Universität Hamburg

DER FORSCHUNG | DER LEHRE | DER BILDUNG

Die psychometrische Erfassung von Körperbeschwerden und ihrer kognitiven, affektiven und behavioralen Begleitsymptome: Rational und Instrumentenentwicklungen

Dissertation

zur Erlangung des Doktorgrades

an der Universität Hamburg,

Fakultät für Psychologie und Bewegungswissenschaft,

Institut für Psychologie

vorgelegt von

Dipl.-Psych. Benjamin Gierk

aus Hannover

Hamburg, 2017

Vorsitz:

Prof. Dr. Alexander Redlich

Dissertationsgutachten:

1. Prof. Dr. Yvonne Nestoriuc

2. Prof. Dr. Bernd Löwe

Disputationsgutachten:

1. Dr. Marcel Riehle

2. Prof. Dr. Steffen Moritz

Tag der mündlichen Prüfung: 31.08.2017

Danksagungen

Ein großer Dank gebührt Herrn Prof. Dr. Bernd Löwe, ohne den diese Doktorarbeit nicht entstanden wäre. Ich bin dankbar für Ihre wohlwollende Unterstützung bei meinen wissenschaftlichen Bemühungen, für Ihr stets „offenes Ohr“ und für die klinischen Einblicke, die Sie mir eröffnet haben. Sie haben mir ermöglicht, in Wissenschaft und Klinik „Wurzeln zu schlagen“ und mir Orientierung gegeben, wenn ich sie brauchte.

Weiterer Dank gilt Frau Prof. Dr. Yvonne Nestoriuc, ohne die diese Arbeit niemals ein Ende gefunden hätte. Sie haben mir ermöglicht, dieses Projekt doch noch abzuschließen.

Danken möchte ich auch Herrn Dr. Sebastian Kohlmann, mit dem ich drei Jahre ein Büro am Universitätsklinikum Hamburg-Eppendorf teilte. Während dieser Zeit entstand ein Großteil der hier versammelten Forschungsarbeiten. Ich bin dankbar für die kreativen, arbeitsreichen Stunden und deine freundschaftliche Begleitung.

Meiner Familie und meinen FreundInnen gilt ein besonderer Dank. Ihr habt unverbrüchlich daran geglaubt, dass ich diese Doktorarbeit abschließen werde und wolltet euch nicht damit abfinden, dass ich sie unvollendet „ad acta“ lege. Danke dafür!

Ebenfalls danken möchte ich allen KollegInnen, die gemeinsam mit mir an den hier versammelten Studien gearbeitet haben. Dankbar bin ich auch allen Personen, die an unseren Studien teilgenommen und damit einen unerlässlichen Beitrag geleistet haben.

Inhaltsverzeichnis

1 Zusammenfassung....1
2 Theoretischer Hintergrund.... 4
2.1 Begriffsbestimmung.... 4
2.2 Epidemiologie von Körperbeschwerden.... 4
2.2.1 Häufigkeit von Körperbeschwerden.... 4
2.2.2 Art der Körperbeschwerden....6
2.2.3 Komorbidität von Körperbeschwerden, Depressivität und Ängstlichkeit7
2.2.4 Zeitlicher Verlauf von Körperbeschwerden....8
2.2.5 Zusammenfassung der Epidemiologie.... 10
2.3 Ableitung der Fragestellung.... 11
2.3.1 Einsatz von Körperbeschwerdenfragebögen....11
2.3.2 Kritische Übersicht über Körperbeschwerdenfragebögen....14
2.3.3 Beschreibung des PHQ-15....16
2.3.4 Notwendigkeit einer Kurzversion des PHQ-15....19
2.3.5 Erfassung des Somatic Symptom Disorder B-Kriteriums....20
3 Darstellung der Einzelstudien.... 22
3.1 Studie 1: Psychometrische Überprüfung der SSS-8....22
3.2 Studie 2: Empirischer Vergleich von PHQ-15 und SSS-8....25
3.3 Studie 3: Verlaufsmessungen mit der SSS-8....28
3.4 Studie 4: Entwicklung der Somatic Symptom Disorder B-Criteria Scale….32
4 Diskussion.... 37
4.1 Zusammenfassung der Einzelstudien....37
4.2 Limitationen....39
4.3 Schlussfolgerung und Ausblick.... 40
5 Literaturverzeichnis.... 43
6 Anhang.... 54

1 Zusammenfassung

Körperbeschwerden treten häufig auf und sind mit Funktionseinschränkungen, psychischer Belastung, vermehrten Krankheitstagen, erhöhter Inanspruchnahme medizinischer Versorgung und dem Verlust an Lebensqualität verbunden. Ihre genaue Erhebung und Verlaufsbeurteilung sind daher von besonderer Bedeutung für die klinische Versorgung und Forschung. Diese Dissertation versammelt vier Einzelstudien zur psychometrischen Erfassung von Körperbeschwerden und ihrer kognitiven, affektiven und behavioralen Begleitsymptome.

