

Psychometrische Überprüfung einer verkürzten Version der „Hamburger Module zur Erfassung allgemeiner Aspekte psychosozialer Gesundheit für die therapeutische Praxis“ (HEALTH-49)

Sven Rabung¹, Timo Harfst², Stephan Kowski¹, Uwe Koch¹, Hans-Ullrich Wittchen³,
Holger Schulz¹

¹ Institut und Poliklinik für Medizinische Psychologie,
Zentrum für Psychosoziale Medizin, Universitätsklinikum Hamburg-Eppendorf

² Bundespsychotherapeutenkammer, Berlin

³ Institut für Klinische Psychologie und Psychotherapie, Technische Universität Dresden

erschienen in:

Zeitschrift für Psychosomatische Medizin und Psychotherapie 55/2009, 162-179

Korrespondenzadresse:

Dr. phil. Sven Rabung, Dipl.-Psych.

Institut und Poliklinik für Medizinische Psychologie

Universitätsklinikum Hamburg Eppendorf

Martinistr. 52, W26

20246 Hamburg

Tel.: 040 / 7410-58138, Fax: 040 / 7410-54940

E-Mail: srabung@uke.uni-hamburg.de

Die vorliegende Untersuchung wurde durch die Gesetzliche Krankenversicherung (QS-Reha[®]-Verfahren, www.qs-reha.de) und eine Forschungszuwendung im Rahmen des DETECT Projekts (www.detect-studie.de) unterstützt.

Zusammenfassung

Fragestellung: Der *HEALTH*-Fragebogen, der in seiner Ursprungsversion 79 Items umfasste, soll der multidimensionalen Erfassung generischer Aspekte von psychosozialer Gesundheit dienen.

Methode: Auf der Grundlage umfangreicher klinischer und gesunder Stichproben (n = 1.548 Patienten aus stationärer psychosomatischer Rehabilitation, n = 5.630 Patienten aus der primärärztlichen Versorgung) wurde er verkürzt und einer umfassenden psychometrischen Überprüfung unterzogen.

Ergebnisse: Der verkürzte Fragebogen umfasst 49 Items („*HEALTH-49*“), die sechs eigenständigen Modulen mit insgesamt neun Skalen zugeordnet sind (Somatoforme Beschwerden, Depressivität, Phobische Ängste, Psychisches Wohlbefinden, Interaktionale Schwierigkeiten, Selbstwirksamkeit, Aktivität und Partizipation, Soziale Unterstützung und Soziale Beeinträchtigung). Er erweist sich als praktikables Instrument. Die dimensionale Struktur und die relative Eigenständigkeit der entwickelten Module lassen sich faktorenanalytisch gut bestätigen. Die einzelnen Skalen zeichnen sich durch hohe Zuverlässigkeit, zufrieden stellende Validität und sehr gute Veränderungssensitivität aus.

Diskussion: Mit dem *HEALTH-49* liegt ein Selbstbeurteilungs-Instrument vor, das die angemessene Berücksichtigung psychosozialer Aspekte in Diagnostik und Therapieplanung wie auch in der Evaluation und Qualitätssicherung psychotherapeutischer wie auch medizinischer Behandlungen im Allgemeinen zugleich umfassend und ökonomisch ermöglicht. Der Fragebogen kann kostenfrei im Internet abgerufen werden (<http://www.hamburger-module.de>).

Schlüsselwörter: Psychosoziale Gesundheit, Psychosoziale Beeinträchtigung, Selbstbeurteilungsinstrument, Testentwicklung, Testgütekriterien

Psychometric analysis of a short form of the “Hamburg Modules for the Assessment of Psychosocial Health” (HEALTH-49)

Abstract

Objectives: The *HEALTH*-questionnaire, originally consisting of 79 items, aims to assess generic aspects of psychosocial health.

Methods: Based on large clinical and healthy samples (n = 1.548 psychotherapy inpatients, n = 5.630 primary care patients), the questionnaire has been shortened and psychometrically analyzed.

Results: The resulting 49-item questionnaire (*HEALTH-49*) comprises six discrete modules with nine scales (somatoform complaints, depressiveness, phobic anxiety, psychological well-being, interactional problems, self-efficacy, activity and participation, social support, and social stress). It proves to be well accepted and feasible under routine conditions. Factor analyses confirm the intended dimensional configuration and the relative independence of modules. The scales reveal high reliability. Evidence of their validity and sensitivity to change are demonstrated.

Conclusions: The *HEALTH-49* is a self-rating instrument that allows for the comprehensive and economic assessment of generic aspects of psychosocial health. It is highly suitable for use in clinical practice. The questionnaire is available as a free download from the internet (<http://www.hamburger-module.de>).

Key words: Psychosocial Health, Psychosocial Disorders, Questionnaires, Psychological Tests, Health Care

Einleitung

Der Einsatz standardisierter Selbstbeurteilungsverfahren hat sich im Bereich der Psychotherapie etabliert, da diese es nicht nur erlauben, benötigte diagnostische Informationen ökonomisch erheben und bewerten zu können, sondern sie – bei wiederholter Erhebung – zugleich zur Evaluation und Qualitätssicherung der durchgeführten Maßnahmen zu nutzen (Schumacher u. Brähler 2005). Bis heute liegen allerdings nur wenige Selbsteinschätzungsinstrumente vor, die geeignet wären, psychosoziale Gesundheit zugleich *umfassend*, also in ihren verschiedenen Dimensionen, *generisch*, also nicht auf einen spezifischen Kontext begrenzt, und zudem *ökonomisch* zu erfassen. Üblicherweise werden deshalb zum Zwecke einer mehrdimensionalen Erfassung psychosozialer Faktoren verschiedene Instrumente zu sog. Testbatterien zusammengestellt (Barkham et al. 1998; Hahlweg et al. 1999; Kraus et al. 2005). Die Kombination mehrerer Instrumente aus verschiedenen Entwicklungskontexten führt jedoch häufig zu Inkonsistenzen (z.B. bezüglich der Itemformulierungen, der Antwortformate usw.) sowie zu Redundanzen (z.B. durch ähnliche Items oder Skalen in verschiedenen Fragebögen). Neben einem uneinheitlichen Erscheinungsbild resultiert häufig ein unnötig umfangreiches Gesamtinstrumentarium, was sich insbesondere im Falle des routinemäßigen Einsatzes in erheblichen zusätzlichen Kosten niederschlagen kann (Druckkosten, Lizenzgebühren usw.). Vor diesem Hintergrund wurden die „*Hamburger Module zur Erfassung allgemeiner Aspekte psychosozialer Gesundheit für die therapeutische Praxis*“ (kurz: *HEALTH*) entwickelt, die zentrale Aspekte psychosozialer Gesundheit zugleich umfassend, generisch und ökonomisch erfassen sollten (Rabung et al. 2007). Der *HEALTH*-Fragebogen sollte einerseits als Gesamtinstrument eingesetzt werden können, der modulare Aufbau sollte andererseits aber auch die Zusammenstellung einzelner Module zu einem verkürzten Erhebungsinstrument ermöglichen. Weitere Vorzüge des neuen Fragebogens liegen im Vergleich zu herkömmlichen Fragebogenbatterien im vereinheitlichten Antwortformat, dem Ausschluss von Redundanzen und Inkonsistenzen, einer stärkeren Fokussierung auf die häufig vernachlässigten Bereiche von Ressourcen, Aktivität und Partizipation sowie – nicht zuletzt – seiner lizenzfreien Anwendbarkeit.

Der *HEALTH*-Fragebogen deckt in seiner Ursprungsversion mit 79 Items folgende Bereiche psychosozialer Gesundheit ab: (a) *Psychische und somatoforme Beschwerden*, (b) *Interaktionelle Beeinträchtigungen*, (c) *Psychisches Wohlbefinden*, (d) *Selbstwirksamkeit*, (e) *Lebensqualität*, (f) *Aktivität und Teilhabe* sowie (g) *Soziale Unterstützung und Belastung*.

Eine erste Überprüfung des Fragebogens belegte die hohe Reliabilität und die überwiegend zufrieden stellende Validität der Skalen (Rabung et al. 2007). Allerdings ergaben sich auch Hinweise auf möglichen Überarbeitungsbedarf: So ließ sich beispielsweise die Skala „*Lebensqualität*“ nicht als eigenständiger Faktor belegen (Hinweis auf eingeschränkte faktorielle Validität), es ergaben sich auffällig erhöhte Fehlwertquoten für die Items der Skala „*Selbstwirksamkeit*“ (Hinweis auf eingeschränkte Praktikabilität) und es zeichneten sich durchgängig Optionen zur Verkürzung der Skalen ab (Hinweis auf optimierbare Ökonomie).

Im Folgenden wird daher die psychometrische Überprüfung der anhand einer breiteren Datenbasis entwickelten Kurzform des *HEALTH*-Fragebogens dargestellt.

