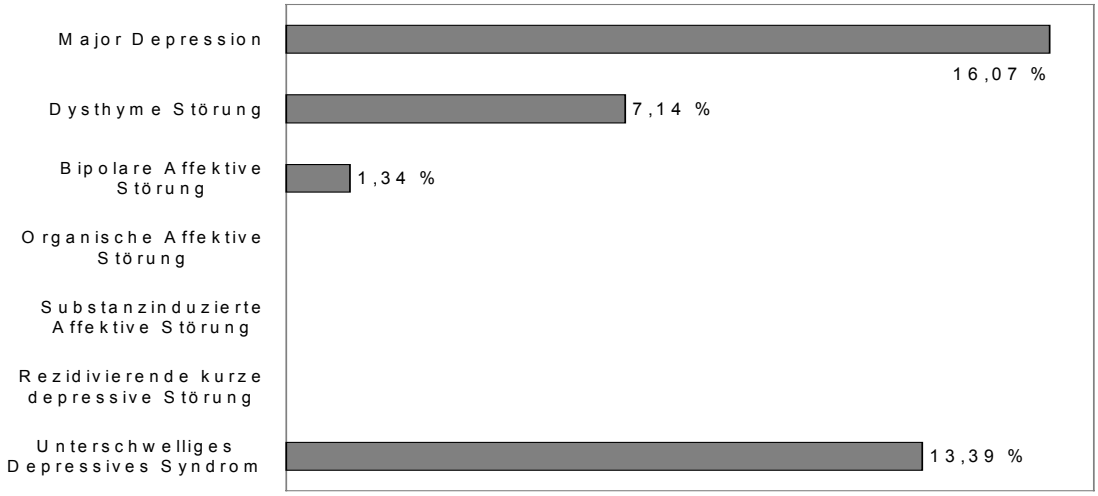


Abbildung 34: Prozentuale Verteilung der Depressionsdiagnosen für Probanden mit der Variable Alter < 56,21 Jahre (N = 224).

Abbildung 35 zeigt die prozentuale Verteilung der Depressionsdiagnosen für Probanden mit der Variable Alter > 56.20. Die Prävalenz ist 22,77%.

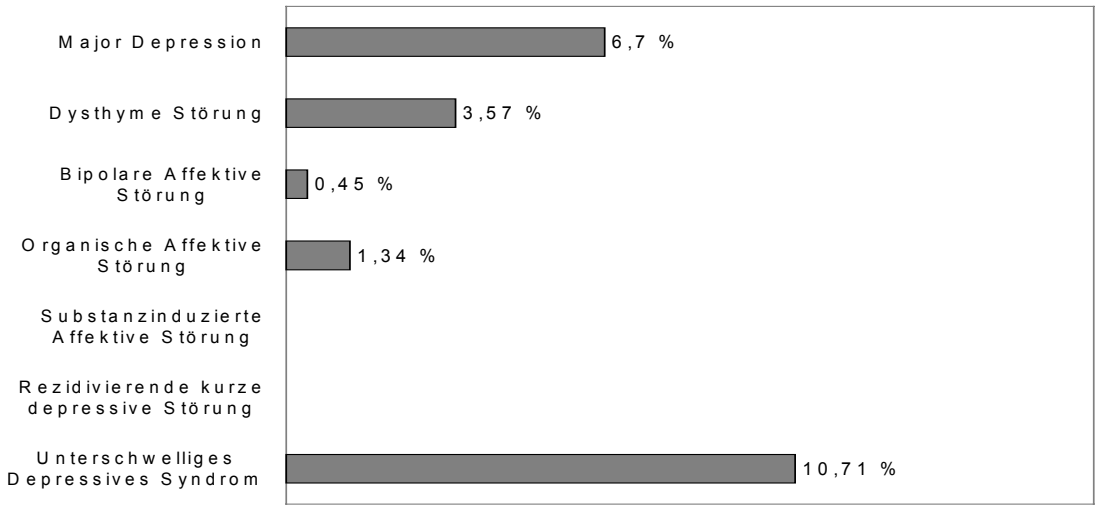


Abbildung 35: Prozentuale Verteilung der Depressionsdiagnosen für Probanden mit der Variable Alter > 56,20 Jahre (N = 224).

Tabelle 8 gibt eine Beschreibung des Durchschnittsalters und der Geschlechtsverteilung der folgenden fünf Untergruppen wieder:

1. Patienten mit einer Standarddepressionsdiagnose im DIA-X Interview.
2. Patienten mit der Diagnose *unterschwelliges depressives Syndrom.*
3. Patienten mit einer Standarddepressionsdiagnose oder der Diagnose *unterschwelliges depressives Syndrom,*
4. Patienten ohne Depressionsdiagnose, auch keine unterschwellige Diagnose.
5. Patienten ohne Standarddepressionsdiagnose.

Tabelle 8: Alter und Geschlecht der Gesunden und der Probanden mit einer Depressionsdiagnose (N = 448).

Diagnose	Anzahl	Alter	Geschlecht
Standarddepressionsdiagnose (ICD-10)	82	47,22 Jahre (SD = 15,02)	57 (69,5%) Frauen 25 (30,5%) Männer
Diagnose unterschwelliges depressives Syndrom	54	50,33 Jahre (SD = 17,25)	30 (56%) Frauen 24 (44%) Männer
Standarddepressionsdiagnose oder Diagnose unterschwelliges depressives Syndrom	136	48,59 Jahre (SD = 16,01)	87 (64%) Frauen 49 (36%) Männer
Keine Depressionsdiagnose, auch keine unterschwellige	312	54,94 Jahre (SD = 16,48)	196 (63%) Frauen 116 (37%) Männer
Keine Standarddepressionsdiagnose	366	54,26 Jahre (SD = 16,68)	226 (62%) Frauen 140 (38%) Männer

4.2 Sensitivität und Spezifität

4.2.1 Sensitivität und Spezifität des GHQ-12

Sensitivität und Spezifität des GHQ-12 für die gesamte Stichprobe

In Tabelle 9 sind die Sensitivität und die Spezifität der gesamten Stichprobe für die Cut-Off-Werte des GHQ-12, validiert anhand der Standarddepressionsdiagnosen des DIA-X, dargestellt.

Tabelle 9: Sensitivität und Spezifität des GHQ-12 für die gesamte Stichprobe validiert an den Standardepressionsdiagnosen des DIA-X (N = 448).

Cut-Off-Wert	**0/1**	**1/2**	**2/3**	**3/4**	**4/5**	**5/6**	**6/7**	**7/8**	**8/9**	**9/10**	**10/11**	**11/12**
Sensitivität	0.89	0.85	0.79	0.77	0.73	0.72	0.67	0.63	0.57	0.49	0.41	0.29
Spezifität	0.53	0.63	0.70	0.73	0.80	0.85	0.88	0.91	0.92	0.95	0.96	0.97

In Tabelle 10 sind die Sensitivität und die Spezifität der gesamten Stichprobe für die Cut-Off-Werte des GHQ-12, validiert an den Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom, abgebildet.

Tabelle 10: Sensitivität und Spezifität des GHQ-12 für die gesamte Stichprobe validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 448).

Cut-Off-Wert	**0/1**	**1/2**	**2/3**	**3/4**	**4/5**	**5/6**	**6/7**	**7/8**	**8/9**	**9/10**	**10/11**	**11/12**
Sensitivität	0.86	0.81	0.74	0.69	0.63	0.58	0.54	0.50	0.44	0.36	0.31	0.21
Spezifität	0.59	0.69	0.77	0.79	0.85	0.89	0.92	0.94	0.95	0.96	0.98	0.98

Sensitivität und Spezifität des GHQ-12 für die Subgruppe der Frauen

Tabelle 11 zeigt die Sensitivität und die Spezifität für die Subgruppe der Frauen für die Cut-Off-Werte des GHQ-12, validiert an den Standarddepressionsdiagnosen des DIA-X.

Tabelle 11: Sensitivität und Spezifität des GHQ-12 für die Subgruppe der Frauen validiert anhand der Standarddepressionsdiagnosen des DIA-X (N = 283).

Cut-Off-Wert	**0/1**	**1/2**	**2/3**	**3/4**	**4/5**	**5/6**	**6/7**	**7/8**	**8/9**	**9/10**	**10/11**	**11/12**
Sensitivität	0.89	0.88	0.81	0.77	0.74	0.72	0.67	0.63	0.58	0.49	0.42	0.28
Spezifität	0.53	0.63	0.72	0.74	0.82	0.88	0.90	0.92	0.92	0.95	0.96	0.98

In Tabelle 12 sind die Sensitivität und die Spezifität der Subgruppe der Frauen für die Cut-Off-Werte des GHQ-12, validiert an den Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom, wiedergegeben.

Tabelle 12: Sensitivität und Spezifität des GHQ-12 für die Subgruppe der Frauen validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 283).

Cut-Off-Wert	**0/1**	**1/2**	**2/3**	**3/4**	**4/5**	**5/6**	**6/7**	**7/8**	**8/9**	**9/10**	**10/11**	**11/12**
Sensitivität	0.87	0.83	0.77	0.71	0.63	0.60	0.55	0.51	0.46	0.37	0.32	0.18
Spezifität	0.58	0.69	0.78	0.80	0.86	0.91	0.93	0.94	0.94	0.96	0.97	0.97

Sensitivität und Spezifität des GHQ-12 für die Subgruppe der Männer

In Tabelle 13 sind die Sensitivität und die Spezifität für die Subgruppe der Männer für die Cut-Off-Werte des GHQ-12, validiert anhand der Standarddepressionsdiagnosen des DIA-X, dargestellt.

Tabelle 13: Sensitivität und Spezifität des GHQ-12 für die Subgruppe der Männer validiert an den Standarddepressionsdiagnosen des DIA-X (N = 165).

Cut-Off-Wert	**0/1**	**1/2**	**2/3**	**3/4**	**4/5**	**5/6**	**6/7**	**7/8**	**8/9**	**9/10**	**10/11**	**11/12**
Sensitivität	0.88	0.80	0.76	0.76	0.72	0.72	0.68	0.64	0.56	0.48	0.40	0.32
Spezifität	0.54	0.62	0.69	0.72	0.77	0.81	0.84	0.89	0.93	0.94	0.96	0.96

In Tabelle 14 sind die Sensitivität und die Spezifität der Subgruppe der Männer für die Cut-Off-Werte des GHQ-12, validiert an den Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom, abgebildet.

Tabelle 14: Sensitivität und Spezifität des GHQ-12 für die Subgruppe der Männer validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 165).

Cut-Off-Wert	**0/1**	**1/2**	**2/3**	**3/4**	**4/5**	**5/6**	**6/7**	**7/8**	**8/9**	**9/10**	**10/11**	**11/12**
Sensitivität	0.84	0.78	0.69	0.65	0.61	0.55	0.53	0.49	0.41	0.35	0.29	0.24
Spezifität	0.60	0.70	0.75	0.78	0.83	0.85	0.89	0.94	0.97	0.97	0.99	0.99

Sensitivität und Spezifität des GHQ-12 für die Subgruppe Alter < 56,21 Jahre

Tabelle 15 zeigt die Sensitivität und die Spezifität für die Subgruppe Alter < 56,21 Jahre für die Cut-Off-Werte des GHQ-12, validiert anhand der Standarddepressionsdiagnosen des DIA-X.

Tabelle 15: Sensitivität und Spezifität des GHQ-12 für die Subgruppe Alter < 56,21 Jahre validiert an den Standarddepressionsdiagnosen des DIA-X (N = 224).

Cut-Off-Wert	**0/1**	**1/2**	**2/3**	**3/4**	**4/5**	**5/6**	**6/7**	**7/8**	**8/9**	**9/10**	**10/11**	**11/12**
Sensitivität	0.93	0.91	0.85	0.84	0.78	0.78	0.75	0.73	0.64	0.55	0.45	0.31
Spezifität	0.42	0.53	0.64	0.66	0.76	0.83	0.86	0.90	0.91	0.93	0.95	0.97

In Tabelle 16 sind die Sensitivität und die Spezifität der Subgruppe Alter < 56,21 Jahre für die Cut-Off-Werte des GHQ-12, validiert an den Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom, abgebildet.

Tabelle 16: Sensitivität und Spezifität des GHQ-12 für die Subgruppe Alter < 56,21 Jahre validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 224).

Cut-Off-Wert	**0/1**	**1/2**	**2/3**	**3/4**	**4/5**	**5/6**	**6/7**	**7/8**	**8/9**	**9/10**	**10/11**	**11/12**
Sensitivität	0.91	0.88	0.80	0.79	0.69	0.65	0.62	0.59	0.51	0.44	0.38	0.25
Spezifität	0.48	0.61	0.71	0.74	0.82	0.88	0.92	0.95	0.95	0.96	0.99	0.99

Sensitivität und Spezifität des GHQ-12 für die Subgruppe Alter > 56,20 Jahre

In Tabelle 17 sind die Sensitivität und die Spezifität für die Subgruppe Alter > 56,20 Jahre für die Cut-Off-Werte des GHQ-12, validiert anhand der Standarddepressionsdiagnosen des DIA-X, dargestellt.

Tabelle 17: Sensitivität und Spezifität des GHQ-12 für die Subgruppe Alter > 56,20 Jahre validiert an den Standarddepressionsdiagnosen des DIA-X (N = 224).

Cut-Off-Wert	**0/1**	**1/2**	**2/3**	**3/4**	**4/5**	**5/6**	**6/7**	**7/8**	**8/9**	**9/10**	**10/11**	**11/12**
Sensitivität	0.81	0.74	0.67	0.63	0.63	0.59	0.52	0.44	0.44	0.37	0.33	0.26
Spezifität	0.62	0.71	0.76	0.80	0.84	0.87	0.89	0.91	0.93	0.96	0.96	0.97

In Tabelle 18 sind die Sensitivität und die Spezifität der Subgruppe Alter > 56,20 Jahre für die Cut-Off-Werte des GHQ-12, validiert an den Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom, abgebildet.

Tabelle 18: Sensitivität und Spezifität des GHQ-12 für die Subgruppe Alter > 56,20 Jahre validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 224).

Cut-Off-Wert	**0/1**	**1/2**	**2/3**	**3/4**	**4/5**	**5/6**	**6/7**	**7/8**	**8/9**	**9/10**	**10/11**	**11/12**
Sensitivität	0.78	0.69	0.65	0.53	0.51	0.47	0.41	0.35	0.33	0.24	0.20	0.14
Spezifität	0.68	0.76	0.82	0.83	0.87	0.90	0.91	0.94	0.95	0.97	0.97	0.97

4.2.2 Sensitivität und Spezifität des B-PHQ Auswertungsalgorithmus A

Sensitivität und Spezifität des B-PHQ Auswertungsalgorithmus A für die gesamte Stichprobe

In Tabelle 19 sind die Sensitivität und die Spezifität der gesamten Stichprobe für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus A, validiert anhand der Standarddepressionsdiagnosen des DIA-X, dargestellt.

Tabelle 19: Sensitivität und Spezifität des B-PHQ Auswertungsalgorithmus A für die gesamte Stichprobe validiert an den Standarddepressionsdiagnosen des DIA-X (N = 448).

Cut-Off-Wert	mindestens das Item 1a oder 1b ist positiv	mindestens 2 Items sind positiv darunter 1a oder 1b	mindestens 3 Items sind positiv darunter 1a oder 1b	mindestens 4 Items sind positiv darunter 1a oder 1b	mindestens 5 Items sind positiv darunter 1a oder 1b	mindestens 6 Items sind positiv darunter 1a oder 1b	mindestens 7 Items sind positiv darunter 1a oder 1b	mindestens 8 Items sind positiv darunter 1a oder 1b	alle 9 Items sind positiv
Sensitivität	0.79	0.79	0.77	0.71	0.61	0.45	0.35	0.23	0.09
Spezifität	0.83	0.86	0.89	0.91	0.93	0.95	0.97	0.99	0.99

In Tabelle 20 sind die Sensitivität und die Spezifität der gesamten Stichprobe für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus A, validiert an den Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom, abgebildet.

Tabelle 20: Sensitivität und Spezifität des B-PHQ Auswertungsalgorithmus A für die gesamte Stichprobe validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 448).

Cut-Off-Wert	mindestens das Item 1a oder 1b ist positiv	mindestens 2 Items sind positiv darunter 1a oder 1b	mindestens 3 Items sind positiv darunter 1a oder 1b	mindestens 4 Items sind positiv darunter 1a oder 1b	mindestens 5 Items sind positiv darunter 1a oder 1b	mindestens 6 Items sind positiv darunter 1a oder 1b	mindestens 7 Items sind positiv darunter 1a oder 1b	mindestens 8 Items sind positiv darunter 1a oder 1b	alle 9 Items sind positiv
Sensitivität	0.65	0.65	0.60	0.54	0.46	0.35	0.27	0.17	0.07
Spezifität	0.87	0.90	0.92	0.94	0.96	0.98	0.99	1.00	1.00

Sensitivität und Spezifität des B-PHQ Auswertungsalgorithmus A für die Subgruppe der Frauen

Tabelle 21 zeigt die Sensitivität und die Spezifität für die Subgruppe der Frauen für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus A, validiert anhand der Standarddepressionsdiagnosen des DIA-X.

Tabelle 21: Sensitivität und Spezifität des B-PHQ Auswertungsalgorithmus A für die Subgruppe der Frauen validiert an den Standarddepressionsdiagnosen des DIA-X (N = 283).

Cut-Off-Wert	mindestens das Item 1a oder 1b ist positiv	mindestens 2 Items sind positiv darunter 1a oder 1b	mindestens 3 Items sind positiv darunter 1a oder 1b	mindestens 4 Items sind positiv darunter 1a oder 1b	mindestens 5 Items sind positiv darunter 1a oder 1b	mindestens 6 Items sind positiv darunter 1a oder 1b	mindestens 7 Items sind positiv darunter 1a oder 1b	mindestens 8 Items sind positiv darunter 1a oder 1b	alle 9 Items sind positiv
Sensitivität	0.77	0.77	0.74	0.68	0.60	0.46	0.35	0.26	0.09
Spezifität	0.85	0.88	0.89	0.92	0.94	0.96	0.97	0.99	1.00

In Tabelle 22 sind die Sensitivität und die Spezifität der Subgruppe der Frauen für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus A, validiert an den

Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom, wiedergegeben.

Tabelle 22: Sensitivität und Spezifität des B-PHQ Auswertungsalgorithmus A für die Subgruppe der Frauen validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 283).

Cut-Off-Wert	mindestens das Item 1a oder 1b ist positiv	mindestens 2 Items sind positiv darunter 1a oder 1b	mindestens 3 Items sind positiv darunter 1a oder 1b	mindestens 4 Items sind positiv darunter 1a oder 1b	mindestens 5 Items sind positiv darunter 1a oder 1b	mindestens 6 Items sind positiv darunter 1a oder 1b	mindestens 7 Items sind positiv darunter 1a oder 1b	mindestens 8 Items sind positiv darunter 1a oder 1b	alle 9 Items sind positiv
Sensitivität	0.63	0.63	0.60	0.54	0.48	0.37	0.28	0.18	0.06
Spezifität	0.89	0.91	0.93	0.94	0.97	0.98	0.98	0.99	0.99

Sensitivität und Spezifität des B-PHQ Auswertungsalgorithmus A für die Subgruppe der Männer

In Tabelle 23 sind die Sensitivität und die Spezifität für die Subgruppe der Männer für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus A, validiert anhand der Standarddepressionsdiagnosen des DIA-X, dargestellt.

Tabelle 23: Sensitivität und Spezifität des B-PHQ Auswertungsalgorithmus A für die Subgruppe der Männer validiert an den Standarddepressionsdiagnosen des DIA-X (N = 165).

Cut-Off-Wert	mindestens das Item 1a oder 1b ist positiv	mindestens 2 Items sind positiv darunter 1a oder 1b	mindestens 3 Items sind positiv darunter 1a oder 1b	mindestens 4 Items sind positiv darunter 1a oder 1b	mindestens 5 Items sind positiv darunter 1a oder 1b	mindestens 6 Items sind positiv darunter 1a oder 1b	mindestens 7 Items sind positiv darunter 1a oder 1b	mindestens 8 Items sind positiv darunter 1a oder 1b	alle 9 Items sind positiv
Sensitivität	0.84	0.84	0.84	0.76	0.64	0.44	0.36	0.16	0.08
Spezifität	0.79	0.82	0.87	0.89	0.92	0.94	0.96	0.98	0.99

In Tabelle 24 sind die Sensitivität und die Spezifität der Subgruppe der Männer für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus A, validiert an den Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom, abgebildet.

Tabelle 24: Sensitivität und Spezifität des B-PHQ Auswertungsalgorithmus A für die Subgruppe der Männer validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 165).

Cut-Off-Wert	mindestens das Item 1a oder 1b ist positiv	mindestens 2 Items sind positiv darunter 1a oder 1b	mindestens 3 Items sind positiv darunter 1a oder 1b	mindestens 4 Items sind positiv darunter 1a oder 1b	mindestens 5 Items sind positiv darunter 1a oder 1b	mindestens 6 Items sind positiv darunter 1a oder 1b	mindestens 7 Items sind positiv darunter 1a oder 1b	mindestens 8 Items sind positiv darunter 1a oder 1b	alle 9 Items sind positiv
Sensitivität	0.67	0.67	0.59	0.53	0.43	0.33	0.27	0.14	0.08
Spezifität	0.84	0.89	0.91	0.93	0.95	0.97	0.99	1.00	1.00

Sensitivität und Spezifität des B-PHQ Auswertungsalgorithmus A für die Subgruppe Alter < 56,21 Jahre

In Tabelle 25 sind die Sensitivität und die Spezifität für die Subgruppe Alter < 56,21 Jahre für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus A, validiert anhand der Standarddepressionsdiagnosen des DIA-X, wiedergegeben.

Tabelle 25: Sensitivität und Spezifität des B-PHQ Auswertungsalgorithmus A für die Subgruppe Alter < 56,21 Jahre validiert an den Standarddepressionsdiagnosen des DIA-X (N = 224).

Cut-Off-Wert	mindestens das Item 1a oder 1b ist positiv	mindestens 2 Items sind positiv darunter 1a oder 1b	mindestens 3 Items sind positiv darunter 1a oder 1b	mindestens 4 Items sind positiv darunter 1a oder 1b	mindestens 5 Items sind positiv darunter 1a oder 1b	mindestens 6 Items sind positiv darunter 1a oder 1b	mindestens 7 Items sind positiv darunter 1a oder 1b	mindestens 8 Items sind positiv darunter 1a oder 1b	alle 9 Items sind positiv
Sensitivität	0.87	0.87	0.84	0.76	0.65	0.51	0.40	0.25	0.09
Spezifität	0.76	0.79	0.83	0.86	0.89	0.92	0.94	0.98	0.99

Tabelle 26 zeigt die Sensitivität und die Spezifität der Subgruppe Alter < 56,21 Jahre für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus A, validiert an den Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom.

Tabelle 26: Sensitivität und Spezifität des B-PHQ Auswertungsalgorithmus A für die Subgruppe Alter < 56,21 Jahre validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 224).

Cut-Off-Wert	mindestens das Item 1a oder 1b ist positiv	mindestens 2 Items sind positiv darunter 1a oder 1b	mindestens 3 Items sind positiv darunter 1a oder 1b	mindestens 4 Items sind positiv darunter 1a oder 1b	mindestens 5 Items sind positiv darunter 1a oder 1b	mindestens 6 Items sind positiv darunter 1a oder 1b	mindestens 7 Items sind positiv darunter 1a oder 1b	mindestens 8 Items sind positiv darunter 1a oder 1b	alle 9 Items sind positiv
Sensitivität	0.75	0.75	0.69	0.64	0.55	0.44	0.34	0.21	0.08
Spezifität	0.82	0.86	0.88	0.91	0.95	0.97	0.98	1.00	1.00

Sensitivität und Spezifität des B-PHQ Auswertungsalgorithmus A für die Subgruppe Alter > 56,20 Jahre

In Tabelle 27 sind die Sensitivität und die Spezifität für die Subgruppe Alter > 56,20 Jahre für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus A, validiert anhand der Standarddepressionsdiagnosen des DIA-X, dargestellt.

Tabelle 27: Sensitivität und Spezifität des B-PHQ Auswertungsalgorithmus A für die Subgruppe Alter > 56,20 Jahre validiert an den Standarddepressionsdiagnosen des DIA-X (N = 224).

Cut-Off-Wert	mindestens das Item 1a oder 1b ist positiv	mindestens 2 Items sind positiv darunter 1a oder 1b	mindestens 3 Items sind positiv darunter 1a oder 1b	mindestens 4 Items sind positiv darunter 1a oder 1b	mindestens 5 Items sind positiv darunter 1a oder 1b	mindestens 6 Items sind positiv darunter 1a oder 1b	mindestens 7 Items sind positiv darunter 1a oder 1b	mindestens 8 Items sind positiv darunter 1a oder 1b	alle 9 Items sind positiv
Sensitivität	0.63	0.63	0.63	0.59	0.52	0.33	0.26	0.19	0.07
Spezifität	0.89	0.91	0.93	0.95	0.97	0.97	0.99	0.99	0.99

In Tabelle 28 sind die Sensitivität und die Spezifität der Subgruppe Alter > 56,20 Jahre für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus A, validiert an den Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom, abgebildet.

Tabelle 28: Sensitivität und Spezifität des B-PHQ Auswertungsalgorithmus A für die Subgruppe Alter > 56,20 Jahre validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 224).

Cut-Off-Wert	mindestens das Item 1a oder 1b ist positiv	mindestens 2 Items sind positiv darunter 1a oder 1b	mindestens 3 Items sind positiv darunter 1a oder 1b	mindestens 4 Items sind positiv darunter 1a oder 1b	mindestens 5 Items sind positiv darunter 1a oder 1b	mindestens 6 Items sind positiv darunter 1a oder 1b	mindestens 7 Items sind positiv darunter 1a oder 1b	mindestens 8 Items sind positiv darunter 1a oder 1b	alle 9 Items sind positiv
Sensitivität	0.47	0.47	0.43	0.37	0.31	0.22	0.14	0.10	0.04
Spezifität	0.91	0.94	0.95	0.96	0.98	0.98	0.99	0.99	0.99

4.2.3 Sensitivität und Spezifität des B-PHQ Auswertungsalgorithmus B

Sensitivität und Spezifität des B-PHQ Auswertungsalgorithmus B für die gesamte Stichprobe

In Tabelle 29 sind die Sensitivität und die Spezifität der gesamten Stichprobe für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus B, validiert anhand der Standarddepressionsdiagnosen des DIA-X, dargestellt.

Tabelle 29: Sensitivität und Spezifität des B-PHQ Auswertungsalgorithmus B für die gesamte Stichprobe validiert an den Standarddepressionsdiagnosen des DIA-X (N = 448).

Cut-Off-Wert	mindestens 1 Item ist positiv	mindestens 2 Items sind positiv	mindestens 3 Items sind positiv	mindestens 4 Items sind positiv	mindestens 5 Items sind positiv	mindestens 6 Items sind positiv	mindestens 7 Items sind positiv	mindestens 8 Items sind positiv	alle 9 Items sind positiv
Sensitivität	0.91	0.87	0.78	0.71	0.61	0.45	0.35	0.23	0.09
Spezifität	0.54	0.74	0.83	0.89	0.93	0.95	0.97	0.99	0.99

In Tabelle 30 sind die Sensitivität und die Spezifität der gesamten Stichprobe für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus B, validiert an den Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom, abgebildet.

Tabelle 30: Sensitivität und Spezifität des B-PHQ Auswertungsalgorithmus B für die gesamte Stichprobe validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 448).

Cut-Off-Wert	mindestens 1 Item ist positiv	mindestens 2 Items sind positiv	mindestens 3 Items sind positiv	mindestens 4 Items sind positiv	mindestens 5 Items sind positiv	mindestens 6 Items sind positiv	mindestens 7 Items sind positiv	mindestens 8 Items sind positiv	alle 9 Items sind positiv
Sensitivität	0.87	0.79	0.65	0.54	0.47	0.35	0.27	0.17	0.07
Spezifität	0.60	0.81	0.88	0.92	0.96	0.98	0.99	1.00	1.00

Sensitivität und Spezifität des B-PHQ Auswertungsalgorithmus B für die Subgruppe der Frauen

Tabelle 31 zeigt die Sensitivität und die Spezifität für die Subgruppe der Frauen für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus B, validiert anhand der Standarddepressionsdiagnosen des DIA-X.

Tabelle 31: Sensitivität und Spezifität des B-PHQ Auswertungsalgorithmus B für die Subgruppe der Frauen validiert an den Standarddepressionsdiagnosen des DIA-X (N = 283).

Cut-Off-Wert	mindestens 1 Item ist positiv	mindestens 2 Items sind positiv	mindestens 3 Items sind positiv	mindestens 4 Items sind positiv	mindestens 5 Items sind positiv	mindestens 6 Items sind positiv	mindestens 7 Items sind positiv	mindestens 8 Items sind positiv	alle 9 Items sind positiv
Sensitivität	0.93	0.86	0.75	0.68	0.60	0.46	0.35	0.26	0.09
Spezifität	0.54	0.75	0.84	0.91	0.94	0.96	0.97	0.99	1.00

In Tabelle 32 sind die Sensitivität und die Spezifität der Subgruppe der Frauen für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus B, validiert an den Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom, wiedergegeben.

Tabelle 32: Sensitivität und Spezifität des B-PHQ Auswertungsalgorithmus B für die Subgruppe der Frauen validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 283).

Cut-Off-Wert	mindestens 1 Item ist positiv	mindestens 2 Items sind positiv	mindestens 3 Items sind positiv	mindestens 4 Items sind positiv	mindestens 5 Items sind positiv	mindestens 6 Items sind positiv	mindestens 7 Items sind positiv	mindestens 8 Items sind positiv	alle 9 Items sind positiv
Sensitivität	0.89	0.79	0.66	0.54	0.48	0.37	0.28	0.18	0.06
Spezifität	0.60	0.82	0.88	0.93	0.97	0.98	0.98	0.99	0.99

Sensitivität und Spezifität des B-PHQ Auswertungsalgorithmus B für die Subgruppe der Männer

In Tabelle 33 sind die Sensitivität und die Spezifität für die Subgruppe der Männer für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus B, validiert anhand der Standarddepressionsdiagnosen des DIA-X, dargestellt.

Tabelle 33: Sensitivität und Spezifität des B-PHQ Auswertungsalgorithmus B für die Subgruppe der Männer validiert an den Standarddepressionsdiagnosen des DIA-X (N = 165).

Cut-Off-Wert	mindestens 1 Item ist positiv	mindestens 2 Items sind positiv	mindestens 3 Items sind positiv	mindestens 4 Items sind positiv	mindestens 5 Items sind positiv	mindestens 6 Items sind positiv	mindestens 7 Items sind positiv	mindestens 8 Items sind positiv	alle 9 Items sind positiv
Sensitivität	0.88	0.88	0.84	0.76	0.64	0.44	0.36	0.16	0.08
Spezifität	0.54	0.72	0.81	0.86	0.91	0.94	0.96	0.98	0.99

In Tabelle 34 sind die Sensitivität und die Spezifität der Subgruppe der Männer für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus B, validiert an den Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom, abgebildet.

Tabelle 34: Sensitivität und Spezifität des B-PHQ Auswertungsalgorithmus B für die Subgruppe der Männer validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 165).

Cut-Off-Wert	mindestens 1 Item ist positiv	mindestens 2 Items sind positiv	mindestens 3 Items sind positiv	mindestens 4 Items sind positiv	mindestens 5 Items sind positiv	mindestens 6 Items sind positiv	mindestens 7 Items sind positiv	mindestens 8 Items sind positiv	alle 9 Items sind positiv
Sensitivität	0.84	0.78	0.63	0.55	0.45	0.33	0.27	0.14	0.08
Spezifität	0.60	0.80	0.86	0.91	0.94	0.97	0.99	1.00	1.00

Sensitivität und Spezifität des B-PHQ Auswertungsalgorithmus B für die Subgruppe Alter < 56,21 Jahre

In Tabelle 35 sind die Sensitivität und die Spezifität für die Subgruppe Alter < 56,21 Jahre für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus B, validiert anhand der Standarddepressionsdiagnosen des DIA-X, wiedergegeben.

Tabelle 35: Sensitivität und Spezifität des B-PHQ Auswertungsalgorithmus B für die Subgruppe Alter < 56,21 Jahre validiert an den Standarddepressionsdiagnosen des DIA-X (N = 224).

Cut-Off-Wert	mindestens 1 Item ist positiv	mindestens 2 Items sind positiv	mindestens 3 Items sind positiv	mindestens 4 Items sind positiv	mindestens 5 Items sind positiv	mindestens 6 Items sind positiv	mindestens 7 Items sind positiv	mindestens 8 Items sind positiv	alle 9 Items sind positiv
Sensitivität	0.95	0.89	0.84	0.76	0.65	0.51	0.40	0.25	0.09
Spezifität	0.50	0.67	0.77	0.83	0.89	0.92	0.94	0.98	0.99

Tabelle 36 zeigt die Sensitivität und die Spezifität der Subgruppe Alter < 56,21 Jahre für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus B, validiert an den Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom.

Tabelle 36: Sensitivität und Spezifität des B-PHQ Auswertungsalgorithmus B für die Subgruppe Alter < 56,21 Jahre validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 224).

Cut-Off-Wert	mindestens 1 Item ist positiv	mindestens 2 Items sind positiv	mindestens 3 Items sind positiv	mindestens 4 Items sind positiv	mindestens 5 Items sind positiv	mindestens 6 Items sind positiv	mindestens 7 Items sind positiv	mindestens 8 Items sind positiv	alle 9 Items sind positiv
Sensitivität	0.92	0.84	0.74	0.65	0.56	0.44	0.34	0.21	0.08
Spezifität	0.58	0.76	0.84	0.89	0.95	0.97	0.98	1.00	1.00

Sensitivität und Spezifität des B-PHQ Auswertungsalgorithmus B für die Subgruppe Alter > 56,20 Jahre

In Tabelle 37 sind die Sensitivität und die Spezifität für die Subgruppe Alter > 56,20 Jahre für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus B, validiert anhand der Standarddepressionsdiagnosen des DIA-X, dargestellt.

Tabelle 37: Sensitivität und Spezifität des B-PHQ Auswertungsalgorithmus B für die Subgruppe Alter > 56,20 Jahre validiert an den Standarddepressionsdiagnosen des DIA-X (N = 224).

Cut-Off-Wert	mindestens 1 Item ist positiv	mindestens 2 Items sind positiv	mindestens 3 Items sind positiv	mindestens 4 Items sind positiv	mindestens 5 Items sind positiv	mindestens 6 Items sind positiv	mindestens 7 Items sind positiv	mindestens 8 Items sind positiv	alle 9 Items sind positiv
Sensitivität	0.85	0.81	0.67	0.59	0.52	0.33	0.26	0.19	0.07
Spezifität	0.57	0.80	0.88	0.94	0.96	0.97	0.99	0.99	0.99

In Tabelle 38 sind die Sensitivität und die Spezifität der Subgruppe Alter > 56,20 Jahre für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus B, validiert an den Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom, abgebildet.

Tabelle 38: Sensitivität und Spezifität des B-PHQ Auswertungsalgorithmus B für die Subgruppe Alter > 56,20 Jahre validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 224).

Cut-Off-Wert	mindestens 1 Item ist positiv	mindestens 2 Items sind positiv	mindestens 3 Items sind positiv	mindestens 4 Items sind positiv	mindestens 5 Items sind positiv	mindestens 6 Items sind positiv	mindestens 7 Items sind positiv	mindestens 8 Items sind positiv	alle 9 Items sind positiv
Sensitivität	0.78	0.71	0.49	0.37	0.31	0.22	0.16	0.10	0.04
Spezifität	0.61	0.86	0.90	0.95	0.97	0.98	0.99	0.99	0.99

4.3 Prädiktive Werte

4.3.1 Prädiktive Werte des GHQ-12

Prädiktive Werte des GHQ-12 für die gesamte Stichprobe

In Tabelle 39 sind die prädiktiven Werte der gesamten Stichprobe für die Cut-Off-Werte des GHQ-12, validiert anhand der Standarddepressionsdiagnosen des DIA-X, dargestellt.

Tabelle 39: Prädiktive Werte des GHQ-12 für die gesamte Stichprobe validiert an den Standardepressionsdiagnosen des DIA-X (N = 448).

Cut-Off-Wert	**0/1**	**1/2**	**2/3**	**3/4**	**4/5**	**5/6**	**6/7**	**7/8**	**8/9**	**9/10**	**10/11**	**11/12**
positiver prädiktiver Wert	0.30	0.34	0.38	0.39	0.45	0.52	0.55	0.60	0.63	0.67	0.69	0.71
negativer prädiktiver Wert	0.96	0.95	0.94	0.93	0.93	0.93	0.92	0.92	0.91	0.89	0.88	0.86

In Tabelle 40 sind die prädiktiven Werte der gesamten Stichprobe für die Cut-Off-Werte des GHQ-12, validiert an den Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom, abgebildet.

Tabelle 40: Prädiktive Werte des GHQ-12 für die gesamte Stichprobe validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 448).

Cut-Off-Wert	**0/1**	**1/2**	**2/3**	**3/4**	**4/5**	**5/6**	**6/7**	**7/8**	**8/9**	**9/10**	**10/11**	**11/12**
positiver prädiktiver Wert	0.48	0.53	0.58	0.59	0.64	0.70	0.74	0.79	0.80	0.82	0.86	0.82
negativer prädiktiver Wert	0.91	0.89	0.87	0.85	0.84	0.83	0.82	0.81	0.80	0.78	0.76	0.74

Prädiktive Werte des GHQ-12 für die Subgruppe der Frauen

Tabelle 41 zeigt die prädiktiven Werte für die Subgruppe der Frauen für die Cut-Off-Werte des GHQ-12, validiert anhand der Standarddepressionsdiagnosen des DIA-X.

Tabelle 41: Prädiktive Werte des GHQ-12 für die Subgruppe der Frauen validiert an den Standarddepressionsdiagnosen des DIA-X (N = 283).

Cut-Off-Wert	**0/1**	**1/2**	**2/3**	**3/4**	**4/5**	**5/6**	**6/7**	**7/8**	**8/9**	**9/10**	**10/11**	**11/12**
positiver prädiktiver Wert	0.32	0.38	0.42	0.43	0.51	0.59	0.62	0.65	0.65	0.70	0.71	0.76
negativer prädiktiver Wert	0.95	0.95	0.94	0.93	0.93	0.93	0.91	0.91	0.90	0.88	0.87	0.84

In Tabelle 42 sind die prädiktiven Werte der Subgruppe der Frauen für die Cut-Off-Werte des GHQ-12, validiert an den Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom, wiedergegeben.

Tabelle 42: Prädiktive Werte des GHQ-12 für die Subgruppe der Frauen validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 283).

Cut-Off-Wert	**0/1**	**1/2**	**2/3**	**3/4**	**4/5**	**5/6**	**6/7**	**7/8**	**8/9**	**9/10**	**10/11**	**11/12**
positiver prädiktiver Wert	0.48	0.54	0.61	0.61	0.66	0.75	0.79	0.80	0.78	0.80	0.82	0.76
negativer prädiktiver Wert	0.91	0.90	0.88	0.86	0.84	0.84	0.82	0.81	0.80	0.77	0.76	0.73

Prädiktive Werte des GHQ-12 für die Subgruppe der Männer

In Tabelle 43 sind die prädiktiven Werte für die Subgruppe der Männer für die Cut-Off-Werte des GHQ-12, validiert anhand der Standarddepressionsdiagnosen des DIA-X, dargestellt.