Die erste Studie stellt die Entwicklung und psychometrische Überprüfung einer Kurzversion des weit verbreiteten Körperbeschwerdenfragebogens PHQ-15 vor. Sie zeigt, dass die Kurzversion (SSS-8) eine hohe Reliabilität aufweist und erbringt Nachweise für verschiedene Aspekte der Validität. Überdies ermittelt sie Cut-off-Werte zur Schweregradklassifikation und Normwerte aus der deutschen Allgemeinbevölkerung, die die Interpretation der Messwerte ermöglichen.

Daran anknüpfend nimmt die zweite Studie einen direkten empirischen Vergleich der Langversion PHQ-15 und der Kurzversion SSS-8 in einer klinischen Stichprobe vor. Sie legt dar, dass die SSS-8 eine adäquate Kurzversion des PHQ-15 darstellt und zieht bei dieser Beurteilung die Reliabilität, verschiedene Aspekte der Validität sowie die Prozentrangverteilung

der Messwerte heran. Weiteren Forschungsbedarf veranschlagt sie für die Schweregradklassifikation, bei der sie Abweichungen zwischen den beiden Instrumenten feststellt.

Die dritte Studie untersucht, ob sich die SSS-8 für die Messung von Veränderungen der Körperbeschwerden eignet und ermittelt die minimale klinisch bedeutsame Veränderung. Hierfür analysiert sie die Messwertveränderungen bei PatientInnen mit psychischen Störungen, die eine evidenzbasierte Behandlung erhielten. Sie weist nach, dass es sich bei der SSS-8 um ein änderungssensitives Messinstrument handelt, bei dem ab einer Reduzierung von 4 Punkten von einem klinisch bedeutsamen Ergebnis gesprochen werden kann.

Die vierte Studie greift die im DSM-5 neu etablierte Diagnose Somatic Symptom Disorder auf. Das für diese Diagnose notwendige Vorliegen von Körperbeschwerden (A-Kriterium) kann mit der SSS-8 beurteilt werden. Jedoch benötigt es für die Beurteilung des B-Kriteriums (kognitive, affektive und behaviorale Symptome, die auf Körperbeschwerden bezogen sind) ein weiteres Messinstrument. Die Studie stellt daher die Entwicklung und psychometrische Überprüfung eines solchen Fragebogens (SSD-12) dar. Die Ergebnisse zeigen eine hohe Reliabilität und erste Befunde zur Validität. Weiterhin werden Vergleichswerte von PatientInnen aus einer psychosomatischen Ambulanz berichtet.

Insgesamt stellt die Dissertation die Entwicklung und Überprüfung von zwei neuen Fragebögen zu Körperbeschwerden und deren kognitiven, affektiven und behavioralen Begleitsymptomen vor. Die Erfassung dieser Konstrukte ist von zentraler Bedeutung in vielen Bereichen der klinischen Versorgung und Forschung, im besonderen Maße bei der Diagnose Somatic Symptom Disorder.

Stichworte: Körperbeschwerden, SSS-8, SSD-12, Somatic Symptom Disorder

2 Theoretischer Hintergrund

2.1 Begriffsbestimmung

Das Auftreten somatischer Symptome ist ein weit verbreitetes Phänomen. Somatische Symptome können als von Betroffenen wahrgenommene Abweichungen vom körperlichen „Normalzustand“ verstanden werden (Sharpe, 2013), wobei der Normalzustand keine zeitstabile, allgemeingültige, physikalisch-objektive und zweifelsfrei bestimmbare Größe darstellt. Im medizinischen und psychologischen Kontext handelt es sich bei diesen wahrgenommenen Abweichungen in der Regel um Einschränkungen der Körperfunktionen, Veränderungen der Körperstruktur und um Schmerz- oder Missempfindungen (Kroenke, Arrington, & Mangelsdorff, 1990; Sharpe, 2013). Da somatische Symptome von den betroffenen Personen meist als subjektiv belastend erlebt werden, wird in der Folge der Begriff „Körperbeschwerden“ verwendet. Körperbeschwerden sind also von den Betroffenen wahrgenommene und als belastend, einschränkend, beeinträchtigend und unerwünscht bewertete Körperempfindungen (Kroenke, 2014).