Methode

Stichproben

Die Ursprungsversion des *HEALTH*-Fragebogens wurde in zwei groß angelegten Studien an umfangreichen klinischen und gesunden Stichproben untersucht:

Studie 1: Zum einen wurde der Fragebogen im Rahmen der *Follow-up*-Untersuchung zur DETECT-Studie (Pieper et al. 2005; Wittchen et al. 2005) als Teil einer umfangreicheren Fragebogenbatterie eingesetzt. In der DETECT-Studie (DETECT: “Diabetes Cardiovascular Risk-Evaluation: Targets and Essential Data for Commitment of Treatment”; www.detect-studie.de) wurde durch 3.188 bundesweit repräsentativ ausgewählte Hausärzte das Diagnose- und Behandlungsprofil von insgesamt 55.518 unselektierten, konsekutiven Patienten dokumentiert. Eine randomisierte Teilstichprobe dieser Patienten (n=7.519 aus 851 Hausarztpraxen) wurde danach über ein Jahr hinweg klinisch verfolgt und nach 12 Monaten erneut untersucht. Insgesamt 5.630 dieser zufällig ausgewählten Patienten (59,9 % weiblich) aus 798 Hausarztpraxen haben den *HEALTH*-

Fragebogen bearbeitet, als sie zur *Follow-up*-Untersuchung in die Praxis ihres Hausarztes einbestellt worden waren. Die Patienten waren zum Untersuchungs-Zeitpunkt (September bis Dezember 2004) im Mittel 58,6 Jahre alt (SD = 13,9; Range: 19-96 Jahre). 38,9 % waren aktuell berufstätig, 47,4 % bereits berentet. 68,9 % waren verheiratet. In Bezug auf die vier in der DETECT-Studie fokussierten „Kerndiagnosen“ fanden sich am häufigsten Hypertonien (46,3 %), gefolgt von Hyperlipidaemien (42,3 %) und seltener Diabetes mellitus (19,4 %) sowie Koronare Herzerkrankungen (13,0 %). Insgesamt wurde ein Großteil (44,6 %) der Patienten von den beteiligten Hausärzten als gesund beschrieben. Lediglich bei 16,4 % der Patienten wurde die körperliche respektive psychische Morbidität als deutlich oder schwerst auffällig eingestuft.

Studie 2: Zum zweiten konnte der *HEALTH*-Fragebogen im Rahmen des Qualitätssicherungsprogramms der Gesetzlichen Krankenversicherung in der medizinischen Rehabilitation im Indikationsbereich psychische und psychosomatische Erkrankungen eingesetzt werden (*QS-Reha*[®]-Verfahren; www.qs-reha.de). Im Untersuchungszeitraum (Mai 2005 bis Dezember 2006) wurden in 12 Fachkliniken für psychosomatische Medizin und Psychotherapie konsekutive Patientenstichproben jeweils zu Beginn und am Ende der stationären Behandlung sowie in einer 6-Monats-Katamnese untersucht. Insgesamt konnten Aufnahmedaten von 1.548 Patienten erhoben werden. Zu 1.357 von ihnen lagen außerdem verwertbare Daten zum Entlassungszeitpunkt vor.

Die *QS-Reha*[®]-Patienten (67 % weiblich) waren zu Beginn der stationären Rehabilitation im Mittel 48,3 Jahre alt (SD = 13,0; Range: 16-85 Jahre). 50,4 % waren aktuell berufstätig, 12,3 % arbeitslos und 4,5 % berentet. 64,0 % lebten in einer festen Partnerschaft. Als häufigste primäre Behandlungsdiagnosen wurden Depressive Störungen (ICD-10: F32-33; F34.1; 36,4 %), Anpassungsstörungen und Neurasthenien (F43.2-9; F48; 18,7 %), Angststörungen (F40-41; 9,6 %) sowie Somatoforme Störungen (F45; 8,1 %) dokumentiert. Bei 11,1 % der Patienten wurde zusätzlich eine Persönlichkeitsstörung (F60) diagnostiziert. Durchschnittlich wiesen die Patienten insgesamt 2,9 komorbide Behandlungsdiagnosen nach ICD-10 auf (SD = 1,5).

Instrumente

Neben dem *HEALTH* wurden in Studie 1 (DETECT) außerdem der Lebensqualitätsfragebogen „EuroQoL“ (EQ-5D, Hinz et al. 2006; Schulenburg et al. 1998) sowie das „Depression Screening Questionnaire“ (DSQ, Wittchen u. Perkonig 1997) eingesetzt. Im Rahmen von Studie 2 (QS-Reha[®]) kamen ergänzend Kurzversionen verschiedener standardisierter Fragebögen zum Einsatz: Die 14-Item-Kurzform der „Symptomcheckliste SCL-90-R“ (Derogatis 1977; Franke 1995; 2002; SCL-14: Harfst et al. 2002), die 25-Item-Version des „Inventars zur Erfassung Interpersonaler Probleme IIP-D“ (Horowitz et al. 1988; 1994; IIP-25: Harfst et al. 2004) sowie die 8-Item-Kurzform der „MOS Health-Survey Short-Form SF-36“ (Bullinger u. Kirchberger 1998; Ware u. Sherbourne 1992; SF-8: Ware et al. 1999).

Entwicklung der HEALTH-Kurzform

Drei explorative Faktorenanalysen, die auf der Basis von drei verschiedenen Datensätzen (DETECT-Follow-up, QS-Reha[®]-Aufnahme, QS-Reha[®]-Entlassung) durchgeführt wurden, legten einheitlich eine 9-Faktoren-Lösung nahe. Da die in der Ursprungsversion vorgesehene Skala „Lebensqualität“ nicht als eigenständiger Faktor bestätigt werden konnte, wurde das entsprechende Modul komplett eliminiert.

Bei der weiteren Itemselektion wurden vorrangig die über alle drei Stichproben hinweg ladungsstärksten Items berücksichtigt. Die Auswahl der Items orientierte sich darüber hinaus an folgenden Kriterien:

- a) das Gesamtinstrument sowie die einzelnen Skalen bzw. Module sollten, um für den Routineeinsatz geeignet zu sein, möglichst kurz sein;
- b) die Skalen sollten eine befriedigende Reliabilität besitzen (mindestens $r > 0,70$, möglichst $r > 0,80$);
- c) die Ladungsmuster der Items auf den neun Skalen sollten eindeutig ausfallen (jeweils $r > 0,40$ auf der Index-Skala, $r < 0,40$ auf allen anderen Skalen);
- d) die erfassten Merkmale sollten angemessen heterogen erfasst werden.

Psychometrische Analysen zur HEALTH-Kurzform

Nach vorgeschalteten Analysen zur Praktikabilität wurden alle Fälle mit weniger als 70 % gültigen Angaben von der weiteren Auswertung ausgeschlossen. In den resultieren-

den Datensätzen wurden einzelne fehlende Werte (pro Fall im Mittel ein Fehlwert im *HEALTH*-Fragebogen) mittels der Expectation Maximization Methode (EM-Methode) ersetzt (Little u. Rubin 2002). Die Dimensionalität des verkürzten Instrumentariums wurde mittels konfirmatorischer und explorativer Faktorenanalysen überprüft. Die Reliabilität der Skalen wurde über die Berechnung von internen Konsistenzen (Cronbachs Alpha) und Trennschärfe-Koeffizienten bestimmt. Zur Untersuchung der Verteilungseigenschaften wurden Decken- und Bodeneffekte ermittelt sowie Schiefe und Kurtosis als Hinweise auf mögliche Abweichungen von der Normalverteilung bestimmt. Hinweise auf die Validität der *HEALTH*-Skalen wurden über Pearson-Korrelationen zu den weiteren im Rahmen der beiden Studien (DETECT und QS-Reha[®]) eingesetzten Instrumenten sowie durch den Vergleich umschriebener Subgruppen ermittelt. Um das Differenzierungspotential der verkürzten Skalen zu überprüfen, wurden auch ROC-Kurven verwendet.¹ Anhand der Verlaufsdaten aus Studie 2 (QS-Reha[®]-Verfahren) konnten zudem Prä-post-Effektstärken zur Beurteilung der Veränderungssensitivität bestimmt werden.

Ergebnisse

Ökonomie und Praktikabilität

Die gemäß den oben beschriebenen Kriterien entwickelte Kurzform des *HEALTH*-Fragebogens umfasst nur noch 49 Items („*HEALTH-49*“), deren Bearbeitung im Mittel etwa 10-15 Minuten in Anspruch nimmt.