Tabelle 43: Prädiktive Werte des GHQ-12 für die Subgruppe der Männer validiert an den Standarddepressionsdiagnosen des DIA-X (N = 165).

Cut-Off-Wert	**0/1**	**1/2**	**2/3**	**3/4**	**4/5**	**5/6**	**6/7**	**7/8**	**8/9**	**9/10**	**10/11**	**11/12**
positiver prädiktiver Wert	0.25	0.27	0.30	0.33	0.36	0.41	0.44	0.52	0.58	0.60	0.67	0.62
negativer prädiktiver Wert	0.96	0.95	0.94	0.94	0.94	0.94	0.94	0.93	0.92	0.91	0.90	0.89

In Tabelle 44 sind die prädiktiven Werte der Subgruppe der Männer für die Cut-Off-Werte des GHQ-12, validiert an den Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom, abgebildet.

Tabelle 44: Prädiktiven Werte des GHQ-12 für die Subgruppe der Männer validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 165).

Cut-Off-Wert	**0/1**	**1/2**	**2/3**	**3/4**	**4/5**	**5/6**	**6/7**	**7/8**	**8/9**	**9/10**	**10/11**	**11/12**
positiver prädiktiver Wert	0.47	0.52	0.54	0.55	0.60	0.61	0.67	0.77	0.83	0.85	0.93	0.92
negativer prädiktiver Wert	0.90	0.88	0.85	0.84	0.83	0.82	0.82	0.81	0.79	0.78	0.77	0.76

Prädiktive Werte des GHQ-12 für die Subgruppe Alter < 56,21 Jahre

In Tabelle 45 sind die prädiktiven Werte für die Subgruppe Alter < 56,21 Jahre für die Cut-Off-Werte des GHQ-12, validiert anhand der Standarddepressionsdiagnosen des DIA-X, wiedergegeben.

Tabelle 45: Prädiktive Werte des GHQ-12 für die Subgruppe Alter < 56,21 Jahre validiert an den Standarddepressionsdiagnosen des DIA-X (N = 224).

Cut-Off-Wert	**0/1**	**1/2**	**2/3**	**3/4**	**4/5**	**5/6**	**6/7**	**7/8**	**8/9**	**9/10**	**10/11**	**11/12**
positiver prädiktiver Wert	0.34	0.39	0.44	0.45	0.51	0.60	0.64	0.70	0.70	0.71	0.76	0.77
negativer prädiktiver Wert	0.95	0.95	0.93	0.93	0.91	0.92	0.91	0.91	0.89	0.86	0.84	0.81

Tabelle 46 zeigt die prädiktiven Werte der Subgruppe Alter < 56,21 Jahre für die Cut-Off-Werte des GHQ-12, validiert an den Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom.

Tabelle 46: Prädiktiven Werte des GHQ-12 für die Subgruppe Alter < 56,21 Jahre validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 224).

Cut-Off-Wert	**0/1**	**1/2**	**2/3**	**3/4**	**4/5**	**5/6**	**6/7**	**7/8**	**8/9**	**9/10**	**10/11**	**11/12**
positiver prädiktiver Wert	0.52	0.58	0.63	0.65	0.70	0.76	0.83	0.88	0.86	0.88	0.97	0.95
negativer prädiktiver Wert	0.89	0.89	0.85	0.85	0.81	0.80	0.80	0.79	0.76	0.74	0.72	0.68

Prädiktive Werte des GHQ-12 für die Subgruppe Alter > 56,20 Jahre

In Tabelle 47 sind die prädiktiven Werte für die Subgruppe Alter > 56,20 Jahre für die Cut-Off-Werte des GHQ-12, validiert anhand der Standarddepressionsdiagnosen des DIA-X, dargestellt.

Tabelle 47: Prädiktive Werte des GHQ-12 für die Subgruppe Alter > 56,20 Jahre validiert an den Standarddepressionsdiagnosen des DIA-X (N = 224).

Cut-Off-Wert	**0/1**	**1/2**	**2/3**	**3/4**	**4/5**	**5/6**	**6/7**	**7/8**	**8/9**	**9/10**	**10/11**	**11/12**
positiver prädiktiver Wert	0.23	0.26	0.28	0.30	0.35	0.39	0.39	0.41	0.48	0.56	0.56	0.58
negativer prädiktiver Wert	0.96	0.95	0.94	0.94	0.94	0.94	0.93	0.92	0.92	0.92	0.91	0.91

In Tabelle 48 sind die prädiktiven Werte der Subgruppe Alter > 56,20 Jahre für die Cut-Off-Werte des GHQ-12, validiert an den Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom, abgebildet.

Tabelle 48: Prädiktive Werte des GHQ-12 für die Subgruppe Alter > 56,20 Jahre validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 224).

Cut-Off-Wert	**0/1**	**1/2**	**2/3**	**3/4**	**4/5**	**5/6**	**6/7**	**7/8**	**8/9**	**9/10**	**10/11**	**11/12**
positiver prädiktiver Wert	0.42	0.45	0.51	0.47	0.53	0.59	0.58	0.62	0.68	0.67	0.63	0.58
negativer prädiktiver Wert	0.91	0.89	0.89	0.86	0.86	0.85	0.84	0.83	0.83	0.81	0.80	0.79

4.3.2 Prädiktive Werte des B-PHQ Auswertungsalgorithmus A

Prädiktive Werte des B-PHQ Auswertungsalgorithmus A für die gesamte Stichprobe

In Tabelle 49 sind die prädiktiven Werte der gesamten Stichprobe für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus A, validiert anhand der Standarddepressionsdiagnosen des DIA-X, dargestellt.

Tabelle 49: Prädiktive Werte des B-PHQ Auswertungsalgorithmus A für die gesamte Stichprobe validiert an den Standarddepressionsdiagnosen des DIA-X (N = 448).

Cut-Off-Wert	mindestens das Item 1a oder 1b ist positiv	mindestens 2 Items sind positiv darunter 1a oder 1b	mindestens 3 Items sind positiv darunter 1a oder 1b	mindestens 4 Items sind positiv darunter 1a oder 1b	mindestens 5 Items sind positiv darunter 1a oder 1b	mindestens 6 Items sind positiv darunter 1a oder 1b	mindestens 7 Items sind positiv darunter 1a oder 1b	mindestens 8 Items sind positiv darunter 1a oder 1b	alle 9 Items sind positiv
positiver prädiktiver Wert	0.51	0.55	0.60	0.63	0.68	0.67	0.71	0.79	0.70
negativer prädiktiver Wert	0.95	0.95	0.94	0.93	0.91	0.89	0.87	0.85	0.83

In Tabelle 50 sind die prädiktiven Werte der gesamten Stichprobe für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus A, validiert an den Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom, abgebildet.

Tabelle 50: Prädiktive Werte des B-PHQ Auswertungsalgorithmus A für die gesamte Stichprobe validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 448).

Cut-Off-Wert	mindestens das Item 1a oder 1b ist positiv	mindestens 2 Items sind positiv darunter 1a oder 1b	mindestens 3 Items sind positiv darunter 1a oder 1b	mindestens 4 Items sind positiv darunter 1a oder 1b	mindestens 5 Items sind positiv darunter 1a oder 1b	mindestens 6 Items sind positiv darunter 1a oder 1b	mindestens 7 Items sind positiv darunter 1a oder 1b	mindestens 8 Items sind positiv darunter 1a oder 1b	alle 9 Items sind positiv
positiver prädiktiver Wert	0.69	0.75	0.77	0.79	0.85	0.87	0.90	0.96	0.90
negativer prädiktiver Wert	0.85	0.85	0.84	0.82	0.80	0.78	0.76	0.73	0.71

Prädiktive Werte des B-PHQ Auswertungsalgorithmus A für die Subgruppe der Frauen

Tabelle 51 zeigt die prädiktiven Werte für die Subgruppe der Frauen für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus A, validiert anhand der Standarddepressionsdiagnosen des DIA-X.

Tabelle 51: Prädiktive Werte des B-PHQ Auswertungsalgorithmus A für die Subgruppe der Frauen validiert an den Standarddepressionsdiagnosen des DIA-X (N = 283).

Cut-Off-Wert	mindestens das Item 1a oder 1b ist positiv	mindestens 2 Items sind positiv darunter 1a oder 1b	mindestens 3 Items sind positiv darunter 1a oder 1b	mindestens 4 Items sind positiv darunter 1a oder 1b	mindestens 5 Items sind positiv darunter 1a oder 1b	mindestens 6 Items sind positiv darunter 1a oder 1b	mindestens 7 Items sind positiv darunter 1a oder 1b	mindestens 8 Items sind positiv darunter 1a oder 1b	alle 9 Items sind positiv
positiver prädiktiver Wert	0.57	0.61	0.64	0.67	0.72	0.74	0.74	0.88	0.83
negativer prädiktiver Wert	0.94	0.94	0.93	0.92	0.90	0.88	0.86	0.84	0.81

In Tabelle 52 sind die prädiktiven Werte der Subgruppe der Frauen für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus A, validiert an den Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom, wiedergegeben.

Tabelle 52: Prädiktive Werte des B-PHQ Auswertungsalgorithmus A für die Subgruppe der Frauen validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 283).

Cut-Off-Wert	mindestens das Item 1a oder 1b ist positiv	mindestens 2 Items sind positiv darunter 1a oder 1b	mindestens 3 Items sind positiv darunter 1a oder 1b	mindestens 4 Items sind positiv darunter 1a oder 1b	mindestens 5 Items sind positiv darunter 1a oder 1b	mindestens 6 Items sind positiv darunter 1a oder 1b	mindestens 7 Items sind positiv darunter 1a oder 1b	mindestens 8 Items sind positiv darunter 1a oder 1b	alle 9 Items sind positiv
positiver prädiktiver Wert	0.71	0.76	0.79	0.81	0.89	0.91	0.89	0.94	0.83
negativer prädiktiver Wert	0.84	0.85	0.84	0.82	0.81	0.78	0.75	0.73	0.70

Prädiktive Werte des B-PHQ Auswertungsalgorithmus A für die Subgruppe der Männer

In Tabelle 53 sind die prädiktiven Werte für die Subgruppe der Männer für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus A, validiert anhand der Standarddepressionsdiagnosen des DIA-X, dargestellt.

Tabelle 53: Prädiktive Werte des B-PHQ Auswertungsalgorithmus A für die Subgruppe der Männer validiert an den Standarddepressionsdiagnosen des DIA-X (N = 165).

Cut-Off-Wert	mindestens das Item 1a oder 1b ist positiv	mindestens 2 Items sind positiv darunter 1a oder 1b	mindestens 3 Items sind positiv darunter 1a oder 1b	mindestens 4 Items sind positiv darunter 1a oder 1b	mindestens 5 Items sind positiv darunter 1a oder 1b	mindestens 6 Items sind positiv darunter 1a oder 1b	mindestens 7 Items sind positiv darunter 1a oder 1b	mindestens 8 Items sind positiv darunter 1a oder 1b	alle 9 Items sind positiv
positiver prädiktiver Wert	0.41	0.46	0.54	0.56	0.59	0.55	0.64	0.57	0.50
negativer prädiktiver Wert	0.96	0.97	0.97	0.95	0.93	0.90	0.89	0.87	0.86

In Tabelle 54 sind die prädiktiven Werte der Subgruppe der Männer für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus A, validiert an den Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom, abgebildet.

Tabelle 54: Prädiktive Werte des B-PHQ Auswertungsalgorithmus A für die Subgruppe der Männer validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 165).

Cut-Off-Wert	mindestens das Item 1a oder 1b ist positiv	mindestens 2 Items sind positiv darunter 1a oder 1b	mindestens 3 Items sind positiv darunter 1a oder 1b	mindestens 4 Items sind positiv darunter 1a oder 1b	mindestens 5 Items sind positiv darunter 1a oder 1b	mindestens 6 Items sind positiv darunter 1a oder 1b	mindestens 7 Items sind positiv darunter 1a oder 1b	mindestens 8 Items sind positiv darunter 1a oder 1b	alle 9 Items sind positiv
positiver prädiktiver Wert	0.65	0.72	0.74	0.76	0.78	0.80	0.93	1.00	1.00
negativer prädiktiver Wert	0.86	0.87	0.84	0.82	0.80	0.77	0.76	0.73	0.72

Prädiktive Werte des B-PHQ Auswertungsalgorithmus A für die Subgruppe Alter < 56,21 Jahre

In Tabelle 55 sind die prädiktiven Werte für die Subgruppe Alter < 56,21 Jahre für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus A, validiert anhand der Standarddepressionsdiagnosen des DIA-X, wiedergegeben.

Tabelle 55: Prädiktive Werte des B-PHQ Auswertungsalgorithmus A für die Subgruppe Alter < 56,21 Jahre validiert an den Standarddepressionsdiagnosen des DIA-X (N = 224).

Cut-Off-Wert	mindestens das Item 1a oder 1b ist positiv	mindestens 2 Items sind positiv darunter 1a oder 1b	mindestens 3 Items sind positiv darunter 1a oder 1b	mindestens 4 Items sind positiv darunter 1a oder 1b	mindestens 5 Items sind positiv darunter 1a oder 1b	mindestens 6 Items sind positiv darunter 1a oder 1b	mindestens 7 Items sind positiv darunter 1a oder 1b	mindestens 8 Items sind positiv darunter 1a oder 1b	alle 9 Items sind positiv
positiver prädiktiver Wert	0.54	0.57	0.61	0.64	0.67	0.68	0.69	0.78	0.71
negativer prädiktiver Wert	0.95	0.95	0.94	0.92	0.89	0.85	0.83	0.80	0.77

Tabelle 56 zeigt die prädiktiven Werte der Subgruppe Alter < 56,21 Jahre für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus A, validiert an den Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom.

Tabelle 56: Prädiktive Werte des B-PHQ Auswertungsalgorithmus A für die Subgruppe Alter < 56,21 Jahre validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 224).

Cut-Off-Wert	mindestens das Item 1a oder 1b ist positiv	mindestens 2 Items sind positiv darunter 1a oder 1b	mindestens 3 Items sind positiv darunter 1a oder 1b	mindestens 4 Items sind positiv darunter 1a oder 1b	mindestens 5 Items sind positiv darunter 1a oder 1b	mindestens 6 Items sind positiv darunter 1a oder 1b	mindestens 7 Items sind positiv darunter 1a oder 1b	mindestens 8 Items sind positiv darunter 1a oder 1b	alle 9 Items sind positiv
positiver prädiktiver Wert	0.72	0.76	0.79	0.82	0.87	0.90	0.91	1.00	1.00
negativer prädiktiver Wert	0.84	0.85	0.83	0.80	0.78	0.74	0.71	0.67	0.64

Prädiktive Werte des B-PHQ Auswertungsalgorithmus A für die Subgruppe Alter > 56,20 Jahre

In Tabelle 57 sind die prädiktiven Werte für die Subgruppe Alter > 56,20 Jahre für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus A, validiert anhand der Standarddepressionsdiagnosen des DIA-X, dargestellt.

Tabelle 57: Prädiktive Werte des B-PHQ Auswertungsalgorithmus A für die Subgruppe Alter > 56,20 Jahre validiert an den Standarddepressionsdiagnosen des DIA-X (N = 224).

Cut-Off-Wert	mindestens das Item 1a oder 1b ist positiv	mindestens 2 Items sind positiv darunter 1a oder 1b	mindestens 3 Items sind positiv darunter 1a oder 1b	mindestens 4 Items sind positiv darunter 1a oder 1b	mindestens 5 Items sind positiv darunter 1a oder 1b	mindestens 6 Items sind positiv darunter 1a oder 1b	mindestens 7 Items sind positiv darunter 1a oder 1b	mindestens 8 Items sind positiv darunter 1a oder 1b	alle 9 Items sind positiv
positiver prädiktiver Wert	0.44	0.50	0.57	0.62	0.70	0.64	0.78	0.83	0.67
negativer prädiktiver Wert	0.95	0.95	0.95	0.94	0.94	0.91	0.91	0.90	0.89

In Tabelle 58 sind die prädiktiven Werte der Subgruppe Alter > 56,20 Jahre für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus A, validiert an den Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom, abgebildet.

Tabelle 58: Prädiktive Werte des B-PHQ Auswertungsalgorithmus A für die Subgruppe Alter > 56,20 Jahre validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 224).

Cut-Off-Wert	mindestens das Item 1a oder 1b ist positiv	mindestens 2 Items sind positiv darunter 1a oder 1b	mindestens 3 Items sind positiv darunter 1a oder 1b	mindestens 4 Items sind positiv darunter 1a oder 1b	mindestens 5 Items sind positiv darunter 1a oder 1b	mindestens 6 Items sind positiv darunter 1a oder 1b	mindestens 7 Items sind positiv darunter 1a oder 1b	mindestens 8 Items sind positiv darunter 1a oder 1b	alle 9 Items sind positiv
positiver prädiktiver Wert	0.62	0.71	0.73	0.73	0.80	0.79	0.89	0.83	0.67
negativer prädiktiver Wert	0.85	0.86	0.85	0.84	0.83	0.81	0.80	0.79	0.78

4.3.3 Prädiktive Werte des B-PHQ Auswertungsalgorithmus B

Prädiktive Werte des B-PHQ Auswertungsalgorithmus B für die gesamte Stichprobe

In Tabelle 59 sind die prädiktiven Werte der gesamten Stichprobe für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus B, validiert anhand der Standarddepressionsdiagnosen des DIA-X, dargestellt.

Tabelle 59: Prädiktive Werte des B-PHQ Auswertungsalgorithmus B für die gesamte Stichprobe validiert an den Standarddepressionsdiagnosen des DIA-X (N = 448).

Cut-Off-Wert	mindestens 1 Item ist positiv	mindestens 2 Items sind positiv	mindestens 3 Items sind positiv	mindestens 4 Items sind positiv	mindestens 5 Items sind positiv	mindestens 6 Items sind positiv	mindestens 7 Items sind positiv	mindestens 8 Items sind positiv	alle 9 Items sind positiv
positiver prädiktiver Wert	0.31	0.43	0.50	0.59	0.66	0.67	0.71	0.79	0.70
negativer prädiktiver Wert	0.97	0.96	0.94	0.93	0.91	0.89	0.87	0.85	0.83

In Tabelle 60 sind die prädiktiven Werte der gesamten Stichprobe für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus B, validiert an den Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom, abgebildet.

Tabelle 60: Prädiktive Werte des B-PHQ Auswertungsalgorithmus B für die gesamte Stichprobe validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 448).

Cut-Off-Wert	mindestens 1 Item ist positiv	mindestens 2 Items sind positiv	mindestens 3 Items sind positiv	mindestens 4 Items sind positiv	mindestens 5 Items sind positiv	mindestens 6 Items sind positiv	mindestens 7 Items sind positiv	mindestens 8 Items sind positiv	alle 9 Items sind positiv
positiver prädiktiver Wert	0.49	0.64	0.69	0.76	0.84	0.87	0.90	0.96	0.90
negativer prädiktiver Wert	0.91	0.90	0.85	0.82	0.81	0.78	0.76	0.73	0.71

Prädiktive Werte des B-PHQ Auswertungsalgorithmus B für die Subgruppe der Frauen

Tabelle 61 zeigt die prädiktiven Werte für die Subgruppe der Frauen für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus B, validiert anhand der Standarddepressionsdiagnosen des DIA-X.

Tabelle 61: Prädiktive Werte des B-PHQ Auswertungsalgorithmus B für die Subgruppe der Frauen validiert an den Standarddepressionsdiagnosen des DIA-X (N = 283).

Cut-Off-Wert	mindestens 1 Item ist positiv	mindestens 2 Items sind positiv	mindestens 3 Items sind positiv	mindestens 4 Items sind positiv	mindestens 5 Items sind positiv	mindestens 6 Items sind positiv	mindestens 7 Items sind positiv	mindestens 8 Items sind positiv	alle 9 Items sind positiv
positiver prädiktiver Wert	0.34	0.47	0.54	0.65	0.72	0.74	0.74	0.88	0.83
negativer prädiktiver Wert	0.97	0.96	0.93	0.92	0.90	0.88	0.86	0.84	0.81

In Tabelle 62 sind die prädiktiven Werte der Subgruppe der Frauen für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus B, validiert an den Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom, wiedergegeben.

Tabelle 62: Prädiktive Werte des B-PHQ Auswertungsalgorithmus B für die Subgruppe der Frauen validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 283).

Cut-Off-Wert	mindestens 1 Item ist positiv	mindestens 2 Items sind positiv	mindestens 3 Items sind positiv	mindestens 4 Items sind positiv	mindestens 5 Items sind positiv	mindestens 6 Items sind positiv	mindestens 7 Items sind positiv	mindestens 8 Items sind positiv	alle 9 Items sind positiv
positiver prädiktiver Wert	0.49	0.66	0.71	0.78	0.89	0.91	0.89	0.94	0.83
negativer prädiktiver Wert	0.92	0.90	0.85	0.82	0.81	0.78	0.75	0.73	0.70

Prädiktive Werte des B-PHQ Auswertungsalgorithmus B für die Subgruppe der Männer

In Tabelle 63 sind die prädiktiven Werte für die Subgruppe der Männer für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus B, validiert anhand der Standarddepressionsdiagnosen des DIA-X, dargestellt.

Tabelle 63: Prädiktive Werte des B-PHQ Auswertungsalgorithmus B für die Subgruppe der Männer validiert an den Standarddepressionsdiagnosen des DIA-X (N = 165).

Cut-Off-Wert	mindestens 1 Item ist positiv	mindestens 2 Items sind positiv	mindestens 3 Items sind positiv	mindestens 4 Items sind positiv	mindestens 5 Items sind positiv	mindestens 6 Items sind positiv	mindestens 7 Items sind positiv	mindestens 8 Items sind positiv	alle 9 Items sind positiv
positiver prädiktiver Wert	0.25	0.36	0.45	0.50	0.55	0.55	0.64	0.57	0.50
negativer prädiktiver Wert	0.96	0.97	0.97	0.95	0.93	0.90	0.89	0.87	0.86

In Tabelle 64 sind die prädiktiven Werte der Subgruppe der Männer für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus B, validiert an den Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom, abgebildet.

Tabelle 64: Prädiktive Werte des B-PHQ Auswertungsalgorithmus B für die Subgruppe der Männer validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 165).

Cut-Off-Wert	mindestens 1 Item ist positiv	mindestens 2 Items sind positiv	mindestens 3 Items sind positiv	mindestens 4 Items sind positiv	mindestens 5 Items sind positiv	mindestens 6 Items sind positiv	mindestens 7 Items sind positiv	mindestens 8 Items sind positiv	alle 9 Items sind positiv
positiver prädiktiver Wert	0.47	0.62	0.66	0.71	0.76	0.80	0.93	1.00	1.00
negativer prädiktiver Wert	0.90	0.89	0.85	0.83	0.80	0.77	0.76	0.73	0.72

Prädiktive Werte des B-PHQ Auswertungsalgorithmus B für die Subgruppe Alter < 56,21 Jahre

In Tabelle 65 sind die prädiktiven Werte für die Subgruppe Alter < 56,21 Jahre für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus B, validiert anhand der Standarddepressionsdiagnosen des DIA-X, wiedergegeben.

Tabelle 65: Prädiktive Werte des B-PHQ Auswertungsalgorithmus B für die Subgruppe Alter < 56,21 Jahre validiert an den Standarddepressionsdiagnosen des DIA-X (N = 224).

Cut-Off-Wert	mindestens 1 Item ist positiv	mindestens 2 Items sind positiv	mindestens 3 Items sind positiv	mindestens 4 Items sind positiv	mindestens 5 Items sind positiv	mindestens 6 Items sind positiv	mindestens 7 Items sind positiv	mindestens 8 Items sind positiv	alle 9 Items sind positiv
positiver prädiktiver Wert	0.38	0.47	0.54	0.60	0.65	0.68	0.69	0.78	0.71
negativer prädiktiver Wert	0.97	0.95	0.94	0.92	0.89	0.85	0.83	0.80	0.77

Tabelle 66 zeigt die prädiktiven Werte der Subgruppe Alter < 56,21 Jahre für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus B, validiert an den Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom.

Tabelle 66: Prädiktive Werte des B-PHQ Auswertungsalgorithmus B für die Subgruppe Alter < 56,21 Jahre validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 224).

Cut-Off-Wert	mindestens 1 Item ist positiv	mindestens 2 Items sind positiv	mindestens 3 Items sind positiv	mindestens 4 Items sind positiv	mindestens 5 Items sind positiv	mindestens 6 Items sind positiv	mindestens 7 Items sind positiv	mindestens 8 Items sind positiv	alle 9 Items sind positiv
positiver prädiktiver Wert	0.57	0.68	0.74	0.79	0.87	0.90	0.91	1.00	1.00
negativer prädiktiver Wert	0.92	0.88	0.84	0.81	0.78	0.74	0.71	0.67	0.64

Prädiktive Werte des B-PHQ Auswertungsalgorithmus B für die Subgruppe Alter > 56,20 Jahre

In Tabelle 67 sind die prädiktiven Werte für die Subgruppe Alter > 56,20 Jahre für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus B, validiert anhand der Standarddepressionsdiagnosen des DIA-X, dargestellt.

Tabelle 67: Prädiktive Werte des B-PHQ Auswertungsalgorithmus B für die Subgruppe Alter > 56,20 Jahre validiert an den Standarddepressionsdiagnosen des DIA-X (N = 224).

Cut-Off-Wert	mindestens 1 Item ist positiv	mindestens 2 Items sind positiv	mindestens 3 Items sind positiv	mindestens 4 Items sind positiv	mindestens 5 Items sind positiv	mindestens 6 Items sind positiv	mindestens 7 Items sind positiv	mindestens 8 Items sind positiv	alle 9 Items sind positiv
positiver prädiktiver Wert	0.21	0.36	0.43	0.57	0.67	0.64	0.78	0.83	0.67
negativer prädiktiver Wert	0.97	0.97	0.95	0.94	0.94	0.91	0.91	0.90	0.89

In Tabelle 68 sind die prädiktiven Werte der Subgruppe Alter > 56,20 Jahre für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus B, validiert an den Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom, abgebildet.

Tabelle 68: Prädiktive Werte des B-PHQ Auswertungsalgorithmus B für die Subgruppe Alter > 56,20 Jahre validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 224).

Cut-Off-Wert	mindestens 1 Item ist positiv	mindestens 2 Items sind positiv	mindestens 3 Items sind positiv	mindestens 4 Items sind positiv	mindestens 5 Items sind positiv	mindestens 6 Items sind positiv	mindestens 7 Items sind positiv	mindestens 8 Items sind positiv	alle 9 Items sind positiv
positiver prädiktiver Wert	0.37	0.59	0.60	0.68	0.76	0.79	0.89	0.83	0.67
negativer prädiktiver Wert	0.91	0.91	0.86	0.84	0.83	0.81	0.80	0.79	0.78

4.4 Overall Accuracy, Kappa Index und Likelihood Quotienten

4.4.1 Overall Accuracy, Kappa Index und Likelihood Quotienten des GHQ-12

Overall Accuracy, Kappa Index und Likelihood Quotienten des GHQ-12 für die gesamte Stichprobe

In Tabelle 69 sind die Overall Accuracy, der Kappa Index und die Likelihood Quotienten der gesamten Stichprobe für die Cut-Off-Werte des GHQ-12, validiert anhand der Standarddepressionsdiagnosen des DIA-X, dargestellt.

Tabelle 69: Overall Accuracy, Kappa Index und Likelihood Quotienten des GHQ-12 für die gesamte Stichprobe validiert an den Standdardepressionsdiagnosen des DIA-X (N = 448).

Cut-Off-Wert	**0/1**	**1/2**	**2/3**	**3/4**	**4/5**	**5/6**	**6/7**	**7/8**	**8/9**	**9/10**	**10/11**	**11/12**
Overall Accuracy	0.60	0.67	0.72	0.74	0.79	0.83	0.84	0.86	0.86	0.86	0.86	0.85
Kappa Index	0.24	0.30	0.35	0.37	0.43	0.50	0.50	0.53	0.51	0.48	0.44	0.34
positiver Likelihood Quotient	1.89	2.30	2.69	2.90	3.67	4.88	5.46	6.83	7.49	8.93	10.12	10.71
negativer Likelihood Quotient	0.21	0.23	0.29	0.32	0.34	0.33	0.38	0.40	0.46	0.54	0.61	0.73

In Tabelle 70 sind die Overall Accuracy, der Kappa Index und die Likelihood Quotienten der gesamten Stichprobe für die Cut-Off-Werte des GHQ-12, validiert an den Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom, abgebildet.

Tabelle 70: Overall Accuracy, Kappa Index und Likelihood Quotienten des GHQ-12 für die gesamte Stichprobe validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 448).

Cut-Off-Wert	**0/1**	**1/2**	**2/3**	**3/4**	**4/5**	**5/6**	**6/7**	**7/8**	**8/9**	**9/10**	**10/11**	**11/12**
Overall Accuracy	0.67	0.73	0.76	0.76	0.78	0.80	0.80	0.81	0.80	0.78	0.77	0.75
Kappa Index	0.37	0.44	0.48	0.46	0.47	0.50	0.50	0.49	0.45	0.39	0.35	0.24
positiver Likelihood Quotient	2.10	2.63	3.22	3.27	4.06	5.33	6.53	8.67	9.18	10.22	13.76	10.71
negativer Likelihood Quotient	0.24	0.28	0.33	0.39	0.44	0.47	0.50	0.53	0.59	0.66	0.71	0.81

Overall Accuracy, Kappa Index und Likelihood Quotienten des GHQ-12 für die Subgruppe der Frauen

Tabelle 71 zeigt die Overall Accuracy, den Kappa Index und die Likelihood Quotienten für die Subgruppe der Frauen für die Cut-Off-Werte des GHQ-12, validiert anhand der Standarddepressionsdiagnosen des DIA-X.

Tabelle 71: Overall Accuracy, Kappa Index und Likelihood Quotienten des GHQ-12 für die Subgruppe der Frauen validiert an den Standarddepressionsdiagnosen des DIA-X (N = 283).

Cut-Off-Wert	**0/1**	**1/2**	**2/3**	**3/4**	**4/5**	**5/6**	**6/7**	**7/8**	**8/9**	**9/10**	**10/11**	**11/12**
Overall Accuracy	0.60	0.68	0.73	0.75	0.80	0.84	0.85	0.86	0.85	0.86	0.85	0.84
Kappa Index	0.25	0.34	0.39	0.40	0.47	0.55	0.55	0.55	0.52	0.49	0.44	0.34
positiver Likelihood Quotient	1.89	2.39	2.85	3.01	4.06	5.81	6.55	7.51	7.27	9.25	9.52	12.69
negativer Likelihood Quotient	0.20	0.19	0.27	0.31	0.32	0.32	0.37	0.40	0.46	0.54	0.61	0.74

In Tabelle 72 sind die Overall Accuracy, der Kappa Index und die Likelihood Quotienten der Subgruppe der Frauen für die Cut-Off-Werte des GHQ-12, validiert an den Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom, wiedergegeben.

Tabelle 72: Overall Accuracy, Kappa Index und Likelihood Quotienten des GHQ-12 für die Subgruppe der Frauen validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 283).

Cut-Off-Wert	**0/1**	**1/2**	**2/3**	**3/4**	**4/5**	**5/6**	**6/7**	**7/8**	**8/9**	**9/10**	**10/11**	**11/12**
Overall Accuracy	0.67	0.73	0.78	0.77	0.79	0.82	0.82	0.81	0.80	0.78	0.77	0.73
Kappa Index	0.37	0.45	0.51	0.49	0.50	0.54	0.53	0.50	0.46	0.38	0.35	0.20
positiver Likelihood Quotient	2.09	2.66	3.51	3.49	4.43	6.89	8.32	9.01	8.19	9.01	10.51	7.21
negativer Likelihood Quotient	0.22	0.25	0.29	0.36	0.43	0.44	0.48	0.52	0.57	0.66	0.70	0.84

Overall Accuracy, Kappa Index und Likelihood Quotienten des GHQ-12 für die Subgruppe der Männer

In Tabelle 73 sind die Overall Accuracy, der Kappa Index und die Likelihood Quotienten für die Subgruppe der Männer für die Cut-Off-Werte des GHQ-12, validiert anhand der Standarddepressionsdiagnosen des DIA-X, dargestellt.

Tabelle 73: Overall Accuracy, Kappa Index und Likelihood Quotienten des GHQ-12 für die Subgruppe der Männer validiert an den Standarddepressionsdiagnosen des DIA-X (N = 165).

Cut-Off-Wert	**0/1**	**1/2**	**2/3**	**3/4**	**4/5**	**5/6**	**6/7**	**7/8**	**8/9**	**9/10**	**10/11**	**11/12**
Overall Accuracy	0.59	0.65	0.70	0.73	0.76	0.80	0.82	0.85	0.87	0.87	0.88	0.87
Kappa Index	0.21	0.24	0.27	0.31	0.35	0.41	0.43	0.49	0.50	0.46	0.44	0.35
positiver Likelihood Quotient	1.90	2.11	2.42	2.73	3.15	3.88	4.33	5.97	7.84	8.40	11.20	8.96
negativer Likelihood Quotient	0.22	0.32	0.35	0.33	0.36	0.34	0.38	0.40	0.47	0.55	0.62	0.71

In Tabelle 74 sind die Overall Accuracy, der Kappa Index und die Likelihood Quotienten der Subgruppe der Männer für die Cut-Off-Werte des GHQ-12, validiert an den Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom, abgebildet.

Tabelle 74: Overall Accuracy, Kappa Index und Likelihood Quotienten des GHQ-12 für die Subgruppe der Männer validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 165).

Cut-Off-Wert	**0/1**	**1/2**	**2/3**	**3/4**	**4/5**	**5/6**	**6/7**	**7/8**	**8/9**	**9/10**	**10/11**	**11/12**
Overall Accuracy	0.67	0.72	0.73	0.74	0.76	0.76	0.78	0.81	0.80	0.79	0.78	0.77
Kappa Index	0.36	0.42	0.41	0.41	0.44	0.42	0.44	0.48	0.44	0.39	0.35	0.30
positiver Likelihood Quotient	2.11	2.57	2.78	2.91	3.55	3.76	4.73	8.12	11.84	13.41	33.14	28.41
negativer Likelihood Quotient	0.27	0.32	0.41	0.45	0.47	0.53	0.53	0.54	0.61	0.67	0.72	0.76

Overall Accuracy, Kappa Index und Likelihood Quotienten des GHQ-12 für die Subgruppe Alter < 56,21 Jahre

In Tabelle 75 sind die Overall Accuracy, der Kappa Index und die Likelihood Quotienten für die Subgruppe Alter < 56,21 Jahre für die Cut-Off-Werte des GHQ-12, validiert anhand der Standarddepressionsdiagnosen des DIA-X, wiedergegeben.

Tabelle 75: Overall Accuracy, Kappa Index und Likelihood Quotienten des GHQ-12 für die Subgruppe Alter < 56,21 Jahre validiert an den Standarddepressionsdiagnosen des DIA-X (N = 224).

Cut-Off-Wert	**0/1**	**1/2**	**2/3**	**3/4**	**4/5**	**5/6**	**6/7**	**7/8**	**8/9**	**9/10**	**10/11**	**11/12**
Overall Accuracy	0.54	0.63	0.69	0.71	0.76	0.82	0.83	0.86	0.84	0.83	0.83	0.81
Kappa Index	0.22	0.30	0.37	0.39	0.46	0.55	0.58	0.62	0.56	0.52	0.47	0.35
positiver Likelihood Quotient	1.60	1.94	2.37	2.48	3.22	4.56	5.48	7.23	7.17	7.68	9.60	10.45
negativer Likelihood Quotient	0.17	0.17	0.23	0.25	0.29	0.26	0.29	0.30	0.40	0.49	0.57	0.71

Tabelle 76 zeigt die Overall Accuracy, den Kappa Index und die Likelihood Quotienten der Subgruppe Alter < 56,21 Jahre für die Cut-Off-Werte des GHQ-12, validiert an den Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom.

Tabelle 76: Overall Accuracy, Kappa Index und Likelihood Quotienten des GHQ-12 für die Subgruppe Alter < 56,21 Jahre validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 224).

Cut-Off-Wert	**0/1**	**1/2**	**2/3**	**3/4**	**4/5**	**5/6**	**6/7**	**7/8**	**8/9**	**9/10**	**10/11**	**11/12**
Overall Accuracy	0.64	0.71	0.75	0.76	0.77	0.79	0.81	0.81	0.78	0.76	0.76	0.71
Kappa Index	0.34	0.45	0.49	0.51	0.52	0.54	0.57	0.57	0.50	0.44	0.42	0.28
positiver Likelihood Quotient	1.75	2.27	2.78	3.04	3.86	5.29	7.88	11.68	10.05	12.10	52.33	34.34
negativer Likelihood Quotient	0.20	0.19	0.28	0.29	0.37	0.40	0.41	0.43	0.52	0.59	0.63	0.76

Overall Accuracy, Kappa Index und Likelihood Quotienten des GHQ-12 für die Subgruppe Alter > 56,20 Jahre

In Tabelle 77 sind die Overall Accuracy, der Kappa Index und die Likelihood Quotienten für die Subgruppe Alter > 56,20 Jahre für die Cut-Off-Werte des GHQ-12, validiert anhand der Standarddepressionsdiagnosen des DIA-X, dargestellt.

Tabelle 77: Overall Accuracy, Kappa Index und Likelihood Quotienten des GHQ-12 für die Subgruppe Alter > 56,20 Jahre validiert an den Standarddepressionsdiagnosen des DIA-X (N = 224).

Cut-Off-Wert	**0/1**	**1/2**	**2/3**	**3/4**	**4/5**	**5/6**	**6/7**	**7/8**	**8/9**	**9/10**	**10/11**	**11/12**
Overall Accuracy	0.65	0.71	0.75	0.78	0.81	0.84	0.84	0.86	0.88	0.89	0.89	0.89
Kappa Index	0.21	0.25	0.27	0.29	0.35	0.38	0.36	0.35	0.39	0.39	0.36	0.31
positiver Likelihood Quotient	2.17	2.56	2.79	3.10	3.88	4.67	4.64	5.15	6.74	9.12	9.38	10.21
negativer Likelihood Quotient	0.30	0.36	0.44	0.46	0.44	0.47	0.54	0.61	0.59	0.66	0.69	0.76

In Tabelle 78 sind die Overall Accuracy, der Kappa Index und die Likelihood Quotienten der Subgruppe Alter > 56,20 Jahre für die Cut-Off-Werte des GHQ-12, validiert an den Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom, abgebildet.

Tabelle 78: Overall Accuracy, Kappa Index und Likelihood Quotienten des GHQ-12 für die Subgruppe Alter > 56,20 Jahre validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 224).

Cut-Off-Wert	**0/1**	**1/2**	**2/3**	**3/4**	**4/5**	**5/6**	**6/7**	**7/8**	**8/9**	**9/10**	**10/11**	**11/12**
Overall Accuracy	0.70	0.74	0.78	0.76	0.79	0.80	0.80	0.80	0.81	0.80	0.79	0.78
Kappa Index	0.35	0.38	0.42	0.34	0.38	0.40	0.36	0.34	0.35	0.26	0.21	0.15
positiver Likelihood Quotient	2.42	2.83	3.50	3.05	3.83	4.79	4.75	5.55	7.21	6.78	5.65	4.75
negativer Likelihood Quotient	0.32	0.41	0.43	0.57	0.57	0.59	0.64	0.69	0.70	0.79	0.83	0.89

4.4.2 Overall Accuracy, Kappa Index und Likelihood Quotienten des B-PHQ Auswertungsalgorithmus A

Overall Accuracy, Kappa Index und Likelihood Quotienten des B-PHQ Auswertungsalgorithmus A für die gesamte Stichprobe

In Tabelle 79 sind die Overall Accuracy, der Kappa Index und die Likelihood Quotienten der gesamten Stichprobe für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus A, validiert anhand der Standarddepressionsdiagnosen des DIA-X, dargestellt.