Durchschnittlich beantworteten die beteiligten Patienten in Studie 1 (DETECT) 45,6 der 49 Items (SD = 8,6; 93,1 %). In Studie 2 (QS-Reha[®]) lag der mittlere Anteil beantworteter Items bei 97,2 % (47,6 von 49 Items, SD = 3,9). Die höchsten Anteile fehlender Werte fanden sich in Studie 1 in den Modulen „*Selbstwirksamkeit*“ (11,4 %) und „*Soziale Unterstützung/Soziale Belastung*“ (10,3 %), in Studie 2 – wenngleich in deutlich geringerer Ausprägung – im Modul „*Soziale Unterstützung/Soziale Belastung*“ (4,4 %)

¹ ROC-Kurven (ROC = Receiver Operating Characteristics) sind eine grafische Methode zur Ermittlung optimaler Skalengrenzwerte für die Unterscheidung dichotomer Merkmale (z.B. Gruppenzugehörigkeit „Kranke“ vs. „Gesunde“). Die Fläche unter der ROC-Kurve (AUC = Area Under the Curve) kann dabei als Wahrscheinlichkeit interpretiert werden, dass ein Fall mit einer positiven Merkmalsausprägung (z.B. „Kranke“) tatsächlich einen höheren Skalenwert aufweist als ein Fall mit negativer Merkmalsausprägung (z.B. „Gesunder“).

sowie im Modul „*Aktivität und Partizipation*“ (3,7 %). Innerhalb der einzelnen Module erwies sich die Akzeptanz der jeweiligen Items in beiden Studien als homogen.

Dimensionalität

In den Faktorenanalysen und allen weiterführenden psychometrischen Untersuchungen wurden nur noch die Patienten berücksichtigt, deren Daten ausreichend vollständig vorlagen (Studie 1: N = 5.125/5.630, Studie 2: N = 1.357/1.548).

Der in konfirmatorischen Faktorenanalysen (CFA) bestimmte “Standardized Root Mean Square of Residuals” Index (SRMR, Bentler 1995), der angibt, wie genau das spezifizierte Modell die Daten reproduziert, liegt mit Werten von 0,043 (Studie 1) bzw. 0,056 (Studie 2) unterhalb des von Hu und Bentler (1998) vorgeschlagenen Cut-off-Wertes von 0,08. Auch der „Root Mean Square Error of Approximation“ (RMSEA, z.B. Steiger 1990), ein Fit-Index, der gleichzeitig den absoluten Modell-Fit, die Sparsamkeit des Modells und die Stichprobengröße berücksichtigt, unterschreitet mit Werten von 0,046 (Studie 1) bzw. 0,054 (Studie 2) den ebenfalls von Hu und Bentler (1998) angegebenen Cut-off-Wert von 0,06. SRMR und RMSEA sprechen somit für die Angemessenheit des Modells. Der „Comparative Fit Index“ (CFI, Bentler 1990), der den inkrementellen Zuwachs an Modell-Fit im Vergleich zu einem Modell, in dem die Kovarianzen der manifesten Variablen auf Null fixiert und nur die Varianzen frei geschätzt werden, vergleicht, liegt mit Werten von 0,915 (Studie 1) bzw. 0,868 (Studie 2) jedoch etwas unterhalb des für ein akzeptables Modell geforderten Mindestwertes von 0,95 (Hu u. Bentler 1998).

Der in den ergänzend durchgeführten explorativen Faktorenanalysen (EFA) ermittelte Eigenwerteverlauf (Studie 1: 15,13; 3,08; 2,73; 2,31; 2,10; 1,88; 1,66; 1,30; 1,20; 0,87; 0,81; 0,76; 0,70 usw.; Studie 2: 14,40; 3,06; 2,91; 2,58; 2,07; 1,75; 1,36; 1,23; 1,20; 0,94; 0,84; 0,80; 0,76 usw.) legt in beiden Stichproben jedoch wiederum einheitlich die Extraktion von neun Faktoren nahe, welche zusammen 64,0 % (Studie 1) bzw. 62,4 % (Studie 2) der Gesamtvarianz aufklären. Die 9-Faktoren-Lösung der EFA ergibt für alle 49 Items eine eindeutige Item-Skalen-Zuordnung. Die drei Skalen „*Somatoforme Beschwerden*“ (SOM), „*Depressivität*“ (DEP) und „*Phobische Ängste*“ (PHO) repräsentieren das Modul A „*Psychische und somatoforme Beschwerden*“ (PSB) und lassen sich

zudem inhaltlich sinnvoll zu einem Gesamtwert zusammenfassen. Die *Module B* „*Psychisches Wohlbefinden*“ (*WOHL*), *C* „*Interaktionelle Schwierigkeiten*“ (*INT*), *D* „*Selbstwirksamkeit*“ (*SELB*) und *E* „*Aktivität und Partizipation*“ (*A&P*) sind jeweils durch eine entsprechende Skala repräsentiert. Dem *Modul F* („*Soziale Unterstützung/Soziale Belastung*“) können zwei entsprechende Skalen zugeordnet werden (*SOZU* u. *SOZB*), die aufgrund ihrer inhaltlichen Unvereinbarkeit jedoch nicht zu einem Gesamtwert aggregiert werden sollten.

Reliabilität

In Tabelle 1 sind interne Konsistenzen (Cronbachs Alpha) sowie mittlere Trennschärfe-koeffizienten als Indikatoren für die Reliabilität der Skalen des *HEALTH-49* aufgeführt. Die interne Konsistenz der Skalen ist mit Werten zwischen 0,76 und 0,91 (Studie 1) bzw. 0,73 und 0,90 (Studie 2) als überwiegend hoch zu bewerten und auch die interne Konsistenz für den Gesamtwert *PSB* liegt mit Alpha = 0,89 im hohen Bereich. Die Trennschärfe der einzelnen Items liegt durchgehend im mittleren bis hohen Bereich (alle $r_{it} > 0,40$).

- Tabelle 1 etwa hier einfügen -

Verteilungseigenschaften

Die Verteilungskennwerte sind ebenfalls in Tabelle 1 wiedergegeben. Während sich für keine der neun Skalen nennenswerte Bodeneffekte, also Häufungen in Richtung der ungünstigsten Ausprägung, ergeben, finden sich vor allem in der Hausarztstichprobe (DETECT-Studie) deutliche Deckeneffekte (Anteile von Fällen in der bestmöglichen Antwortkategorie von mehr als 20 %) für die Skalen *DEP*, *PHO* und *INT*. In Studie 1 (DETECT) weisen alle Skalen einheitlich eine zumindest leicht rechtsschiefe Verteilung auf (positive Schiefe, Deckeneffekte), die zumeist steiler als die Normalverteilung ausfällt (positive Kurtosis; vgl. Tab. 1). Demgegenüber zeigen sich die Werte der meisten Skalen in Studie 2 (QS-Reha[®]) deutlich normaler verteilt (kaum Boden- und Deckeneffekte, geringe Verteilungsschiefe), die Verteilung fällt für die meisten Skalen lediglich etwas flacher als die Normalverteilung aus (negative Kurtosis, vgl. Tab. 1).

Faktorielle und Konstrukt-Validität

Die Korrelationen zwischen den *HEALTH-49*-Skalen sowie zu den weiteren eingesetzten Skalen sind den Tabellen 2 und 3 zu entnehmen. Die einzelnen *HEALTH-49*-Skalen sind untereinander durchschnittlich mit $r = 0,38$ korreliert (Studie 1: $r = 0,37$, Studie 2: $r = 0,38$), was für ihre relative Eigenständigkeit spricht. Etwas höhere Interkorrelationen finden sich insbesondere im Zusammenhang mit der Skala *DEP*, die niedrigsten Interkorrelationen hingegen für die Skala *SOZU* (vgl. Tab. 2).

- Tabelle 2 etwa hier einfügen -

Wie der Tabelle 3 zu entnehmen ist, ergibt sich in Studie 1 (DETECT) die höchste Korrelation zum Lebensqualitätsfragebogen EuroQoL (EQ-5D) für die Skala *SOM* ($r = 0,62$), die höchste Korrelation zum Depressionsfragebogen DSQ für die Skala *DEP* ($r = 0,73$), was als Hinweis auf die Konstruktvalidität der Skalen gewertet werden kann. In Studie 2 (QS-Reha[®]) finden sich sehr hohe Korrelationen ($r > 0,80$) zwischen den *HEALTH-49*-Skalen *SOM*, *DEP*, *PHO* sowie dem *Modul-A*-Gesamtwert *PSB* und den jeweils korrespondierenden Skalen der SCL-14 (*SOM*, *DEP*, *PHO*, *GSI*). Für die Skala *SOM* zeigt sich außerdem eine hohe Korrelation zur Körperlichen Summenskala der SF-8 ($r = 0,71$). Die Skala *INT* korreliert besonders hoch mit dem ihr inhaltlich verwandten Gesamtwert des IIP-25 ($r = 0,75$). Zu den übrigen Skalen wurden keine direkt vergleichbaren Instrumente eingesetzt. Die höchsten Korrelationen zu den Skalen *WOHL*, *SELB* und *A&P* finden sich jeweils für die Psychische Summenskala der SF-8 ($r = 0,64$; $r = 0,65$; $r = 0,68$), was für die inhaltliche Nähe dieser Konstrukte spricht. Die Skala *SOZU* zeigt sich von allen anderen Skalen weitestgehend unabhängig (alle $r \leq 0,26$). Auch für die Skala *SOZB* zeigen sich höchstens mittelhohe Korrelationen zu den anderen eingesetzten Skalen ($r \leq 0,44$).

- Tabelle 3 etwa hier einfügen -

Differenzielle Validität

Im Rahmen der DETECT-Studie beurteilten die beteiligten Hausärzte die körperliche und psychische Mortalität, den psychosozialen Funktionsstatus und die autonome Selbstversorgung ihrer Patienten auf einer vierstufigen Skala (Skalierung von „unauffällig/gesund“ bis „schwerst auffällig/krank“). Um die Eignung des *HEALTH-49* zur Unterscheidung verschiedener Patientengruppen zu überprüfen, wurde anhand dieser Angaben zum aktuellen Gesundheitszustand der Patienten eine Subgruppe „gesunder“ Patienten identifiziert, welche in allen beurteilten Aspekten als unauffällig bzw. gesund eingeschätzt wurden ($n = 2.248$). Als „psychisch Kranke“ wurden demgegenüber alle Patienten der DETECT-Stichprobe definiert, deren psychische Morbidität durch den behandelnden Allgemeinarzt als deutlich oder schwerst auffällig eingestuft wurde ($n = 425$). Als dritte Vergleichsgruppe konnten daneben alle Patienten der QS-Reha[®]-Stichprobe herangezogen werden, die in Abgrenzung zu den DETECT-Patienten als stationär behandlungsbedürftig „psychisch Kranke“ zu charakterisieren sind ($n = 1.357$). In Abbildung 1 sind die Skalenwerte für die drei Subgruppen aufgeführt. Es zeigt sich, dass die Gesunden in sämtlichen Skalen jeweils signifikant geringere Beeinträchtigungswerte aufweisen als die Gesamtgruppe der psychisch Kranken (alle $p < 0,001$). Mit Ausnahme der Skala *SOZU* (Cohen's $d = 0,34$) erreichen diese Unterschiede durchwegs das Ausmaß großer Effektstärken (alle $d > 0,80$). Auch die Unterschiede zwischen den Subgruppen psychisch kranker Patienten aus der Primärversorgung (DETECT-Studie) und stationär behandlungsbedürftig psychisch Kranker (QS-Reha[®]) erweisen sich durchwegs als signifikant ($p_{(PHO)} < 0,05$; $p_{(SOM)} < 0,01$; alle anderen $p < 0,001$), wobei hier erwartungsgemäß jeweils die stationären Patienten als schwerer beeinträchtigt sind. Allerdings erreichen diese Unterschiede nur noch für die Skalen *DEP* und *INT* immerhin annähernd große Effektstärken ($d_{(DEP)} = 0,72$; $d_{(INT)} = 0,76$). Die geringsten Unterschiede zwischen den beiden Gruppen psychisch Kranker finden sich bezüglich somatoformer und phobischer Beschwerden (beide $d = 0,16$).

- Abbildung 1 etwa hier einfügen -

Die weiterführende Analyse mittels ROC-Kurven ergab, dass die meisten Skalen auch im Einzelfall sehr gut geeignet sind, um zwischen Gesunden und psychisch Kranken zu differenzieren (AUC > 80 %). Eine besonders hohe Diskriminationsfähigkeit lässt sich für die Skala *DEP* (AUC = 91,1 %) sowie den Gesamtwert *PSB* (AUC = 90,5 %) konstatieren. Die Differenzierungsleistungen der Skalen *SOZU* (AUC = 61,5 %) und *SOZB* (AUC = 73,6 %) liegen hingegen erwartungsgemäß deutlich niedriger.

Die Eignung des *HEALTH-49* zur Differenzierung zwischen stationär behandlungsbedürftig psychisch Kranken (QS-Reha[®]) und psychisch erkrankten Patienten aus dem Bereich der primärärztlichen Versorgung (DETECT) fällt demgegenüber deutlich niedriger aus (mittlere AUC = 62,9 % gegenüber durchschnittlich 81,8 % bei der Unterscheidung von Gesunden und psychisch Kranken). Hierbei weisen die Skalen *INT* (AUC = 71,2 %) und *DEP* (AUC = 71,1 %) die höchste Diskriminationsleistung auf. Die Differenzierungsfähigkeit der Skalen *SOM* (AUC = 54,2 %), *PHO* (AUC = 54,3 %), *SOZU* (AUC = 57,1 %) und *SOZB* (AUC = 57,8 %) bewegt sich hingegen nahe an der Zufallswahrscheinlichkeit von 50 %.

Im Gruppenvergleich weisen weibliche Patienten zwar in den meisten Skalen signifikant höhere Beeinträchtigungswerte auf als männliche Patienten, diese Unterschiede entsprechen jedoch höchstens kleinen Effektstärken (DETECT: alle $d \leq 0,23$; QS-Reha[®]: alle $d \leq 0,33$). Überhaupt keine signifikanten Geschlechtsunterschiede finden sich in der DETECT-Studie bezüglich der Skalen *SELB*, *A&P* und *SOZB*, in der QS-Reha[®]-Stichprobe bezüglich der Skalen *SELB* und *SOZU*. Auch in den durchgeführten ROC-Analysen lassen sich kaum systematische Geschlechtsunterschiede nachweisen (alle AUC < 60 %).

Desgleichen finden sich höchstens kleine Alterseffekte: Die größten Zusammenhänge zum Alter finden sich in der Hausarzt-Stichprobe für die Skalen *A&P* ($r = 0,18$), *WOHL* ($r = -0,16$) und *SOM* ($r = 0,14$; alle $p < 0,001$). Im QS-Reha[®]-Datensatz finden sich sogar noch geringere Alterseffekte (maximal $r = -0,11$ für die Skala *WOHL*).

Veränderungssensitivität

Zur Beurteilung der Veränderungssensitivität wurden anhand der Verlaufsdaten aus Studie 2 (QS-Reha[®]-Verfahren) für alle eingesetzten Skalen standardisierte Prä-post-

Effektstärken (SES) bestimmt, die in Abbildung 2 dargestellt sind. Es zeigt sich, dass sich in sechs der neun *HEALTH-49*-Skalen sowie im Gesamtwert *PSB* Veränderungen im Ausmaß einer mindestens mittleren Effektstärke nachweisen lassen ($SES > 0,50$). Als besonders veränderungssensitiv erweist sich die Skala *WOHL* ($SES = 1,21$). Die Veränderungen in der Skala *PHO* erreichen lediglich das Ausmaß einer kleinen Effektstärke ($SES = 0,34$), in den Skalen *SOZU* und *SOZB* finden sich nur noch geringere als kleine Veränderungseffekte ($SES < 0,20$). Im Vergleich zwischen den *HEALTH-49*-Skalen und den inhaltlich korrespondierenden Skalen der sonstigen eingesetzten Instrumente (z.B. *HEALTH-49-SOM* vs. *SCL-14-SOM* usw.) zeigt sich, dass die *HEALTH-49*-Skalen durchwegs größere SES aufweisen, was für ihre vergleichsweise höhere Veränderungssensitivität spricht. Besonders deutlich fällt dieser Unterschied im Vergleich der Skala *INT* und dem IIP-25-Gesamtwert „*Interpersonelle Probleme*“ (IIP-GES) aus: Hier beträgt die Differenz zwischen den SES 0,22 Standardabweichungen, was für sich genommen bereits einer kleinen Effektstärke entspricht.

- Abbildung 2 etwa hier einfügen -

Diskussion

Die Kurzform des *HEALTH*-Fragebogens (*HEALTH-49*) ist angesichts einer Verkürzung der Ursprungsversion um 38 % (von 79 auf 49 Items) im Sinne der optimierten Ökonomie positiv zu bewerten. Mit der vorliegenden 49-Item-Lösung scheint jedoch die untere Grenze vertretbarer Kürzungen erreicht, da bei Eliminierung weiterer Items sowohl die interne Konsistenz (Reliabilität) als auch die erforderliche Heterogenität der einzelnen Skaleninhalte (Validität) gefährdet wären. Bei Bedarf (z.B. bei wiederholter Verlaufsmessung) lässt sich der Fragebogen jedoch durch die Auswahl der für die jeweilige Fragestellung relevantesten Module weiter verschlanken.