Tabelle 79: Overall Accuracy, Kappa Index und Likelihood Quotienten des B-PHQ Auswertungsalgorithmus A für die gesamte Stichprobe validiert an den Standarddepressionsdiagnosen des DIA-X (N = 448).

Cut-Off-Wert	mindestens das Item 1a oder 1b ist positiv	mindestens 2 Items sind positiv darunter 1a oder 1b	mindestens 3 Items sind positiv darunter 1a oder 1b	mindestens 4 Items sind positiv darunter 1a oder 1b	mindestens 5 Items sind positiv darunter 1a oder 1b	mindestens 6 Items sind positiv darunter 1a oder 1b	mindestens 7 Items sind positiv darunter 1a oder 1b	mindestens 8 Items sind positiv darunter 1a oder 1b	alle 9 Items sind positiv
Overall Accuracy	0.82	0.84	0.86	0.87	0.88	0.86	0.85	0.85	0.83
Kappa Index	0.51	0.55	0.59	0.59	0.57	0.46	0.40	0.30	0.12
positiver Likelihood Quotient	4.61	5.47	6.70	7.61	9.30	9.17	10.79	16.96	10.41
negativer likelihood Quotient	0.25	0.24	0.26	0.32	0.42	0.58	0.67	0.78	0.92

Tabelle 80 zeigt die Overall Accuracy, den Kappa Index und die Likelihood Quotienten der gesamten Stichprobe für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus A, validiert an den Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom.

Tabelle 80: Overall Accuracy, Kappa Index und Likelihood Quotienten des B-PHQ Auswertungsalgorithmus A für die gesamte Stichprobe validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 448).

Cut-Off-Wert	mindestens das Item 1a oder 1b ist positiv	mindestens 2 Items sind positiv darunter 1a oder 1b	mindestens 3 Items sind positiv darunter 1a oder 1b	mindestens 4 Items sind positiv darunter 1a oder 1b	mindestens 5 Items sind positiv darunter 1a oder 1b	mindestens 6 Items sind positiv darunter 1a oder 1b	mindestens 7 Items sind positiv darunter 1a oder 1b	mindestens 8 Items sind positiv darunter 1a oder 1b	alle 9 Items sind positiv
Overall Accuracy	0.80	0.83	0.82	0.82	0.81	0.79	0.77	0.75	0.71
Kappa Index	0.53	0.57	0.55	0.52	0.49	0.40	0.32	0.22	0.09
positiver Likelihood Quotient	5.05	6.73	7.74	8.81	13.14	15.73	21.22	52.76	20.65
negativer likelihood Quotient	0.40	0.39	0.44	0.49	0.56	0.66	0.74	0.83	0.94

Overall Accuracy, Kappa Index und Likelihood Quotienten des B-PHQ Auswertungsalgorithmus A für die Subgruppe der Frauen

Tabelle 81 zeigt die Overall Accuracy, den Kappa Index und die Likelihood Quotienten für die Subgruppe der Frauen für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus A, validiert anhand der Standarddepressionsdiagnosen des DIA-X.

Tabelle 81: Overall Accuracy, Kappa Index und Likelihood Quotienten des B-PHQ Auswertungsalgorithmus A für die Subgruppe der Frauen validiert an den Standarddepressionsdiagnosen des DIA-X (N = 283).

Cut-Off-Wert	mindestens das Item 1a oder 1b ist positiv	mindestens 2 Items sind positiv darunter 1a oder 1b	mindestens 3 Items sind positiv darunter 1a oder 1b	mindestens 4 Items sind positiv darunter 1a oder 1b	mindestens 5 Items sind positiv darunter 1a oder 1b	mindestens 6 Items sind positiv darunter 1a oder 1b	mindestens 7 Items sind positiv darunter 1a oder 1b	mindestens 8 Items sind positiv darunter 1a oder 1b	alle 9 Items sind positiv
Overall Accuracy	0.84	0.86	0.86	0.87	0.87	0.86	0.84	0.84	0.81
Kappa Index	0.55	0.59	0.60	0.60	0.58	0.49	0.40	0.34	0.13
positiver Likelihood Quotient	5.29	6.23	6.94	8.14	10.37	11.45	11.33	29.74	19.82
negativer likelihood Quotient	0.27	0.26	0.29	0.34	0.43	0.57	0.67	0.74	0.92

In Tabelle 82 sind die Overall Accuracy, der Kappa Index und die Likelihood Quotienten der Subgruppe der Frauen für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus A, validiert an den Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom, wiedergegeben.

Tabelle 82: Overall Accuracy, Kappa Index und Likelihood Quotienten des B-PHQ Auswertungsalgorithmus A für die Subgruppe der Frauen validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 283).

Cut-Off-Wert	mindestens das Item 1a oder 1b ist positiv	mindestens 2 Items sind positiv darunter 1a oder 1b	mindestens 3 Items sind positiv darunter 1a oder 1b	mindestens 4 Items sind positiv darunter 1a oder 1b	mindestens 5 Items sind positiv darunter 1a oder 1b	mindestens 6 Items sind positiv darunter 1a oder 1b	mindestens 7 Items sind positiv darunter 1a oder 1b	mindestens 8 Items sind positiv darunter 1a oder 1b	alle 9 Items sind positiv
Overall Accuracy	0.81	0.83	0.83	0.82	0.82	0.80	0.77	0.75	0.71
Kappa Index	0.54	0.57	0.56	0.53	0.52	0.42	0.32	0.23	0.07
positiver Likelihood Quotient	5.63	7.29	8.37	9.63	18.92	24.03	18.02	36.05	11.26
negativer likelihood Quotient	0.41	0.40	0.43	0.49	0.53	0.64	0.74	0.82	0.95

Overall Accuracy, Kappa Index und Likelihood Quotienten des B-PHQ Auswertungsalgorithmus A für die Subgruppe der Männer

In Tabelle 83 sind die Overall Accuracy, der Kappa Index und die Likelihood Quotienten für die Subgruppe der Männer für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus A, validiert anhand der Standarddepressionsdiagnosen des DIA-X, dargestellt.

Tabelle 83: Overall Accuracy, Kappa Index und Likelihood Quotienten des B-PHQ Auswertungsalgorithmus A für die Subgruppe der Männer validiert an den Standarddepressionsdiagnosen des DIA-X (N = 165).

Cut-Off-Wert	mindestens das Item 1a oder 1b ist positiv	mindestens 2 Items sind positiv darunter 1a oder 1b	mindestens 3 Items sind positiv darunter 1a oder 1b	mindestens 4 Items sind positiv darunter 1a oder 1b	mindestens 5 Items sind positiv darunter 1a oder 1b	mindestens 6 Items sind positiv darunter 1a oder 1b	mindestens 7 Items sind positiv darunter 1a oder 1b	mindestens 8 Items sind positiv darunter 1a oder 1b	alle 9 Items sind positiv
Overall Accuracy	0.79	0.82	0.87	0.87	0.88	0.86	0.87	0.85	0.85
Kappa Index	0.44	0.49	0.58	0.57	0.54	0.41	0.40	0.20	0.10
positiver Likelihood Quotient	3.92	4.70	6.53	7.09	8.15	6.84	10.08	7.47	5.60
negativer likelihood Quotient	0.20	0.19	0.18	0.27	0.39	0.60	0.66	0.86	0.93

Tabelle 84 zeigt die Overall Accuracy, den Kappa Index und die Likelihood Quotienten der Subgruppe der Männer für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus A, validiert an den Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom.

Tabelle 84: Overall Accuracy, Kappa Index und Likelihood Quotienten des B-PHQ Auswertungsalgorithmus A für die Subgruppe der Männer validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 165).

Cut-Off-Wert	mindestens das Item 1a oder 1b ist positiv	mindestens 2 Items sind positiv darunter 1a oder 1b	mindestens 3 Items sind positiv darunter 1a oder 1b	mindestens 4 Items sind positiv darunter 1a oder 1b	mindestens 5 Items sind positiv darunter 1a oder 1b	mindestens 6 Items sind positiv darunter 1a oder 1b	mindestens 7 Items sind positiv darunter 1a oder 1b	mindestens 8 Items sind positiv darunter 1a oder 1b	alle 9 Items sind positiv
Overall Accuracy	0.79	0.82	0.82	0.81	0.79	0.78	0.78	0.75	0.73
Kappa Index	0.51	0.57	0.54	0.51	0.43	0.35	0.32	0.19	0.11
positiver Likelihood Quotient	4.34	6.01	6.87	7.69	8.29	9.47	30.78	∞	∞
negativer likelihood Quotient	0.39	0.37	0.45	0.50	0.60	0.70	0.74	0.86	0.92

Overall Accuracy, Kappa Index und Likelihood Quotienten des B-PHQ Auswertungsalgorithmus A für die Subgruppe Alter < 56,21 Jahre

In Tabelle 85 sind die Overall Accuracy, der Kappa Index und die Likelihood Quotienten für die Subgruppe Alter < 56,21 Jahre für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus A, validiert anhand der Standarddepressionsdiagnosen des DIA-X, wiedergegeben.

Tabelle 85: Overall Accuracy, Kappa Index und Likelihood Quotienten des B-PHQ Auswertungsalgorithmus A für die Subgruppe Alter < 56,21 Jahre validiert an den Standarddepressionsdiagnosen des DIA-X (N = 224).

Cut-Off-Wert	mindestens das Item 1a oder 1b ist positiv	mindestens 2 Items sind positiv darunter 1a oder 1b	mindestens 3 Items sind positiv darunter 1a oder 1b	mindestens 4 Items sind positiv darunter 1a oder 1b	mindestens 5 Items sind positiv darunter 1a oder 1b	mindestens 6 Items sind positiv darunter 1a oder 1b	mindestens 7 Items sind positiv darunter 1a oder 1b	mindestens 8 Items sind positiv darunter 1a oder 1b	alle 9 Items sind positiv
Overall Accuracy	0.79	0.81	0.83	0.83	0.83	0.82	0.81	0.80	0.77
Kappa Index	0.52	0.56	0.59	0.58	0.55	0.47	0.40	0.30	0.11
positiver Likelihood Quotient	3.60	4.10	4.87	5.38	6.15	6.62	6.76	10.75	7.68
negativer likelihood Quotient	0.17	0.16	0.20	0.28	0.39	0.53	0.64	0.76	0.92

Tabelle 86 zeigt die Overall Accuracy, den Kappa Index und die Likelihood Quotienten für die Subgruppe Variable Alter < 56,21 Jahre für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus A, validiert an den Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom.

Tabelle 86: Overall Accuracy, Kappa Index und Likelihood Quotienten des B-PHQ Auswertungsalgorithmus A für die Subgruppe Alter < 56,21 Jahre validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 224).

Cut-Off-Wert	mindestens das Item 1a oder 1b ist positiv	mindestens 2 Items sind positiv darunter 1a oder 1b	mindestens 3 Items sind positiv darunter 1a oder 1b	mindestens 4 Items sind positiv darunter 1a oder 1b	mindestens 5 Items sind positiv darunter 1a oder 1b	mindestens 6 Items sind positiv darunter 1a oder 1b	mindestens 7 Items sind positiv darunter 1a oder 1b	mindestens 8 Items sind positiv darunter 1a oder 1b	alle 9 Items sind positiv
Overall Accuracy	0.79	0.82	0.81	0.81	0.80	0.77	0.74	0.70	0.65
Kappa Index	0.57	0.61	0.59	0.57	0.54	0.45	0.36	0.25	0.10
positiver Likelihood Quotient	4.19	5.23	6.03	7.36	10.98	15.13	15.81	∞	∞
negativer likelihood Quotient	0.30	0.29	0.35	0.40	0.47	0.58	0.67	0.79	0.92

Overall Accuracy, Kappa Index und Likelihood Quotienten des B-PHQ Auswertungsalgorithmus A für die Subgruppe Alter > 56,20 Jahre

In Tabelle 87 sind die Overall Accuracy, der Kappa Index und die Likelihood Quotienten für die Subgruppe Variable Alter > 56,20 Jahre für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus A, validiert anhand der Standarddepressionsdiagnosen des DIA-X, dargestellt.

Tabelle 87: Overall Accuracy, Kappa Index und Likelihood Quotienten des B-PHQ Auswertungsalgorithmus A für die Subgruppe Alter > 56,20 Jahre validiert an den Standarddepressionsdiagnosen des DIA-X (N = 224).

Cut-Off-Wert	mindestens das Item 1a oder 1b ist positiv	mindestens 2 Items sind positiv darunter 1a oder 1b	mindestens 3 Items sind positiv darunter 1a oder 1b	mindestens 4 Items sind positiv darunter 1a oder 1b	mindestens 5 Items sind positiv darunter 1a oder 1b	mindestens 6 Items sind positiv darunter 1a oder 1b	mindestens 7 Items sind positiv darunter 1a oder 1b	mindestens 8 Items sind positiv darunter 1a oder 1b	alle 9 Items sind positiv
Overall Accuracy	0.86	0.88	0.90	0.91	0.92	0.90	0.90	0.90	0.88
Kappa Index	0.43	0.49	0.54	0.55	0.55	0.39	0.35	0.27	0.11
positiver Likelihood Quotient	5.64	7.30	9.54	11.67	17.02	13.13	25.54	36.48	14.59
negativer likelihood Quotient	0.42	0.41	0.40	0.43	0.50	0.68	0.75	0.82	0.93

In Tabelle 88 sind die Overall Accuracy, der Kappa Index und die Likelihood Quotienten der Subgruppe Variable Alter > 56,20 Jahre für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus A, validiert an den Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom, abgebildet.

Tabelle 88: Overall Accuracy, Kappa Index und Likelihood Quotienten des B-PHQ Auswertungsalgorithmus A für die Subgruppe Alter > 56,20 Jahre validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 224).

Cut-Off-Wert	mindestens das Item 1a oder 1b ist positiv	mindestens 2 Items sind positiv darunter 1a oder 1b	mindestens 3 Items sind positiv darunter 1a oder 1b	mindestens 4 Items sind positiv darunter 1a oder 1b	mindestens 5 Items sind positiv darunter 1a oder 1b	mindestens 6 Items sind positiv darunter 1a oder 1b	mindestens 7 Items sind positiv darunter 1a oder 1b	mindestens 8 Items sind positiv darunter 1a oder 1b	alle 9 Items sind positiv
Overall Accuracy	0.81	0.83	0.83	0.83	0.83	0.81	0.80	0.79	0.78
Kappa Index	0.42	0.47	0.45	0.40	0.37	0.27	0.21	0.13	0.05
positiver Likelihood Quotient	5.43	8.14	9.33	9.21	13.57	12.44	27.14	16.96	6.78
negativer likelihood Quotient	0.58	0.56	0.60	0.65	0.70	0.80	0.85	0.91	0.97

4.4.3 Overall Accuracy, Kappa Index und Likelihood Quotienten des B-PHQ Auswertungsalgorithmus B

Overall Accuracy, Kappa Index und Likelihood Quotienten des B-PHQ Auswertungsalgorithmus B für die gesamte Stichprobe

In Tabelle 89 sind die Overall Accuracy, der Kappa Index und die Likelihood Quotienten der gesamten Stichprobe für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus B, validiert anhand der Standarddepressionsdiagnosen des DIA-X, dargestellt.

Tabelle 89: Overall Accuracy, Kappa Index und Likelihood Quotienten des B-PHQ Auswertungsalgorithmus B für die gesamte Stichprobe validiert an den Standarddepressionsdiagnosen des DIA-X (N = 448).

Cut-Off-Wert	mindestens 1 Item ist positiv	mindestens 2 Items sind positiv	mindestens 3 Items sind positiv	mindestens 4 Items sind positiv	mindestens 5 Items sind positiv	mindestens 6 Items sind positiv	mindestens 7 Items sind positiv	mindestens 8 Items sind positiv	alle 9 Items sind positiv
Overall Accuracy	0.61	0.76	0.82	0.86	0.87	0.86	0.85	0.85	0.83
Kappa Index	0.26	0.43	0.50	0.56	0.55	0.46	0.40	0.30	0.12
positiver Likelihood Quotient	1.99	3.34	4.53	6.47	8.58	9.17	10.79	16.96	10.41
negativer likelihood Quotient	0.16	0.18	0.27	0.33	0.42	0.58	0.67	0.78	0.92

In Tabelle 90 sind die Overall Accuracy, der Kappa Index und die Likelihood Quotienten der gesamten Stichprobe für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus B, validiert an den Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom, abgebildet.

Tabelle 90: Overall Accuracy, Kappa Index und Likelihood Quotienten des B-PHQ Auswertungsalgorithmus B für die gesamte Stichprobe validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 448).

Cut-Off-Wert	mindestens 1 Item ist positiv	mindestens 2 Items sind positiv	mindestens 3 Items sind positiv	mindestens 4 Items sind positiv	mindestens 5 Items sind positiv	mindestens 6 Items sind positiv	mindestens 7 Items sind positiv	mindestens 8 Items sind positiv	alle 9 Items sind positiv
Overall Accuracy	0.68	0.80	0.81	0.81	0.81	0.79	0.77	0.75	0.71
Kappa Index	0.38	0.56	0.53	0.51	0.49	0.40	0.32	0.22	0.09
positiver Likelihood Quotient	2.17	4.16	5.18	7.07	12.24	15.73	21.22	52.76	20.65
negativer likelihood Quotient	0.22	0.26	0.40	0.49	0.55	0.66	0.74	0.83	0.94

Overall Accuracy, Kappa Index und Likelihood Quotienten des B-PHQ Auswertungsalgorithmus B für die Subgruppe der Frauen

Tabelle 91 zeigt die Overall Accuracy, den Kappa Index und die Likelihood Quotienten für die Subgruppe der Frauen für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus B, validiert anhand der Standarddepressionsdiagnosen des DIA-X.

Tabelle 91: Overall Accuracy, Kappa Index und Likelihood Quotienten des B-PHQ Auswertungsalgorithmus B für die Subgruppe der Frauen validiert an den Standarddepressionsdiagnosen des DIA-X (N = 283).

Cut-Off-Wert	mindestens 1 Item ist positiv	mindestens 2 Items sind positiv	mindestens 3 Items sind positiv	mindestens 4 Items sind positiv	mindestens 5 Items sind positiv	mindestens 6 Items sind positiv	mindestens 7 Items sind positiv	mindestens 8 Items sind positiv	alle 9 Items sind positiv
Overall Accuracy	0.62	0.77	0.82	0.86	0.87	0.86	0.84	0.84	0.81
Kappa Index	0.29	0.47	0.51	0.58	0.58	0.49	0.40	0.34	0.13
positiver Likelihood Quotient	2.04	3.47	4.61	7.36	10.37	11.45	11.33	29.74	19.82
negativer likelihood Quotient	0.13	0.19	0.29	0.35	0.43	0.57	0.67	0.74	0.92

In Tabelle 92 sind die Overall Accuracy, der Kappa Index und die Likelihood Quotienten der Subgruppe der Frauen für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus B, validiert an den Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom, wiedergegeben.

Tabelle 92: Overall Accuracy, Kappa Index und Likelihood Quotienten des B-PHQ Auswertungsalgorithmus B für die Subgruppe der Frauen validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 283).

Cut-Off-Wert	mindestens 1 Item ist positiv	mindestens 2 Items sind positiv	mindestens 3 Items sind positiv	mindestens 4 Items sind positiv	mindestens 5 Items sind positiv	mindestens 6 Items sind positiv	mindestens 7 Items sind positiv	mindestens 8 Items sind positiv	alle 9 Items sind positiv
Overall Accuracy	0.69	0.81	0.81	0.81	0.82	0.80	0.77	0.75	0.71
Kappa Index	0.39	0.58	0.55	0.52	0.52	0.42	0.32	0.23	0.07
positiver Likelihood Quotient	2.20	4.32	5.58	8.15	18.92	24.03	18.02	36.05	11.26
negativer likelihood Quotient	0.19	0.25	0.39	0.49	0.53	0.64	0.74	0.82	0.95

Overall Accuracy, Kappa Index und Likelihood Quotienten des B-PHQ Auswertungsalgorithmus B für die Subgruppe der Männer

In Tabelle 93 sind die Overall Accuracy, der Kappa Index und die Likelihood Quotienten für die Subgruppe der Männer für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus B, validiert anhand der Standarddepressionsdiagnosen des DIA-X, dargestellt.

Tabelle 93: Overall Accuracy, Kappa Index und Likelihood Quotienten des B-PHQ Auswertungsalgorithmus B für die Subgruppe der Männer validiert an den Standarddepressionsdiagnosen des DIA-X (N = 165).

Cut-Off-Wert	mindestens 1 Item ist positiv	mindestens 2 Items sind positiv	mindestens 3 Items sind positiv	mindestens 4 Items sind positiv	mindestens 5 Items sind positiv	mindestens 6 Items sind positiv	mindestens 7 Items sind positiv	mindestens 8 Items sind positiv	alle 9 Items sind positiv
Overall Accuracy	0.59	0.75	0.82	0.85	0.87	0.86	0.87	0.85	0.85
Kappa Index	0.21	0.38	0.48	0.51	0.51	0.41	0.40	0.20	0.10
positiver Likelihood Quotient	1.90	3.16	4.52	5.60	6.89	6.84	10.08	7.47	5.60
negativer likelihood Quotient	0.22	0.17	0.20	0.28	0.40	0.60	0.66	0.86	0.93

In Tabelle 94 sind die Overall Accuracy, der Kappa Index und die Likelihood Quotienten der Subgruppe der Männer für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus B, validiert an den Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom, abgebildet.

Tabelle 94: Overall Accuracy, Kappa Index und Likelihood Quotienten des B-PHQ Auswertungsalgorithmus B für die Subgruppe der Männer validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 165).

Cut-Off-Wert	mindestens 1 Item ist positiv	mindestens 2 Items sind positiv	mindestens 3 Items sind positiv	mindestens 4 Items sind positiv	mindestens 5 Items sind positiv	mindestens 6 Items sind positiv	mindestens 7 Items sind positiv	mindestens 8 Items sind positiv	alle 9 Items sind positiv
Overall Accuracy	0.67	0.79	0.79	0.80	0.79	0.78	0.78	0.75	0.73
Kappa Index	0.36	0.54	0.50	0.49	0.44	0.35	0.32	0.19	0.11
positiver Likelihood Quotient	2.11	3.91	4.59	5.81	7.44	9.47	30.78	∞	∞
negativer likelihood Quotient	0.27	0.28	0.43	0.50	0.59	0.70	0.74	0.86	0.92

Overall Accuracy, Kappa Index und Likelihood Quotienten des B-PHQ Auswertungsalgorithmus B für die Subgruppe Alter < 56,21 Jahre

In Tabelle 95 sind die Overall Accuracy, der Kappa Index und die Likelihood Quotienten für die Subgruppe Alter < 56,21 Jahre für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus B, validiert anhand der Standarddepressionsdiagnosen des DIA-X, wiedergegeben.

Tabelle 95: Overall Accuracy, Kappa Index und Likelihood Quotienten des B-PHQ Auswertungsalgorithmus B für die Subgruppe Alter < 56,21 Jahre validiert an den Standarddepressionsdiagnosen des DIA-X (N = 224).

Cut-Off-Wert	mindestens 1 Item ist positiv	mindestens 2 Items sind positiv	mindestens 3 Items sind positiv	mindestens 4 Items sind positiv	mindestens 5 Items sind positiv	mindestens 6 Items sind positiv	mindestens 7 Items sind positiv	mindestens 8 Items sind positiv	alle 9 Items sind positiv
Overall Accuracy	0.61	0.72	0.79	0.82	0.83	0.82	0.81	0.80	0.77
Kappa Index	0.30	0.43	0.51	0.55	0.54	0.47	0.40	0.30	0.11
positiver Likelihood Quotient	1.90	2.69	3.62	4.61	5.82	6.62	6.76	10.75	7.68
negativer likelihood Quotient	0.11	0.16	0.21	0.28	0.39	0.53	0.64	0.76	0.92

Tabelle 96 zeigt die Overall Accuracy, den Kappa Index und die Likelihood Quotienten der Subgruppe Alter < 56,21 Jahre für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus B, validiert an den Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom.

Tabelle 96: Overall Accuracy, Kappa Index und Likelihood Quotienten des B-PHQ Auswertungsalgorithmus B für die Subgruppe Alter < 56,21 Jahre validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 224).

Cut-Off-Wert	mindestens 1 Item ist positiv	mindestens 2 Items sind positiv	mindestens 3 Items sind positiv	mindestens 4 Items sind positiv	mindestens 5 Items sind positiv	mindestens 6 Items sind positiv	mindestens 7 Items sind positiv	mindestens 8 Items sind positiv	alle 9 Items sind positiv
Overall Accuracy	0.71	0.79	0.80	0.80	0.80	0.77	0.74	0.70	0.65
Kappa Index	0.45	0.56	0.58	0.56	0.55	0.45	0.36	0.25	0.10
positiver Likelihood Quotient	2.20	3.41	4.68	6.00	11.21	15.13	15.81	∞	∞
negativer likelihood Quotient	0.14	0.22	0.31	0.40	0.46	0.58	0.67	0.79	0.92

Overall Accuracy, Kappa Index und Likelihood Quotienten des B-PHQ Auswertungsalgorithmus B für die Subgruppe Alter > 56,20 Jahre

In Tabelle 97 sind die Overall Accuracy, der Kappa Index und die Likelihood Quotienten für die Subgruppe Alter > 56,20 Jahre für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus B, validiert anhand der Standarddepressionsdiagnosen des DIA-X, dargestellt.

Tabelle 97: Overall Accuracy, Kappa Index und Likelihood Quotienten des B-PHQ Auswertungsalgorithmus B für die Subgruppe Alter > 56,20 Jahre validiert an den Standarddepressionsdiagnosen des DIA-X (N = 224).

Cut-Off-Wert	mindestens 1 Item ist positiv	mindestens 2 Items sind positiv	mindestens 3 Items sind positiv	mindestens 4 Items sind positiv	mindestens 5 Items sind positiv	mindestens 6 Items sind positiv	mindestens 7 Items sind positiv	mindestens 8 Items sind positiv	alle 9 Items sind positiv
Overall Accuracy	0.61	0.80	0.85	0.90	0.91	0.90	0.90	0.90	0.88
Kappa Index	0.19	0.40	0.44	0.52	0.53	0.39	0.35	0.27	0.11
positiver Likelihood Quotient	2.00	4.12	5.47	9.73	14.59	13.13	25.54	36.48	14.59
negativer likelihood Quotient	0.26	0.23	0.38	0.43	0.50	0.68	0.75	0.82	0.93

In Tabelle 98 sind die Overall Accuracy, der Kappa Index und die Likelihood Quotienten der Subgruppe Alter > 56,20 Jahre für die Cut-Off-Werte des B-PHQ Auswertungsalgorithmus B, validiert an den Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom, abgebildet.

Tabelle 98: Overall Accuracy, Kappa Index und Likelihood Quotienten des B-PHQ Auswertungsalgorithmus B für die Subgruppe Alter > 56,20 Jahre validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 224).

Cut-Off-Wert	mindestens 1 Item ist positiv	mindestens 2 Items sind positiv	mindestens 3 Items sind positiv	mindestens 4 Items sind positiv	mindestens 5 Items sind positiv	mindestens 6 Items sind positiv	mindestens 7 Items sind positiv	mindestens 8 Items sind positiv	alle 9 Items sind positiv
Overall Accuracy	0.65	0.82	0.81	0.82	0.82	0.81	0.80	0.79	0.78
Kappa Index	0.29	0.53	0.42	0.38	0.36	0.27	0.21	0.13	0.05
positiver Likelihood Quotient	2.03	4.88	4.99	7.16	10.85	12.44	27.14	16.96	6.78
negativer likelihood Quotient	0.35	0.34	0.57	0.66	0.71	0.80	0.85	0.91	0.97

4.5 Receiver Operating Characteristica (ROC) Kurven

4.5.1 ROC-Kurven des GHQ-12, des B-PHQ Auswertungsalgorithmus A und des B-PHQ Auswertungsalgorithmus B für die gesamte Stichprobe

In Abbildung 36 sind die ROC-Kurven für die gesamte Stichprobe für die Cut-Off-Werte des GHQ-12, des B-PHQ Auswertungsalgorithmus A und des B-PHQ Auswertungsalgorithmus B, validiert anhand der Standarddepressionsdiagnosen des DIA-X, dargestellt. Tabelle 99 enthält die zugehörigen Werte für die Flächen unter den ROC-Kurven (AUC) sowie die Standardfehler (SE) und die asymptotischen 95%-Konfidenzintervalle. Die Berechnung dieser Werte wurde mit Hilfe des Statistikprogramms *Statistical Package for Social Sciences (SPSS, 10.0,* SPSS Inc., 1999) durchgeführt. Die Schätzung der Standardfehler erfolgte anhand der nichtparametrischen Methode.

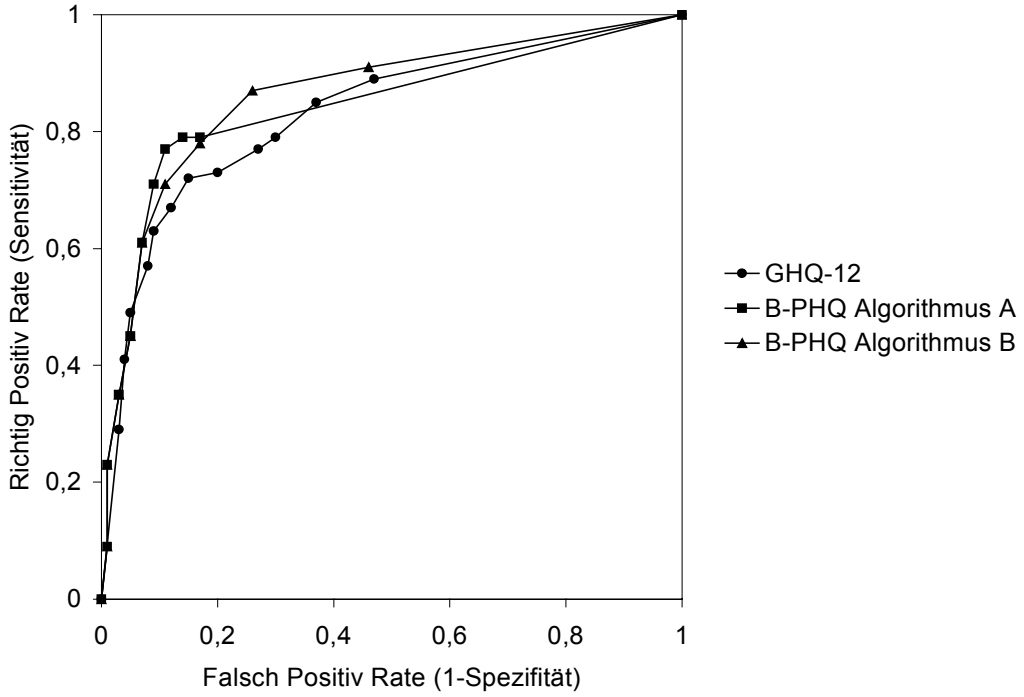


Abbildung 36: ROC-Kurven des GHQ-12, des B-PHQ Auswertungsalgorithmus A und des B-PHQ Auswertungsalgorithmus B für die gesamte Stichprobe validiert anhand der Standarddepressionsdiagnosen des DIA-X (N= 448).

Tabelle 99: Flächen unter den ROC-Kurven, Standardfehler und asymptotische 95%-Konfidenzintervalle des GHQ-12, des B-PHQ Auswertungsalgorithmus A und des B-PHQ Auswertungsalgorithmus B für die gesamte Stichprobe validiert anhand der Standarddepressionsdiagnosen des DIA-X (N= 448).

Screeninginstrument	Fläche unter der ROC-Kurve (AUC)	Standardfehler (SE)	Asymptotisches 95%-Konfidenzintervall
GHQ-12	0.833	0.027	[0.780, 0.887]
B-PHQ Algorithmus A	0.843	0.029	[0.787, 0.899]
B-PHQ Algorithmus B	0.864	0.025	[0.816, 0.912]

Die Vergleiche der Flächen unter der ROC-Kurve (AUC) wurden nach der von DeLong, DeLong und Clark-Pearson (1988) beschriebenen nichtparametrischen Methode für korrelierte ROC-Kurven mit Hilfe der SAS Software (Release 8.1, SAS Institute) durchgeführt. Die Testung der statistischen Signifikanz erfolgte über die Chi-Quadrat-Verteilung. In Tabelle 100 sind die zugehörigen Chi-Quadrat-Werte, Freiheitsgrade (DF) und P-Werte abgebildet.

Tabelle 100: Vergleich der Flächen unter den ROC-Kurven (AUC) der Screeninginstrumente für die gesamte Stichprobe validiert anhand der Standarddepressionsdiagnosen des DIA-X (N= 448). (α = .05, zweiseitig, nicht adjustiert).

	GHQ-12 versus B-PHQ Algorithmus A	GHQ-12 versus B-PHQ Algorithmus B	B-PHQ Algorithmus A versus B-PHQ Algorithmus B
Chi-Quadrat-Wert	0.162	1.550	2.245
DF	1	1	1
P-Wert	0.687	0.213	0.134

In Abbildung 37 sind die ROC-Kurven für die gesamte Stichprobe für die Cut-Off-Werte des GHQ-12, des B-PHQ Auswertungsalgorithmus A und des B-PHQ Auswertungsalgorithmus B validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom, abgebildet.

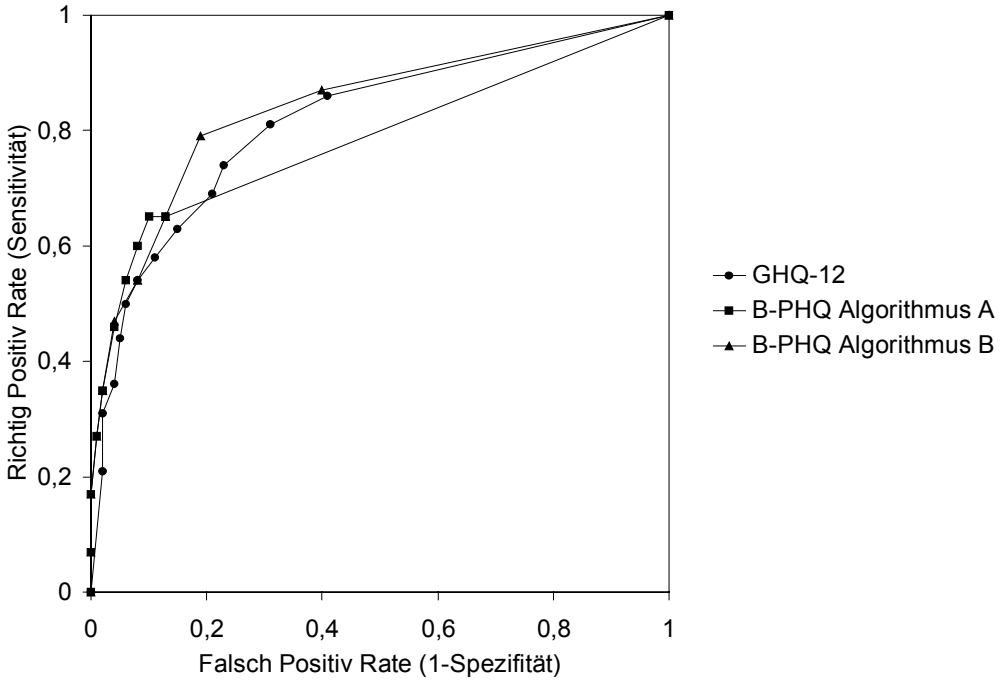


Abbildung 37: ROC-Kurven des GHQ-12, des B-PHQ Auswertungsalgorithmus A und des B-PHQ Auswertungsalgorithmus B für die gesamte Stichprobe validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N= 448).

Tabelle 101 enthält die zu Abbildung 37 gehörigen Werte der Flächen unter den ROC-Kurven (AUC), sowie die Standardfehler (SE) und die asymptotischen 95%-Konfidenzintervalle. In Tabelle 102 sind die Ergebnisse der AUC-Vergleiche abgebildet.

Tabelle 101: Flächen unter den ROC-Kurven, Standardfehler und asymptotische 95%-Konfidenzintervalle des GHQ-12, des B-PHQ Auswertungsalgorithmus A und des B-PHQ Auswertungsalgorithmus B für die gesamte Stichprobe validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N= 448).

Screeninginstrument	Fläche unter der ROC-Kurve (AUC)	Standardfehler (SE)	Asymptotisches 95%-Konfidenzintervall
GHQ-12	0.817	0.023	[0.772, 0.863]
B-PHQ Algorithmus A	0.783	0.027	[0.730, 0.836]
B-PHQ Algorithmus B	0.841	0.022	[0.797, 0.884]

Tabelle 102: Vergleich der Flächen unter den ROC-Kurven (AUC) der Screeninginstrumente für die gesamte Stichprobe validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N= 448). (α = .05, zweiseitig, nicht adjustiert).

	GHQ-12 versus B-PHQ Algorithmus A	GHQ-12 versus B-PHQ Algorithmus B	B-PHQ Algorithmus A versus B-PHQ Algorithmus B
Chi-Quadrat-Wert	2.498	1.340	12.628
DF	1	1	1
P-Wert	0.114	0.247	0.000 **

4.5.2 ROC-Kurven des GHQ-12, des B-PHQ Auswertungsalgorithmus A und des B-PHQ Auswertungsalgorithmus B getrennt nach Geschlecht

ROC-Kurven des GHQ-12, des B-PHQ Auswertungsalgorithmus A und des B-PHQ Auswertungsalgorithmus B für die Subgruppe der Frauen

Abbildung 38 zeigt die ROC-Kurven für die Subgruppe der Frauen für die Cut-Off-Werte des GHQ-12, des B-PHQ Auswertungsalgorithmus A und des B-PHQ Auswertungsalgorithmus B validiert anhand der Standarddepressionsdiagnosen des DIA-X. Tabelle 103 enthält die zugehörigen Werte für die Flächen unter den ROC-Kurven (AUC) sowie die Standardfehler (SE) und die asymptotischen 95%-Konfidenzintervalle. In Tabelle 104 sind die Ergebnisse der AUC-Vergleiche wiedergegeben.

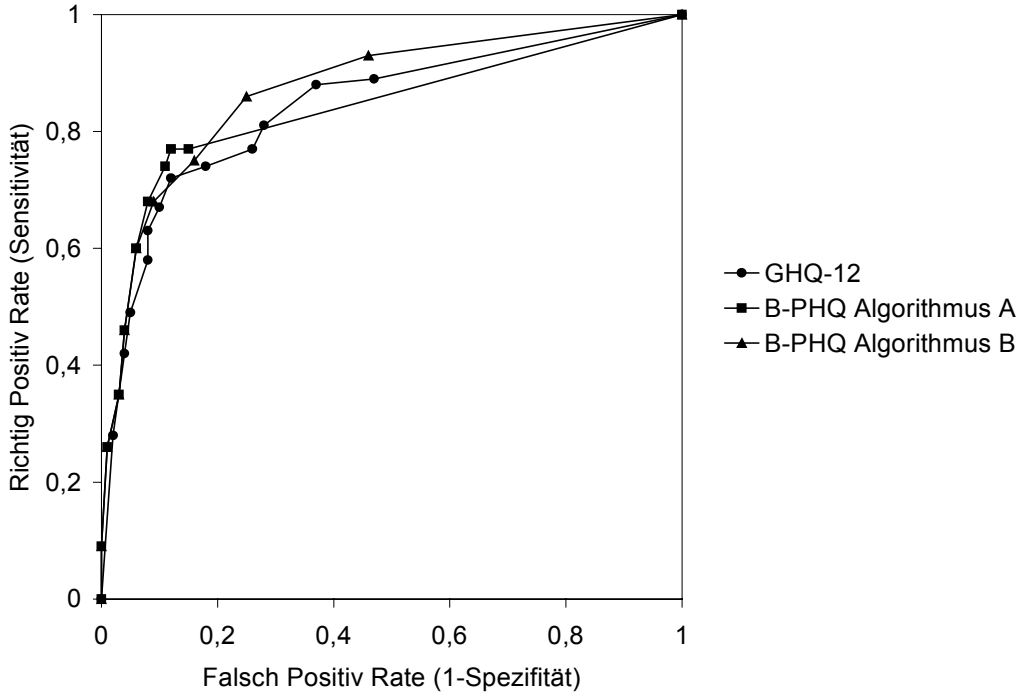


Abbildung 38: ROC-Kurven des GHQ-12, des B-PHQ Auswertungsalgorithmus A und des B-PHQ Auswertungsalgorithmus B für die Subgruppe der Frauen validiert anhand der Standarddepressionsdiagnosen des DIA-X (N= 283).

Tabelle 103: Flächen unter den ROC-Kurven, Standardfehler und asymptotische 95%-Konfidenzintervalle des GHQ-12, des B-PHQ Auswertungsalgorithmus A und des B-PHQ Auswertungsalgorithmus B für die Subgruppe der Frauen validiert anhand der Standarddepressionsdiagnosen des DIA-X (N= 283).

Screeninginstrument	Fläche unter der ROC-Kurve (AUC)	Standardfehler (SE)	Asymptotisches 95%-Konfidenzintervall
GHQ-12	0.842	0.032	[0.779, 0.905]
B-PHQ Algorithmus A	0.840	0.035	[0.771, 0.910]
B-PHQ Algorithmus B	0.873	0.028	[0.818, 0.928]

Tabelle 104: Vergleich der Flächen unter den ROC-Kurven (AUC) der Screeninginstrumente für die Subgruppe der Frauen validiert anhand der Standarddepressionsdiagnosen des DIA-X (N= 283). (α = .05, zweiseitig, nicht adjustiert).

	GHQ-12 versus B-PHQ Algorithmus A	GHQ-12 versus B-PHQ Algorithmus B	B-PHQ Algorithmus A versus B-PHQ Algorithmus B
Chi-Quadrat-Wert	0.002	0.891	3.268
DF	1	1	1
P-Wert	0.964	0.345	0.071

In Abbildung 39 sind die ROC-Kurven für die Subgruppe der Frauen für die Cut-Off-Werte des GHQ-12, des B-PHQ Auswertungsalgorithmus A und des B-PHQ Auswertungsalgorithmus B, validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom, abgebildet.

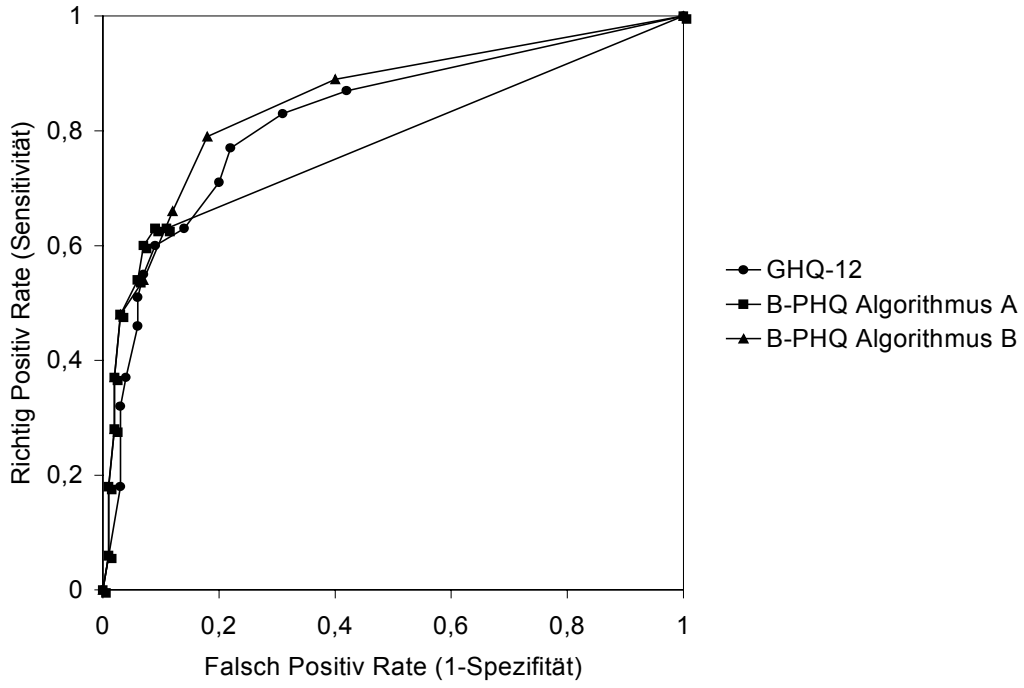


Abbildung 39: ROC-Kurven des GHQ-12, des B-PHQ Auswertungsalgorithmus A und des B-PHQ Auswertungsalgorithmus B für die Subgruppe der Frauen validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N= 283).

Tabelle 105 enthält die zu Abbildung 39 gehörigen Werte für die Flächen unter den ROC-Kurven (AUC) sowie die Standardfehler (SE) und die asymptotischen 95%-Konfidenzintervalle. In Tabelle 106 sind die Ergebnisse der AUC-Vergleiche dargestellt.

Tabelle 105: Flächen unter den ROC-Kurven, Standardfehler und asymptotische 95%-Konfidenzintervalle des GHQ-12, des B-PHQ Auswertungsalgorithmus A und des B-PHQ Auswertungsalgorithmus B für die Subgruppe der Frauen validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N= 283).

Screeninginstrument	Fläche unter der ROC-Kurve (AUC)	Standardfehler (SE)	Asymptotisches 95%-Konfidenzintervall
GHQ-12	0.827	0.028	[0.772, 0.882]
B-PHQ Algorithmus A	0.781	0.034	[0.714, 0.848]
B-PHQ Algorithmus B	0.853	0.027	[0.801, 0.905]

Tabelle 106: Vergleich der Flächen unter den ROC-Kurven (AUC) der Screeninginstrumente für die Subgruppe der Frauen validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N= 283). (α = .05, zweiseitig, nicht adjustiert).

	GHQ-12 versus B-PHQ Algorithmus A	GHQ-12 versus B-PHQ Algorithmus B	B-PHQ Algorithmus A versus B-PHQ Algorithmus B
Chi-Quadrat-Wert	2.426	0.822	11.899
DF	1	1	1
P-Wert	0.119	0.364	0.000 **

ROC-Kurven des GHQ-12, des B-PHQ Auswertungsalgorithmus A und des B-PHQ Auswertungsalgorithmus B für die Subgruppe der Männer

Abbildung 40 zeigt die ROC-Kurven für die Subgruppe der Männer für die Cut-Off-Werte des GHQ-12, des B-PHQ Auswertungsalgorithmus A und des B-PHQ Auswertungsalgorithmus B validiert anhand der Standarddepressionsdiagnosen des DIA-X. Tabelle 107 enthält die zugehörigen Werte für die Flächen unter den ROC-Kurven (AUC) sowie die Standardfehler (SE) und die asymptotischen 95%-Konfidenzintervalle. Tabelle 108 zeigt die Ergebnisse der zugehörigen AUC-Vergleiche.

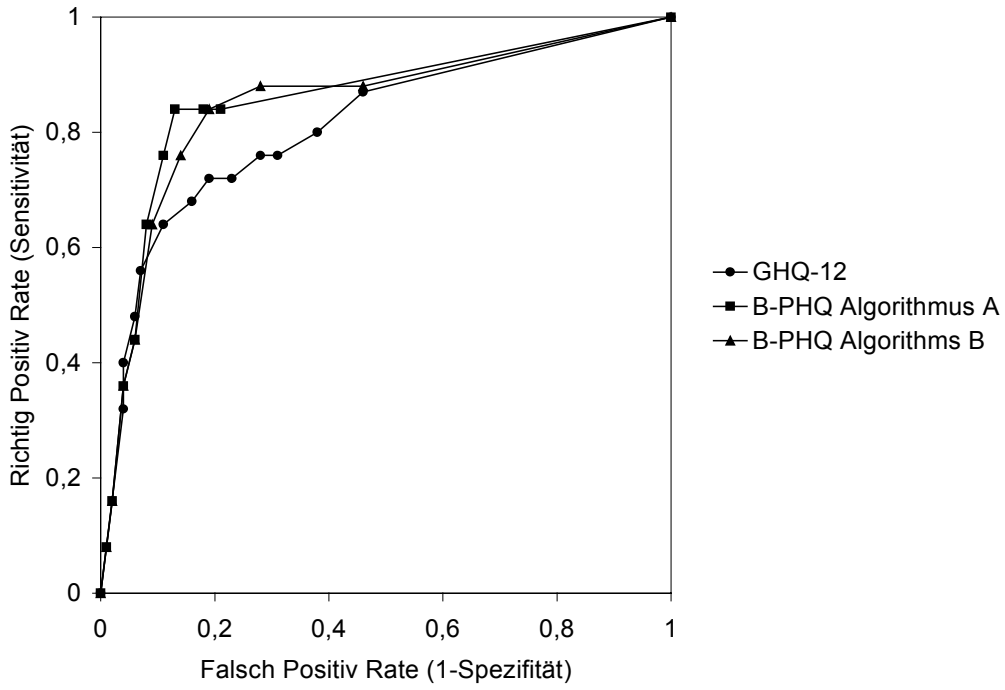


Abbildung 40: ROC-Kurven des GHQ-12, des B-PHQ Auswertungsalgorithmus A und des B-PHQ Auswertungsalgorithmus B für die Subgruppe der Männer validiert anhand der Standarddepressionsdiagnosen des DIA-X (N= 165).

Tabelle 107: Flächen unter den ROC-Kurven, Standardfehler und asymptotische 95%-Konfidenzintervalle des GHQ-12, des B-PHQ Auswertungsalgorithmus A und des B-PHQ Auswertungsalgorithmus B für die Subgruppe der Männer validiert anhand der Standarddepressionsdiagnosen des DIA-X (N= 165).

Screeninginstrument	Fläche unter der ROC-Kurve (AUC)	Standardfehler (SE)	Asymptotisches 95%-Konfidenzintervall
GHQ-12	0.818	0.051	[0.718, 0.918]
B-PHQ Algorithmus A	0.856	0.047	[0.763, 0.949]
B-PHQ Algorithmus B	0.849	0.048	[0.754, 0.943]

Tabelle 108: Vergleich der Flächen unter den ROC-Kurven (AUC) der Screeninginstrumente für die Subgruppe der Männer validiert anhand der Standarddepressionsdiagnosen des DIA-X (N= 165). (α = .05, zweiseitig, nicht adjustiert).

	GHQ-12 versus B-PHQ Algorithmus A	GHQ-12 versus B-PHQ Algorithmus B	B-PHQ Algorithmus A versus B-PHQ Algorithmus B
Chi-Quadrat-Wert	1.212	0.795	0.159
DF	1	1	1
P-Wert	0.271	0.373	0.610

In Abbildung 41 sind die ROC-Kurven für die Subgruppe der Männer für die Cut-Off-Werte des GHQ-12, des B-PHQ Auswertungsalgorithmus A und des B-PHQ Auswertungsalgorithmus B, validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom, abgebildet.

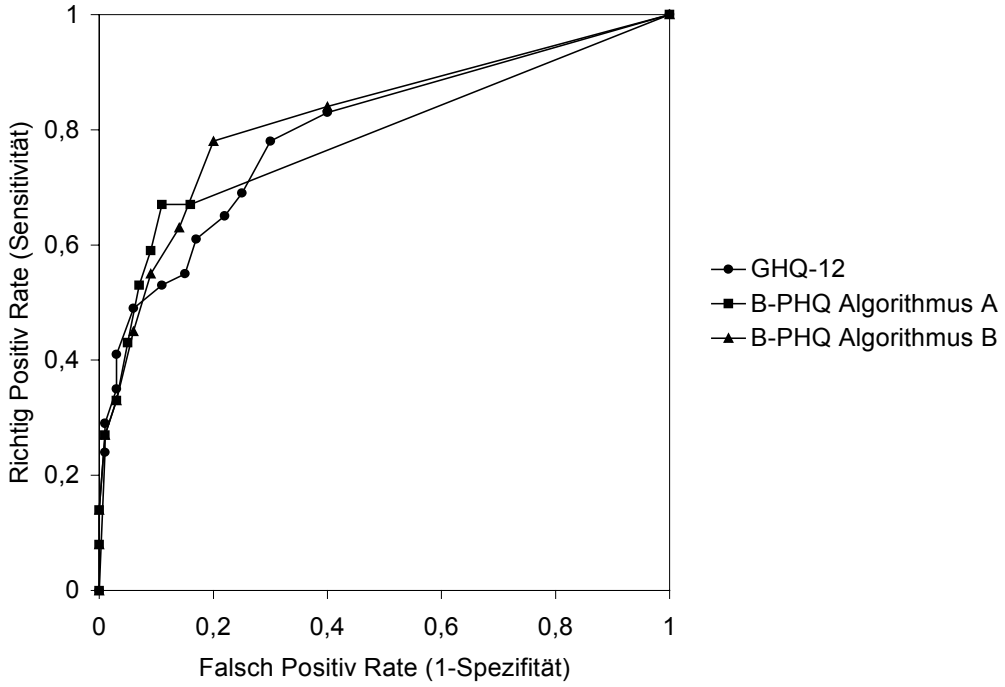


Abbildung 41: ROC-Kurven des GHQ-12, des B-PHQ Auswertungsalgorithmus A und des B-PHQ Auswertungsalgorithmus B für die Subgruppe der Männer validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N= 165).

Tabelle 109 enthält die zu Abbildung 41 gehörigen Werte für die Flächen unter den ROC-Kurven (AUC) sowie die Standardfehler (SE) und die asymptotischen 95%-Konfidenzintervalle. In Tabelle 110 sind die Werte der AUC-Vergleiche dargestellt.

Tabelle 109: Flächen unter den ROC-Kurven, Standardfehler und asymptotische 95%-Konfidenzintervalle des GHQ-12, des B-PHQ Auswertungsalgorithmus A und des B-PHQ Auswertungsalgorithmus B für die Subgruppe der Männer validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N= 165).

Screeninginstrument	Fläche unter der ROC-Kurve (AUC)	Standardfehler (SE)	Asymptotisches 95%-Konfidenzintervall
GHQ-12	0.800	0.040	[0.721, 0.879]
B-PHQ Algorithmus A	0.786	0.044	[0.700, 0.873]
B-PHQ Algorithmus B	0.819	0.039	[0.741, 0.896]

Tabelle 110: Vergleich der Flächen unter den ROC-Kurven (AUC) der Screeninginstrumente für die Subgruppe der Männer validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N= 165). (α = .05, zweiseitig, nicht adjustiert).

	GHQ-12 versus B-PHQ Algorithmus A	GHQ-12 versus B-PHQ Algorithmus B	B-PHQ Algorithmus A versus B-PHQ Algorithmus B
Chi-Quadrat-Wert	0.227	0.540	1.635
DF	1	1	1
P-Wert	0.634	0.462	0.201

4.5.3 ROC-Kurven des GHQ-12, des B-PHQ Auswertungsalgorithmus A und des B-PHQ Auswertungsalgorithmus B getrennt nach Altersgruppen

ROC-Kurven des GHQ-12, des B-PHQ Auswertungsalgorithmus A und des B-PHQ Auswertungsalgorithmus B für die Subgruppe Alter < 56,21 Jahre

Abbildung 42 zeigt die ROC-Kurven für die Subgruppe Alter < 56,21 für die Cut-Off-Werte des GHQ-12, des B-PHQ Auswertungsalgorithmus A und des B-PHQ Auswertungsalgorithmus B validiert anhand der Standarddepressionsdiagnosen des DIA-X.

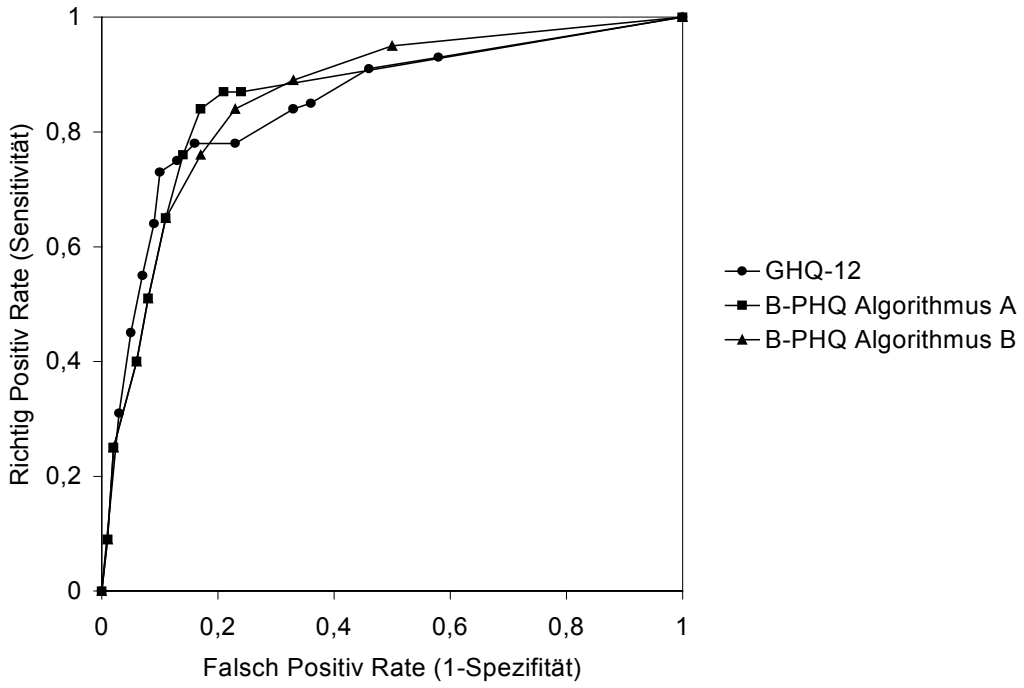


Abbildung 42: ROC-Kurven des GHQ-12, des B-PHQ Auswertungsalgorithmus A und des B-PHQ Auswertungsalgorithmus B für die Subgruppe Alter < 56,21 Jahre validiert anhand der Standarddepressionsdiagnosen des DIA-X (N= 224).

Tabelle 111 enthält die zu Abbildung 42 gehörigen Werte für die Flächen unter den ROC-Kurven (AUC) sowie die Standardfehler (SE) und die asymptotischen 95%-Konfidenzintervalle. In Tabelle 112 sind die Ergebnisse der AUC-Vergleiche wiedergegeben.

Tabelle 111: Flächen unter den ROC-Kurven, Standardfehler und asymptotische 95%-Konfidenzintervalle des GHQ-12, des B-PHQ Auswertungsalgorithmus A und des B-PHQ Auswertungsalgorithmus B für die Subgruppe Alter < 56,21 Jahre validiert anhand der Standarddepressionsdiagnosen des DIA-X (N= 224).

Screeninginstrument	Fläche unter der ROC-Kurve (AUC)	Standardfehler (SE)	Asymptotisches 95%-Konfidenzintervall
GHQ-12	0.849	0.033	[0.785, 0.913]
B-PHQ Algorithmus A	0.859	0.031	[0.797, 0.920]
B-PHQ Algorithmus B	0.861	0.029	[0.804, 0.918]

Tabelle 112: Vergleich der Flächen unter den ROC-Kurven (AUC) der Screeninginstrumente für die Subgruppe Alter < 56,21 Jahre validiert anhand der Standarddepressionsdiagnosen des DIA-X (N= 224). (α = .05, zweiseitig, nicht adjustiert).

	GHQ-12 versus B-PHQ Algorithmus A	GHQ-12 versus B-PHQ Algorithmus B	B-PHQ Algorithmus A versus B-PHQ Algorithmus B
Chi-Quadrat-Wert	0.100	0.176	0.040
DF	1	1	1
P-Wert	0.752	0.675	0.841

In Abbildung 43 sind die ROC-Kurven für die Subgruppe Alter < 56,21 Jahre für die Cut-Off-Werte des GHQ-12, des B-PHQ Auswertungsalgorithmus A und des B-PHQ Auswertungsalgorithmus B, validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom, abgebildet.

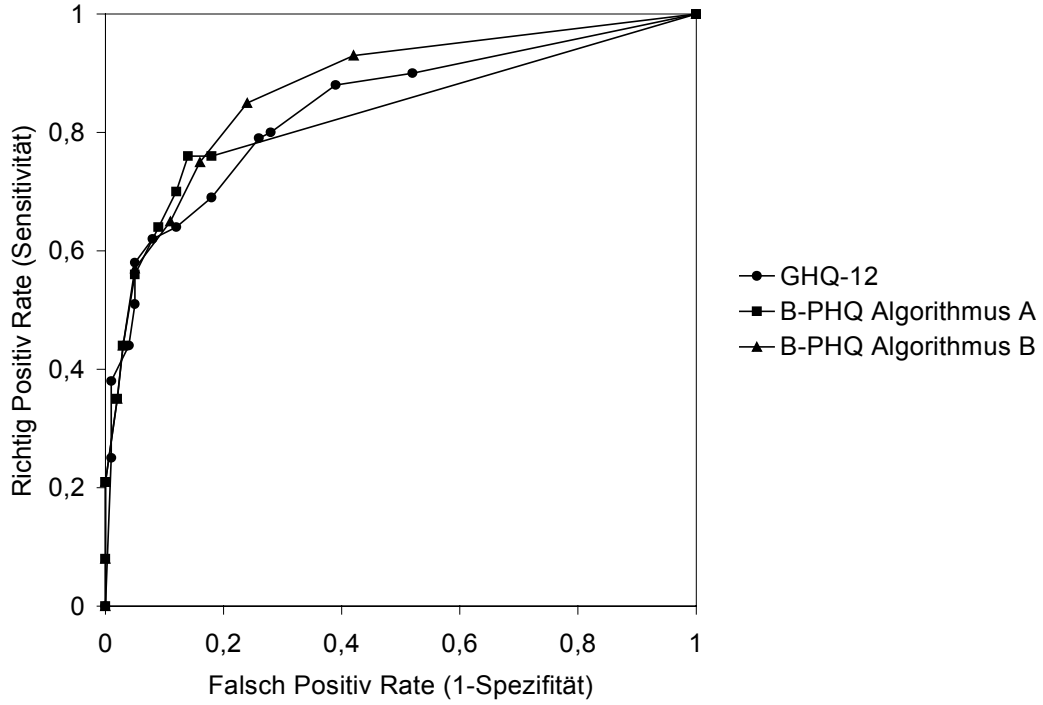


Abbildung 43: ROC-Kurven des GHQ-12, des B-PHQ Auswertungsalgorithmus A und des B-PHQ Auswertungsalgorithmus B für die Subgruppe Alter < 56,21 Jahre validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N= 224).

Tabelle 113 enthält die zu Abbildung 43 gehörigen Werte für die Flächen unter den ROC-Kurven (AUC) sowie die Standardfehler (SE) und die asymptotischen 95%-Konfidenzintervalle. In Tabelle 114 sind die Ergebnisse der AUC-Vergleiche abgebildet.

Tabelle 113: Flächen unter den ROC-Kurven, Standardfehler und asymptotische 95%-Konfidenzintervalle des GHQ-12, des B-PHQ Auswertungsalgorithmus A und des B-PHQ Auswertungsalgorithmus B für die Subgruppe Alter < 56,21 validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N= 224).

Screeninginstrument	Fläche unter der ROC-Kurve (AUC)	Standardfehler (SE)	Asymptotisches 95%-Konfidenzintervall
GHQ-12	0.844	0.028	[0.789, 0.900]
B-PHQ Algorithmus A	0.827	0.032	[0.765, 0.888]
B-PHQ Algorithmus B	0.867	0.026	[0.817, 0.918]

Tabelle 114: Vergleich der Flächen unter den ROC-Kurven (AUC) der Screeninginstrumente für die Subgruppe Alter < 56,21 Jahre validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N= 224). (α = .05, zweiseitig, nicht adjustiert).

	GHQ-12 versus B-PHQ Algorithmus A	GHQ-12 versus B-PHQ Algorithmus B	B-PHQ Algorithmus A versus B-PHQ Algorithmus B
Chi-Quadrat-Wert	0.456	0.988	5.573
DF	1	1	1
P-Wert	0.499	0.320	0.018 *

ROC-Kurven des GHQ-12, des B-PHQ Auswertungsalgorithmus A und des B-PHQ Auswertungsalgorithmus B für die Subgruppe Alter > 56,20 Jahre

Abbildung 44 zeigt die ROC-Kurven für die Subgruppe Alter > 56,20 Jahre für die Cut-Off-Werte des GHQ-12, des B-PHQ Auswertungsalgorithmus A und des B-PHQ Auswertungsalgorithmus B validiert anhand der Standarddepressionsdiagnosen des DIA-X.

Tabelle 115 enthält die zugehörigen Werte für die Flächen unter den ROC-Kurven (AUC) sowie die Standardfehler (SE) und die asymptotischen 95%-Konfidenzintervalle. Tabelle 116 zeigt die Ergebnisse der AUC-Vergleiche.

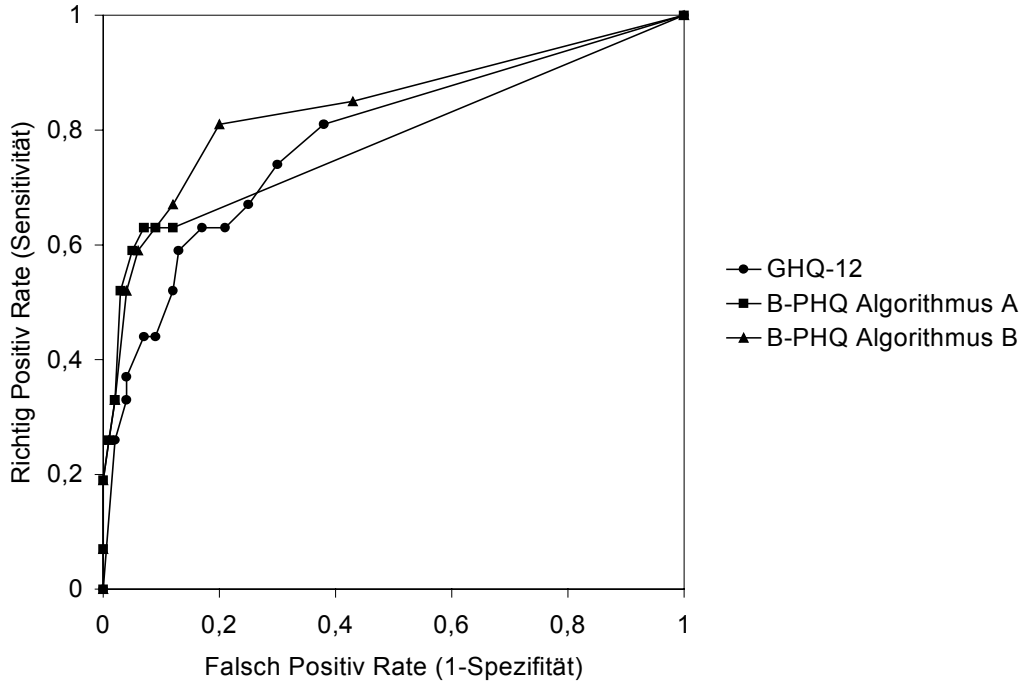


Abbildung 44: ROC-Kurven des GHQ-12, des B-PHQ Auswertungsalgorithmus A und des B-PHQ Auswertungsalgorithmus B für die Subgruppe Alter > 56,20 Jahre validiert anhand der Standarddepressionsdiagnosen des DIA-X (N= 224).

Tabelle 115: Flächen unter den ROC-Kurven, Standardfehler und asymptotische 95%-Konfidenzintervalle des GHQ-12, des B-PHQ Auswertungsalgorithmus A und des B-PHQ Auswertungsalgorithmus B für die Subgruppe Alter > 56,20 validiert anhand der Standarddepressionsdiagnosen des DIA-X (N= 224).

Screeninginstrument	Fläche unter der ROC-Kurve (AUC)	Standardfehler (SE)	Asymptotisches 95%-Konfidenzintervall
GHQ-12	0.788	0.052	[0.686, 0.889]
B-PHQ Algorithmus A	0.781	0.059	[0.665, 0.897]
B-PHQ Algorithmus B	0.840	0.049	[0.744, 0.937]

Tabelle 116: Vergleich der Flächen unter den ROC-Kurven (AUC) der Screeninginstrumente für die Subgruppe Alter > 56,20 Jahre validiert anhand der Standarddepressionsdiagnosen des DIA-X (N= 224). (α = .05, zweiseitig, nicht adjustiert).

	GHQ-12 versus B-PHQ Algorithmus A	GHQ-12 versus B-PHQ Algorithmus B	B-PHQ Algorithmus A versus B-PHQ Algorithmus B
Chi-Quadrat-Wert	0.0201	1.0549	2.639
DF	1	1	1
P-Wert	0.887	0.304	0.104

In Abbildung 45 sind die ROC-Kurven für die Subgruppe Alter > 56,20 für die Cut-Off-Werte des GHQ-12, des B-PHQ Auswertungsalgorithmus A und des B-PHQ Auswertungsalgorithmus B, validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom, abgebildet.

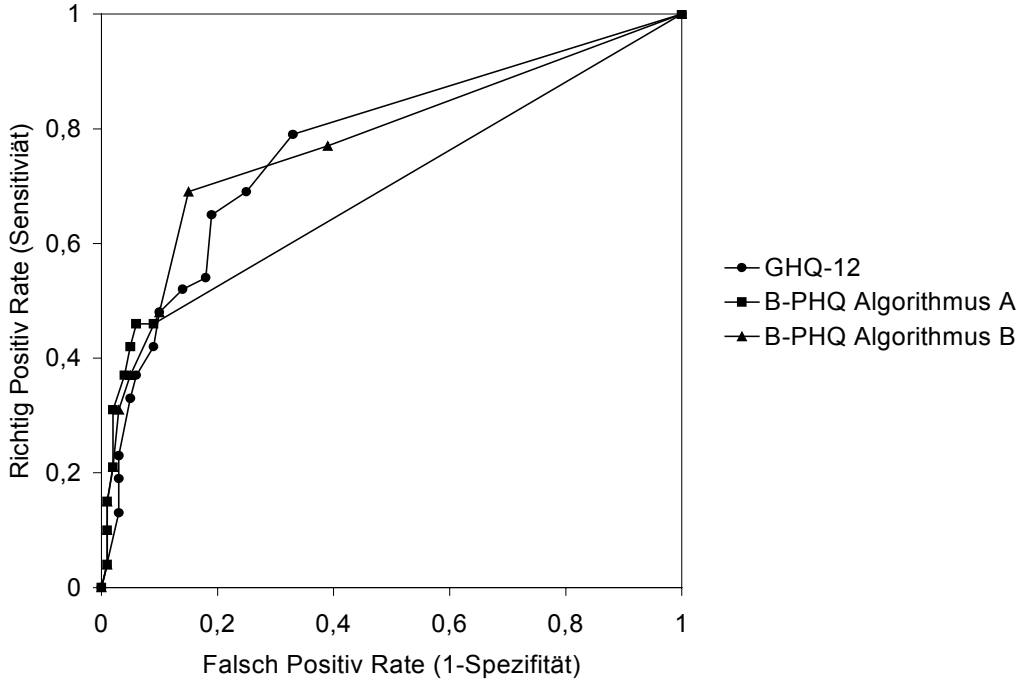


Abbildung 45: ROC-Kurven des GHQ-12, des B-PHQ Auswertungsalgorithmus A und des B-PHQ Auswertungsalgorithmus B für die Subgruppe Alter > 56,20 validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N= 224).

Tabelle 117 enthält die zu Abbildung 45 gehörigen Werte für die Flächen unter den ROC-Kurven (AUC) sowie die Standardfehler (SE) und die asymptotischen 95%-Konfidenzintervalle. In Tabelle 118 sind die Ergebnisse der AUC-Vergleiche wiedergegeben.

Tabelle 117: Flächen unter den ROC-Kurven, Standardfehler und asymptotische 95%-Konfidenzintervalle des GHQ-12, des B-PHQ Auswertungsalgorithmus A und des B-PHQ Auswertungsalgorithmus B für die Subgruppe Alter > 56,20 validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N= 224).

Screeninginstrument	Fläche unter der ROC-Kurve (AUC)	Standardfehler (SE)	Asymptotisches 95%-Konfidenzintervall
GHQ-12	0.772	0.040	[0.694, 0.850]
B-PHQ Algorithmus A	0.702	0.048	[0.609, 0.796]
B-PHQ Algorithmus B	0.788	0.041	[0.708, 0.869]

Tabelle 118: Vergleich der Flächen unter den ROC-Kurven (AUC) der Screeninginstrumente für die Subgruppe Alter > 56,20 Jahre validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N= 224). (α = .05, zweiseitig, nicht adjustiert).

	GHQ-12 versus B-PHQ Algorithmus A	GHQ-12 versus B-PHQ Algorithmus B	B-PHQ Algorithmus A versus B-PHQ Algorithmus B
Chi-Quadrat-Wert	3.430	0.175	6.979
DF	1	1	1
P-Wert	0.064	0.675	0.008 *

In Anhang B sind die Vier-Felder-Tafeln für die gesamte Stichprobe und die vier Untergruppen abgebildet. Die Vier-Felder-Tafeln sind die Grundlage für die Berechnung aller aufgeführten Validitätsmaße. In Anhang C finden sich aus Gründen der Vollständigkeit und für die Vergleichbarkeit mit anderen Studien die Werte für die restlichen, in Tabelle 1 aufgeführten Validitätsmaße: Rate der falsch Positiven, Rate der falsch Negativen, Misclassification Rate und Odds-Ratio.

5 Diskussion

Die eingangs gestellte Hypothese wurde bestätigt, denn die Ergebnisse der Datenauswertung zeigen, dass beide Fragebögen valide Screeninginstrumente zur Erfassung depressiver Syndrome (ICD-10) in der Primärversorgung sind. Mit Hilfe von AUC-Vergleichen lassen sich keine statistisch signifikanten Unterschiede in der Gesamtgenauigkeit zwischen den beiden Screeninginstrumenten nachweisen. Auch zwischen den beiden Auswertungsalgorithmen des B-PHQ ergeben sich keine statistisch signifikanten Unterschiede in der Gesamtgenauigkeit, so dass in der Primärversorgung auch der einfachere und schneller durchzuführende Auswertungsalgorithmus B eingesetzt werden kann.

Die explorative Validierung des GHQ-12 und des B-PHQ anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom zeigt, dass beide Fragebögen valide Instrumente zur Erfassung depressiver Syndrome, einschließlich unterschwelliger depressiver Syndrome, in der Primärversorgung sind. Die Ergebnisse der Flächenvergleiche ergeben keine statistisch signifikanten Unterschiede zwischen dem GHQ-12 und dem B-PHQ (Auswertungsalgorithmus A und B), sie deuten aber auf statistisch signifikante Unterschiede zwischen den beiden Auswertungsalgorithmen des B-PHQ hin. Der Auswertungsalgorithmus B verzichtet, ebenso wie die in der erweiterten Version des DIA-X verwendete Definition des unterschwelligen depressiven Syndroms, auf die Forderung einer hierarchischen Struktur der Symptome einer depressiven Störung. Dies zeigt sich in der größeren Gesamtgenauigkeit des Auswertungsalgorithmus B für die Validierung anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom, so dass auch hier der Auswertungsalgorithmus B eingesetzt werden sollte.

Die explorative Auswertung der Subgruppen ergibt Hinweise darauf, dass bei beiden Fragebögen in den einzelnen Subgruppen verschiedene optimale Cut-Off-Werte vorliegen. Diese Tendenzen müssen jedoch anhand von größeren Stichproben bestätigt werden.

Zusammenfassend kann festgehalten werden, dass beide Fragebögen valide Instrumente zur Erfassung depressiver Syndrome, einschließlich unterschwelli-

ger depressiver Syndrome, in der Primärversorgung sind. Wird nur die Depressionssektion des B-PHQ durchgeführt, so ergibt sich aufgrund der geringeren Anzahl von Fragen eine höhere Praktikabilität für den Einsatz des B-PHQ in der Primärversorgung.

5.1 Stichprobencharakteristika und Prävalenz

Die in der Stichprobe vorliegende Alters- und Geschlechtsverteilung deckt sich weitestgehend mit den Befunden aus anderen Studien im primärärztlichen Versorgungsbereich in der Bundesrepublik Deutschland. So lag das Verhältnis von Frauen zu Männern in der Studie „Depression 2000“ (Winter, Wittchen, Höfler, Spielgel, Ormel, Müller & Pfister, 2000) bei 59% zu 41%. Das der Studie von Kruse et al. (1999) zugrundeliegende Geschlechterverhältnis von 69% Frauen und 31% Männern entspricht nach Angaben der Kassenärztlichen Vereinigung der Verteilung in der ambulanten medizinischen Versorgung (vgl. Kruse et al., 1999). Das Verhältnis von Frauen zu Männern ist in der vorliegenden Stichprobe 63% zu 37%.

Die Altersverteilung der vorliegenden Stichprobe korrespondiert mit der Altersverteilung der von Winter et al. (2000) untersuchten Stichprobe. Bei Winter et al. (2000) liegt der Mittelwert bei 52 Jahren, der Median bei 54 Jahren und 49,4% der Probanden sind 55 Jahre und älter. In der vorliegenden Stichprobe liegt der Mittelwert bei 53 Jahren, der Median bei 56 Jahren und 52,4% der Probanden sind 55 Jahre und älter.

Die Verteilung der GHQ-Scores weist nur geringfügige Unterschiede zwischen Frauen und Männern auf. Bei der Auswertung getrennt nach der Variable Alter zeigt sich, dass in der Subgruppe Variable Alter > 56,20 Jahre deutlich mehr Probanden einen Gesamtscore vom Wert Null haben. Dieses Ergebnis spiegelt sich auch in den niedrigeren Prävalenzraten depressiver Störungen in dieser Subgruppe wider. Das gleiche Bild findet sich in der Verteilung der Scores im B-PHQ für beide Auswertungsalgorithmen.

Schätzungen der Prävalenz depressiver Störungen ergeben, dass bis zu 30% der Patienten in der Primärversorgung unter depressiven Störungen leiden

(Dawson & Tylee, 2001). Unterschiede in der Prävalenz verschiedener Studien sind u.a. auf die Verwendung verschiedener diagnostischer Instrumente sowie auf Unterschiede in der Falldefinition und im Zeitfenster zurückzuführen (vgl. Boyd & Weissman, 1981). Die Prävalenz depressiver Störungen (ICD-10) in den beiden Wochen vor Durchführung des DIA-X Interviews liegt in der hier untersuchten Stichprobe bei 18,30%. Mit Berücksichtigung der Diagnose unterschwelliges depressives Syndrom steigt die Prävalenz auf 30,35%. Die in der vorliegenden Stichprobe gefundenen Prävalenzraten liegen somit im oberen Bereich der in der Literatur gefundenen Werte und sind möglicherweise durch die Freiwilligkeit der Teilnahme der Patienten an der Untersuchung und den Forschungskontext der Studie beeinflusst. Auch können Selektionseffekte bei der Auswahl der Patienten in den Arztpraxen nicht vollständig ausgeschlossen werden.