Bezüglich der Akzeptanz des *HEALTH-49* finden sich keine Unterschiede zur Ursprungsversion (Rabung et al. 2007). Die etwas zu hohe Quote fehlender Werte in der DETECT-Studie (6,9 %) dürfte angesichts einer kaum noch zu optimierenden Bearbeitungsquote von 97,2 Prozent im stationären Setting (QS-Reha[®]-Verfahren) weniger im

Sinne einer Einschränkung der grundsätzlichen Praktikabilität des Fragebogens zu verstehen sein, als vielmehr als wichtiger Hinweis auf eine mögliche Einschränkung seines Anwendungsbereiches. Da der *HEALTH*-Fragebogen im Rahmen der DETECT-Studie als letzter, fakultativer Teil einer umfangreicheren Fragebogenbatterie in den Wartezimmern der beteiligten Hausarztpraxen bearbeitet werden sollte, bleibt jedoch unklar, inwieweit fehlende Angaben dort auf besondere Merkmale des Untersuchungsdesigns (inhaltliche Redundanzen, motivationale Aspekte, begrenzte Zeit) oder auf Besonderheiten der Untersuchungsstichprobe (hoher Anteil „gesunder“ Patienten) zurückzuführen ist. Deutliche Hinweise auf eine Stichprobenabhängigkeit und entsprechenden Entwicklungsbedarf finden sich lediglich im Modul „*Selbstwirksamkeit*“ (SELB): Hier liegt die Fehlwertquote in der DETECT-Studie mit 11,4 % nahezu fünfmal so hoch wie im QS-Reha[®]-Kontext, was sich vermutlich auf die Itemformulierung (alle Items beginnen mit der Einleitung „Trotz meiner Beschwerden bin ich in der Lage...“) zurückführen lässt, die zumindest eine gewisse Beeinträchtigung voraussetzt und dementsprechend von beschwerdefreien Probanden nicht sinnvoll beantwortet werden kann.

Die explorativen Faktorenanalysen legen für den *HEALTH-49* eine 9-Faktoren-Lösung nahe. Trotz der eindeutigen Skalen-Zuordnung aller 49 Items und einer hohen Varianzaufklärung von über 60 % lässt sich die 9-Faktoren-Lösung in konfirmatorischen Faktorenanalysen allerdings nicht abschließend bestätigen. Diese Diskrepanz ist jedoch vermutlich darauf zurückzuführen, dass mit den verschiedenen Skalen eben keine gänzlich voneinander unabhängigen Merkmale erfasst werden, sondern verschiedene Aspekte psychosozialer Gesundheit, die sich doch auch mehr oder weniger ausgeprägt gegenseitig beeinflussen bzw. zumindest überlappen.

Auch wenn sich die interne Konsistenz sämtlicher Skalen in Folge der Kürzungen gegenüber der Originalversion geringfügig verringert hat, ist ihre Reliabilität immer noch als durchwegs gut zu bewerten (Cronbach's Alpha liegt lediglich für die Skala „*Soziale Belastung*“ unter 0,80). Die mittlere Trennschärfe der Skalen hat sich gegenüber der Ursprungsversion etwas erhöht, was auf die Eliminierung weniger trennscharfer Items bei weitgehender Beibehaltung der Heterogenität der Skaleninhalte zurückzuführen ist.

In Bezug auf die Verteilungseigenschaften lässt sich festhalten, dass sich stichprobenu-
nabhängig so gut wie keine Bodeneffekte finden und nennenswerte Deckeneffekte auch
lediglich bei der Untersuchung von gesünderen Probanden auftreten. Schiefe und Kur-
tosis sind gegenüber der Ursprungsversion insgesamt deutlich reduziert und zeigen sich
in der Stichprobe stationärer Psychotherapie-Patienten (QS-Reha[®]) nochmals erheblich
normaler ausgeprägt als in der DETECT-Stichprobe.

Die mittlere Interkorrelation der einzelnen Skalen hat sich im Vergleich zur Ursprungs-
version etwas verringert, was nicht zuletzt auf die Eliminierung der Skala „*Lebensquali-
tät*“ zurückführbar ist, die besonders hoch mit den anderen *HEALTH*-Skalen korreliert
war. Etwas höhere Interkorrelationen finden sich nun nur noch im Zusammenhang mit
der Skala „*Depressivität*“, was jedoch angesichts einer geringen Spezifität depressiver
Beschwerden plausibel erscheint. Insgesamt lässt sich somit eine befriedigende fakto-
rielle Validität konstatieren. Höhere Korrelationen zu konstruktverwandten Skalen an-
derer Fragebögen belegen die Konstruktvalidität einiger *HEALTH-49*-Skalen (*SOM*,
DEP, *PHO*, *PSB*, *INT*), was insbesondere im Hinblick auf die gewünschte Vergleich-
barkeit mit den weit verbreiteten Instrumenten SCL-90-R und IIP (z.B. Jakobsen et al.
2007; Leichsenring et al. 2005; 2007; Schauenburg et al. 2006; Schneider et al. 2006;
Stasch et al. 2007; Wöller et al. 2007) als sehr positiv zu bewerten ist. Für die weiteren
HEALTH-49-Skalen finden sich in den Korrelationen zu den sonstigen eingesetzten
Instrumenten zwar Hinweise auf deren inhaltliche Validität, eine differenziertere Vali-
dierung dieser Skalen an konstrukt näheren externen Kriterien steht jedoch noch aus.
Auf Gruppenebene erwiesen sich alle Skalen zur Unterscheidung von Gesunden und
Psychisch Kranken und sogar zur Differenzierung zwischen Psychisch Kranken aus
dem Gebiet der allgemeinärztlichen Versorgung und Patienten in stationärer psychoso-
matisch/psychotherapeutischer Rehabilitation als geeignet. Und auch im Einzelfall ist
die Differenzierungsleistung der meisten Skalen, insbesondere bei der Unterscheidung
von Gesunden und Psychisch Kranken, als gut zu bewerten. Allerdings sind diese Be-
funde zur differenziellen Validität vor dem Hintergrund der ungeprüften Reliabilität und
Validität – insbesondere – der hausärztlichen Diagnostik psychischer Störungen zu be-
werten. Von besonderem Interesse wäre in diesem Zusammenhang auch die Klassifizie-

rungsleistung im Hinblick auf umschriebene psychische Störungen, welche sich jedoch erst anhand von Daten aus Studien mit einer standardisierten Interviewbasierten Diagnostik beurteilen lassen wird (vgl. z.B. Becker et al. 2006).

Die großen Prä-Post-Effektstärken in sechs der neun Skalen sowie dem Gesamtwert „*Psychische und somatoforme Beschwerden*“ zeigen die hohe Sensitivität der betroffenen Maße in Bezug auf Veränderungen während stationärer psychosomatisch / psychotherapeutischer Rehabilitation auf. Gleichmaßen plausibel erscheinen die geringen Veränderungen in den Skalen „*Soziale Unterstützung*“ und „*Soziale Belastung*“, die eher kontextuelle Aspekte psychosozialer Gesundheit erfassen. Einzig die lediglich kleinen Effekte in der Skala „*Phobische Ängste*“ widersprechen zunächst der Erwartung, lassen sich jedoch auf die in der untersuchten Stichprobe bereits initial verhältnismäßig gering ausgeprägte Beeinträchtigung in diesem Bereich zurückführen. Besonders hervorzuheben ist in diesem Zusammenhang, dass die Effektstärken in den *HEALTH-49*-Skalen durchwegs deutlich höher ausfielen als die entsprechenden Effektstärken auf den jeweils korrespondierenden Skalen der anderen überprüften, bereits bewährten Instrumente, was für eine vergleichsweise höhere Veränderungssensitivität des *HEALTH-49* spricht.

Insgesamt belegen die an umfangreichen klinischen wie auch gesunden Stichproben ermittelten Ergebnisse die gute psychometrische Qualität des *HEALTH-49*. Durch seinen Einsatz in verschiedenen Versorgungskontexten liegen bereits jetzt umfangreiche Referenzdaten vor, die die Einordnung einzelner Patienten bzgl. ihres psychotherapeutischen Behandlungsbedarfs wie auch der im Rahmen einer Behandlung erzielten Verbesserungen unterstützen können. Zur Beantwortung differenzierterer Fragestellungen wäre nun noch die Erweiterung der vorliegenden „Kern“-Module um zusätzliche diagnose- bzw. störungsspezifische Module bzw. Skalen, z.B. zunächst zu den häufigsten psychischen Störungen (z.B. umschriebene affektive und Angst-Störungen, Anpassungsstörungen, Zwangsstörungen, Essstörungen usw.), notwendig und sinnvoll (z.B. Gönner et al. 2007; Tritt et al. 2007). Neben der deutschsprachigen Originalversion liegen zum *HEALTH*-Fragebogen auch schon Versionen in englischer, türkischer und chinesischer Sprache vor, weitere Übersetzungen sind in Vorbereitung.