5.2 Sensitivität und Spezifität

Sensitivität und Spezifität des GHQ-12

Die Auswahl des am besten geeigneten Cut-Off-Wertes hängt von der Fragestellung ab. Häufig wird der Cut-Off-Wert an oder in der Nähe der Schulter der ROC-Kurve gewählt. Die Summe von Sensitivität und Spezifität ist hier maximal. Im Rahmen von Screeningmaßnahmen ist es jedoch häufig sinnvoll einen anderen Cut-Off-Wert zu wählen, da möglichst alle Patienten mit einer depressiven Störung entdeckt werden sollen aber auf der anderen Seite aus Zeit- und Kostengründen die Rate der falsch Positiven möglichst gering bleiben sollte. Eine Möglichkeit dieses Ziel zu erreichen besteht darin, dass die Sensitivität und die Spezifität möglichst ausgewogen sind und, da es sich um Screeningmaßnahmen handelt, die Sensitivität so hoch wie möglich ist.
Nach dieser Regel ist der zu wählende beste Cut-Off-Wert für den GHQ-12, validiert anhand der Standarddepressionsdiagnosen des DIA-X, für die gesamte Stichprobe 2/3. Die Sensitivität beträgt hier 0.79 und die Spezifität 0.70. Die Summe von Sensitivität und Spezifität ist 1.49. Der an der Schulter der

ROC-Kurve gelegene Cut-Off-Wert 5/6 ist hier nicht optimal. Zwar wird hier eine größere Summe von Sensitivität und Spezifität erreicht (∑ = 1.57), allerdings ist die Sensitivität mit .72 kleiner als die Spezifität mit .85. Auch in den Subgruppen ist an der Schulter der ROC-Kurve die Spezifität größer als die Sensitivität.
Die explorative Analyse der Untergruppen ergibt für die Subgruppe der Frauen ebenfalls den Cut-Off-Wert 2/3 (Sensitivität = 0.81; Spezifität = 0.72; ∑ = 1.53).
Für die Subgruppe der Männer liegt der geeignetste Cut-Off-Wert bei 3/4 (Sensitivität = 0.76; Spezifität = 0.72, ∑ = 1.48).
In der Subgruppe Variable Alter < 56,21 Jahre ergibt sich der Cut-Off-Wert 4/5 (Sensitivität = 0.78; Spezifität = 0.76; ∑ = 1.54).
In der Subgruppe Variable Alter > 56,20 Jahre ist der beste Cut-Off-Wert 1/2. (Sensitivität = 0.74; Spezifität = 0.71; ∑ = 1.45).
Die Auswertung der Subgruppen ergibt für die vorliegende Stichprobe Tendenzen, die möglicherweise dafür sprechen, dass für die einzelne Subgruppen verschiedene Cut-Off-Werte gewählt werden sollten. Für eine Bestätigung dieser Tendenzen sind jedoch Folgestudien mit größeren Fallzahlen in den einzelnen Subgruppen erforderlich.
Zusammenfassend kann jedoch festgehalten werden, dass der für die Validierung des GHQ-12 anhand der Standarddepressionsdiagnosen des DIA-X für die gesamte Stichprobe ermittelte Cut-Off-Wert von 2/3 auch in den einzelnen Subgruppen verwendet werden kann, da die Werte für die Sensitivität und Spezifität akzeptabel sind.

Nach der oben formulierten Regel ergibt die Ermittlung des geeignetsten Cut-Off-Wertes anhand der erweiterten Version des DIA-X für die gesamte Stichprobe den Cut-Off-Wert 1/2 (Sensitivität = 0.81; Spezifität = 0.69; ∑ = 1.50).
Für die Subgruppe der Frauen ergibt sich der Cut-Off-Wert 1/2 (Sensitivität = 0.83; Spezifität = 0.69; ∑ = 1.52).
Für die Subgruppe der Männer ergibt sich der Cut-Off-Wert 1/2 (Sensitivität = 0.78; Spezifität = 0.70; ∑ = 1.48).
Für die Subgruppe Alter < 56,21 Jahre ergibt sich der Cut-Off-Wert 2/3 (Sensitivität = 0.80; Spezifität = 0.71; ∑ = 1.51).

Für die Subgruppe Alter > 56,20 Jahre ergibt sich der Cut-Off-Wert 0/1 (Sensitivität = 0.78; Spezifität = 0.68; ∑ = 1.46).
Die explorative Validierung des GHQ-12 an der modifizierten Form des DIA-X ergibt ebenfalls die Tendenz, dass möglicherweise für die Subgruppen verschiedene optimale Cut-Off-Werte existieren.
Zusammenfassend kann festgehalten werden, dass der Cut-Off-Wert 1/2 für die Validierung des GHQ-12 anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom auch in den einzelnen Subgruppen geeignet ist.
Die geringere Anzahl von Symptomen, die für die Diagnose unterschwelliges depressives Syndrom erforderlich ist, spiegelt sich in dem niedrigeren Schwellenwert 1/2 bei der Berechnung der Validitätsmaße anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom wider. Durch die Auswahl des niedrigeren Schwellenwertes ergeben sich ähnliche Werte der Sensitivität und der Spezifität bei der Validierung des GHQ-12 anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom wie für die Validierung anhand der Standarddepressionsdiagnosen.

Nach Furukawa und Goldberg (1998) sowie Goldberg et al. (1997) variiert der beste Cut-Off-Wert für den GHQ-12 bei der verwendeten Methode des GHQ-Scorings zwischen 0/1 und 6/7. Die für die vorliegende Stichprobe empfohlenen Cut-Off-Werte von 2/3 bzw. 1/2 fallen somit in diesen Bereich von Cut-Off-Werten.

Sensitivität und Spezifität des B-PHQ Auswertungsalgorithmus A

Die Berechnung der Sensitivität und der Spezifität anhand der Standarddepressionsdiagnosen des DIA-X ergibt für den Auswertungsalgorithmus A des B-PHQ hohe Werte der Spezifität, während die Sensitivitätswerte bei allen Cut-Off-Werten für die gesamte Stichprobe kleiner bleiben als die Spezifitätswerte. Für die gesamte Stichprobe liegt das Minimum der Spezifität bei 0.83, das Maximum der Sensitivität bei 0.79. Gleiches gilt auch für die Validierung anhand

der modifizierten Version des DIA-X. Hier liegt das Minimum der Spezifität für die gesamte Stichprobe bei 0.87 und das Maximum der Sensitivität bei 0.65. Soll der B-PHQ jedoch als Screeninginstrument eingesetzt werden, so bleibt das primäre Ziel eine möglichst hohe Sensitivität. Eine geeignete Auswahlregel bestünde hier darin, denjenigen Cut-Off-Wert zu wählen, der die höchste Sensitivität erreicht, da die Spezifität für alle Cut-Off-Werte hoch ist. Zusätzlich sollte bei gleicher Sensitivität der Cut-Off-Wert mit der höheren Spezifität gewählt werden, um so die Rate der falsch positiven Zuordnungen zu minimieren.
Nach dieser Regel ist der zu wählende Cut-Off-Wert für den B-PHQ Auswertungsalgorithmus A, validiert anhand der Standarddepressionsdiagnosen des DIA-X für die gesamte Stichprobe: Mindestens zwei Items sind positiv, darunter die Items 1a oder 1b. Die Sensitivität liegt hier bei 0.79 und die Spezifität bei 0.86. Die Summe von Sensitivität und Spezifität beträgt 1.65. An der Schulter der ROC-Kurve liegt der Cut-Off-Wert: Mindestens drei Items sind positiv, darunter die Items 1a oder 1b, da hier der Summenwert höher ist ($\sum$ = 1.66). Die Sensitivität ist hier jedoch mit 0.77 geringer. Die Spezifität beträgt 0.89.
Die explorative Analyse der Untergruppen ergibt für die Subgruppe der Frauen ebenfalls den Cut-Off-Wert: Mindestens zwei Items sind positiv, darunter die Items 1a oder 1b (Sensitivität = 0.77; Spezifität = 0.88; $\sum$ = 1.65).
Für die Subgruppe der Männer ergibt sich der Cut-Off-Wert: Mindestens drei Items sind positiv, darunter die Items 1a oder 1b (Sensitivität = 0.84; Sensitivität = 0.87; $\sum$ = 1.71).
In der Subgruppe Variable Alter < 56,21 ist der geeignetste Cut-Off-Wert: Mindestens zwei Items sind positiv, darunter die Items 1a oder 1b. (Sensitivität = 0.87; Spezifität = 0.79; $\sum$ = 1.66).
In der Subgruppe Alter > 56,20 Jahre ergibt sich der Cut-Off-Wert: Mindestens drei Items sind positiv, darunter die Items 1a oder 1b (Sensitivität = 0.63; Spezifität = 0.93; $\sum$ = 1.56).
Die Auswertung der Subgruppen ergibt auch hier wieder die Tendenz, dass unter Umständen in bestimmten Subgruppen ein anderer Cut-Off-Wert gewählt werden sollte. Die Sensitivität und die Spezifität des für die gesamte Stichprobe empfohlenen Cut-Off-Wertes: Mindestens zwei Items sind positiv, darunter die Items 1a oder 1b sind aber wiederum in allen Subgruppen ausreichend hoch.

Die explorative Ermittlung eines geeigneten Cut-Off-Wertes anhand der erweiterten Version des DIA-X ergibt für die gesamte Stichprobe den Cut-Off-Wert: Mindestens zwei Items sind positiv, darunter die Items 1a oder 1b (Sensitivität = 0.65; Spezifität = 0.90; ∑ = 1.55). Für die vier Subgruppen ergibt sich derselbe Cut-Off-Wert. Dieser Wert liegt in der gesamten Stichprobe und den Subgruppen an der Schulter der ROC-Kurve.

Der Cut-Off-Wert: Mindestens zwei Items sind positiv, darunter die Items 1a oder 1b ist gleich dem niedrigsten Cut-Off-Wert der nach Spitzer et al. (1999) für die Diagnose einer geringgradigen depressiven Störung gefordert wird. Nach Spitzer et al. (1999) wird die Diagnose geringgradige depressive Störung gestellt, wenn zwei bis vier Items von 1a bis 1i positiv beantwortet sind und darunter das Item 1a oder 1b ist.

Sensitivität und Spezifität des B-PHQ Auswertungsalgorithmus B

Die Werte der Sensitivität und Spezifität des Auswertungsalgorithmus B, validiert anhand der Standarddepressionsdiagnosen des DIA-X, ergeben nicht wie beim Auswertungsalgorithmus A ein Überwiegen der Spezifität gegenüber der Sensitivität.

Für den B-PHQ Auswertungsalgorithmus B validiert anhand der Standarddepressionsdiagnosen des DIA-X ergibt sich für die gesamte Stichprobe der Cut-Off-Wert: Mindestens zwei Items sind positiv (Sensitivität = 0.87; Spezifität = 0.74; ∑ = 1.61).

Die explorative Analyse der Subgruppen ergibt, dass in allen vier Untergruppen der an der Schulter der ROC-Kurve gefundene Cut-Off-Wert der Geeignetste ist. Für die Subgruppe der Frauen liegt der Cut-Off-Wert: Mindestens zwei Items sind positiv an der Schulter der ROC-Kurve (Sensitivität = 0.86; Spezifität = 0.75; ∑ = 1.61).

Für die Subgruppe der Männer findet sich der Cut-Off-Wert: Mindestens drei Items sind positiv (Sensitivität = 0.84; Spezifität = 0.81, ∑ = 1.65).

In der Subgruppe Variable Alter < 56,21 Jahre liegt der Cut-Off-Wert: Mindestens drei Items sind positiv an der Schulter der ROC-Kurve (Sensitivität = 0.84; Spezifität = 0.77; ∑ = 1.61).
In der Subgruppe Variable Alter > 56,20 Jahre liegt an der Schulter der ROC-Kurve der Cut-Off-Wert: Mindestens zwei Items sind positiv (Sensitivität = 0.81; Spezifität = 0.80; ∑ = 1.61).
Zusammenfassend kann festgehalten werden, dass der für die gesamte Stichprobe ermittelte Cut-Off-Wert: Mindestens zwei Items sind positiv auch in den einzelnen Subgruppen verwendet werden kann, da die Werte für die Sensitivität und Spezifität akzeptabel hoch sind.

Für die gesamte Stichprobe validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom ergibt sich der Cut-Off-Wert: Mindestens zwei Items sind positiv. Dieser Cut-Off-Wert liegt an der Schulter der ROC-Kurve (Sensitivität = 0.79; Spezifität = 0.81; ∑ = 1.60). Auch in den vier Subgruppen ist dieser Cut-Off-Wert der Geeignetste und liegt an der Schulter der ROC-Kurve.

Der hier gefundene Cut-Off-Wert: Mindestens zwei Items sind positiv entspricht in der Anzahl der Items dem niedrigsten Cut-Off-Wert der nach Spitzer et al. (1999) für die Diagnose einer geringgradigen depressiven Störung gefordert wird. Jedoch ist bei dem hier vorliegenden Algorithmus B die positive Beantwortung der Items 1a oder 1b nicht notwendig.

5.3 Prädiktive Werte

Prädiktive Werte des GHQ-12, B-PHQ Algorithmus A und Algorithmus B

Da die prädiktiven Werte von der Prävalenz abhängig sind, können sie nur auf klinische Situationen mit ähnlich hohen Prävalenzraten übertragen werden. Ein Vergleich prädiktiver Werte für verschieden hohe Prävalenzraten ist nicht aussagekräftig. Daher kann ein Vergleich der prädiktiven Werte zum einen nur für die Validierung der Fragebögen anhand der Standarddepressionsdiagnosen

und zum anderen nur für die Validierung der Fragebögen anhand der Standarddepressionsdiagnosen einschließlich der Diagnose unterschwelliges depressives Syndrom durchgeführt werden.

Der Vergleich der prädiktiven Werte des GHQ-12, des B-PHQ Auswertungsalgorithmus A und des B-PHQ Auswertungsalgorithmus B, validiert anhand der Standarddepressionsdiagnosen des DIA-X, zeigt, dass die negativen prädiktiven Werte für die gewählten Cut-Off-Werte sowohl in der gesamten Stichprobe als auch in den Subgruppen jeweils vergleichbar hoch sind. Die positiven prädiktiven Werte sind für den GHQ-12 verglichen mit den beiden Auswertungsalgorithmen des B-PHQ am niedrigsten. Die positiven prädiktiven Werte des B-PHQ Auswertungsalgorithmus B sind drei bis acht Prozentpunkte größer als die des GHQ-12. Die positiven prädiktiven Werte des B-PHQ Auswertungsalgorithmus A sind im Vergleich zum GHQ-12 sogar 13 bis 22 Prozentpunkte größer. Die positiven prädiktiven Werte des B-PHQ Auswertungsalgorithmus B sind am gewählten Cut-Off-Wert um 10 bis 14 Prozentpunkte niedriger als die des B-PHQ Auswertungsalgorithmus A.

Der Vergleich der prädiktiven Werte des GHQ-12, des B-PHQ Auswertungsalgorithmus A und des B-PHQ Auswertungsalgorithmus B, validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom, zeigt, dass die negativen prädiktiven Werte für die gewählten Cut-Off-Werte in der gesamten Stichprobe und in den Subgruppen des GHQ-12 jeweils vergleichbar hoch sind wie die negativen prädiktiven Werte des B-PHQ Auswertungsalgorithmus B. Die positiven prädiktiven Werte des B-PHQ Auswertungsalgorithmus B sind jedoch um 10 bis 14 Prozentpunkte größer als die des GHQ-12. Die negativen prädiktiven Werte des B-PHQ Auswertungsalgorithmus A sind zwar um ein bis fünf Prozentpunkte kleiner als die des GHQ-12 und um zwei bis fünf Prozentpunkte kleiner als die des B-PHQ Algorithmus B, jedoch sind die positiven prädiktiven Werte um 18 bis 26 Prozentpunkte größer als die des GHQ-12 und um acht bis zwölf Prozentpunkte größer als die des B-PHQ Algorithmus B.

5.4 Overall Accuracy, Kappa Index und Likelihood Quotienten

Overall Accuracy des GHQ-12, des B-PHQ Algorithmus A und des B-PHQ Algorithmus B

Da die Overall Accuracy von der Prävalenz abhängt, erfolgt der Vergleich der Overall Accuracy zum einen für die Validierung der Fragebögen anhand der Standarddepressionsdiagnosen und zum anderen für die Validierung der Fragebögen anhand der Standarddepressionsdiagnosen einschließlich der Diagnose unterschwelliges depressives Syndrom.
In beiden Fällen zeigt sich für die gewählten Cut-Off-Werte, dass der B-PHQ Auswertungsalgorithmus A die höchsten Werte der Overall Accuracy sowohl in der gesamten Stichprobe als auch in den Subgruppen erreicht. Die Overall Accuracy des B-PHQ Auswertungsalgorithmus B ist höher als die des GHQ-12, sowohl in der gesamten Stichprobe als auch in den Subgruppen.

Kappa Index des GHQ-12, des B-PHQ Algorithmus A und des B-PHQ Algorithmus B

Auch der Kappa-Index ist von der Prävalenz abhängig, so dass ein Vergleich des Kappa-Index zum einen für die Validierung der Fragebögen anhand der Standarddepressionsdiagnosen und zum anderen für die Validierung der Fragebögen anhand der Standarddepressionsdiagnosen einschließlich der Diagnose unterschwelliges depressives Syndrom erfolgt.

Für die Validierung der Fragebögen anhand der Standarddepressionsdiagnosen zeigt sich für die gewählten Cut-Off-Werte, dass der B-PHQ Auswertungsalgorithmus A die höchsten Werte des Kappa-Index sowohl in der gesamten Stichprobe als auch in den Subgruppen erreicht. Die Werte des Kappa-Index des B-PHQ Auswertungsalgorithmus B sind höher als die des GHQ-12, sowohl in der gesamten Stichprobe als auch in den Subgruppen.

Für die Validierung der Fragebögen anhand der Standarddepressionsdiagnosen einschließlich der Diagnose unterschwelliges depressives Syndrom zeigt sich für die gewählten Cut-Off-Werte, dass der B-PHQ Auswertungsalgorithmus A und der B-PHQ Auswertungsalgorithmus B vergleichbar hohe Werte des Kappa-Index aufweisen, die über den Werten des GHQ-12 liegen.

Likelihood Quotienten des GHQ-12, des B-PHQ Auswertungsalgorithmus A und B-PHQ Auswertungsalgorithmus B

Likelihood Quotienten des GHQ-12

Nach den Richtwerten von Jaeschke et al. (1994), siehe auch Kapitel 3, S. 48, fallen die Likelihood Quotienten des GHQ-12 am gewählten Cut-Off-Wert in den Bereich der kleineren, aber manchmal wichtigen Veränderungen der Pretestwahrscheinlichkeit. Lediglich der negative Likelihood Quotient der Subgruppe Variable Alter < 56,21 Jahre, validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom, fällt in den Bereich der moderaten Veränderungen. Zusammenfassend kann festgehalten werden, dass die Anwendung des GHQ-12 nach den Richtwerten von Jaeschke et al. (1994) zu kleineren Veränderungen zwischen Pretest- und Posttestwahrscheinlichkeit führt.

Likelihood Quotienten des B-PHQ Auswertungsalgorithmus A

Die positiven Likelihood Quotienten des B-PHQ Auswertungsalgorithmus A fallen für die Validierung anhand der Standarddepressionsdiagnosen am gewählten Cut-Off-Wert für die gesamte Stichprobe, die Subgruppe der Frauen und die Subgruppe Variable Alter > 56,20 Jahre in den Bereich der moderaten Veränderungen. Die positiven Likelihood Quotienten der übrigen Subgruppen liegen im Bereich der kleineren Veränderungen. Bei den negativen Likelihood Quotienten verhält es sich umgekehrt.

Die positiven Likelihood Quotienten fallen für die Validierung anhand der Standarddepressionsdiagnosen einschließlich der Diagnose unterschwelliges de-

pressives Syndrom am gewählten Cut-Off-Wert für die gesamte Stichprobe und die vier Subgruppen in den Bereich der moderaten Veränderungen. Die negativen Likelihood Quotienten liegen im Bereich der kleineren Veränderungen, nur der negative Likelihood Quotient der Subgruppe Variable Alter > 56,20 Jahre fällt in den Bereich der kaum wichtigen Veränderungen.

Likelihood Quotienten des B-PHQ Auswertungsalgorithmus B

Die positiven Likelihood Quotienten fallen für die Validierung anhand der Standarddepressionsdiagnosen am gewählten Cut-Off-Wert für die gesamte Stichprobe und die vier Subgruppen in den Bereich der kleineren Veränderungen. Die negativen Likelihood Quotienten liegen im Bereich der moderaten Veränderungen, lediglich der negative Likelihood Quotient der Subgruppe Variable Alter > 56,20 Jahre fällt in den Bereich der kleineren Veränderungen.
Die positiven und negativen Likelihood Quotienten fallen für die Validierung anhand der Standarddepressionsdiagnosen einschließlich der Diagnose unterschwelliges depressives Syndrom am gewählten Cut-Off-Wert für die gesamte Stichprobe und die vier Subgruppen in den Bereich der kleineren Veränderungen.

5.5 ROC-Kurven

Vergleich der ROC-Kurven des GHQ-12, des B-PHQ Auswertungsalgorithmus A und des B-PHQ Auswertungsalgorithmus B, validiert anhand der Standarddepressionsdiagnosen des DIA-X

Der Verlauf der ROC-Kurven zeigt, dass der GHQ-12, der B-PHQ Auswertungsalgorithmus A und der B-PHQ Auswertungsalgorithmus B gut zwischen kranken und gesunden Personen unterscheiden, da alle ROC-Kurven weit oberhalb der Diagonale verlaufen. Auch unterscheiden sich alle Flächen unter der ROC-Kurve (AUC) statistisch signifikant von 0.5.
Für die Validierung anhand der Standarddepressionsdiagnosen des DIA-X zeigt der Kurvenverlauf des B-PHQ Auswertungsalgorithmus A in der gesamten

Stichprobe und den vier Subgruppen hohe Werte der Spezifität. Die ROC-Kurven des GHQ-12 bleiben in der gesamten Stichprobe und den vier Subgruppen für fast alle Cut-Off-Werte unterhalb den ROC-Kurven des B-PHQ Auswertungsalgorithmus A und des B-PHQ Auswertungsalgorithmus B. Dies spiegelt sich auch in den AUC-Werten wider. Die Fläche unter der ROC-Kurve (AUC) beschreibt die Gesamtgenauigkeit eines Tests. Die AUC-Werte des GHQ-12 bleiben immer geringer als die des B-PHQ Auswertungsalgorithmus B. Im Vergleich zu den AUC-Werten des B-PHQ Auswertungsalgorithmus A bleiben die AUC Werte des GHQ-12 meist geringer. Lediglich in der Subgruppe der Frauen liegt der AUC-Wert des GHQ-12 um 0.002 und in der Subgruppe Alter > 56,20 Jahre um 0.007 über dem AUC-Wert des B-PHQ Auswertungsalgorithmus A. Die AUC-Werte des B-PHQ Auswertungsalgorithmus B sind, mit Ausnahme der Subgruppe der Männer, größer als die des Auswertungsalgorithmus A.

In den explorativen Einzelvergleichen lassen sich jedoch keine statistisch signifikanten Flächenunterschiede nachweisen (α = .05; zweiseitig, nicht adjustiert).

Vergleich der ROC-Kurven des GHQ-12, des B-PHQ Auswertungsalgorithmus A und des B-PHQ Auswertungsalgorithmus B, validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom

Auch für die Validierung anhand der Standarddepressionsdiagnosen einschließlich der Diagnose unterschwelliges depressives Syndrom verlaufen alle ROC-Kurven weit über der Diagonalen. Alle Flächen unter der ROC-Kurve (AUC) unterscheiden sich statistisch signifikant von 0.5.

Der B-PHQ Auswertungsalgorithmus A weist auch hier hohe Werte der Spezifität auf. So verläuft die ROC-Kurve des B-PHQ Auswertungsalgorithmus A in der gesamten Stichprobe und den vier Subgruppen nahe an der y-Achse. Die AUC-Werte des B-PHQ Auswertungsalgorithmus A sind in der gesamten Stichprobe und den vier Subgruppen geringer als die AUC-Werte des GHQ-12 und des B-PHQ Auswertungsalgorithmus B. Die AUC-Werte des GHQ-12 sind in

der gesamten Stichprobe und den Subgruppen niedriger als die des B-PHQ Auswertungsalgorithmus B.

Die explorativen Einzelvergleiche (α = .05; zweiseitig, nicht adjustiert) der Flächen unter den ROC-Kurven ergeben statistisch signifikante Unterschiede zwischen den AUC-Werten des B-PHQ Auswertungsalgorithmus A und des B-PHQ Auswertungsalgorithmus B für die gesamten Stichprobe, die Subgruppe der Frauen, die Subgruppe Variable Alter < 56,21 Jahre und die Subgruppe Variable Alter > 56,20 Jahre.

Die Durchführung der Bonferroni Korrektur ergibt α = .0017. Statistisch signifikant (α = .0017; zweiseitig, adjustiert) sind die Unterschiede zwischen den AUC-Werten des B-PHQ Auswertungsalgorithmus A und des B-PHQ Auswertungsalgorithmus B für die gesamten Stichprobe sowie die Subgruppe der Frauen.

In der Primärversorgung kann daher auch der einfachere Auswertungsalgorithmus B des B-PHQ eingesetzt werden.

5.6 Empfehlungen für die Praxis

Zusammenfassend kann festgehalten werden, dass sowohl der GHQ-12 als auch der B-PHQ valide Screeninginstrumente zur Erfassung depressiver Syndrome (ICD-10) einschließlich der Diagnose unterschwelliges depressives Syndrom in der Primärversorgung sind. Die Auswahl eines bestimmten Instrumentes hängt nicht alleine von der Validität des Instrumentes ab, sondern auch von der Praktikabilität und den Kosten. Beide Fragebögen sind praktikable, kostengünstige und valide Screeninginstrumente. Wird nur die Depressionssektion des B-PHQ durchgeführt, ergibt sich durch die geringere Anzahl von Fragen ein leichter Vorteil in der Praktikabilität des B-PHQ.

Ein einfacher und kostengünstiger Screeningprozess kann in der Primärversorgung wie folgt aussehen: Der GHQ-12 oder der B-PHQ wird routinemäßig von jedem Patienten im Wartezimmer ausgefüllt. Patienten, die ein positives Screeningergebnis aufweisen, sollten vom Arzt weiter diagnostisch untersucht werden. Bei Vorliegen einer depressiven Störung sollten dann körperliche Ursa-

chen der depressiven Störung, das Vorliegen einer normalen Trauerreaktion sowie das Vorhandensein von manischen Episoden in der Vorgeschichte geprüft werden, bevor eine Behandlungsentscheidung gefällt wird.

Sollten zukünftige Bemühungen um eine einheitliche Definition der Diagnose unterschwelliges depressives Syndrom in dieselbe Richtung weisen wie die in der modifizierten Version des DIA-X verwendete Definition, so können ebenfalls beide Screeninginstrumente eingesetzt werden.

Henkel, Braun, Plattner, Maier & Hegerl (2001) konnten zeigen, dass auch der WHO-5 Well-Being-Index ein valides Screeninginstrument zur Erfassung depressiver Störungen in der Primärversorgung ist. Mögliche Vorteile des WHO-5 gegenüber dem GHQ-12 und dem B-PHQ aufgrund seiner Kürze werden im Rahmen des Promotionsvorhabens von Frau Ohlendorf untersucht.

Die Verbesserung der Erkennensrate depressiver Störungen in der Primärversorgung mit Hilfe von Screeningmaßnahmen stellt den ersten und entscheidenden Schritt in Richtung einer bestmöglichen Versorgung depressiver Patienten dar. Eine optimale Diagnostik ist die Basis für einen gezielten Einsatz der vorhanden guten therapeutischen Möglichkeiten.

6 Zusammenfassung

Depressionen gehören weltweit zu den häufigsten psychischen Erkrankungen. Die Mehrheit der Patienten mit einer depressiven Störung befindet sich nicht in fachärztlicher Behandlung, sondern sucht den Hausarzt auf (Coyne et al., 1997; Goldberg, 1999; Linden et al., 1999). Schätzungen zeigen, dass bis zu 30% der Patienten in der Primärversorgung unter depressiven Störungen leiden (Dawson & Tylee, 2001). Damit kommt dem Hausarzt bei der Versorgung depressiver Patienten eine Schlüsselrolle zu (Von Korff et al., 1987). Hier werden die Weichen für die Diagnostik, die Therapie und den weiteren Verlauf der Erkrankung gestellt.

Zahlreiche Studien belegen, dass bei einem großen Teil der Patienten in der Primärversorgung die depressive Störung nicht richtig erkannt oder nicht konsequent behandelt wird (Ballenger et al., 2001; Hegerl, 2000; Rost et al., 2000, Tylee, 2000). So erkennen Allgemeinärzte weltweit über die Hälfte der betroffenen Patienten nicht (Coyne et al., 1995; Coyne et al., 2000; Ormel et al., 1990; Paykel et al., 1997). Die Ursachen für diese Unterdiagnostizierung sind vielfältig. Verschiedene Studien haben gezeigt, dass durch eine routinemäßige Anwendung von Screeninginstrumenten und Rückmeldung der Ergebnisse die Erkennensrate und das Management depressiver Störungen verbessert werden können (Asvall, 2001; Rutz, 1999; Titus et al., 1999). Die vorliegende Arbeit ist die erste Studie in der die Validität von zwei in der Praxis häufig eingesetzten Screeninginstrumenten, dem General Health Questionnaire (GHQ-12, Goldberg, 1972; 1978) und dem Brief Patient Health Questionnaire (B-PHQ, Spitzer, Kroenke & Williams, 1999) in Relation zu einem standardisierten klinisch-diagnostischen Interview, dem DIA-X (Wittchen & Pfister, 1997a), anhand einer einzigen Stichprobe in der Primärversorgung erhoben und verglichen wird.

Im Kontext des Screenings meint Validität, wie gut ein Instrument zwischen Personen unterscheiden kann, die das gesuchte Merkmal besitzen und solchen Personen, die das gesuchte Merkmal nicht besitzen. Zur Beurteilung der Validität stehen eine Reihe von Validierungsmaßen zur Verfügung. Um die Vergleichbarkeit der Untersuchungsergebnisse mit anderen Studien zu ermöglichen, wurden neben den am häufigsten verwendeten Validitätsmaßen Sensiti-

vität, Spezifität, ROC-Kurven und Flächen unter der ROC-Kurve (AUC) auch prädiktive Werte, Overall Accuracy, Kappa-Index und Likelihood Quotienten berechnet.

Der General Health Questionnaire wurde 1972 von Goldberg zur Erfassung psychischer Beeinträchtigung in der Allgemeinbevölkerung entwickelt und ist seitdem in zahlreichen Studien eingesetzt worden. Die Auswertung des GHQ-12 erfolgt nach der Methode des GHQ Scorings (Goldberg et al., 1997).

Mit dem B-PHQ liegt eine reine Selbstbeurteilungsversion des Prime-MD (Primary Care Evaluation of Mental Disorders; Spitzer et al., 1994) zur Diagnostik von affektiven Störungen und Panikstörung in der Primärversorgung anhand von DSM-IV Kriterien vor. Der B-PHQ wurde auf zwei verschiedene Arten ausgewertet. Ein Auswertungsalgorithmus folgt dem Auswertungsschema von Spitzer et al. (1999). Zusätzlich wurde im Rahmen der vorliegenden Arbeit ein neuer, einfacherer Auswertungsalgorithmus eingeführt, der keine hierarchische Struktur der depressiven Symptome beinhaltet.

Das Validierungsinstrument, das DIA-X, ist eine rechnergestützte Version des Composite International Diagnostic Interviews (CIDI, WHO, 1993). In Zusammenarbeit mit den Entwicklern des DIA-X wurde das Interview dahingehend modifiziert, dass auch die Diagnose unterschwelliges depressives Syndrom gestellt werden kann. In explorativen Analysen wurde die Validität der beiden Screeninginstrumente zum einen unter Berücksichtigung der Diagnose unterschwelliges depressives Syndrom und zum anderen für verschiedene Subgruppen, getrennt nach den Variablen Alter und Geschlecht, erhoben und verglichen.

Die Datenerhebung wurde in Kooperation mit 18 hausärztlich tätigen Praxen, Allgemeinärzte und Internisten, in Nürnberg durchgeführt. Das Ausfüllen der Screeningbögen fand in den Arztpraxen statt. Der Goldstandard, das DIA-X, wurde von geschulten Interviewerinnen innerhalb der folgenden sechs Tage telefonisch durchgeführt.

In der vorliegenden Stichprobe (N = 448) sind 283 (63%) Patienten weiblich und 165 (37%) männlich. Das Alter der Patienten liegt zwischen 18 und 88 Jahren. Der Mittelwert ist 53 Jahren, die Standardabweichung beträgt 17 und der Median liegt bei 56 Jahren.

Die Prävalenz depressiver Störungen (ICD-10) beträgt 18,3%, für die Diagnose unterschwelliges depressives Syndrom 12,1%. Es handelt sich hierbei um aktuelle Querschnittsdiagnosen, die sich auf den Zeitraum der letzten beiden Wochen vor der Durchführung des Interviews beziehen.

Die Datenauswertung zeigt, dass sowohl der GHQ-12 als auch der B-PHQ valide Instrumente zur Erfassung depressiver Syndrome (ICD-10) einschließlich unterschwelliger depressiver Syndrome in der Primärversorgung sind. Die Vergleiche der AUC-Werte ergeben für die gesamte Stichprobe aber auch für die Subgruppen keine statistisch signifikanten Unterschiede in der Gesamtgenauigkeit der beiden Screeninginstrumente.

Die Flächenvergleiche der explorativen Validierung unter Berücksichtigung der Diagnose unterschwelliges depressives Syndrom deuten auf statistisch signifikante Unterschiede zwischen den beiden Auswertungsalgorithmen des B-PHQ zugunsten des neu entwickelten einfacheren Auswertungsalgorithmus hin, so dass dieser in der Primärversorgung eingesetzt werden kann.

Die Auswahl eines Screeninginstrumentes hängt jedoch nicht alleine von der Validität des Instrumentes ab, sondern auch von den Kosten und der Praktikabilität. Beide Fragebögen sind kostengünstige, praktikable und valide Screeninginstrumente zur Erfassung depressiver Störungen in der Primärversorgung. Wird nur die Depressionssektion des B-PHQ durchgeführt, ergibt sich durch die geringere Anzahl von Fragen ein leichter Vorteil in der Praktikabilität des B-PHQ.

Zur Verringerung des diagnostischen Defizits kann in der Primärversorgung somit ein einfacher und kostengünstiger Screeningprozess durchgeführt werden, indem der GHQ-12 oder der B-PHQ von jedem Patienten im Wartezimmer ausgefüllt wird. Bei einem positiven Screeningergebnis liegt möglicherweise eine depressive Störung vor, so dass der Arzt diese Patienten diagnostisch weiter untersuchen sollte.

Die Verbesserung der Erkennensrate depressiver Störungen in der Primärversorgung mit Hilfe von Screeningmaßnahmen ist der erste und entscheidende Schritt in Richtung einer bestmöglichen Versorgung depressiver Patienten, denn eine optimale Diagnostik ist die Basis für einen gezielten Einsatz der vorhanden guten therapeutischen Möglichkeiten.

Literaturverzeichnis

Abramson, L.Y., Seligman, M.E. & Teasdale, J.D. (1978). Learned helplessness in humans: critique and reformulation. *Journal of Abnormal Psychology, 87 (1),* 49-74.

Ahrens, B. (2000). Wesentliches zur Erkrankung und Diagnostik affektiver Erkrankungen. In H.-J. Möller, (Hrsg.). *Therapie psychiatrischer Erkrankungen* (S. 300-311). Stuttgart: Thieme Verlag.

Amelang, M. & Zielinski, W. (1994). *Psychologische Diagnostik und Intervention.* Berlin, Heidelberg, New York: Springer.

American Psychiatric Association (1987). *Diagnostic and Statistical Manual of Mental Disorders, Third Edition, Reviced (DSM-III-R).* Washington, D.C.: American Psychiatric Association.

American Psychiatric Association (1994). *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV).* Washington, D.C.: American Psychiatric Association.

Ames, D. (2001). Depression and the elderly. In A. Dawson & A. Tylee (Eds.). *Depression: Social and economic timebomb* (pp. 49-54). London: BMJ Books.

Andrews, G. & Peters, L. (1998). The psychometric properties of the Composite International Diagnostic Interview. *Soc Psychiatry Psychiatr Epidemiol, 33 (2),* 80-88.

Angst, J. (1992). Epidemiology of depression. *Psychopharmacology, 106,* 71-74.

Angst, J., Merikangas, K.R. & Preisig, M. (1997). Subthreshold syndroms of depression and anxiety in the community. *J Clin Psychiatry, 58 (Suppl. 8),* 6-10.

Araya, R., Wynn, R. & Lewis, G. (1992). Comparison of two self administered questionnaires (GHQ-12 and SRQ-20) in primary care in Chile. *Soc Psychiatry Psychiatr Epidemiol, 27(4),* 168-173.

Asvall, J.E. (2001). Can we turn the table on depression? In A. Dawson & A. Tylee (Eds.). *Depression: Social and economic timebomb* (pp. 3-8). London: BMJ Books.

Baldessarini, R.J., Finklestein, S. & Arana, G.W. (1983). The predictive power of dianostic tests and the effect of prevalence of illness. *Arch Gen Psychiatry, 40 (5),* 569-573.

Ballenger, J.C. (2000). Structural issues and policy in the management of depressive illness. Overview and expert panel commentary. *European Neuropsychopharmacology, 10 (Suppl. 4),* 449-453.

Ballenger, J.C., Davidson, J.R.T., Lecrubier,Y., Nutt, D.J., Kirmayer, J., Lépine, J.-P., Lin, K.-M., Tajima, O. & Ono, Y. (2001). Consensus statement on transcultural issues in depression and anxiety from the International Consensus Group on Depression and Anxiety*. J Clin Psychiatry, 62 (Suppl. 13),* 47-55.

Banks, M.H., Clegg, G.W., Jackson, P.R., Kemp, N.J., Stafford, E.M. & Wall, T.D. (1980). The use of the General Health Questionnaire as an indicator of health in occupational studies. *Journal of occupational psychology, 53 (3),* 187-194.

Beck, A.T. (1974) The developement of depression. A cognitive model. In R.J. Friedman & M.M. Katz (Eds.). *The psychology of depression* (pp. 5-20*).* New York, Toronto, London, Sydney: Wiley & Sons.

Beck, D.A. & Koenig, H.G. (1996). Minor depression: a review of the literature. *INT'L. J. Psychiatry in Medicine, 26 (2),* 177-209.

Biggerstaff, B.J. (2000). Comparing diagnostic tests: a simple graphic using likelihood ratios. *Statistics in Medicine, 19 (5),* 649-663.

Botega, N., Mann, A., Blizard, R. & Wilkinson, G. (1992). General practitioners and depression – first use of the Depression Attitude Questionnaire. *International Journal of Methods in Psychiatric Research, 2,* 169-180.

Boyd, J.H. & Weissman, M.M. (1981). Epidemiology of affective disorders. A reexamination and further directions. *Archives of General Psychiatry, 38 (9),* 1039-1046.

Brickenkamp, R. (1997). *Handbuch psychologischer und pädagogischer Tests.* Göttingen, Bern, Toronto, Seattle: Hogrefe.

Broadhead, W.E., Blazer, D., George, L. & Tse, C.K. (1990). Depression, disability days, and days lost from work in a prospective epidemiologic survey. *JAMA, 264 (19),* 2524-2528.

Brodman, K., Erdman, A.J., Lorge, I., Wolff, G. & Broadbent, T.H. (1949). The Cornell Medical Index: an adjunction to the medical interview. *Journal of the American Medical Association, 140,* 530-545. [zitiert nach Goldberg & Williams, 1991].

Callahan, C.M., Hendrie, H.C., Dittus, R.S., Brater, D.C., Hui, S.L. & Tierney, W.M. (1994a). Improving treatment of late life depression in primary care: a randomized clinical trial. *J Am Geriatr Soc, 42 (8),* 839-846.

Callahan, C.M., Hendrie, H.C., Dittus, R.S., Brater, D.C., Hui, S.L. & Tierney, W.M. (1994b). Depression in late life: the use of clinical characteristics to focus screening efforts. *Journal of Gerontology: MEDICAL SCIENCES, 49 (1),* M9-M14.

Caraveo-Anduaga, J.J., Martinez, N.A., Saldivar, G., Lopez, J.L. & Saltijeral, T. (1998). Performance of the GHQ-12 in relation to current and lifetime CIDI psychiatric diagnoses. *Salud Mental, 21(4),* 1-12.

Chisholm, D. (2001). The economic consequences of depression. In A. Dawson & A. Tylee (Eds.). *Depression: Social and economic timebomb* (pp. 121-130). London: BMJ Books.

Choi, B.C.K. (1998). Slopes of a receiver operating characteristic curve and likelihood ratios for a diagnostic test. *American Journal of Epidemiology, 148 (11),* 1127-1132.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, XX (1),* 37-46.

Cohen, J. (1968). Weighted Kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin, 70(4),* 213-220.

Coyne, J., Fechner-Bates, S. & Schwenk, T.L. (1994). Prevalence, nature, and comorbidity of depressive disorders in primary care. *General Hospital Psychiatry, 16 (4),* 267-276.

Coyne, J., Klinkman, M., Gallo, S. & Schwenk, T. (1997). Short-term outcomes of detected and undetected depressed primary care patients and depressed psychiatric patients. *General Hospital Psychiatry, 19 (5),* 333-343.

Coyne, J., Schwenk, M.D. & Fechner-Bates, S. (1995). Nondetection of depression in primary care physicians reconsidered. *General Hospital Psychiatry, 17 (1),* 3-12.

Coyne, J., Thompson, R., Palmer, S., Kagee, A. & Maunsell, E. (2000). Should we screen for depression? Caveats and potential pitfalls. *Applied and Preventive Psychology, 9 (2),* 101-121.

Dawson A. & Tylee A. (2001). Depression and general practice. In A. Dawson & A. Tylee (Eds.). *Depression: Social and economic timebomb* (pp. 75-76). London: BMJ Books.

Dawson A. & Tylee A. (Eds.). (2001). *Depression: Social and economic timebomb.* London: BMJ Books.

DeLong, E.R., DeLong, D.M. & Clarke-Pearson, D.L. (1988). Comparing the areas under two or more correlated receiver operating chracteristica curves: a nonparametric approach. *Biometrics 44,* 837-845.

Dilling, H., Mombour,W. & Schmidt (Hrsg.). (1991). *Internationale Klassifikation psychischer Störungen: ICD-10, Kapitel V (F), Klinisch-Diagnostische Leitlinien, Weltgesundheitsorganisation.* 1. Aufl. Bern, Göttingen, Toronto: Huber.

Dorsch, F. (Hrsg.) (1994). *Psychologisches Wörterbuch*. Bern, Göttingen, Tornto, Seattle: Huber.

Dunn, G. (2000). *Statistics in psychiatry.* New York: Arnold.

Erdreich, L.S. & Lee, E.T. (1981). Use of relative operating characteristic analysis in epidemiology. *American Journal of Epidemiology, 114(5),* 649-662.

Essau, C.A. & Wittchen, H.-U. (1993). An overview of the Composite International Diagnostic Interview (CIDI). *International Journal of Methods in Psychiatric Research, 3 (2),* 79-85.

Fagan, T. J. (1975). Nomogram for Bayes's theorem. *N. Engl. J. Med. (letter), 293 (5),* 257.

Feldman, E.L., Robbins, A. & Jaffe, A. (1998). Selections from current literature: Minor depression. *Family Practice, 15 (3),* 275-281.

Fielding, A.H. & Bell, J.F. (1997). A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation, 24(1),* 38-49.

Fletcher, R.H., Fletcher, S.W. & Wagner, E.H. (1999). *Klinische Epidemiologie: Grundlagen und Anwendung.* Wiesbaden: Ullstein Medical.

Friedman R.J. & Katz M.M. (Eds.). (1974). *The psychology of depression.* New York, Toronto, London, Sydney: Wiley & Sons.

Furukawa, T. & Goldberg, D.P. (1999). Cultural invariance of likelihood ratios for the General Health Questionnaire. *Lancet, 353,* 561-562.

Gastpar, M. (1986). Unterschiedliche Pharmakoeffekte bei Angst und Depression. In H. Helmchen & M. Linden (Hrsg.). *Die Differenzirung von Angst und Depression* (S. 167-176). Berlin, Heidelberg, New York: Springer-Verlag.

Goldberg, D. (1972). *The detection of psychiatric illness by questionnaire. A technique for identification and assessment of non-psychotic psychiatric illness.* Maudsley Monographs 21. Oxford: University Press.

Goldberg, D. (1978). *Manual of the General Health Questionnaire.* Windsor: NFER.

Goldberg, D. (1999). The management of anxious depression in primary care. *J Clin Psychiatry, 60 (Suppl. 7),* 39-42.

Goldberg, D., Gater, R., Sartorius, N., Ustun, T.B., Piccinelli, M., Gureje, O. & Rutter, C. (1997). The validity of two versions of the GHQ in the WHO study of mental illness in general health care. *Psychological Medicine, 27 (1),* 191-197.

Goldberg, D. & Hiller, V.F. (1979). A scaled version of the General Health Questionnaire. *Psychological Medicine, 9,* 139-145.

Goldberg, D., Oldehinkel, T., & Ormel, J. (1998). Why the GHQ threshold varies from one place to another. *Psychological Medicine, 28 (4),* 915-921.

Goldberg, D. & Williams, P. (1991). *A User's Guide to the General Health Questionnaire*: Windsor: NFER-Nelson.

Goodchild, M.E. & Duncan-Jones, P. (1985). Chronicity and the General Health Questionnaire. *British Journal of Psychiatry, 146,* 55-61.

Gottfries, C.G. (1997). Recognition and management of depression in the elderly. *International Clinical Psychopharmacology, 12 (Suppl. 7),* 31-36.

Grant, J.M. (1991). The fetal heart rate trace is normal, isn't it? Observer-agreement of categorial assessments. *Lancet, 337,* 215-218.

Gray, D.P. (2001) Depression in General Practice. In A. Dawson & A. Tylee (Eds.). *Depression: Social and economic timebomb* (pp. 77-84). London: BMJ Books.

Grove, W.M., Andreasen, N.C., McDonald-Scott, P., Keller, M.B. & Shapiro, R.W. (1981). Reliability studies of psychiatric diagnosis. *Arch Gen Psychiatry, 38 (4),* 408-413.

Gureje, O. (1991). Reliability and the factor structure of the Yoruba version of the 12-item General Health Questionnaire. *Acta Psychiatr Scand, 84 (2),* 125-129.

Gureje, O. & Obikoya, B. (1990). The GHQ-12 as a screening tool in primary care setting. *Soc Psychiatry Psychiatr Epidemiol, 25 (5),* 276-280.

Gureje, O., Üstün, T.B. & Simon, G.E. (1997). The syndrome of Hypochondrias: a cross national study in primary care. *Psychological Medicine, 27,* 1001-1010.

Gurin, G., Veroff, J. & Feld, S. (1969). *Americans view their mental health.* New York: Basic Books [zitiert nach Goldberg & Williams, 1991].

Hamilton, M. (1960). A rating scale für depression. *Journal of Neurology Neurosurgery and Psychiatry, 23,* 56-61.

Hance, M., Carney, R.M., Freedland, K.E., Skala, J. (1996). Depression in patients with coronary heart disease. *General Hospital Psychiatry, 18,* 6165.

Hanley, J.A. (1987). Standard error of the kappa-statistic. *Psychological Bulletin, 102 (2),* 315-321.

Hardy, G.E., Shapiro, D.A., Haynes, C.E. & Rick, J.E. (1999). Validation of the General Health Questionnaire 12, using a sample of employees from England's health care services. *Psychological Assesment, 11 (2),* 159-165.

Hautzinger, M. (1995). Psychotherapie und Pharmokotherapie bei Depressionen. *Psychotherapeut, 40,* 373-380.

Hautzinger, M. & de Jong-Meyer, R. (1998). Depressionen. In H. Reinecker (Hrsg.), *Lehrbuch der Klinischen Psychologie. Modelle psychischer Störungen* (S. 207-248). Göttingen: Hogrefe Verlag für Psychologie.

Hegerl, U. (2000). Die Studie Depression 2000 aus der Sicht des Kompetenznetzes*. Fortschritte in der Medizin, Sonderheft I/2000,* 40-41*.*

Hegerl, U. (2001). Antidepressiva. In U. Hegerl, M. Zaudig & H.-J. Möller (Hrsg.). (2001). *Depression und Demenz im Alter. Abgrenzung, Wechselwirkung, Diagnose und Therapie* (S.79-100)*.* Wien, New York: Springer-Verlag.

Hegerl, U. & Möller, H.-J. (2000). Pharmakotherapie der Altersdepression. *Nervenarzt, 71,* 1-8.

Hegerl, U., Zaudig, M. & Möller, H.-J. (Hrsg.). (2001). *Depression und Demenz im Alter. Abgrenzung, Wechselwirkung, Diagnose und Therapie.* Wien, New York: Springer-Verlag

Helmchen, H. & Linden, M. (Hrsg.) (1986). *Die Differenzirung von Angst und Depression.* Berlin, Heidelberg, New York: Springer-Verlag.

Helmchen, H. & Linden, M. (2000). Subthreshold disorders in psychiatry: Clinical reality, methodological artefact, and the double-treshold problem. *Comprehensive Psychiatry, 41 (Suppl. 2),* 1-7.

Henkel, V., Braun, S., Pattner, A., Maier, W. & Hegerl, U. (2001, July). *Screening for depression in primary care.* Poster presented at the 7th World Congress of Biological Psychiatry, Berlin, Germany.

Henkel, V. & Hegerl, U. (2000). Altersdepression – eine unterschätzte Gefahr. *Geburtshilfe und Frauenheilkunde, 60:* M 185-M 188.

Hirschfeld, R.M., Keller, M., Panico, S., Arons, B., Barlow, D., Davidoff, F., Endicott, J., Froom, J., Goldstein, M., Gorman, J., Guthrie, D., Marek, R., Maurer, T., Meyer, R., Phillips, K., Ross, J., Schwenk, T., Sharfstein, S., Thase, M. & Wyatt, R. (1997). The national depressive and manic-depressive association consensus on the undertreatment of depression. *JAMA, 277, 4,* 330-340.

Hobbs, P., Ballinger, C.B., Greenwood, C., Martin, B. & McClure, A. (1983). Factor analysis and validation of the General Health Questionnaire in men: a general practice survey. *Brit. J. Psychiat., 144,* 270-275.

Hobbs, P., Ballinger, C.B., Smith, A.H.W. (1983). Factor analysis and validation of the General Health Questionnaire in women: a general practice survey. *Brit. J. Psychiat., 142,* 257-264.

Hobi, V., Gerhard, U. & Gutzwiller, F. (1989). Mitteilungen über die Erfahrungen mit dem GHQ (General Health Questionnaires) von D. G. Goldberg. *Schweiz. Rundschau Med (Praxis), 78 (9),* 219-225.

Holsboer-Trachsler, E. & Vanoni, C. (2000). *Depression und Schlafstörung in der Allgemeinarztpraxis.* Binningen: Medical Congress GmbH.

Horwath, E., Johnson, J., Klerman, G. & Weissman, M. (1992). Depressive symptoms as relative and attributable risk factors for first-onset major depression. *Arch Gen Psychiatry, 49,* 817-823.

Hotopf, M., Sharp, D. & Lewis, G. (1998). What's in a name? A Comparison of four psychiatric assessments. *Soc Psychiatry Psychiatr Epidemiol, 33,* 27-31.

Jaeschke, R., Guyatt, G.H., & Sackett, D.L. (1994). Users guides to the medical literature. How to use an article about a diagnostic test. What are the results and will they help me in caring for my patients? *JAMA, 271 (9),* 703-707.

Jaffe, A., Froom, J. & Galambos, N. (1994). Minor depression and functional impairment. *Arch Fam Med, 3,* 1081-1086.

Jäger, S. & Petermann, F. (Hrsg.) (1992). *Psychologische Diagnostik: ein Lehrbuch*. Weinheim: Psychologie-Verl.-Union.

Johnson, J.G., Spitzer, R.L., Williams, J.B., Kroenke, K., Linzer, M., Brody, D., deGruy, F. & Hahn, S. (1995). Psychiatric comorbidity, health status, and functional impairment associated with alcohol abuse and dependence in primary care patients: findings of the Prime MD-1000 study. *Journal of consulting and Clinical Psychology, 63(1),* 133-140.

Johnson, J., Weissman, M. & Klerman, G.L. (1992). Service utilization and social morbidity associated with depressive symptoms in the community. *JAMA, 267 (11),* 1478-1483.

Judd, L.L., Akiskal, H. & Paulus M.P. (1997). The role and clinical significance of subsyndromal depressive symptoms in unipolar major depressive disorder. *Journal of Affective Disorders, 45,* 5-18.

Judd, L.L., Rapaport, M.H., Paulus, M.P. & Brown, J.L. (1994). Subsyndromal symptomatic depression: a new mood disorder? *J Clin Psychiatry, 55 (Suppl. 4),* 18-28.

Kasper, S., Möller, H.-J. & Müller-Spahn, F. (1997). *Depression. Diagnose und Pharmakotherapie.* Stuttgart: Thieme Verlag.

Katon, M.D., Von Korff, M., Lin, E., Lipscomb, P., Russo, J., Wagner, E. & Polk, E. (1990). Distressed high utilizers of medical care. DSM-III-R diagnoses and treatment needs. *General Hospital Psychiatry, 12 (6), 355-362.*

Katona, C., Livingston, G., Manela, M., Leek, C., Mullan, E., Orrell, M., D'Ath, P.P. & Zeitlin, D. (1997). The symptomatology of depression in the elderly. *International Clinical Psychopharmacology, 12 (Suppl. 7).* 19-23.

Kessler, R., McGonagle, K., Zhao, S., Nelson, C., Hughes, M., Eshleman, S., Wittchen, H.-U. & Kendler, K., (1994). Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States. Results from the National Comorbidity Survey. *Arch Gen Psych., 51,* 8-19.

Kessler, R., Zhao, S., Blazer, D.G. & Swartz, M. (1997). Prevalence, correlates, and course of Minor Depression and Major Depression in the National Comorbidity Survey. *Journal of Affective Disorders, 45,* 19-30.

Kihç, C., Rezaki, M., Rezaki, B., Kaplan, I., Özgen, G., Sagaduyu, A. & Öztürk, M.O. (1997). General Health Questionnaire (GHQ 12 & GHQ 28): psychometric properties and factor structure of the Scales in a Turkish primary care sample. *Soc Psychiatry Psychiatr Epidemiol, 32,* 372-331.

Kleinman A. & Cohnen A. (2001). A global view of depression from an anthropological perspective. In A. Dawson & A. Tylee (Eds.). *Depression: Social and economic timebomb* (pp. 11-16). London: BMJ Books.

Kobak, K., Taylor, L., Dottl, S., Greist, J., Jefferson, J., Burroughs, D., Katzelnick, D. & Mandell, M. (1997). Computerized screening for psychiatric disorders in an outpatient community mental health clinic. *Psychiatric Services, 48, 8,* 1048-1057.

Kraemer, H.C. (1988). Assessment of 2 x 2 associations: generalisation of the signal-detection methodology. *The American Statistican, 42 (1),* 37-49.

Kramer, M. & Feinstein, A.R. (1981). Clinical biostatistics LIV. The biostatistics of concordance. *Clin. Pharmacol. Ther., 29 (1),* 111-123.

Kruse, J., Heckrath, C., Schmitz, N., Alberti, L. & Tress, W. (1999). Zur hausärztlichen Diagnose und Versorgung psychogen Kranker. *Psychother. Psychosom. med. Psychol., 49,* 14-22.

Le Fevre, P., Devereux, J., Smith, S., Lawrie, S. & Cornbleet, M. (1999). Screening for psychiatric illness in the palliative care inpatient setting: a comparison between the Hospital Anxiety Depression Scale and the General Health Questionnaire-12. *Palliative Medicine, 13,* 399-407.

Lépine, J.-P., Gastpar, M., Mendlewicz, J. & Tylee, A. (1997). Depression in the community: the first pan-european study DEPRES (Depression Research in European Society). *International Clinical Psycho-pharmacology, 12,* 19-29.

Lewinsohn, P.M. (1974). A behavioral approach to depression. In R.J. Friedman & M.M. Katz (Eds.). *The psychology of depression* (pp. 57-178). New York, Toronto, London, Sydney: Wiley & Sons.

Lienert, G.A. (1994). *Testaufbau und Testanalyse.* Weinheim: Beltz, Psychologie- Verl.-Union.

Linden, M., Lecrubier, Y., Bellantuono, C., Benkert, O., Kisley, S. & Simon, G. (1999). The prescribing of psychotropic drugs by primary care physicans: an international collaborative study. *Journal of Clinical Psycho-pharmacology, 19 (2),* 132-140.

Linden, M., Maier, W., Achberger, M., Herr, R., Helmchen, H. & Benkert, O. (1996). Psychische Erkrankungen und ihre Behandlung in Allgemein-arztpraxen in Deutschland. *Nervenarzt, 67,* 205-215.

Loerch, B., Szegedi, A., Kohnen, R. & Benkert, O. (2000.). The Primary Care Evaluation of Mental Disorders (Prime-MD), German version: a comparison with the CIDI. J*ournal of Psychiatric Research, 34,* 211-220.

Lyness, J.M., King, D.A., Cox, C., Yoediono, Z. & Caine, E.D. (1999). The importance of subsyndromal depression in older primary care patients: prevalence and associated functional disability. *JAGS, 47,* 647-652.

Magruder, K.M. & Calderone, G.E. (2000). Public health consequences of different tresholds for the diagnosis of mental disorders. *Comprehensive Psychiatry, 41 (Suppl. 1),* 14-18.

Maier, W., Gänsicke, M. & Weiffenbach, O. (1997). The releationship between major and subtreshold variants of unipolar depression. *Journal of Affective Disorders, 45,* 41-51.

Maier, W., Linden, M. & Sartorius, N. (1996). Psychische Erkrankungen in der Allgemeinarztpraxis. Ergebnisse und Schlußfolgerungen einer WHO-Studie. *Deutsches Ärzteblatt, 93 (18),* B-947-B-950.

Mari, J.J. & Williams, P. (1985). A comparison of the validity of two psychiatric screening questionnaires (GHQ-12 and SRQ-20) in Brazil, using relative operating chracteristic (ROC) analysis. *Psychological Medicine, 15,* 651-659.

Maruish M.-E. (Ed.). (2000). *Handbook of psychological Assessment in Primary Care Settings.* Mahwah, New Jersey: Lawrence Erlbaum Associates.

Maxwell, A.E. (1977). Coefficients of agreement between observers and their interpretation. *Brit. J. Psychiat., 130,* 79-83.

Metz, D.E. (1978). Basic principles of ROC analysis. *Seminars in Nuclear Medicine, VIII (4),* 283-298.

Miranda, J. & Munoz, R. (1994). Intervention for minor depression in primary care patients. *Psychosomatic Medicine, 56,* 136-142.

Möller, H.-J. (Hrsg.). (2000). *Therapie psychiatrischer Erkrankungen*. Stuttgart: Thieme Verlag.

Möller, H.-J., Müller, W.E. & Volz, H.-P. (2000). *Psychopharmakotherapie. Ein Leitfaden für Klinik und Praxis.* Stuttgart: Kohlhammer.

Montano, C.B. (1994). Recognition and treatment of depression in a primary care setting. *J Clin Psychiatry, 55 (Suppl. 12),* 18-34.

Murphy, J.M., Berwick, D.M., Weinstein, M.C., Borus, J.F., Budman, S.H. & Klerman, G.L. (1987). Performance of screening and diagnostic tests. Application of receiver operating characteristic analysis. *Arch Gen Psychiatry, 44,* 550-555.

Murray, C.J.L. & Lopez, A.D. (1997). Global mortality, disability and the contribution of risk factors: Global burden of disease study. *Lancet, 349*,1436-1352.

Newman, S.C., Bland, R.C. & Orn, H. (1988). A comparison of methods of scoring the General Health Questionnaire. *Comprehensive Psychiarty, 29 (4),* 402-408.

NFER Nelson (http//www.nfer-nelson.co.uk/ghq/).

Noack, H. & Petermann, F. (1992). Entscheidungstheorie. In S. Jäger und F. Petermann (Hrsg.). *Psychologische Diagnostik: ein Lehrbuch* (S. 295-310). Weinheim: Psychologische Verlagsunion.

Ohlendorf, P. (unveröffentlichte Dissertation). Der WHO-5 Wellbeing Index. Ludwig-Maximilians-Universität München.

Ormel, J., Van Den Brink, W., Koeter, M.W.J.,Giel, R., Van Der Meer, K., Van De Willige, G. & Wilmink, F.W. (1990). Recognition, management and outcome of psychological disorders in primary care: a naturalistic follow-up study. *Psychological Medicine, 20,* 909-923.

Oxman, T.E., Barrett, J.E. , Barrett, J. & Gerber, P. (1990). Symptomatology of late-life minor depression among primary care patients. *Psychosomatics, 31 (2),* 174-180.

Palsson, S. & Skoog, I. (1997). The epidemioloy of affective disorders in the elderly: a review. *International Clinical Psychopharmacology, 12 (Suppl. 7),* 3-13.

Papassotiropoulos, A. & Heun, R. (1999). Detection of subtreshold depression and subtreshold anxiety in the elderly. *International Journal of Geriatric Psychiatry, 14,* 643-650.

Papassotiropoulos, A., Heun, R. & Maier, W.E. (1997). Age and cognitive impairement influence the performance of the General Health Questionnaire. *Comprehensive Psychiatry, 38 (6),* 335-340.

Paykel, E.S., Tylee, F.R.C., Wright, A., Priest, R.G., Rix, S. & Hart, D. (1997). The Defeat Depression Campaign: psychiatry in the public arena. *Am J Psychiatry, 154 (Suppl. 6),* 59-65.

Peirce, J.C. & Cornell, R.G. (1993). Integrating stratum-specific likelihood ratios with the analysis of ROC curves. *Med Decis Making, 13,* 141-151.

Philipp, M., Delmo, C.D., Buller, R., Schwarze, H., Winter, P., Maier, W. & Benkert, O. (1992). Differentiation between major and minor depression. *Psychopharmacology, 106,* 75-78.

Pincus, H.A., Davis, E.W. & McQueen, L.E. (1999). Subthreshold mental disorders. A review and synthesis of studies on minor depression and other brand names. *Br-J-Psychiatry, 174,* 288-296.

Pini, S., Perkonnig, A., Tansella, M. & Wittchen, H.-U. (1999). Prevalence and 12-month outcome of threshold and subthreshold mental disorders in primary care. *Journal of Affective Disorders, 56,* 37-48.

Politi, P.L., Piccinelli, M. & Wilkinson, G. (1994). Reliability, validity and factor structur of the 12-Item General Health Questionnaire among young males in Italy. *Acta Psychiatr Scand, 90,* 432-437.

Power, M.J. (1988). The „Worst Ever" Version of the General Health Questionnaire (GHQ). *Journal of Clinical Psychology, 44 (2),* 215-216.

Rapaport M.H. & Judd L.L. (1998). Minor depressive disorder and subsyndromal depressive symptoms: functional impairment and response to treatment. *Journal of Affective Disorders, 48,* 227-232.

Reinecker, H. (Hrsg.). (1998). *Lehrbuch der Klinischen Psychologie. Modelle psychischer Störungen.* Göttingen, Bern, Toronto, Seattle: Hogrefe.

Rieger, D.A., Narrow, W., Rae, D., Manderscheid, R., Locke, B. & Goodwin, F. (1993). The de facto US mental and addictive disoders service system. Epidemiologic catchment area prospective 1-year prevalence rates of disorders and services. *Arch Gen Psychiatry, 50,* 85-94.

Ritchie, K. & Fuhrer, R. (1992). A comparative study of the performance of screening tests for senile dementia using receiver operating characteristcs analysis. *J Clin Epidemiol, 45 (6),* 627-637.

Robins, L.N., Helzer, J.E., Croughan, J.L., Williams, J.B.W. & Spitzer, R.L. (1981). *The NIHM Diagnostic Interview Schedule, Version III.* Washington: Public Health Services (HSS) ADM-T-42-3 (5-81, 8-81).

Robins, L.N. & Sartorius N. (1993). Editorial, *International Journal of Methods in Psychiatric Research, 3,* 63-65.

Robins, L.N., Wing, J., Wittchen, H.-U., Helzer, J.E., Babor, T.F., Burke, J., Farmer, A., Jablenski, A., Pickens, R., Regier, D.A., Sartorius, N. & Towle, L.H. (1988). The Composite International Diagnostic Interview. An epidemiologic instrument suitable for use in conjunction with different diagnostic systems and in different cultures. *Arch Gen Psychiatry, 45,* 1069-1077.

Rost, K. Zhang, M., Fortney, J., Smith, J., Coyne, J. & Smith, R. (1998). Persistently poor outcomes of undetected major depression in primary care. *General Hospital Psychiatry, 20,* 12-20.

Ruttimann, U.E (1994). Statistical approaches to developement and validation of predictive instruments. *Critical Care Clinics, 10 (1),* 19-35.

Rutz, W. (1999). Improvement of care for people suffering from depression: The need for comprehensive education. *International Clinical Psychopharmacology, 14 (Suppl. 3),* 27-33.

Sachs, L. (1999). *Angewandte Statistik: Anwendung statistischer Methoden.* Berlin, Heidelberg: Springer.

Sackett, D.L., Haynes, R.B., Guyatt, G.H. & Tugwell, P. (1991). *Clinical epidemiology. A basic science for clinical medicine.* Boston, Toronto, London: Little, Brown and Company.

Sackville, T. (2001). Player and payer: the role of health insurers in improving healthcare provision for people with depression. In A. Dawson & A. Tylee (Eds.). *Depression: Social and economic timebomb* (pp. 9-10). London: BMJ Books.

Salzman, C. (1997). Depressive disorders and other emotional issues in the elderly: current issues. *International Clinical Psychopharmacology, 12 (Suppl. 7),* 37-42.

Sartorius, N., Üstün, T.B., Costa e Silva, J.-A., Goldberg, D., Lecrubier, Y., Ormel, J., Von Korff, M. & Wittchen, H.-U. (1993). An international study of psychological problems in primary care. Preliminary report from the World Health Organisation Collaborative Project on „Psychological Problems in General Health Care“. *Arch Gen Psychiatry, 50,* 819-824.

Sartorius, N., Üstün, T.B., Lecrubier, Y. & Wittchen, H.-U. (1996). Depression comorbid with anxiety: results from the WHO Study on Psychological Disorders in Primary Health Care. *British Journal of Psychiatry, 168 (Suppl. 30),* 38-43.

SAS Software Release 8.1 (1999). SAS Instiutute, Cary North Carolina.

Schmitz, N., Kruse, J., Heckrath, C., Alberti, L. & Tress, W. (1999). Diagnosing mental disorders in primary care: the General Health Questionnaire (GHQ) and the Symptom Check List (SCL-90-R) as screening instruments. *Soc Psychiatry Psychiatr Epidemiol, 34,* 360-366.

Schmitz, N., Kruse, J. & Tress, W. (1999). Psychometric properties of the General Health Questionnaire (GHQ-12) in a German primary care sample. *Acta Psychiatr Scand, 100,* 1-7.

Schmitz, N., Kruse, J. & Tress, W. (2000). Application of stratum-specific likeleehood ratios in mental health screening. *Soc Psychiatry Psychiatr Epidemiol, 35,* 375-379.

Schneider, S. & Margraf, J. (1998). Klassifikatorische Diagnostik, strukturierte Interviews und Therapieindikation. In H. Reinecker (Hrsg.). *Lehrbuch der Klinischen Psychologie. Modelle psychischer Störungen* (S. 47-77). Göttingen, Bern, Toronto, Seattle: Hogrefe.

Schüren, A., Kleinschmidt, S., Müßigbrodt, H. & Schürmann, A. (1999). Diagnostik und Therapie depressiver Störungen in der hausärztlichen Praxis. *Psychiat. Prax., 26,* 163-166.

Seligman, M.E.P. (1974). *Depression and learned helplessness*. In R.J. Friedman & M.M. Katz (Eds.). *The psychology of depression* (pp. 83-113). New York, Toronto, London, Sydney: Wiley & Sons.

Simon, G., Goldberg, D., Tiemens, B.G. & Ustun T.B. (1999). Outcomes of recognized and unrecognized depression in an international primary care study. *General Hospital Psychiatry, 21,* 95-105.

Simon, G. & Von Korff, M. (1995). Recognition, management and outcomes of depression in primary care. *Arch Fam Med, 4,* 99-105.

Skalad, S.R. (1995). Pharmacoeconomic issues in the treatment of depression. *Psychopharmacotherapy, 15 (Suppl. 6),* 76-83.

Skodol, A.E., Schwartz, S., Dohrenwend, B., Levav, I. & Shrout, P. (1994). Minor depression in a cohort of young adults in Israel. *Arch Gen Psychiatry, 51,* 542-551.

Somoza, E., Steer, R.A., Beck, A.T. & Clark, D.A. (1993). Differentiating major depression and panic disorders by self-report and clinical rating scales: ROC analysis and information theory. *Behav. Res. Ther., 32 (7),* 771-782.

Spitzer, R.L., Endicott, J. & Robins, E. (1978). Research Diagnostic Criteria. *Arch Gen Psychiatry, 35,* 773-782.

Spitzer, R.L., Kroenke, K., Linzer, M., Hahn, S., Williams, J., deGruy, F.V., Brody, D. & Davies, M. (1995). Health-related quality of life in primary care patients with mental disorders. *JAMA, 274 (19),* 1511-1517.

Spitzer, R.L., Kroenke, K. & Williams, S. (1999). Validation and utility of a self-report version of Prime-MD. The PHQ Primary Care Study. *JAMA, 282 (18),* 1737-1744.

Spitzer, R.L., Williams J.B., Kroenke, K., Linzer, M., deGruy, F.V., Hahn, S.R., Brody, J. & Johnson, J.G. (1994). Utility of a new procedure for diagnosing mental disorders in primary care. The Prime-MD 1000 Study. *JAMA, 272 (22),* 1749-1756.

Spitznagel, E.L. & Helzer, J.E. (1985). A proposed solution to the base rate problem in the kappa statistic. *Arch Gen Psychiatry, 42,* 725-728.

SPSS (1999), Statistical Package for Social Sciences, SPSS Inc.

Stahl, S.M. (1999). *Psychopharmakologie der Antidepressiva.* Gillingham: Scribe Design.

Stieglitz, R.-D. (1998). Aktueller Stand der syndromalen Diagnostik depressiver Störungen. In R.-D. Stieglitz, E. Fähndrich & H.-J. Möller (Hrsg.). *Syndromale Diagnostik psychischer Störungen* (S. 115-128). Göttingen, Bern, Toronto, Seattle: Hogrefe.

Stieglitz, R.-D. & Baumann, U. (1994). *Psychodiagnostik psychischer Störungen.* Stuttgart: Enke.

Stieglitz, R.-D., Dittmann, V. & Mombour, W. (1992). Erfassungsmethoden und Instrumente zur ICD-10. *Fundamenta Psychiatrica, 6,* 128-136.

Stieglitz, R.-D., Fähndrich, E. & Möller, H.-J. (Hrsg.). (1998). *Syndromale Diagnostik psychischer Störungen.* Göttingen, Bern, Toronto, Seattle: Hogrefe.

Swets, J.A. (1988). Measuring the accuracy of diagnostic systems. *SCIENCE, 240,* 1286-1293.

Szegedi, A., Wetzel, H., Angersbach, D., Philipp, M. & Benkert, O. (1997). Response to treatment in minor and major depression: results of a double-blind comparative study with paroxetine and maprotiline. *Journal of Affective Disoders, 45,* 167-178.

Thangavelu, R. & Martin, R.L. (1995). ICD-10 and DSM-IV: Depiction of the Diagnostic Elephant. *Psychiatric Annals, 25 (1),* 20-28.

Titus, W., van Os, D.P., Ormel, J., van den Brink, R., Jenner, J., Van der Meer, K., Tiemens, B., van der Doorn, W., Smith, A. & van den Brink, W. (1999). Training in primary care physicans improves the management of depression. *General Hospital Psychiatry, 21,* 168-176.

Tylee, A. (2000). Depression in Europe: experience from the DEPRES II survey. *European Neuropsychopharmacology, 10 (Suppl. 4),* 445-448.

Tylee, A. (2001). Management of depression in primary care. In A. Dawson & A. Tylee (Eds.). *Depression: Social and economic timebomb* (pp. 85-92). London: BMJ Books.

Üstün, T.B. & Chatterji (2001). Global burden of depressive disorders and future projections. In A. Dawson & A. Tylee (Eds.). *Depression: Social and economic timebomb* (pp. 31-44). London: BMJ Books.

Van Hook, M.P., Berkman, B. & Dunkle, R. (1996). Assessment tools for general health care settings: Prime-MD, OARS, and SF-36. *Health and Social Work, 21 (3),* 230-234.

Von Korff, M., Shapiro, S., Burke, J.D., Teitlebaum, M., Skinner, E.A., German, P., Turner, R.W., Klein, L. & Burns, B. (1987). Anxiety and depression in primary care. *Arch Gen Psychiatry, 44, 152-156.*

Wagner, H.R., Burns, B.J., Broadhead, W.E., Yarnall, K.S.H., Sigmon, A. & Gaynes, B.N. (2000). Minor depression in family practice: functional morbidity, comorbidity, service utilization and outcomes. *Psychological Medicine, 30,* 1377-1390.

Wells, K., Stewart, A., Hays, R., Burnam, M., Rogers, W., Daniels, M., Berry, S., Greenfield, S. & Ware, J. (1989). The functioning and well-being of depressed patients. Results from the medical outcomes study. *JAMA, 262 (7),* 914-919.

Werneke, U., Goldberg, D.P., Yalcin, I. & Üstün, B.T. (2000). The stability of the factor structure of the General Health Questionnaire. *Pschological Medicine, 30,* 823-829.

Williams, J.B.W., Spitzer, R.L., Linzer, M., Kroenke, K., Hahn, S., deGruy, F.V. & Lazev, L. (1995). Gender differences in depression in primary care. *Am J Obstet Gynecol, 173 (2),* 654-659.

Wilson, J.M. & Junger, F. (1968). *Principles of screening for diseases.* Public Health Papers No. 34. Geneva: WHO.

Wing, J.K., Cooper, J.E. & Sartorius, N. (1974). *Measurement of psychiatric symptoms.* London: University Press.

Winter, S., Wittchen, H.-U., Höfler, M., Spiegel, B., Ormel, H., Müller, N. & Pfister, H. (2000). Häufigkeit und Erkennensrate von Depression in der hausärztlichen Praxis. *Fortschritte in der Medizin, Sonderheft I/2000*, 22-30.

Wittchen, H.-U. (1993). Computer scoring of CIDI diagnosis. *International Journal of Methods in Psychiatric Research, 3,* 101-107.

Wittchen, H.-U. (1994). Reliability and validity studies of the WHO-Composite International Diagnostic Interview (CIDI); a critical review. *J. psychiat. Res., 28 (1),* 57-84.

Wittchen, H.-U. (1997). *Wenn Traurigkeit krank macht. Depressionen erkennen, behandeln, überwinden*. München: Mosaik-Verlag GmbH.

Wittchen, H.-U. (2000). Depression 2000“. Eine Bundesweite Depressions-Screening-Studie in Allgemeinarztpraxen. *Fortschritte in der Medizin, Sonderheft I/2000,* 1-3.

Wittchen, H.-U., Beloch, E., Garczynski, E., Holly, A., Lachner, G., Perkonigg, A., Pfütze, E.-M., Schuster, P., Vodermaier, A., Vosse, A., Wunderlich, U. & Zieglgänsberger, S. (1995). *Manual zum Münchner Composite International Diagnostic Interview (M-CIDI, Paper-Pencil, 2.0 1/95).* München: Max-Planck-Institut für Psychiatrie, Klinische Psychologie (Eigendruck).

Wittchen, H.-U., Essau, C.A., von Zerssen, D., Krieg, J.C. & Zaudig, M. (1992). Lifetime and sixmonth prevalence of mental disorders in the Munich follow-up studie. *Eur Arch Psychiatry Clin Neurosci, 241,* 247-258.

Wittchen, H.-U., Höfler, M. & Meister, W. (2000). *Depressionen in der Allgemeinartzpraxis. Die bundesweite Depressionsstudie.* Stuttgart: Schattauer Verlagsgesellschaft.

Wittchen, H.-U., Höfler, M. & Meister, W. (2001). Prevalence and recognition of depressive syndroms in Germany primary care settings: poorly recognized and treated? *International Clinical Psychopharmacology, 16,* 121-135.

Wittchen, H.-U., Knäuper, B. & Kessler, R. (1994). Lifetime risk of depression. *British Journal of Psychiatry, 165 (Suppl. 26),* 16-22.

Wittchen, H.-U., Lieb, R., Wunderlich, U. & Schuster, P. (1999). Comorbidity in primary care: presentation and consequences. *J Clin Psychiatry, 60 (Suppl. 7),* 29-36.

Wittchen, H.-U., Müller, N., Schmidtkunz, B., Winter, S. & Pfister, H. (2000). Erscheinungsformen, Häufigkeit und Versorgung von Depressionen – Ergebnisse des Bundesweiten *Gesundheitssurveys „Psychische Störungen“. Fortschritte* in der Medizin, Sonderheft I/2000, 4-10.

Wittchen, H.-U. & Pfister, H. (Hrsg.). (1997a). *Diagnostisches Expertensystem für psychische Störungen. DIA-X Interview Längsschnitt.* Frankfurt: Swets & Zeitlinger.

Wittchen, H.-U. & Pfister, H. (Hrsg.). (1997b). *Diagnostisches Expertensystem für psychische Störungen. DIA-X Instruktionsmanual.* Frankfurt: Swets & Zeitlinger.

Wittchen, H.-U., Zaudig, M., Schramm, E., Spengler, P., Mombour, W., Klug, J., Horn, R. (1990). *Strukturiertes Klinisches Interview für DSM-III-R (SKID, Version 2.0).* Weinheim: Beltz.

Wolfersdorf, M. (2000). *Krankheit Depression erkennen, verstehen, behandeln.* Bonn: Psychiatrie-Verlag.

World Health Organisation (1992). *The ICD-10 Classification of Mental and Behavioral Disorders: Clinical Descriptions and Diagnostic Guidelines.* Geneva: WHO.

World Health Organisation (1993). *Composite International Diagnostic Interview,* Geneva: WHO.

World Health Organisation (1998). *Info Package: Mastering Depression in Primary Care, Version 2.2.* Frederiksborg: WHO, Regional Office for Europe, Psychiatric Research Unit.