Literatur

- Barkham, M., Evans, C., Margison, F., McGrath, G., Mellor-Clark, J., Milne, D., Connell, J. (1998): The rationale for developing and implementing core outcome batteries for routine use in service settings and psychotherapy outcome research. *J Ment Health* 7, 35-47.
- Becker, J., Kocalevent, R.-D., Rose, M., Fliege, H., Walter, O.B., Frommer, J., Klapp, B.F. (2006): Standardisierte Diagnosefindung: Computergestützte (CIDI) - versus klinische Diagnosestellungen an einer psychosomatischen Stichprobe. *Psychother Psych Med* 56, 5-14.
- Bentler, P.M. (1990): Comparative fit indexes in structural models. *Psychol Bull* 107, 238-246.
- Bentler, P.M. (1995): EQS: A structural equations program manual. Encino, CA: Multivariate Software, Inc.
- Bullinger, M., Kirchberger, I. (1998): SF-36. Fragebogen zum Gesundheitszustand. Göttingen: Hogrefe.
- Derogatis, L.R. (Hg.) (1977): SCL-90-R, administration, scoring & procedures manual-I for the R(evised) version. Baltimore: John Hopkins University School Medicine.
- Franke, G. (1995): SCL-90-R. Die Symptom-Checkliste von Derogatis - Deutsche Version. Göttingen: Beltz.
- Franke, G. (2002): SCL-90-R. Die Symptom Checkliste von L.R. Derogatis - Manual zur Deutsche Version: Beltz.
- Göner, S., Leonhart, R., Ecker, W. (2007): Das Zwangsinventar OCI-R - die deutsche Version des Obsessive-Compulsive Inventory-Revised. *Psychother Psych Med* 57, 395-404.
- Hahlweg, K., Schmid, H., Selck, A. (1999): Testbatterie zur Qualitätssicherung von Psychotherapie. Göttingen: Hogrefe.
- Hardt, J., Egle, U., Brähler, E. (2006): Die Symptom-Checkliste-27 in Deutschland. *Psychother Psych Med* 56, 276-284.
- Harfst, T., Dirmaier, J., Koch, U., Schulz, H. (2004): Entwicklung und psychometrische Überprüfung einer multidimensionalen Kurzform des Inventars zur Erfassung interpersonaler Probleme. *DRV-Schriften* 52, 53-55.
- Harfst, T., Koch, U., Kurtz von Aschoff, C., Nutzinger, D.O., Rüdell, H., Schulz, H. (2002): Entwicklung und Validierung einer Kurzform der Symptom Checklist-90-R. *DRV-Schriften* 33, 71-73.
- Hinz, A., Klaiberg, A., Brähler, E., König, H.-H. (2006): Der Lebensqualitätsfragebogen EQ-5D: Modelle und Normwerte für die Allgemeinbevölkerung. *Psychother Psych Med* 56, 42-48.
- Horowitz, L.M., Rosenberg, S.E., Baer, B.A., Ureño, G., Villaseñor, V.S. (1988): Inventory of Interpersonal Problems: Psychometric properties and clinical applications. *J Consult Clin Psych* 56, 885-892.
- Horowitz, L.M., Strauß, B., Kordy, H. (1994): Inventar zur Erfassung interpersonaler Probleme - Deutsche Version - (IIP-D). Weinheim: Beltz.
- Hu, L.T., Bentler, P.M. (1998): Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychol Methods* 3, 424-453.
- Jakobsen, T., Rudolf, G., Eckert, J., Huber, D., Klug, G., Grande, T., Keller, W., Staats, H., Leichsenring, F. (2007): Ergebnisse analytischer Langzeitpsychotherapien bei spezifischen psychischen Störungen: Verbesserungen in der Symptomatik und in interpersonellen Beziehungen. *Z Psychosom Med Psychother* 53, 87-110.
- Kraus, D.R., Seligman, D.A., Jordan, J.R. (2005): Validation of a Behavioral Health Treatment Outcome and Assessment Tool Designed for Naturalistic Settings: The Treatment Outcome Package. *J Clin Psychol* 61, 285-314.
- Leichsenring, F., Dümpelmann, M., Berger, J., Jaeger, U., Rabung, S. (2005): Ergebnisse stationärer psychiatrischer und psychotherapeutischer Behandlung von schizophrenen, schizoaffektiven und anderen psychotischen Störungen. *Z Psychosom Med Psychother* 51, 23-37.

- Leichsenring, F., Masuhr, O., Jaeger, U., Dally, A., Streeck, U. (2007): Zur Wirksamkeit psychoanalytisch-interaktioneller Therapie am Beispiel der Borderline-Persönlichkeitsstörung: Daten aus der klinischen Praxis. *Z Psychosom Med Psychother* 53, 129-143.
- Little, R., Rubin, D. (2002): *Statistical Analysis with Missing Data*. New York: John Wiley.
- Pieper, L., Wittchen, H.-U., Glaesmer, H., Klotsche, J., März, W., Stalla, G.K., Lehnert, H., Zeiher, A.M., Silber, S., Koch, U., Böhler, S., Pittrow, D., Ruf, G. (2005): Kardiovaskuläre Hochrisikokonstellationen in der primärärztlichen Versorgung: DETECT Studie 2003. *Bundesgesundheitsblatt* 12, 1374-1382.
- Rabung, S., Harfst, T., Koch, U., Wittchen, H.-U., Schulz, H. (2007): „Hamburger Module zur Erfassung allgemeiner Aspekte psychosozialer Gesundheit für die therapeutische Praxis (HEALTH)“ – psychometrische Überprüfung eines neuen Selbstbeurteilungsinstrumentes zur multidimensionalen Erfassung psychosozialer Gesundheit. *Phys Med Rehab Kuror* 17, 133-140.
- Schauenburg, H., Dinger, U., Buchheim, A. (2006): Bindungsmuster von Psychotherapeuten. *Z Psychosom Med Psychother* 52, 358-372.
- Schneider, G., Schmitz-Moormann, S., Baer, O., Driesch, G., Heuft, G. (2006): Psychodynamische Therapiefoki in einer stationären multimodalen Therapie - Zusammenhänge zu den Selbstauskunftskalen und individuellen Therapiezielen. *Z Psychosom Med Psychother* 52, 259-273.
- Schulenburg, G.v.d.J., Claes, C., Greiner, W., Uber, A. (1998): Die deutsche Version des EuroQOL-Fragebogens. *Z Gesundheitswiss* 6, 3-20.
- Schumacher, J., Brähler, E. (2005): Testdiagnostik in der Psychotherapie. In: Senf, W., Broda, M. (Hg.): *Praxis der Psychotherapie*, S. 169-183. Stuttgart: Thieme.
- Stasch, M., Cierpka, M., Schmal, H., Hillenbrand, E. (2007): Fokusorientierte Interventionen mit der OPD in der stationären Psychotherapie: Effekte auf Ergebnis und Verlauf der Behandlung. *Z Psychosom Med Psychother* 53, 309-323.
- Steiger, J.H. (1990): Structural model evaluation and modification: An interval estimation approach. *Multivar Behav Res* 25, 173-180.
- Tritt, K., von Heymann, F., Zaudig, M., Söllner, W., Loew, T. (2007): Erfassung psychischer Symptomatik mit der ICD-10 Symptom Rating (ISR). *Nervenarzt* 78 (Suppl 2), 467.
- Ware, J.E., Kosinski, M., Dewey, J.E., Gandek, B. (1999): *How to Score and Interpret Single-Item Health Status Measures: A Manual for Users of the SF-8 TM Health Survey*. Lincoln, RI: QualityMetric Incorporated.
- Ware, J.E., Sherbourne, C.D. (1992): The MOS 36-Item short form health-survey (SF-36). I. Conceptual framework and item selection. *Med Care* 30, 473-483.
- Wittchen, H.-U., Glaesmer, H., März, W., Stalla, G., Lehnert, H., Zeiher, A.M., Silber, S., Koch, U., Böhler, S., Pittrow, D., Ruf, G., for the DETECT-Study Group. (2005): Cardiovascular Risk Factors in Primary Care: Methods and Baseline Prevalence Rates. *The DETECT Program. Curr Med Res Opin* 21, 619-629.
- Wittchen, H.-U., Perkonig, A. (1997): DIA-X-Screening Verfahren: Fragebogen DIA-SSQ: Screening für psychische Störungen; Fragebogen DIA-ASQ: Screening für Angststörungen; Fragebogen DIA-DSQ: Screening für Depressionen. Frankfurt: Swets & Zeitlinger bv.
- Wöller, W., Hartkamp, N., Tress, W. (2007): Einfluss von Persönlichkeitsmerkmalen auf die erlebte Hilfe und Belastung durch Gruppentherapie und Mitpatienten in der stationären psychodynamischen Psychotherapie. *Z Psychosom Med Psychother* 53, 163-176.