Wu, J., Buchsbaum, M., Gillin, C., Tang, C., Cadwell, S., Wiegand, M., Najafi, A., Klein, E., Hazen, K. & Bunney, W.E. (1999). Prediction of antidepressant effects of sleep deprivation by metabolic rates in the ventral anterior cingulate and medial prefrontal cortex. *Am J Psychiatry, 156 (8),* 1149-1158.

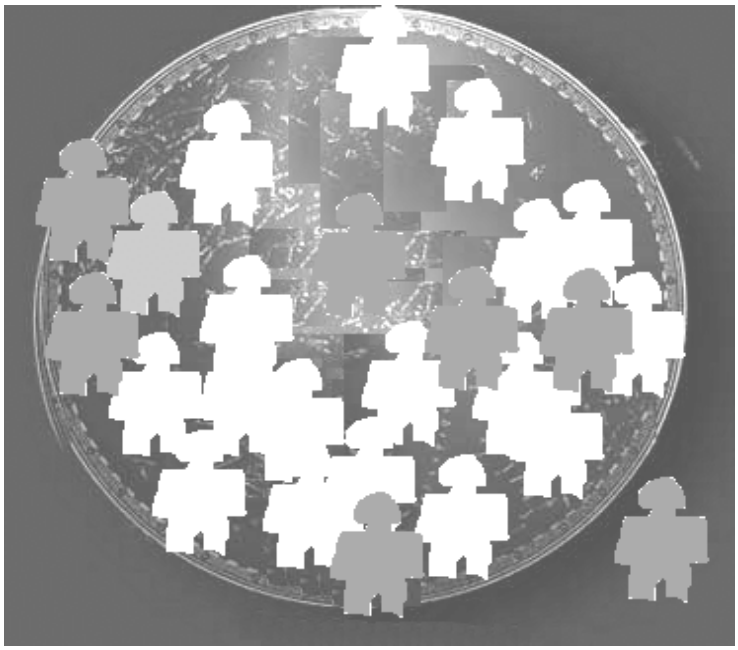
Zaudig, M. (2001). Diagnose und Differentialdiagnose der Depression und Demenz im Alter. In U. Hegerl, M. Zaudig & H.-J. Möller (Hrsg.). *Depression und Demenz im Alter. Abgrenzung, Wechselwirkung, Diagnose und Therapie* (S. 19-38). Wien, New York: Springer-Verlag.

Zweig, M.H. & Campbell, G. (1993). Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clinical Chemistry, 39 (4),* 561-577.

Anhang A

Poster zur Patienteninformation im Wartezimmer

Was uns alle angeht ....



Die **Weltgesundheitsorganisation** hat den weltweit hohen medizinischen Stellenwert der Depression herausgestellt und nationale Programme zur Bekämpfung dieser Volkskrankheit gefordert.

Bundesweite Programme sind geschaffen worden, um die Zusammenarbeit von Universitäten, Fachkliniken, Krankenkassen und insbesondere der Hausärzte zum Wohle der Patienten zu verbessern.

Ziel ist es, die **Krankheit Depression**, ihre Formen, Symptome und Behandlungsmethoden bei den Betroffenen und in der Öffentlichkeit bekannter zu machen. Denn oft wird die Depression **nicht erkannt**, insbesondere wenn sie sich nur in **körperlichen Symptomen** äußert.

Für das Gelingen dieser Aufgabe sollten sich nicht nur depressive Patienten, sondern auch eine **möglichst große Zahl von nicht depressiven Interessierten** an den in dieser Praxis durchgeführten Befragungen beteiligen.

Weitere Informationen zu diesem Forschungsvorhaben erhalten Sie im
Studienzentrum Nürnberg der
Ludwig- Maximilians- Universität München.
Kesslerplatz 5
90489 Nürnberg
Tel: 0911/53 01 809 oder 53 01 810

Kompetenznetz Depression — LMU Ludwig Maximilians Universität — Praxisnetz Nürnberg Nord — DEPRESSION

Weitere Informationen zum Kompetenznetz Depression erhalten Sie bei der
Ludwig-Maximilians-Universität München.
Klinik für Psychiatrie und Psychotherapie,
Kompetenznetz Depression
Nußbaumstr. 7, 80336 München
Tel: 089/5160-5541 oder 51605558

Anhang B

Vierfeldertafeln für die Cut-Off-Werte des GHQ-12, des B-PHQ Algorithmus A und des B-PHQ Algorithmus B

Vierfeldertafeln des GHQ-12 für die gesamte Stichprobe validiert anhand der Standarddepressionsdiagnosen des DIA-X (N = 448).

Cut-Off-Wert 0/1

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	73	172	245
	negativ	9	194	203
		82	366	448

Cut-Off-Wert 1/2

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	70	136	206
	negativ	12	230	242
		82	366	448

Cut-Off-Wert 2/3

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	65	108	173
	negativ	17	258	275
		82	366	448

Cut-Off-Wert 3/4

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	63	97	160
	negativ	19	269	288
		82	366	448

Cut-Off-Wert 4/5

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	60	73	133
	negativ	22	293	315
		82	366	448

Cut-Off-Wert 5/6

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	59	54	113
	negativ	23	312	335
		82	366	448

Cut-Off-Wert 6/7

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	55	45	100
	negativ	27	321	348
		82	366	448

Cut-Off-Wert 7/8

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	52	34	86
	negativ	30	332	362
		82	366	448

Cut-Off-Wert 8/9

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	47	28	75
	negativ	35	338	373
		82	366	448

Cut-Off-Wert 9/10

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	40	20	60
	negativ	42	346	388
		82	366	448

Cut-Off-Wert 10/11

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	34	15	49
	negativ	48	351	399
		82	366	448

Cut-Off-Wert 11/12

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	24	10	34
	negativ	58	356	414
		82	366	448

Vierfeldertafeln des GHQ-12 für die gesamte Stichprobe validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 448)

Cut-Off-Wert 0/1

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	117	128	245
	negativ	19	184	203
		136	312	448

Cut-Off-Wert 1/2

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	110	96	206
	negativ	26	216	242
		136	312	448

Cut-Off-Wert 2/3

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	101	72	173
	negativ	35	240	275
		136	312	448

Cut-Off-Wert 3/4

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	94	66	160
	negativ	42	246	288
		136	312	448

Cut-Off-Wert 4/5

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	85	48	133
	negativ	51	264	315
		136	312	448

Cut-Off-Wert 5/6

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	79	34	113
	negativ	57	278	335
		136	312	448

Cut-Off-Wert 6/7

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	74	26	100
	negativ	62	286	348
		136	312	448

Cut-Off-Wert 7/8

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	68	18	86
	negativ	68	294	362
		136	312	448

Cut-Off-Wert 8/9

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	60	15	75
	negativ	76	297	373
		136	312	448

Cut-Off-Wert 9/10

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	49	11	60
	negativ	87	301	388
		136	312	448

Cut-Off-Wert 10/11

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	42	7	49
	negativ	94	305	399
		136	312	448

Cut-Off-Wert 11/12

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	28	6	34
	negativ	108	306	414
		136	312	448

Vierfeldertafeln des GHQ-12 für die Subgruppe der Frauen validiert anhand der Standarddepressionsdiagnosen des DIA-X (N = 283)

Cut-Off-Wert 0/1

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	51	107	158
	negativ	6	119	125
		57	226	283

Cut-Off-Wert 1/2

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	50	83	133
	negativ	7	143	150
		57	226	283

Cut-Off-Wert 2/3

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	46	64	110
	negativ	11	162	173
		57	226	283

Cut-Off-Wert 3/4

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	44	58	102
	negativ	13	168	181
		57	226	283

Cut-Off-Wert 4/5

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	42	41	83
	negativ	15	185	200
		57	226	283

Cut-Off-Wert 5/6

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	41	28	69
	negativ	16	198	214
		57	226	283

Cut-Off-Wert 6/7

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	38	23	61
	negativ	19	203	222
		57	226	283

Cut-Off-Wert 7/8

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	36	19	55
	negativ	21	207	228
		57	226	283

Cut-Off-Wert 8/9

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	33	18	51
	negativ	24	208	232
		57	226	283

Cut-Off-Wert 9/10

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	28	12	40
	negativ	29	214	243
		57	226	283

Cut-Off-Wert 10/11

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	24	10	34
	negativ	33	216	249
		57	226	283

Cut-Off-Wert 11/12

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	16	5	21
	negativ	41	221	262
		57	226	283

Vierfeldertafeln des GHQ-12 für die Subgruppe der Frauen validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 283)

Cut-Off-Wert 0/1

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	76	82	158
	negativ	11	114	125
		87	196	283

Cut-Off-Wert 1/2

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	72	61	133
	negativ	15	135	150
		87	196	283

Cut-Off-Wert 2/3

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	67	43	110
	negativ	20	153	173
		87	196	283

Cut-Off-Wert 3/4

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	62	40	102
	negativ	25	156	181
		87	196	283

Cut-Off-Wert 4/5

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	55	28	83
	negativ	32	168	200
		87	196	283

Cut-Off-Wert 5/6

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	52	17	69
	negativ	35	179	214
		87	196	283

Cut-Off-Wert 6/7

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	48	13	61
	negativ	39	183	222
		87	196	283

Cut-Off-Wert 7/8

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	44	11	55
	negativ	43	185	228
		87	196	283

Cut-Off-Wert 8/9

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	40	11	51
	negativ	47	185	232
		87	196	283

Cut-Off-Wert 9/10

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	32	8	40
	negativ	55	188	243
		87	196	283

Cut-Off-Wert 10/11

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	28	6	34
	negativ	59	190	249
		87	196	283

Cut-Off-Wert 11/12

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	16	5	21
	negativ	71	191	262
		87	196	283

Vierfeldertafeln des GHQ-12 für die Subgruppe der Männer validiert anhand der Standarddepressionsdiagnosen des DIA-X (N = 165)

Cut-Off-Wert 0/1

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	22	65	87
	negativ	3	75	78
		25	140	165

Cut-Off-Wert 1/2

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	20	53	73
	negativ	5	87	92
		25	140	165

Cut-Off-Wert 2/3

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	19	44	63
	negativ	6	96	102
		25	140	165

Cut-Off-Wert 3/4

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	19	39	58
	negativ	6	101	107
		25	140	165

Cut-Off-Wert 4/5

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	18	32	50
	negativ	7	108	115
		25	140	165

Cut-Off-Wert 5/6

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	18	26	44
	negativ	7	114	121
		25	140	165

Cut-Off-Wert 6/7

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	17	22	39
	negativ	8	118	126
		25	140	165

Cut-Off-Wert 7/8

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	16	15	31
	negativ	9	125	134
		25	140	165

Cut-Off-Wert 8/9

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	14	10	24
	negativ	11	130	141
		25	140	165

Cut-Off-Wert 9/10

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	12	8	20
	negativ	13	132	145
		25	140	165

Cut-Off-Wert 10/11

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	10	5	15
	negativ	15	135	150
		25	140	165

Cut-Off-Wert 11/12

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	8	5	13
	negativ	17	135	152
		25	140	165

Vierfeldertafeln des GHQ-12 für die Subgruppe der Männer validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 165)

Cut-Off-Wert 0/1

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	41	46	87
	negativ	8	70	78
		49	116	165

Cut-Off-Wert 1/2

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	38	35	73
	negativ	11	81	92
		49	116	165

Cut-Off-Wert 2/3

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	34	29	63
	negativ	15	87	102
		49	116	165

Cut-Off-Wert 3/4

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	32	26	58
	negativ	17	90	107
		49	116	165

Cut-Off-Wert 4/5

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	30	20	50
	negativ	19	96	115
		49	116	165

Cut-Off-Wert 5/6

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	27	17	44
	negativ	22	99	121
		49	116	165

Cut-Off-Wert 6/7

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	26	13	39
	negativ	23	103	126
		49	116	165

Cut-Off-Wert 7/8

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	24	7	31
	negativ	25	109	134
		49	116	165

Cut-Off-Wert 8/9

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	20	4	24
	negativ	29	112	141
		49	116	165

Cut-Off-Wert 9/10

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	17	3	20
	negativ	32	113	145
		49	116	165

Cut-Off-Wert 10/11

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	14	1	15
	negativ	35	115	150
		49	116	165

Cut-Off-Wert 11/12

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	12	1	13
	negativ	37	115	152
		49	116	165

Vierfeldertafeln des GHQ-12 für die Subgruppe Alter < 56,21 Jahre validiert anhand der Standarddepressionsdiagnosen des DIA-X (N = 224)

Cut-Off-Wert 0/1

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	51	98	149
	negativ	4	71	75
		55	169	224

Cut-Off-Wert 1/2

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	50	79	129
	negativ	5	90	95
		55	169	224

Cut-Off-Wert 2/3

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	47	61	108
	negativ	8	108	116
		55	169	224

Cut-Off-Wert 3/4

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	46	57	103
	negativ	9	112	121
		55	169	224

Cut-Off-Wert 4/5

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	43	41	84
	negativ	12	128	140
		55	169	224

Cut-Off-Wert 5/6

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	43	29	72
	negativ	12	140	152
		55	169	224

Cut-Off-Wert 6/7

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	41	23	64
	negativ	14	146	160
		55	169	224

Cut-Off-Wert 7/8

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	40	17	57
	negativ	15	152	167
		55	169	224

Cut-Off-Wert 8/9

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	35	15	50
	negativ	20	154	174
		55	169	224

Cut-Off-Wert 9/10

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	30	12	42
	negativ	25	157	182
		55	169	224

Cut-Off-Wert 10/11

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	25	8	33
	negativ	30	161	191
		55	169	224

Cut-Off-Wert 11/12

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	17	5	22
	negativ	38	164	202
		55	169	224

Vierfeldertafeln des GHQ-12 für die Subgruppe Alter < 56,21 Jahre validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 224)

Cut-Off-Wert 0/1

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	77	72	149
	negativ	8	67	75
		85	139	224

Cut-Off-Wert 1/2

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	75	54	129
	negativ	10	85	95
		85	139	224

Cut-Off-Wert 2/3

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	68	40	108
	negativ	17	99	116
		85	139	224

Cut-Off-Wert 3/4

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	67	36	103
	negativ	18	103	121
		85	139	224

Cut-Off-Wert 4/5

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	59	25	84
	negativ	26	114	140
		85	139	224

Cut-Off-Wert 5/6

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	55	17	72
	negativ	30	122	152
		85	139	224

Cut-Off-Wert 6/7

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	53	11	64
	negativ	32	128	160
		85	139	224

Cut-Off-Wert 7/8

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	50	7	57
	negativ	35	132	167
		85	139	224

Cut-Off-Wert 8/9

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	43	7	50
	negativ	42	132	174
		85	139	224

Cut-Off-Wert 9/10

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	37	5	42
	negativ	48	134	182
		85	139	224

Cut-Off-Wert 10/11

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	32	1	33
	negativ	53	138	191
		85	139	224

Cut-Off-Wert 11/12

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	21	1	22
	negativ	64	138	202
		85	139	224

Vierfeldertafeln des GHQ-12 für die Subgruppe Alter > 56,21 Jahre validiert anhand der Standarddepressionsdiagnosen des DIA-X (N = 224)

Cut-Off-Wert 0/1

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	22	74	96
	negativ	5	123	128
		27	197	224

Cut-Off-Wert 1/2

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	20	57	77
	negativ	7	140	147
		27	197	224

Cut-Off-Wert 2/3

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	18	47	65
	negativ	9	150	159
		27	197	224

Cut-Off-Wert 3/4

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	17	40	57
	negativ	10	157	167
		27	197	224

Cut-Off-Wert 4/5

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	17	32	49
	negativ	10	165	175
		27	197	224

Cut-Off-Wert 5/6

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	16	25	41
	negativ	11	172	183
		27	197	224

Cut-Off-Wert 6/7

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	14	22	36
	negativ	13	175	188
		27	197	224

Cut-Off-Wert 7/8

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	12	17	29
	negativ	15	180	195
		27	197	224

Cut-Off-Wert 8/9

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	12	13	25
	negativ	15	184	199
		27	197	224

Cut-Off-Wert 9/10

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	10	8	18
	negativ	17	189	206
		27	197	224

Cut-Off-Wert 10/11

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	9	7	16
	negativ	18	190	208
		27	197	224

Cut-Off-Wert 11/12

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	7	5	12
	negativ	20	192	212
		27	197	224

Vierfeldertafeln des GHQ-12 für die Subgruppe Alter > 56,20 Jahre validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 224)

Cut-Off-Wert 0/1

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	40	56	96
	negativ	11	117	128
		51	173	224

Cut-Off-Wert 1/2

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	35	42	77
	negativ	16	131	147
		51	173	224

Cut-Off-Wert 2/3

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	33	32	65
	negativ	18	141	159
		51	173	224

Cut-Off-Wert 3/4

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	27	30	57
	negativ	24	143	167
		51	173	224

Cut-Off-Wert 4/5

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	26	23	49
	negativ	25	150	175
		51	173	224

Cut-Off-Wert 5/6

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	24	17	41
	negativ	27	156	183
		51	173	224

Cut-Off-Wert 6/7

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	21	15	36
	negativ	30	158	188
		51	173	224

Cut-Off-Wert 7/8

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	18	11	29
	negativ	33	162	195
		51	173	224

Cut-Off-Wert 8/9

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	17	8	25
	negativ	34	165	199
		51	173	224

Cut-Off-Wert 9/10

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	12	6	18
	negativ	39	167	206
		51	173	224

Cut-Off-Wert 10/11

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	10	6	16
	negativ	41	167	208
		51	173	224

Cut-Off-Wert 11/12

		CIDI positiv	CIDI negativ	
GHQ-12	positiv	7	5	12
	negativ	44	168	212
		51	173	224

Vierfeldertafeln des B-PHQ Auswertungsalgorithmus A für die gesamte Stichprobe validiert anhand der Standarddepressionsdiagnosen des DIA-X (N = 448)

Cut-Off-Wert: Mindestens das Item 1a oder 1b ist positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	65	63	128
	negativ	17	303	320
		82	366	448

Cut-Off-Wert: Mindestens 2 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	65	53	118
	negativ	17	313	330
		82	366	448

Cut-Off-Wert: Mindestens 3 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	63	42	105
	negativ	19	324	343
		82	366	448

Cut-Off-Wert: Mindestens 4 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	58	34	92
	negativ	24	332	356
		82	366	448

Cut-Off-Wert: Mindestens 5 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	50	24	74
	negativ	32	342	374
		82	366	448

Cut-Off-Wert: Mindestens 6 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	37	18	55
	negativ	45	348	393
		82	366	448

Cut-Off-Wert: Mindestens 7 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	29	12	41
	negativ	53	354	407
		82	366	448

Cut-Off-Wert: Mindestens 8 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	19	5	24
	negativ	63	361	424
		82	366	448

Cut-Off-Wert: Alle 9 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	7	3	10
	negativ	75	363	438
		82	366	448

Vierfeldertafeln des B-PHQ Auswertungsalgorithmus A für die gesamte Stichprobe validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 448)

Cut-Off-Wert: Mindestens das Item 1a oder 1b ist positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	88	40	128
	negativ	48	272	320
		136	312	448

Cut-Off-Wert: Mindestens 2 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	88	30	118
	negativ	48	282	330
		136	312	448

Cut-Off-Wert: Mindestens 3 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	81	24	105
	negativ	55	288	343
		136	312	448

Cut-Off-Wert: Mindestens 4 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	73	19	92
	negativ	63	293	356
		136	312	448

Cut-Off-Wert: Mindestens 5 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	63	11	74
	negativ	73	301	374
		136	312	448

Cut-Off-Wert: Mindestens 6 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	48	7	55
	negativ	88	305	393
		136	312	448

Cut-Off-Wert: Mindestens 7 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	37	4	41
	negativ	99	308	407
		136	312	448

Cut-Off-Wert: Mindestens 8 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	23	1	24
	negativ	113	311	424
		136	312	448

Cut-Off-Wert: Alle 9 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	9	1	10
	negativ	127	311	438
		136	312	448

Vierfeldertafeln des B-PHQ Auswertungsalgorithmus A für die Subgruppe der Frauen validiert anhand der Standarddepressionsdiagnosen des DIA-X (N = 283)

Cut-Off-Wert: Mindestens das Item 1a oder 1b ist positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	44	33	77
	negativ	13	193	206
		57	226	283

Cut-Off-Wert: Mindestens 2 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	44	28	72
	negativ	13	198	211
		57	226	283

Cut-Off-Wert: Mindestens 3 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	42	24	66
	negativ	15	202	217
		57	226	283

Cut-Off-Wert: Mindestens 4 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	39	19	58
	negativ	18	207	225
		57	226	283

Cut-Off-Wert: Mindestens 5 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	34	13	47
	negativ	23	213	236
		57	226	283

Cut-Off-Wert: Mindestens 6 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	26	9	35
	negativ	31	217	248
		57	226	283

Cut-Off-Wert: Mindestens 7 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	20	7	27
	negativ	37	219	256
		57	226	283

Cut-Off-Wert: Mindestens 8 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	15	2	17
	negativ	42	224	266
		57	226	283

Cut-Off-Wert: Alle 9 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	5	1	6
	negativ	52	225	277
		57	226	283

Vierfeldertafeln des B-PHQ Auswertungsalgorithmus A für die Subgruppe der Frauen validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 283)

Cut-Off-Wert: Mindestens das Item 1a oder 1b ist positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	55	22	77
	negativ	32	174	206
		87	196	283

Cut-Off-Wert: Mindestens 2 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	55	17	72
	negativ	32	179	211
		87	196	283

Cut-Off-Wert: Mindestens 3 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	52	14	66
	negativ	35	182	217
		87	196	283

Cut-Off-Wert: Mindestens 4 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	47	11	58
	negativ	40	185	225
		87	196	283

Cut-Off-Wert: Mindestens 5 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	42	5	47
	negativ	45	191	236
		87	196	283

Cut-Off-Wert: Mindestens 6 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	32	3	35
	negativ	55	193	248
		87	196	283

Cut-Off-Wert: Mindestens 7 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	24	3	27
	negativ	63	193	256
		87	196	283

Cut-Off-Wert: Mindestens 8 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	16	1	17
	negativ	71	195	266
		87	196	283

Cut-Off-Wert: Alle 9 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	5	1	6
	negativ	82	195	277
		87	196	283

Vierfeldertafeln des B-PHQ Auswertungsalgorithmus A für die Subgruppe der Männer validiert anhand der Standarddepressionsdiagnosen des DIA-X (N = 165)

Cut-Off-Wert: Mindestens das Item 1a oder 1b ist positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	21	30	51
	negativ	4	110	114
		25	140	165

Cut-Off-Wert: Mindestens 2 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	21	25	46
	negativ	4	115	119
		25	140	165

Cut-Off-Wert: Mindestens 3 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	21	18	39
	negativ	4	122	126
		25	140	165

Cut-Off-Wert: Mindestens 4 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	19	15	34
	negativ	6	125	131
		25	140	165

Cut-Off-Wert: Mindestens 5 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	16	11	27
	negativ	9	129	138
		25	140	165

Cut-Off-Wert: Mindestens 6 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	11	9	20
	negativ	14	131	145
		25	140	165

Cut-Off-Wert: Mindestens 7 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	9	5	14
	negativ	16	135	151
		25	140	165

Cut-Off-Wert: Mindestens 8 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	4	3	7
	negativ	21	137	158
		25	140	165

Cut-Off-Wert: Alle 9 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	2	2	4
	negativ	23	138	161
		25	140	165

Vierfeldertafeln des B-PHQ Auswertungsalgorithmus A für die Subgruppe der Männern validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 165)

Cut-Off-Wert: Mindestens das Item 1a oder 1b ist positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	33	18	51
	negativ	16	98	114
		49	116	165

Cut-Off-Wert: Mindestens 2 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	33	13	46
	negativ	16	103	119
		49	116	165

Cut-Off-Wert: Mindestens 3 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	29	10	39
	negativ	20	106	126
		49	116	165

Cut-Off-Wert: Mindestens 4 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	26	8	34
	negativ	23	108	131
		49	116	165

Cut-Off-Wert: Mindestens 5 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	21	6	27
	negativ	28	110	138
		49	116	165

Cut-Off-Wert: Mindestens 6 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	16	4	20
	negativ	33	112	145
		49	116	165

Cut-Off-Wert: Mindestens 7 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	13	1	14
	negativ	36	115	151
		49	116	165

Cut-Off-Wert: Mindestens 8 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	7	0	7
	negativ	42	116	158
		49	116	165

Cut-Off-Wert: Alle 9 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	4	0	4
	negativ	45	116	161
		49	116	165

Vierfeldertafeln des B-PHQ Auswertungsalgorithmus A für die Subgruppe Alter < 56,21 validiert anhand der Standarddepressionsdiagnosen des DIA-X (N = 224)

Cut-Off-Wert: Mindestens das Item 1a oder 1b ist positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	48	41	89
	negativ	7	128	135
		55	169	224

Cut-Off-Wert: Mindestens 2 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	48	36	84
	negativ	7	133	140
		55	169	224

Cut-Off-Wert: Mindestens 3 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	46	29	75
	negativ	9	140	149
		55	169	224

Cut-Off-Wert: Mindestens 4 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	42	24	66
	negativ	13	145	158
		55	169	224

Cut-Off-Wert: Mindestens 5 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	36	18	54
	negativ	19	151	170
		55	169	224

Cut-Off-Wert: Mindestens 6 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	28	13	41
	negativ	27	156	183
		55	169	224

Cut-Off-Wert: Mindestens 7 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	22	10	32
	negativ	33	159	192
		55	169	224

Cut-Off-Wert: Mindestens 8 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	14	4	18
	negativ	41	165	206
		55	169	224

Cut-Off-Wert: Alle 9 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	5	2	7
	negativ	50	167	217
		55	169	224

Vierfeldertafeln des B-PHQ Auswertungsalgorithmus A für die Subgruppe Alter < 56,21 validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 224)

Cut-Off-Wert: Mindestens das Item 1a oder 1b ist positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	64	25	89
	negativ	21	114	135
		85	139	224

Cut-Off-Wert: Mindestens 2 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	64	20	84
	negativ	21	119	140
		85	139	224

Cut-Off-Wert: Mindestens 3 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	59	16	75
	negativ	26	123	149
		85	139	224

Cut-Off-Wert: Mindestens 4 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	54	12	66
	negativ	31	127	158
		85	139	224

Cut-Off-Wert: Mindestens 5 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	47	7	54
	negativ	38	132	170
		85	139	224

Cut-Off-Wert: Mindestens 6 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	37	4	41
	negativ	48	135	183
		85	139	224

Cut-Off-Wert: Mindestens 7 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	29	3	32
	negativ	56	136	192
		85	139	224

Cut-Off-Wert: Mindestens 8 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	18	0	18
	negativ	67	139	206
		85	139	224

Cut-Off-Wert: Alle 9 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	7	0	7
	negativ	78	139	217
		85	139	224

Vierfeldertafeln des B-PHQ Auswertungsalgorithmus A für die Subgruppe Alter > 56,20 validiert anhand der Standarddepressionsdiagnosen des DIA-X (N = 224)

Cut-Off-Wert: Mindestens das Item 1a oder 1b ist positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	17	22	39
	negativ	10	175	185
		27	197	224

Cut-Off-Wert: Mindestens 2 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	17	17	34
	negativ	10	180	190
		27	197	224

Cut-Off-Wert: Mindestens 3 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	17	13	30
	negativ	10	184	194
		27	197	224

Cut-Off-Wert: Mindestens 4 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	16	10	26
	negativ	11	187	198
		27	197	224

Cut-Off-Wert: Mindestens 5 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	14	6	20
	negativ	13	191	204
		27	197	224

Cut-Off-Wert: Mindestens 6 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	9	5	14
	negativ	18	192	210
		27	197	224

Cut-Off-Wert: Mindestens 7 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	7	2	9
	negativ	20	195	215
		27	197	224

Cut-Off-Wert: Mindestens 8 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	5	1	6
	negativ	22	196	218
		27	197	224

Cut-Off-Wert: Alle 9 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	2	1	3
	negativ	25	196	221
		27	197	224

Vierfeldertafeln des B-PHQ Auswertungsalgorithmus A für die Subgruppe Alter > 56,20 validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 224)

Cut-Off-Wert: Mindestens das Item 1a oder 1b ist positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	24	15	39
	negativ	27	158	185
		51	173	224

Cut-Off-Wert: Mindestens 2 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	24	10	34
	negativ	27	163	190
		51	173	224

Cut-Off-Wert: Mindestens 3 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	22	8	30
	negativ	29	165	194
		51	173	224

Cut-Off-Wert: Mindestens 4 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	19	7	26
	negativ	32	166	198
		51	173	224

Cut-Off-Wert: Mindestens 5 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	16	4	20
	negativ	35	169	204
		51	173	224

Cut-Off-Wert: Mindestens 6 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	11	3	14
	negativ	40	170	210
		51	173	224

Cut-Off-Wert: Mindestens 7 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	8	1	9
	negativ	43	172	215
		51	173	224

Cut-Off-Wert: Mindestens 8 Items sind positiv, darunter sind 1a oder 1b

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	5	1	6
	negativ	46	172	218
		51	173	224

Cut-Off-Wert: Alle 9 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus A)	positiv	2	1	3
	negativ	49	172	221
		51	173	224

Vierfeldertafeln des B-PHQ Auswertungsalgorithmus B für die gesamte Stichprobe validiert anhand der Standarddepressionsdiagnosen des DIA-X (N = 448)

Cut-Off-Wert: Mindestens 1 Item ist positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	75	168	243
	negativ	7	198	205
		82	366	448

Cut-Off-Wert: Mindestens 2 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	71	95	166
	negativ	11	271	282
		82	366	448

Cut-Off-Wert: Mindestens 3 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	64	63	127
	negativ	18	303	321
		82	366	448

Cut-Off-Wert: Mindestens 4 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	58	40	98
	negativ	24	326	350
		82	366	448

Cut-Off-Wert: Mindestens 5 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	50	26	76
	negativ	32	340	372
		82	366	448

Cut-Off-Wert: Mindestens 6 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	37	18	55
	negativ	45	348	393
		82	366	448

Cut-Off-Wert: Mindestens 7 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	29	12	41
	negativ	53	354	407
		82	366	448

Cut-Off-Wert: Mindestens 8 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	19	5	24
	negativ	63	361	424
		82	366	448

Cut-Off-Wert: Alle 9 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	7	3	10
	negativ	75	363	438
		82	366	448

Vierfeldertafeln des B-PHQ Auswertungsalgorithmus B für die gesamte Stichprobe validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 448)

Cut-Off-Wert: Mindestens 1 Item ist positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	118	125	243
	negativ	18	187	205
		136	312	448

Cut-Off-Wert: Mindestens 2 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	107	59	166
	negativ	29	253	282
		136	312	448

Cut-Off-Wert: Mindestens 3 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	88	39	127
	negativ	48	273	321
		136	312	448

Cut-Off-Wert: Mindestens 4 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	74	24	98
	negativ	62	288	350
		136	312	448

Cut-Off-Wert: Mindestens 5 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	64	12	76
	negativ	72	300	372
		136	312	448

Cut-Off-Wert: Mindestens 6 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	48	7	55
	negativ	88	305	393
		136	312	448

Cut-Off-Wert: Mindestens 7 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	37	4	41
	negativ	99	308	407
		136	312	448

Cut-Off-Wert: Mindestens 8 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	23	1	24
	negativ	113	311	424
		136	312	448

Cut-Off-Wert: Alle 9 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	9	1	10
	negativ	127	311	438
		136	312	448

Vierfeldertafeln des B-PHQ Auswertungsalgorithmus B für die Subgruppe der Frauen validiert anhand der Standarddepressionsdiagnosen des DIA-X (N = 283)

Cut-Off-Wert: Mindestens 1 Item ist positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	53	103	156
	negativ	4	123	127
		57	226	283

Cut-Off-Wert: Mindestens 2 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	49	56	105
	negativ	8	170	178
		57	226	283

Cut-Off-Wert: Mindestens 3 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	43	37	80
	negativ	14	189	203
		57	226	283

Cut-Off-Wert: Mindestens 4 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	39	21	60
	negativ	18	205	223
		57	226	283

Cut-Off-Wert: Mindestens 5 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	34	13	47
	negativ	23	213	236
		57	226	283

Cut-Off-Wert: Mindestens 6 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	26	9	35
	negativ	31	217	248
		57	226	283

Cut-Off-Wert: Mindestens 7 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	20	7	27
	negativ	37	219	256
		57	226	283

Cut-Off-Wert: Mindestens 8 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	15	2	17
	negativ	42	224	266
		57	226	283

Cut-Off-Wert: Alle 9 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	5	1	6
	negativ	52	225	277
		57	226	283

Vierfeldertafeln des B-PHQ Auswertungsalgorithmus B für die Subgruppe der Frauen validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 283)

Cut-Off-Wert: Mindestens 1 Item ist positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	77	79	156
	negativ	10	117	127
		87	196	283

Cut-Off-Wert: Mindestens 2 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	69	36	105
	negativ	18	160	178
		87	196	283

Cut-Off-Wert: Mindestens 3 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	57	23	80
	negativ	30	173	203
		87	196	283

Cut-Off-Wert: Mindestens 4 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	47	13	60
	negativ	40	183	223
		87	196	283

Cut-Off-Wert: Mindestens 5 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	42	5	47
	negativ	45	191	236
		87	196	283

Cut-Off-Wert: Mindestens 6 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	32	3	35
	negativ	55	193	248
		87	196	283

Cut-Off-Wert: Mindestens 7 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	24	3	27
	negativ	63	193	256
		87	196	283

Cut-Off-Wert: Mindestens 8 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	16	1	17
	negativ	71	195	266
		87	196	283

Cut-Off-Wert: Alle 9 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	5	1	6
	negativ	82	195	277
		87	196	283

Vierfeldertafeln des B-PHQ Auswertungsalgorithmus B für die Subgruppe der Männer validiert anhand der Standarddepressionsdiagnosen des DIA-X (N = 165)

Cut-Off-Wert: Mindestens 1 Item ist positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	22	65	87
	negativ	3	75	78
		25	140	165

Cut-Off-Wert: Mindestens 2 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	22	39	61
	negativ	3	101	104
		25	140	165

Cut-Off-Wert: Mindestens 3 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	21	26	47
	negativ	4	114	118
		25	140	165

Cut-Off-Wert: Mindestens 4 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	19	19	38
	negativ	6	121	127
		25	140	165

Cut-Off-Wert: Mindestens 5 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	16	13	29
	negativ	9	127	136
		25	140	165

Cut-Off-Wert: Mindestens 6 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	11	9	20
	negativ	14	131	145
		25	140	165

Cut-Off-Wert: Mindestens 7 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	9	5	14
	negativ	16	135	151
		25	140	165

Cut-Off-Wert: Mindestens 8 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	4	3	7
	negativ	21	137	158
		25	140	165

Cut-Off-Wert: Alle 9 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	2	2	4
	negativ	23	138	161
		25	140	165

Vierfeldertafeln des B-PHQ Auswertungsalgorithmus B für die Subgruppe der Männern validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 165)

Cut-Off-Wert: Mindestens 1 Item ist positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	41	46	87
	negativ	8	70	78
		49	116	165

Cut-Off-Wert: Mindestens 2 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	38	23	61
	negativ	11	93	104
		49	116	165

Cut-Off-Wert: Mindestens 3 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	31	16	47
	negativ	18	100	118
		49	116	165

Cut-Off-Wert: Mindestens 4 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	27	11	38
	negativ	22	105	127
		49	116	165

Cut-Off-Wert: Mindestens 5 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	22	7	29
	negativ	27	109	136
		49	116	165

Cut-Off-Wert: Mindestens 6 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	16	4	20
	negativ	33	112	145
		49	116	165

Cut-Off-Wert: Mindestens 7 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	13	1	14
	negativ	36	115	151
		49	116	165

Cut-Off-Wert: Mindestens 8 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	7	0	7
	negativ	42	116	158
		49	116	165

Cut-Off-Wert: Alle 9 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	4	0	4
	negativ	45	116	161
		49	116	165

Vierfeldertafeln des B-PHQ (Auswertungsalgorithmus B) für die Subgruppe Alter < 56,21 validiert anhand der Standarddepressionsdiagnosen des DIA-X (N = 224)

Cut-Off-Wert: Mindestens 1 Item ist positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	52	84	136
	negativ	3	85	88
		55	169	224

Cut-Off-Wert: Mindestens 2 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	49	56	105
	negativ	6	113	119
		55	169	224

Cut-Off-Wert: Mindestens 3 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	46	39	85
	negativ	9	130	139
		55	169	224

Cut-Off-Wert: Mindestens 4 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	42	28	70
	negativ	13	141	154
		55	169	224

Cut-Off-Wert: Mindestens 5 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	36	19	55
	negativ	19	150	169
		55	169	224

Cut-Off-Wert: Mindestens 6 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	28	13	41
	negativ	27	156	183
		55	169	224

Cut-Off-Wert: Mindestens 7 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	22	10	32
	negativ	33	159	192
		55	169	224

Cut-Off-Wert: Mindestens 8 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	14	4	18
	negativ	41	165	206
		55	169	224

Cut-Off-Wert: Alle 9 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	5	2	7
	negativ	50	167	217
		55	169	224

Vierfeldertafeln des B-PHQ Auswertungsalgorithmus B für die Subgruppe Alter < 56,21 validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 224)

Cut-Off-Wert: Mindestens 1 Item ist positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	78	58	136
	negativ	7	81	88
		85	139	224

Cut-Off-Wert: Mindestens 2 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	71	34	105
	negativ	14	105	119
		85	139	224

Cut-Off-Wert: Mindestens 3 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	63	22	85
	negativ	22	117	139
		85	139	224

Cut-Off-Wert: Mindestens 4 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	55	15	70
	negativ	30	124	154
		85	139	224

Cut-Off-Wert: Mindestens 5 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	48	7	55
	negativ	37	132	169
		85	139	224

Cut-Off-Wert: Mindestens 6 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	37	4	41
	negativ	48	135	183
		85	139	224

Cut-Off-Wert: Mindestens 7 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	29	3	32
	negativ	56	136	192
		85	139	224

Cut-Off-Wert: Mindestens 8 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	18	0	18
	negativ	67	139	206
		85	139	224

Cut-Off-Wert: Alle 9 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	7	0	7
	negativ	78	139	217
		85	139	224

Vierfeldertafeln des B-PHQ Auswertungsalgorithmus B für die Subgruppe Alter > 56,20 validiert anhand der Standarddepressionsdiagnosen des DIA-X (N = 224)

Cut-Off-Wert: Mindestens 1 Item ist positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	23	84	107
	negativ	4	113	117
		27	197	224

Cut-Off-Wert: Mindestens 2 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	22	39	61
	negativ	5	158	163
		27	197	224

Cut-Off-Wert: Mindestens 3 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	18	24	42
	negativ	9	173	182
		27	197	224

Cut-Off-Wert: Mindestens 4 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	16	12	28
	negativ	11	185	196
		27	197	224

Cut-Off-Wert: Mindestens 5 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	14	7	21
	negativ	13	190	203
		27	197	224

Cut-Off-Wert: Mindestens 6 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	9	5	14
	negativ	18	192	210
		27	197	224

Cut-Off-Wert: Mindestens 7 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	7	2	9
	negativ	20	195	215
		27	197	224

Cut-Off-Wert: Mindestens 8 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	5	1	6
	negativ	22	196	218
		27	197	224

Cut-Off-Wert: Alle 9 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	2	1	3
	negativ	25	196	221
		27	197	224

Vierfeldertafeln des B-PHQ (Auswertungsalgorithmus B) für die Subgruppe Alter > 56,20 validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 224)

Cut-Off-Wert: Mindestens 1 Item ist positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	40	67	107
	negativ	11	106	117
		51	173	224

Cut-Off-Wert: Mindestens 2 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	36	25	61
	negativ	15	148	163
		51	173	224

Cut-Off-Wert: Mindestens 3 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	25	17	42
	negativ	26	156	182
		51	173	224

Cut-Off-Wert: Mindestens 4 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	19	9	28
	negativ	32	164	196
		51	173	224

Cut-Off-Wert: Mindestens 5 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	16	5	21
	negativ	35	168	203
		51	173	224

Cut-Off-Wert: Mindestens 6 Items sind positiv

		CIDI positiv	CIDI negativ	
B-PHQ (Algorithmus B)	positiv	11	3	14
	negativ	40	170	210
		51	173	224

Cut-Off-Wert: Mindestens 7 Items sind positiv

		CIDI		
		positiv	negativ	
B-PHQ (Algorithmus B)	positiv	8	1	9
	negativ	43	172	215
		51	173	224

Cut-Off-Wert: Mindestens 8 Items sind positiv

		CIDI		
		positiv	negativ	
B-PHQ (Algorithmus B)	positiv	5	1	6
	negativ	46	172	218
		51	173	224

Cut-Off-Wert: Alle 9 Items sind positiv

		CIDI		
		positiv	negativ	
B-PHQ (Algorithmus B)	positiv	2	1	3
	negativ	49	172	221
		51	173	224

Anhang C

Weitere Kenngrößen der Validität: Rate der falsch Positiven, Rate der falsch Negativen, Misclassification Rate und Odds-Ratio

Rate der falsch Positiven, Rate der falsch Negativen, Misclassification Rate und Odds-Ratio des GHQ-12 für die gesamte Stichprobe

Tabelle C-1: Rate der falsch Positiven, Rate der falsch Negativen, Misclassification Rate und Odds-Ratio des GHQ-12 der gesamten Stichprobe validiert an den Standarddepressionsdiagnosen des DIA-X (N = 448).