Tabellen und Abbildungen

Tab. 1 Interne Konsistenzen, Trennschärfe und Verteilungskennwerte der HEALTH-49-Skalen in den untersuchten Stichproben (Studie 1: DETECT-Studie; Studie 2: QS-Reha[®]-Verfahren)*

| Modul bzw. Skala (Abkürzung, Anzahl Items) | Interne Konsistenz (Cronbachs α) | | Mittlere Trennschärfe (korrigiert) [#] | | Decke (%) | | Boden (%) | | Schiefe | | Kurtosis | |
|---|--|-------------|---|-------------|-----------|-------------|-----------|-----------|---------|--------------|----------|--------------|
| | Studie: | 1 / 2 | 1 / 2 | 1 / 2 | 1 / 2 | 1 / 2 | 1 / 2 | 1 / 2 | 1 / 2 | 1 / 2 | 1 / 2 | |
| Modul A: „Psychische und somatoforme Beschwerden“ | | | | | | | | | | | | |
| - Subskala „Somatoforme Beschwerden“ (SOM, 7) | | 0,82 / 0,82 | | 0,57 / 0,56 | | 10,2 / 2,1 | | 0,0 / 0,1 | | 1,05 / 0,31 | | 0,77 / -0,67 |
| - Subskala „Depressivität“ (DEP, 6) | | 0,88 / 0,88 | | 0,70 / 0,69 | | 45,6 / 2,4 | | 0,1 / 1,6 | | 2,47 / 0,16 | | 7,06 / -0,78 |
| - Subskala „Phobische Ängste“ (PHO, 5) | | 0,82 / 0,86 | | 0,63 / 0,70 | | 71,3 / 34,3 | | 0,1 / 0,2 | | 3,61 / 1,26 | | 15,75 / 0,81 |
| Modul B: „Psychisches Wohlbefinden“ (WOHL, 5) | | 0,91 / 0,87 | | 0,77 / 0,69 | | 9,0 / 0,1 | | 0,8 / 3,7 | | 0,71 / -0,37 | | 0,32 / 0,04 |
| Modul C: „Interaktionelle Schwierigkeiten“ (INT, 7) | | 0,90 / 0,90 | | 0,70 / 0,71 | | 28,4 / 3,0 | | 0,1 / 1,0 | | 1,53 / 0,00 | | 2,30 / -0,98 |
| Modul D: „Selbstwirksamkeit“ (SELB, 5) | | 0,91 / 0,87 | | 0,78 / 0,70 | | 18,7 / 0,7 | | 1,0 / 2,8 | | 1,23 / -0,12 | | 1,69 / -0,54 |
| Modul F: „Aktivität und Partizipation“ (A&P, 6) | | 0,85 / 0,83 | | 0,64 / 0,61 | | 19,8 / 1,9 | | 0,1 / 0,6 | | 1,01 / -0,08 | | 0,58 / -0,51 |
| Modul G: „Soziale Unterstützung/Belastung“ | | | | | | | | | | | | |
| - Subskala „Soziale Unterstützung“ (SOZU, 4) | | 0,87 / 0,83 | | 0,73 / 0,65 | | 11,1 / 2,6 | | 3,6 / 0,7 | | 0,80 / 0,25 | | 0,10 / -0,49 |
| - Subskala „Soziale Belastung“ (SOZB, 4) | | 0,76 / 0,73 | | 0,56 / 0,53 | | 10,8 / 1,7 | | 0,2 / 0,1 | | 0,52 / 0,46 | | 0,00 / -0,44 |

* Der HEALTH-Fragebogen ist im Internet abrufbar (<http://www.hamburger-module.de>).

[#] Alle Trennschärfekoeffizienten für die einzelnen Items > 0,40.

Tab. 2 Pearson-Korrelationen zwischen den HEALTH-49-Skalen in den untersuchten Stichproben (Studie 1: DETECT-Studie; Studie 2: QS-Reha[®]-Verfahren)

| <i>Skalen</i> | <i>DEP</i> | <i>PHO</i> | <i>WOHL</i> | <i>INT</i> | <i>SELB</i> | <i>A&P</i> | <i>SOZU</i> | <i>SOZB</i> |
|----------------|--------------|--------------|--------------|--------------|--------------|----------------|--------------|--------------|
| <i>Studie:</i> | <i>1 / 2</i> | <i>1 / 2</i> | <i>1 / 2</i> | <i>1 / 2</i> | <i>1 / 2</i> | <i>1 / 2</i> | <i>1 / 2</i> | <i>1 / 2</i> |
| <i>SOM</i> | 0,52 / 0,42 | 0,45 / 0,39 | 0,48 / 0,40 | 0,40 / 0,34 | 0,41 / 0,32 | 0,53 / 0,47 | 0,05 / 0,15 | 0,30 / 0,31 |
| <i>DEP</i> | | 0,60 / 0,48 | 0,61 / 0,63 | 0,61 / 0,61 | 0,55 / 0,64 | 0,48 / 0,53 | 0,12 / 0,22 | 0,40 / 0,36 |
| <i>PHO</i> | | | 0,38 / 0,32 | 0,45 / 0,42 | 0,41 / 0,34 | 0,40 / 0,39 | 0,06 / 0,14 | 0,26 / 0,27 |
| <i>WOHL</i> | | | | 0,50 / 0,47 | 0,57 / 0,64 | 0,44 / 0,50 | 0,13 / 0,19 | 0,37 / 0,28 |
| <i>INT</i> | | | | | 0,48 / 0,49 | 0,39 / 0,41 | 0,11 / 0,24 | 0,43 / 0,44 |
| <i>SELB</i> | | | | | | 0,50 / 0,55 | 0,18 / 0,23 | 0,31 / 0,29 |
| <i>A&P</i> | | | | | | | 0,06 / 0,16 | 0,40 / 0,35 |
| <i>SOZU</i> | | | | | | | | 0,04 / 0,36 |

HEALTH-49-Skalen – SOM: Somatoforme Beschwerden; DEP: Depressivität; PHO: Phobische Ängste; WOHL: Psychisches Wohlbefinden; INT: Interaktionelle Schwierigkeiten; SELB: Selbstwirksamkeit; A&P: Aktivität und Partizipation; SOZU: Soziale Unterstützung; SOZB: Soziale Belastung.

Höhere Skalenwerte bedeuten jeweils größere Belastung/Beeinträchtigung.

Alle Korrelationen signifikant mit $p < 0,001$ (Ausnahme: SOZU-SOZB in Studie 1, $p = 0,009$).

Tab. 3 Pearson-Korrelationen zwischen HEALTH-49-Skalen und Skalen anderer eingesetzter Fragebögen in den untersuchten Stichproben (Studie 1: DETECT-Studie; Studie 2: QS-Reha[®]-Verfahren)

| Skalen | Studie 1 | | Studie 2 | | | | | | |
|--------------------|--------------------|----------------|----------------|----------------|----------------|----------------|-------------------|--------------|--------------|
| | EuroQoL: GESAMT | DSQ: GESAMT | SCL-14: SOM | SCL-14: DEP | SCL-14: PHO | SCL-14: GSI | IIP-25: GESAMT | SF-8: PSK | SF-8: KSK |
| HEALTH-49- | | | | | | | | | |
| SOM | 0,62 | 0,59 | 0,85 | 0,39 | 0,37 | 0,68 | 0,31 | 0,47 | 0,71 |
| DEP | 0,49 | 0,73 | 0,36 | 0,87 | 0,37 | 0,77 | 0,56 | 0,69 | 0,35 |
| PHO | 0,41 | 0,51 | 0,34 | 0,45 | 0,84 | 0,66 | 0,39 | 0,43 | 0,36 |
| (PSB) [#] | (0,64) | (0,74) | (0,68) | (0,73) | (0,63) | (0,89) | (0,54) | (0,68) | (0,62) |
| WOHL | 0,42 | 0,59 | 0,33 | 0,61 | 0,27 | 0,58 | 0,44 | 0,64 | 0,38 |
| INT | 0,33 | 0,51 | 0,28 | 0,63 | 0,31 | 0,58 | 0,75 | 0,50 | 0,26 |
| SELB | 0,44 | 0,53 | 0,30 | 0,64 | 0,27 | 0,58 | 0,47 | 0,65 | 0,35 |
| A&P | 0,56 | 0,56 | 0,45 | 0,55 | 0,36 | 0,61 | 0,42 | 0,68 | 0,55 |
| SOZU | 0,05 | 0,09 | 0,14 | 0,25 | 0,10 | 0,24 | 0,26 | 0,21 | 0,17 |
| SOZB | 0,23 | 0,36 | 0,28 | 0,40 | 0,23 | 0,42 | 0,44 | 0,37 | 0,27 |

[#] Da es sich beim Gesamtwert „Psychische und Somatoforme Beschwerden“ (PSB) um einen über mehrere HEALTH-Skalen aggregierten Score und nicht um eine eigenständige Skala handelt, werden die Korrelationen zum PSB in Klammern dargestellt.

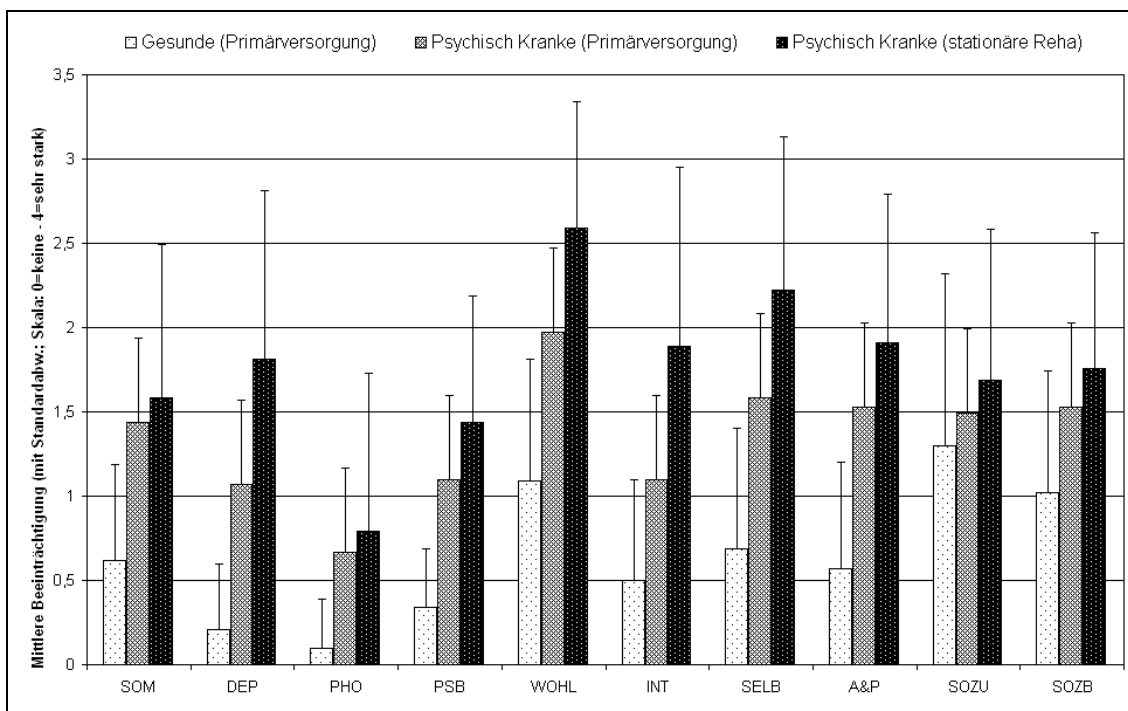
HEALTH-49-Skalen – SOM: Somatoforme Beschwerden; DEP: Depressivität; PHO: Phobische Ängste; PSB: Psychische und Somatoforme Beschwerden (Gesamtwert Modul A); WOHL: Psychisches Wohlbefinden; INT: Interaktionelle Schwierigkeiten; SELB: Selbstwirksamkeit; A&P: Aktivität und Partizipation; SOZU: Soziale Unterstützung; SOZB: Soziale Belastung.

Weitere Skalen – EuroQoL: Skala zur Lebensqualität (multiplikativer Gesamtwert); DSQ: Depression Screening Questionnaire (Depressive Symptomatik); SCL-14: Symptom-Checkliste (14-Item-Kurzversion; Skalen SOM ‚Somatisierung‘, DEP ‚Depressivität‘, PHO ‚Phobische Ängste‘, GSI ‚Globaler Symptom-Schwere-Index‘); SF-8: Health Survey Short Form (8-Item-Kurzversion; Skalen PSK ‚Psychische Summenskala‘, KSK ‚Körperliche Summenskala‘); IIP-25: Inventar Interpersonaler Probleme (25-Item-Kurzversion).

Höhere Skalenwerte bedeuten jeweils größere Belastung/Beeinträchtigung.

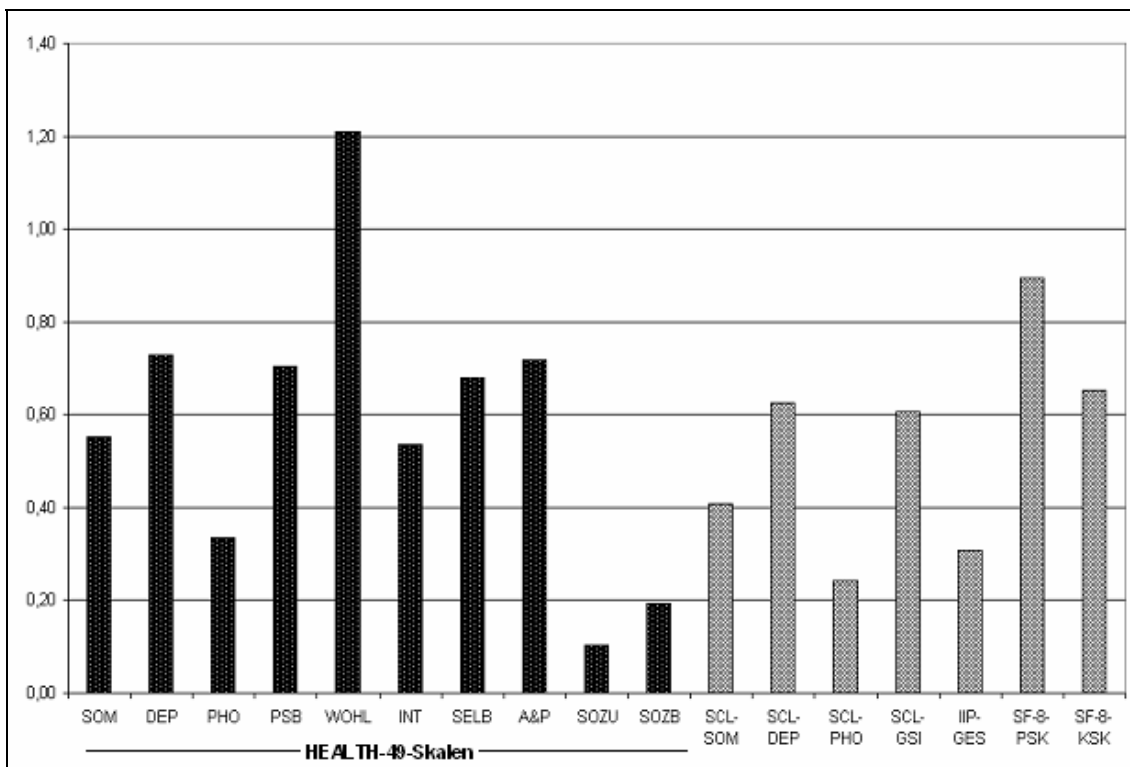
Alle Korrelationen signifikant mit $p < 0,001$.

Abb. 1 Mittlere Ausprägung der HEALTH-49-Skalen in verschiedenen Subgruppen



HEALTH-49-Skalen - SOM: Somatoforme Beschwerden; DEP: Depressivität; PHO: Phobische Ängste; PSB: Psychische und somatoforme Beschwerden (Gesamtwert *Modul A*); WOHL: Psychisches Wohlbefinden; INT: Interaktionelle Schwierigkeiten; SELB: Selbstwirksamkeit; A&P: Aktivität und Partizipation; SOZU: Soziale Unterstützung; SOZB: Soziale Belastung. Höhere Skalenwerte bedeuten jeweils größere Belastung/Beeinträchtigung.

Abb. 2 Standardisierte Effektstärken der Prä-post-Veränderung während stationärer psychosomatisch/psychotherapeutischer Rehabilitation in den Skalen von HEALTH-49, SCL-14, IIP-25 und SF-8



HEALTH-49-Skalen - SOM: Somatoforme Beschwerden; DEP: Depressivität; PHO: Phobische Ängste; PSB: Psychische und somatoforme Beschwerden (Gesamtwert *Modul A*); WOHL: Psychisches Wohlbefinden; INT: Interaktionelle Schwierigkeiten; SELB: Selbstwirksamkeit; A&P: Aktivität und Partizipation; SOZU: Soziale Unterstützung; SOZB: Soziale Belastung.

Weitere Skalen – SCL-14: Symptom-Checkliste (14-Item-Kurzversion; Skalen SOM ‚Somatisierung‘, DEP ‚Depressivität‘, PHO ‚Phobische Ängste‘, GSI ‚Globaler Symptom-Schwere-Index‘); IIP-25: Inventar Interpersonaler Probleme (25-Item-Kurzversion; GES: Gesamtwert); SF-8: Health Survey Short Form (8-Item-Kurzversion; Skalen PSK ‚Psychische Summenskala‘, KSK ‚Körperliche Summenskala‘).

Höhere Skalenwerte bedeuten jeweils größere Belastung/Beeinträchtigung.

Alle Prä-post-Veränderungen statistisch signifikant mit $p < 0,001$.