Cut-Off-Wert	**0/1**	**1/2**	**2/3**	**3/4**	**4/5**	**5/6**	**6/7**	**7/8**	**8/9**	**9/10**	**10/11**	**11/12**
Rate der falsch Positiven	0.47	0.37	0.30	0.27	0.20	0.15	0.12	0.09	0.08	0.05	0.04	0.03
Rate der falsch Negativen	0.11	0.15	0.21	0.23	0.27	0.28	0.33	0.37	0.43	0.51	0.59	0.71
Misclassification Rate	0.40	0.33	0.28	0.26	0.21	0.17	0.16	0.14	0.14	0.14	0.14	0.15
Odds-Ratio	9.15	9.87	9.13	9.20	10.95	14.82	14.53	16.93	16.21	16.48	16.58	14.73

Tabelle C-2: Rate der falsch Positiven, Rate der falsch Negativen, Misclassification Rate und Odds-Ratio des GHQ-12 der gesamten Stichprobe validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 448).

Cut-Off-Wert	**0/1**	**1/2**	**2/3**	**3/4**	**4/5**	**5/6**	**6/7**	**7/8**	**8/9**	**9/10**	**10/11**	**11/12**
Rate der falsch Positiven	0.41	0.31	0.23	0.21	0.15	0.11	0.08	0.06	0.05	0.04	0.02	0.02
Rate der falsch Negativen	0.14	0.19	0.26	0.31	0.38	0.42	0.46	0.50	0.56	0.64	0.69	0.79
Misclassification Rate	0.33	0.27	0.24	0.24	0.22	0.20	0.20	0.19	0.20	0.22	0.23	0.25
Odds-Ratio	8.85	9.52	9.62	8.34	9.17	11.33	13.13	16.33	15.63	15.41	19.47	13.22

Rate der falsch Positiven, Rate der falsch Negativen, Misclassification Rate und Odds-Ratio des GHQ-12 für die Subgruppe der Frauen

Tabelle C-3: Rate der falsch Positiven, Rate der falsch Negativen, Misclassification Rate und Odds-Ratio des GHQ-12 für die Subgruppe der Frauen validiert an den Standarddepressionsdiagnosen des DIA-X (N = 283).

Cut-Off-Wert	**0/1**	**1/2**	**2/3**	**3/4**	**4/5**	**5/6**	**6/7**	**7/8**	**8/9**	**9/10**	**10/11**	**11/12**
Rate der falsch Positiven	0.47	0.37	0.28	0.26	0.18	0.12	0.10	0.08	0.08	0.05	0.04	0.02
Rate der falsch Negativen	0.11	0.12	0.19	0.23	0.26	0.28	0.33	0.37	0.42	0.51	0.58	0.72
Misclassi-fication Rate	0.40	0.32	0.27	0.25	0.20	0.16	0.15	0.14	0.15	0.14	0.15	0.16
Odds-Ratio	9.45	12.31	10.59	9.80	12.63	18.12	17.65	18.68	15.89	17.22	15.71	17.25

Tabelle C-4: Rate der falsch Positiven, Rate der falsch Negativen, Misclassification Rate und Odds-Ratio des GHQ-12 für die Subgruppe der Frauen validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 283).

Cut-Off-Wert	**0/1**	**1/2**	**2/3**	**3/4**	**4/5**	**5/6**	**6/7**	**7/8**	**8/9**	**9/10**	**10/11**	**11/12**
Rate der falsch Positiven	0.42	0.31	0.22	0.20	0.14	0.09	0.07	0.06	0.06	0.04	0.03	0.03
Rate der falsch Negativen	0.13	0.17	0.23	0.29	0.37	0.40	0.45	0.49	0.54	0.63	0.68	0.82
Misclassi-fication Rate	0.33	0.27	0.22	0.23	0.21	0.18	0.18	0.19	0.20	0.22	0.23	0.27
Odds-Ratio	9.61	10.62	11.92	9.67	10.31	15.64	17.33	17.21	14.31	13.67	15.03	8.61

Rate der falsch Positiven, Rate der falsch Negativen, Misclassification Rate und Odds-Ratio des GHQ-12 für die Subgruppe der Männer

Tabelle C-5: Rate der falsch Positiven, Rate der falsch Negativen, Misclassification Rate und Odds-Ratio des GHQ-12 für die Subgruppe der Männer validiert an den Standarddepressionsdiagnosen des DIA-X (N = 165).

Cut-Off-Wert	**0/1**	**1/2**	**2/3**	**3/4**	**4/5**	**5/6**	**6/7**	**7/8**	**8/9**	**9/10**	**10/11**	**11/12**
Rate der falsch Positiven	0.46	0.38	0.31	0.28	0.23	0.19	0.16	0.11	0.07	0.06	0.04	0.04
Rate der falsch Negativen	0.12	0.20	0.24	0.24	0.28	0.28	0.32	0.36	0.44	0.52	0.60	0.68
Misclassi-fication Rate	0.41	0.35	0.30	0.27	0.24	0.20	0.18	0.15	0.13	0.13	0.12	0.13
Odds-Ratio	8.46	6.57	6.91	8.20	8.68	11.27	11.40	14.81	16.55	15.23	18.00	12.71

Tabelle C-6: Rate der falsch Positiven, Rate der falsch Negativen, Misclassification Rate und Odds-Ratio des GHQ-12 für die Subgruppe der Männer validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 165).

Cut-Off-Wert	**0/1**	**1/2**	**2/3**	**3/4**	**4/5**	**5/6**	**6/7**	**7/8**	**8/9**	**9/10**	**10/11**	**11/12**
Rate der falsch Positiven	0.40	0.30	0.25	0.22	0.17	0.15	0.11	0.06	0.03	0.03	0.01	0.01
Rate der falsch Negativen	0.16	0.22	0.31	0.35	0.39	0.45	0.47	0.51	0.59	0.65	0.71	0.76
Misclassi-fication Rate	0.33	0.28	0.27	0.26	0.24	0.24	0.22	0.19	0.20	0.21	0.22	0.23
Odds-Ratio	7.80	7.99	6.80	6.52	7.58	7.15	8.96	14.95	19.31	20.01	46.00	37.30

Rate der falsch Positiven, Rate der falsch Negativen, Misclassification Rate und Odds-Ratio des GHQ-12 für die Subgruppe Alter < 56,21 Jahre

Tabelle C-7: Rate der falsch Positiven, Rate der falsch Negativen, Misclassification Rate und Odds-Ratio des GHQ-12 für die Subgruppe Alter < 56,21 Jahre validiert an den Standarddepressionsdiagnosen des DIA-X (N = 224).

Cut-Off-Wert	**0/1**	**1/2**	**2/3**	**3/4**	**4/5**	**5/6**	**6/7**	**7/8**	**8/9**	**9/10**	**10/11**	**11/12**
Rate der falsch Positiven	0.58	0.47	0.36	0.34	0.24	0.17	0.14	0.10	0.09	0.07	0.05	0.03
Rate der falsch Negativen	0.07	0.09	0.15	0.16	0.22	0.22	0.25	0.27	0.36	0.45	0.55	0.69
Misclassi-fication Rate	0.46	0.38	0.31	0.29	0.24	0.18	0.17	0.14	0.16	0.17	0.17	0.19
Odds-Ratio	9.24	11.39	10.40	10.04	11.19	17.30	18.59	23.84	17.97	15.70	16.77	14.67

Tabelle C-8: Rate der falsch Positiven, Rate der falsch Negativen, Misclassification Rate und Odds-Ratio des GHQ-12 für die Subgruppe Alter < 56,21 Jahre validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 224).

Cut-Off-Wert	**0/1**	**1/2**	**2/3**	**3/4**	**4/5**	**5/6**	**6/7**	**7/8**	**8/9**	**9/10**	**10/11**	**11/12**
Rate der falsch Positiven	0.52	0.39	0.29	0.26	0.18	0.12	0.08	0.05	0.05	0.04	0.01	0.01
Rate der falsch Negativen	0.09	0.12	0.20	0.21	0.31	0.35	0.38	0.41	0.49	0.56	0.62	0.75
Misclassi-fication Rate	0.36	0.29	0.25	0.24	0.23	0.21	0.19	0.19	0.22	0.24	0.24	0.29
Odds-Ratio	8.96	11.81	9.90	10.65	10.35	13.16	19.27	26.94	19.31	20.66	83.32	45.28

Rate der falsch Positiven, Rate der falsch Negativen, Misclassification Rate und Odds-Ratio des GHQ-12 für die Subgruppe Alter > 56,20 Jahre

Tabelle C-9: Rate der falsch Positiven, Rate der falsch Negativen, Misclassification Rate und Odds-Ratio des GHQ-12 für die Subgruppe Alter > 56,20 Jahre validiert an den Standarddepressionsdiagnosen des DIA-X (N = 224).

Cut-Off-Wert	**0/1**	**1/2**	**2/3**	**3/4**	**4/5**	**5/6**	**6/7**	**7/8**	**8/9**	**9/10**	**10/11**	**11/12**
Rate der falsch Positiven	0.38	0.29	0.24	0.20	0.16	0.13	0.11	0.09	0.07	0.04	0.04	0.03
Rate der falsch Negativen	0.19	0.26	0.33	0.37	0.37	0.41	0.48	0.56	0.56	0.63	0.67	0.74
Misclassi-fication Rate	0.35	0.29	0.25	0.22	0.19	0.16	0.16	0.14	0.13	0.11	0.11	0.11
Odds-Ratio	7.31	7.02	6.38	6.67	8.77	10.01	8.57	8.47	11.32	13.90	13.57	13.44

Tabelle C-10: Rate der falsch Positiven, Rate der falsch Negativen, Misclassification Rate und Odds-Ratio des GHQ-12 für die Subgruppe Alter > 56,20 Jahre validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 224).

Cut-Off-Wert	**0/1**	**1/2**	**2/3**	**3/4**	**4/5**	**5/6**	**6/7**	**7/8**	**8/9**	**9/10**	**10/11**	**11/12**
Rate der falsch Positiven	0.32	0.24	0.18	0.17	0.13	0.10	0.09	0.06	0.05	0.03	0.03	0.03
Rate der falsch Negativen	0.22	0.31	0.35	0.47	0.49	0.53	0.59	0.65	0.67	0.76	0.80	0.86
Misclassi-fication Rate	0.30	0.26	0.22	0.24	0.21	0.20	0.20	0.20	0.19	0.20	0.21	0.22
Odds-Ratio	7.60	6.82	8.08	5.36	6.78	8.16	7.37	8.03	10.31	8.56	6.79	5.35

Rate der falsch Positiven, Rate der falsch Negativen, Misclassification Rate und Odds-Ratio des B-PHQ Auswertungsalgorithmus A für die gesamte Stichprobe

Tabelle C-11: Rate der falsch Positiven, Rate der falsch Negativen, Misclassification Rate und Odds-Ratio des B-PHQ Auswertungsalgorithmus A der gesamten Stichprobe validiert an den Standarddepressionsdiagnosen des DIA-X (N = 448).

Cut-Off-Wert	mindestens das Item 1a oder 1b ist positiv	mindestens 2 Items sind positiv darunter 1a oder 1b	mindestens 3 Items sind positiv darunter 1a oder 1b	mindestens 4 Items sind positiv darunter 1a oder 1b	mindestens 5 Items sind positiv darunter 1a oder 1b	mindestens 6 Items sind positiv darunter 1a oder 1b	mindestens 7 Items sind positiv darunter 1a oder 1b	mindestens 8 Items sind positiv darunter 1a oder 1b	alle 9 Items sind positiv
Rate der falsch Positiven	0.17	0.14	0.11	0.09	0.07	0.05	0.03	0.01	0.01
Rate der falsch Negativen	0.21	0.21	0.23	0.29	0.39	0.55	0.65	0.77	0.91
Misclassification Rate	0.18	0.16	0.14	0.13	0.13	0.14	0.15	0.15	0.17
Odds-Ratio	18.39	22.58	25.58	23.60	22.27	15.90	16.14	21.77	11.29

Tabelle C-12: Rate der falsch Positiven, Rate der falsch Negativen, Misclassification Rate und Odds-Ratio des B-PHQ Auswertungsalgorithmus A der gesamten Stichprobe validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 448).

Cut-Off-Wert	mindestens das Item 1a oder 1b ist positiv	mindestens 2 Items sind positiv darunter 1a oder 1b	mindestens 3 Items sind positiv darunter 1a oder 1b	mindestens 4 Items sind positiv darunter 1a oder 1b	mindestens 5 Items sind positiv darunter 1a oder 1b	mindestens 6 Items sind positiv darunter 1a oder 1b	mindestens 7 Items sind positiv darunter 1a oder 1b	mindestens 8 Items sind positiv darunter 1a oder 1b	alle 9 Items sind positiv
Rate der falsch Positiven	0.13	0.10	0.08	0.06	0.04	0.02	0.01	0.00	0.00
Rate der falsch Negativen	0.35	0.35	0.40	0.46	0.54	0.65	0.73	0.83	0.93
Misclassification Rate	0.20	0.17	0.18	0.18	0.19	0.21	0.23	0.25	0.29
Odds-Ratio	12.47	17.23	17.67	17.87	23.62	23.77	28.78	63.30	22.04

Rate der falsch Positiven, Rate der falsch Negativen, Misclassification Rate und Odds-Ratio des B-PHQ Auswertungsalgorithmus A für die Subgruppe der Frauen

Tabelle C-13: Rate der falsch Positiven, Rate der falsch Negativen, Misclassification Rate und Odds-Ratio des B-PHQ Auswertungsalgorithmus A für die Subgruppe der Frauen validiert an den Standarddepressionsdiagnosen des DIA-X (N = 283).

Cut-Off-Wert	mindestens das Item 1a oder 1b ist positiv	mindestens 2 Items sind positiv darunter 1a oder 1b	mindestens 3 Items sind positiv darunter 1a oder 1b	mindestens 4 Items sind positiv darunter 1a oder 1b	mindestens 5 Items sind positiv darunter 1a oder 1b	mindestens 6 Items sind positiv darunter 1a oder 1b	mindestens 7 Items sind positiv darunter 1a oder 1b	mindestens 8 Items sind positiv darunter 1a oder 1b	alle 9 Items sind positiv
Rate der falsch Positiven	0.15	0.12	0.11	0.08	0.06	0.04	0.03	0.01	0.00
Rate der falsch Negativen	0.23	0.23	0.26	0.32	0.40	0.54	0.65	0.74	0.91
Misclassification Rate	0.16	0.14	0.14	0.13	0.13	0.14	0.16	0.16	0.19
Odds-Ratio	19.79	23.93	23.57	23.61	24.22	20.22	16.91	40.00	21.63

Tabelle C-14: Rate der falsch Positiven, Rate der falsch Negativen, Misclassification Rate und Odds-Ratio des B-PHQ Auswertungsalgorithmus A für die Subgruppe der Frauen validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 283).

Cut-Off-Wert	mindestens das Item 1a oder 1b ist positiv	mindestens 2 Items sind positiv darunter 1a oder 1b	mindestens 3 Items sind positiv darunter 1a oder 1b	mindestens 4 Items sind positiv darunter 1a oder 1b	mindestens 5 Items sind positiv darunter 1a oder 1b	mindestens 6 Items sind positiv darunter 1a oder 1b	mindestens 7 Items sind positiv darunter 1a oder 1b	mindestens 8 Items sind positiv darunter 1a oder 1b	alle 9 Items sind positiv
Rate der falsch Positiven	0.11	0.09	0.07	0.06	0.03	0.02	0.02	0.01	0.01
Rate der falsch Negativen	0.37	0.37	0.40	0.46	0.52	0.63	0.72	0.82	0.94
Misclassification Rate	0.19	0.17	0.17	0.18	0.18	0.20	0.23	0.25	0.29
Odds-Ratio	13.59	18.10	19.31	19.76	35.65	37.43	24.51	43.94	11.89

Rate der falsch Positiven, Rate der falsch Negativen, Misclassification Rate und Odds-Ratio des B-PHQ Auswertungsalgorithmus A für die Subgruppe der Männer

Tabelle C-15: Rate der falsch Positiven, Rate der falsch Negativen, Misclassification Rate und Odds-Ratio des B-PHQ Auswertungsalgorithmus A für die Subgruppe der Männer validiert an den Standarddepressionsdiagnosen des DIA-X (N = 165).

Cut-Off-Wert	mindestens das Item 1a oder 1b ist positiv	mindestens 2 Items sind positiv darunter 1a oder 1b	mindestens 3 Items sind positiv darunter 1a oder 1b	mindestens 4 Items sind positiv darunter 1a oder 1b	mindestens 5 Items sind positiv darunter 1a oder 1b	mindestens 6 Items sind positiv darunter 1a oder 1b	mindestens 7 Items sind positiv darunter 1a oder 1b	mindestens 8 Items sind positiv darunter 1a oder 1b	alle 9 Items sind positiv
Rate der falsch Positiven	0.21	0.18	0.13	0.11	0.08	0.06	0.04	0.02	0.01
Rate der falsch Negativen	0.16	0.16	0.16	0.24	0.36	0.56	0.64	0.84	0.92
Misclassification Rate	0.21	0.18	0.13	0.13	0.12	0.14	0.13	0.15	0.15
Odds-Ratio	19.25	24.15	35.58	26.39	20.85	11.44	15.19	8.70	6.00

Tabelle C-16: Rate der falsch Positiven, Rate der falsch Negativen, Misclassification Rate und Odds-Ratio des B-PHQ Auswertungsalgorithmus A für die Subgruppe der Männer validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 165).

Cut-Off-Wert	mindestens das Item 1a oder 1b ist positiv	mindestens 2 Items sind positiv darunter 1a oder 1b	mindestens 3 Items sind positiv darunter 1a oder 1b	mindestens 4 Items sind positiv darunter 1a oder 1b	mindestens 5 Items sind positiv darunter 1a oder 1b	mindestens 6 Items sind positiv darunter 1a oder 1b	mindestens 7 Items sind positiv darunter 1a oder 1b	mindestens 8 Items sind positiv darunter 1a oder 1b	alle 9 Items sind positiv
Rate der falsch Positiven	0.16	0.11	0.09	0.07	0.05	0.03	0.01	0.00	0.00
Rate der falsch Negativen	0.33	0.33	0.41	0.47	0.57	0.67	0.73	0.86	0.92
Misclassification Rate	0.21	0.18	0.18	0.19	0.21	0.22	0.22	0.25	0.27
Odds-Ratio	11.23	16.34	15.37	15.26	13.75	13.58	41.53	∞	∞

Rate der falsch Positiven, Rate der falsch Negativen, Misclassification Rate und Odds-Ratio des B-PHQ Auswertungsalgorithmus A für die Subgruppe Alter < 56,21 Jahre

Tabelle C-17: Rate der falsch Positiven, Rate der falsch Negativen, Misclassification Rate, Odds-Ratio des B-PHQ Auswertungsalgorithmus A für die Subgruppe Alter < 56,21 Jahre validiert an den Standarddepressionsdiagnosen des DIA-X (N = 224).

Cut-Off-Wert	mindestens das Item 1a oder 1b ist positiv	mindestens 2 Items sind positiv darunter 1a oder 1b	mindestens 3 Items sind positiv darunter 1a oder 1b	mindestens 4 Items sind positiv darunter 1a oder 1b	mindestens 5 Items sind positiv darunter 1a oder 1b	mindestens 6 Items sind positiv darunter 1a oder 1b	mindestens 7 Items sind positiv darunter 1a oder 1b	mindestens 8 Items sind positiv darunter 1a oder 1b	alle 9 Items sind positiv
Rate der falsch Positiven	0.24	0.21	0.17	0.14	0.11	0.08	0.06	0.02	0.01
Rate der falsch Negativen	0.13	0.13	0.16	0.24	0.35	0.49	0.60	0.75	0.91
Misclassification Rate	0.21	0.19	0.17	0.17	0.17	0.18	0.19	0.20	0.23
Odds-Ratio	21.41	25.33	24.67	19.52	15.89	12.44	10.60	14.09	8.35

Tabelle C-18: Rate der falsch Positiven, Rate der falsch Negativen, Misclassification Rate und Odds-Ratio des B-PHQ Auswertungsalgorithmus A für die Subgruppe Alter < 56,21 Jahre validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 224).

Cut-Off-Wert	mindestens das Item 1a oder 1b ist positiv	mindestens 2 Items sind positiv darunter 1a oder 1b	mindestens 3 Items sind positiv darunter 1a oder 1b	mindestens 4 Items sind positiv darunter 1a oder 1b	mindestens 5 Items sind positiv darunter 1a oder 1b	mindestens 6 Items sind positiv darunter 1a oder 1b	mindestens 7 Items sind positiv darunter 1a oder 1b	mindestens 8 Items sind positiv darunter 1a oder 1b	alle 9 Items sind positiv
Rate der falsch Positiven	0.18	0.14	0.12	0.09	0.05	0.03	0.02	0.00	0.00
Rate der falsch Negativen	0.25	0.25	0.31	0.36	0.45	0.56	0.66	0.79	0.92
Misclassification Rate	0.21	0.18	0.19	0.19	0.20	0.23	0.26	0.30	0.35
Odds-Ratio	13.90	18.13	17.44	18.44	23.32	26.02	23.48	∞	∞

Rate der falsch Positiven, Rate der falsch Negativen, Misclassification Rate und Odds-Ratio des B-PHQ Auswertungsalgorithmus A für die Subgruppe Alter > 56,20 Jahre

Tabelle C-19: Rate der falsch Positiven, Rate der falsch Negativen, Misclassification Rate, Odds-Ratio des B-PHQ Auswertungsalgorithmus A für die Subgruppe Alter > 56,20 Jahre validiert an den Standarddepressionsdiagnosen des DIA-X (N = 224).

Cut-Off-Wert	mindestens das Item 1a oder 1b ist positiv	mindestens 2 Items sind positiv darunter 1a oder 1b	mindestens 3 Items sind positiv darunter 1a oder 1b	mindestens 4 Items sind positiv darunter 1a oder 1b	mindestens 5 Items sind positiv darunter 1a oder 1b	mindestens 6 Items sind positiv darunter 1a oder 1b	mindestens 7 Items sind positiv darunter 1a oder 1b	mindestens 8 Items sind positiv darunter 1a oder 1b	alle 9 Items sind positiv
Rate der falsch Positiven	0.11	0.09	0.07	0.05	0.03	0.03	0.01	0.01	0.01
Rate der falsch Negativen	0.37	0.37	0.37	0.41	0.48	0.67	0.74	0.81	0.93
Misclassification Rate	0.14	0.12	0.10	0.09	0.08	0.10	0.10	0.10	0.12
Odds-Ratio	13.52	18.00	24.06	27.20	34.28	19.20	34.13	44.55	15.68

Tabelle C-20: Rate der falsch Positiven, Rate der falsch Negativen, Misclassification Rate und Odds-Ratio des B-PHQ Auswertungsalgorithmus A für die Subgruppe Alter > 56,20 Jahre validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 224).

Cut-Off-Wert	mindestens das Item 1a oder 1b ist positiv	mindestens 2 Items sind positiv darunter 1a oder 1b	mindestens 3 Items sind positiv darunter 1a oder 1b	mindestens 4 Items sind positiv darunter 1a oder 1b	mindestens 5 Items sind positiv darunter 1a oder 1b	mindestens 6 Items sind positiv darunter 1a oder 1b	mindestens 7 Items sind positiv darunter 1a oder 1b	mindestens 8 Items sind positiv darunter 1a oder 1b	alle 9 Items sind positiv
Rate der falsch Positiven	0.09	0.06	0.05	0.04	0.02	0.02	0.01	0.01	0.01
Rate der falsch Negativen	0.53	0.53	0.57	0.63	0.69	0.78	0.84	0.90	0.96
Misclassification Rate	0.19	0.17	0.17	0.17	0.17	0.19	0.20	0.21	0.22
Odds-Ratio	9.36	14.49	15.65	14.08	19.31	15.58	32.00	18.70	7.02

Rate der falsch Positiven, Rate der falsch Negativen, Misclassification Rate und Odds-Ratio des B-PHQ Auswertungsalgorithmus B für die gesamte Stichprobe

Tabelle C-21: Rate der falsch Positiven, Rate der falsch Negativen, Misclassification Rate und Odds-Ratio des B-PHQ Auswertungsalgorithmus B der gesamten Stichprobe validiert an den Standarddepressionsdiagnosen des DIA-X (N = 448).

Cut-Off-Wert	mindestens 1 Item ist positiv	mindestens 2 Items sind positiv	mindestens 3 Items sind positiv	mindestens 4 Items sind positiv	mindestens 5 Items sind positiv	mindestens 6 Items sind positiv	mindestens 7 Items sind positiv	mindestens 8 Items sind positiv	alle 9 Items sind positiv
Rate der falsch Positiven	0.46	0.26	0.17	0.11	0.07	0.05	0.03	0.01	0.01
Rate der falsch Negativen	0.09	0.13	0.22	0.29	0.39	0.55	0.65	0.77	0.91
Misclassification Rate	0.39	0.24	0.18	0.14	0.13	0.14	0.15	0.15	0.17
Odds-Ratio	12.63	18.41	17.10	19.70	20.43	15.90	16.14	21.77	11.29

Tabelle C-22: Rate der falsch Positiven, Rate der falsch Negativen, Misclassification Rate und Odds-Ratio Auswertungsalgorithmus B der gesamten Stichprobe validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 448).

Cut-Off-Wert	mindestens 1 Item ist positiv	mindestens 2 Items sind positiv	mindestens 3 Items sind positiv	mindestens 4 Items sind positiv	mindestens 5 Items sind positiv	mindestens 6 Items sind positiv	mindestens 7 Items sind positiv	mindestens 8 Items sind positiv	alle 9 Items sind positiv
Rate der falsch Positiven	0.40	0.19	0.13	0.08	0.04	0.02	0.01	0.00	0.00
Rate der falsch Negativen	0.13	0.21	0.35	0.46	0.53	0.65	0.73	0.83	0.93
Misclassification Rate	0.32	0.20	0.19	0.19	0.19	0.21	0.23	0.25	0.29
Odds-Ratio	9.81	15.82	12.83	14.32	22.22	23.77	28.78	63.30	22.04

Rate der falsch Positiven, Rate der falsch Negativen, Misclassification Rate und Odds-Ratio des B-PHQ Auswertungsalgorithmus B für die Subgruppe der Frauen

Tabelle C-23: Rate der falsch Positiven, Rate der falsch Negativen, Misclassification Rate und Odds-Ratio des B-PHQ Auswertungsalgorithmus B für die Subgruppe der Frauen validiert an den Standarddepressionsdiagnosen des DIA-X (N = 283).

Cut-Off-Wert	mindestens 1 Item ist positiv	mindestens 2 Items sind positiv	mindestens 3 Items sind positiv	mindestens 4 Items sind positiv	mindestens 5 Items sind positiv	mindestens 6 Items sind positiv	mindestens 7 Items sind positiv	mindestens 8 Items sind positiv	alle 9 Items sind positiv
Rate der falsch Positiven	0.46	0.25	0.16	0.09	0.06	0.04	0.03	0.01	0.00
Rate der falsch Negativen	0.07	0.14	0.25	0.32	0.40	0.54	0.65	0.74	0.91
Misclassification Rate	0.38	0.23	0.18	0.14	0.13	0.14	0.16	0.16	0.19
Odds-Ratio	15.82	18.59	15.69	21.15	24.22	20.22	16.91	40.00	21.63

Tabelle C-24: Rate der falsch Positiven, Rate der falsch Negativen, Misclassification Rate und Odds-Ratio des B-PHQ Auswertungsalgorithmus B für die Subgruppe der Frauen validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 283).

Cut-Off-Wert	mindestens 1 Item ist positiv	mindestens 2 Items sind positiv	mindestens 3 Items sind positiv	mindestens 4 Items sind positiv	mindestens 5 Items sind positiv	mindestens 6 Items sind positiv	mindestens 7 Items sind positiv	mindestens 8 Items sind positiv	alle 9 Items sind positiv
Rate der falsch Positiven	0.40	0.18	0.12	0.07	0.03	0.02	0.02	0.01	0.01
Rate der falsch Negativen	0.11	0.21	0.34	0.46	0.52	0.63	0.72	0.82	0.94
Misclassification Rate	0.31	0.19	0.19	0.19	0.18	0.20	0.23	0.25	0.29
Odds-Ratio	11.40	17.04	14.29	16.54	35.65	37.43	24.51	43.94	11.89

Rate der falsch Positiven, Rate der falsch Negativen, Misclassification Rate und Odds-Ratio des B-PHQ Auswertungsalgorithmus B für die Subgruppe der Männer

Tabelle C-25: Rate der falsch Positiven, Rate der falsch Negativen, Misclassification Rate und Odds-Ratio des B-PHQ Auswertungsalgorithmus B für die Subgruppe der Männer validiert an den Standarddepressionsdiagnosen des DIA-X (N = 165).

Cut-Off-Wert	mindestens 1 Item ist positiv	mindestens 2 Items sind positiv	mindestens 3 Items sind positiv	mindestens 4 Items sind positiv	mindestens 5 Items sind positiv	mindestens 6 Items sind positiv	mindestens 7 Items sind positiv	mindestens 8 Items sind positiv	alle 9 Items sind positiv
Rate der falsch Positiven	0.46	0.28	0.19	0.14	0.09	0.06	0.04	0.02	0.01
Rate der falsch Negativen	0.12	0.12	0.16	0.24	0.36	0.56	0.64	0.84	0.92
Misclassification Rate	0.41	0.25	0.18	0.15	0.13	0.14	0.13	0.15	0.15
Odds-Ratio	8.46	18.99	23.02	20.17	17.37	11.44	15.19	8.70	6.00

Tabelle C-26: Rate der falsch Positiven, Rate der falsch Negativen, Misclassification Rate und Odds-Ratio des B-PHQ Auswertungsalgorithmus B für die Subgruppe der Männer validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 165).

Cut-Off-Wert	mindestens 1 Item ist positiv	mindestens 2 Items sind positiv	mindestens 3 Items sind positiv	mindestens 4 Items sind positiv	mindestens 5 Items sind positiv	mindestens 6 Items sind positiv	mindestens 7 Items sind positiv	mindestens 8 Items sind positiv	alle 9 Items sind positiv
Rate der falsch Positiven	0.40	0.20	0.14	0.09	0.06	0.03	0.01	0.00	0.00
Rate der falsch Negativen	0.16	0.22	0.37	0.45	0.55	0.67	0.73	0.86	0.92
Misclassification Rate	0.33	0.21	0.21	0.20	0.21	0.22	0.22	0.25	0.27
Odds-Ratio	7.80	13.97	10.76	11.71	12.69	13.58	41.53	∞	∞

Rate der falsch Positiven, Rate der falsch Negativen, Misclassification Rate und Odds-Ratio des B-PHQ Auswertungsalgorithmus B für die Subgruppe Alter < 56,21 Jahre

Tabelle C-27: Rate der falsch Positiven, Rate der falsch Negativen, Misclassification Rate und Odds-Ratio des B-PHQ Auswertungsalgorithmus B für die Subgruppe Variable Alter < 56,21 Jahre validiert an den Standarddepressionsdiagnosen des DIA-X (N = 224).

Cut-Off-Wert	mindestens 1 Item ist positiv	mindestens 2 Items sind positiv	mindestens 3 Items sind positiv	mindestens 4 Items sind positiv	mindestens 5 Items sind positiv	mindestens 6 Items sind positiv	mindestens 7 Items sind positiv	mindestens 8 Items sind positiv	alle 9 Items sind positiv
Rate der falsch Positiven	0.50	0.33	0.23	0.17	0.11	0.08	0.06	0.02	0.01
Rate der falsch Negativen	0.05	0.11	0.16	0.24	0.35	0.49	0.60	0.75	0.91
Misclassification Rate	0.39	0.28	0.21	0.18	0.17	0.18	0.19	0.20	0.23
Odds-Ratio	17.54	16.48	17.04	16.27	14.96	12.44	10.60	14.09	8.35

Tabelle C-28: Rate der falsch Positiven, Rate der falsch Negativen, Misclassification Rate und Odds-Ratio des B-PHQ Auswertungsalgorithmus B für die Subgruppe Variable Alter < 56,21 Jahre validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 224).

Cut-Off-Wert	mindestens 1 Item ist positiv	mindestens 2 Items sind positiv	mindestens 3 Items sind positiv	mindestens 4 Items sind positiv	mindestens 5 Items sind positiv	mindestens 6 Items sind positiv	mindestens 7 Items sind positiv	mindestens 8 Items sind positiv	alle 9 Items sind positiv
Rate der falsch Positiven	0.42	0.24	0.16	0.11	0.05	0.03	0.02	0.00	0.00
Rate der falsch Negativen	0.08	0.16	0.26	0.35	0.44	0.56	0.66	0.79	0.92
Misclassification Rate	0.29	0.21	0.20	0.20	0.20	0.23	0.26	0.30	0.35
Odds-Ratio	15.56	15.66	15.23	15.16	24.46	26.02	23.48	∞	∞

Rate der falsch Positiven, Rate der falsch Negativen, Misclassification Rate und Odds-Ratio des B-PHQ Auswertungsalgorithmus B für die Subgruppe Alter > 56,20 Jahre

Tabelle C-29: Rate der falsch Positiven, Rate der falsch Negativen, Misclassification Rate und Odds-Ratio des B-PHQ Auswertungsalgorithmus B für die Subgruppe Variable Alter > 56,20 Jahre validiert an den Standarddepressionsdiagnosen des DIA-X (N = 224).

Cut-Off-Wert	mindestens 1 Item ist positiv	mindestens 2 Items sind positiv	mindestens 3 Items sind positiv	mindestens 4 Items sind positiv	mindestens 5 Items sind positiv	mindestens 6 Items sind positiv	mindestens 7 Items sind positiv	mindestens 8 Items sind positiv	alle 9 Items sind positiv
Rate der falsch Positiven	0.43	0.20	0.12	0.06	0.04	0.03	0.01	0.01	0.01
Rate der falsch Negativen	0.15	0.19	0.33	0.41	0.48	0.67	0.74	0.81	0.93
Misclassification Rate	0.39	0.20	0.15	0.10	0.09	0.10	0.10	0.10	0.12
Odds-Ratio	7.74	17.83	14.42	22.42	29.23	19.20	34.13	44.55	15.68

Tabelle C-30: Rate der falsch Positiven, Rate der falsch Negativen, Misclassification Rate und Odds-Ratio des B-PHQ Auswertungsalgorithmus B für die Subgruppe Alter > 56,20 Jahre validiert anhand der Standarddepressionsdiagnosen des DIA-X einschließlich der Diagnose unterschwelliges depressives Syndrom (N = 224).

Cut-Off-Wert	mindestens 1 Item ist positiv	mindestens 2 Items sind positiv	mindestens 3 Items sind positiv	mindestens 4 Items sind positiv	mindestens 5 Items sind positiv	mindestens 6 Items sind positiv	mindestens 7 Items sind positiv	mindestens 8 Items sind positiv	alle 9 Items sind positiv
Rate der falsch Positiven	0.39	0.14	0.10	0.05	0.03	0.02	0.01	0.01	0.01
Rate der falsch Negativen	0.22	0.29	0.51	0.63	0.69	0.78	0.84	0.90	0.96
Misclassification Rate	0.35	0.18	0.19	0.18	0.18	0.19	0.20	0.21	0.22
Odds-Ratio	5.75	14.21	8.82	10.82	15.36	15.58	32.00	18.70	7.02

Curriculum Vitae

Simone Braun	geboren am 02. April 1973 in Heilbronn verheiratet

Promotion

seit 01/00	Promotion an der Klinik für Psychiatrie und Psychotherapie der Ludwig-Maximilians-Universität München, Ärztlicher Direktor: Prof. Dr. H.-J. Möller Abteilungsleiter und Betreuer: Prof. Dr. U. Hegerl Thema der Dissertation: Erfassung depressiver Störungen in der Primärversorgung: Validität des Brief Patient Health Questionnaire (B-PHQ) und des General Health Questionnaire (GHQ-12)
seit 07/01	Stipendium im Rahmen des Hochschul- und Wissenschaftsprogramms
05/01	Promotionsvorprüfung (bestanden; unbenotet)

Studium

10/99	*Diplomprüfung (Gesamtnote 1,1)*
10/94 bis 10/99	Diplom-Psychologie
10/96 bis 10/99	Hauptstudium an der Katholischen Universität Eichstätt Vertiefungsfach: Umweltpsychologie Nachbarfach: Erwachsenenbildung Thema der Diplomarbeit: Der Irrelevant-Speech-Effect: Eine experimentelle Untersuchung zum Einfluß der drei Faktoren Entfernung, Lautstärke und Sprache (Note 1,0)
10/94 bis 09/96	Grundstudium an der Universität Koblenz Landau Diplomvorprüfung (Note 1,5)

Praktika

09/97 bis 10/97	Wissenschaftliches Forschungspraktikum am Humanwissenschaftlichen Filmarchiv der Max-Planck-Gesellschaft und Humanwissenschaftlichen Zentrum der Ludwig-Maximilians-Universität München, Andechs
08/97 bis 09/97	Klinisches Praktikum im Sozialpsychiatrischen Dienst, Caritas Kreisstelle Eichstätt
07/94 bis 10/94	Wissenschaftliches Forschungspraktikum an der Forschungsstelle für Humanethologie in der Max-Planck-Gesellschaft, Andechs

Schulbildung

09/84 bis 07/94	Gymnasium Weilheim i.OB mathematisch-naturwissenschaftlicher Zweig Abschluss: Allgemeine Hochschulreife (Note 1,5)
08/91 bis 08/92	Auslandsschuljahr: Tauranga Girls' College, Neuseeland

Fremdsprachen	Englisch (fließend in Wort und Schrift) Französisch (Grundkenntnisse)
EDV-Kenntnisse	MS Office, SPSS, SAS