2.2 Epidemiologie von Körperbeschwerden

2.2.1 Häufigkeit von Körperbeschwerden

Körperbeschwerden treten im Rahmen vieler somatischer Erkrankungen wie Krebs, Herzerkrankungen, Diabetes, Lungenerkrankungen und Arthritis und psychischer Störungen wie Depressionen, Angststörungen und somatoforme

Störungen auf, wo sie im Regelfall mit Einschränkungen bei der Alltagsbewältigung, ökonomischen Einbußen und einem Verlust von Lebensqualität verbunden sind (Katon, Lin, & Kroenke, 2007; Kohlmann, Gierk, Hümmelgen, Blankenberg, & Löwe, 2013; Kroenke, 2003; Kroenke, Zhong, et al., 2010; Kroenke, Jackson, & Chamberlin, 1997; Rief, Hessel, & Brähler, 2001; Rief, Martin, Klaiberg, & Brähler, 2005; Roelen, Koopmans, & Groothoff, 2010; Simon, von Korff, Piccinelli, Fullerton, & Ormel, 1999). Sie sind der Hauptanlass für die Kontaktaufnahme zum medizinischen Versorgungssystem. So berichten Khan, Khan, Harezlak, Wanzhu und Kroenke (2003), dass etwa die Hälfte der Kontakte zum medizinischen Versorgungssystem der USA aufgrund von Körperbeschwerden erfolgt. Kroenke und Mangelsdorff (1989) beobachteten im Rahmen einer Studie drei Jahre lang eintausend PatientInnen, die in einer ambulanten allgemeinärztlichen Versorgungseinrichtung für Armeeangehörige aktenkundig waren. Sie fanden, dass 38% dieser Personen innerhalb des dreijährigen Studienzeitraums mindestens eine Körperbeschwerde ärztlich schilderten. Eine aktuelle repräsentative Allgemeinbevölkerungsstudie aus Dänemark (Eliasen et al., 2016) zeigt eine sehr hohe Zweiwochenprävalenz für Körperbeschwerden: Sie berichten, dass 95% von den etwa 36.000 Teilnehmenden innerhalb der letzten 14 Tage durch mindestens eine Körperbeschwerde „leicht“ belastet waren. 40% der Untersuchten berichteten sogar, dass sie von einer oder mehreren Körperbeschwerden „deutlich“ belastet waren und 76% gaben an, dass sie in dem Zweiwochenzeitraum drei oder mehr verschiedene Körperbeschwerden erlebten.

2.2.2 Art der Körperbeschwerden

Häufige Körperbeschwerden sind Erschöpfung, muskuloskelettale Beschwerden (z.B. Rücken- und Gelenkschmerzen), Kopfschmerzen, kardiopulmonale Beschwerden (z.B. Herzrasen, Kurzatmigkeit), Schlafstörungen und gastrointestinale Beschwerden (z.B. Magenschmerzen, Verdauungsbeschwerden) (Eliasen et al., 2016; Kocalevent, Hinz, & Brähler, 2013; Kroenke et al., 1990; Rief et al., 2001), wobei etwa Dreiviertel der PatientInnen in der Primärmedizin mehr als ein Symptom zum selben Zeitpunkt beklagen (Kroenke et al., 1990). Kroenke (2003, 2014) und Nimnuan, Hotopf und Wessely (2001) geben an, dass etwa 50-70% der von PatientInnen vorgetragenen Körperbeschwerden durch eine organische Ätiologie erklärt werden können. Der Rest der Körperbeschwerden gelte als „medizinisch nicht erklärbar“ (englisch „medically unexplained symptoms”), „idiopathisch“, „somatisiert“, „funktionell“, „psychisch bedingt“ oder trage je nach Fachdisziplin unterschiedliche syndromale Bezeichnungen, wie „Fibromyalgiesyndrom“, „Reizdarmsyndrom“ oder „chronisches Erschöpfungssyndrom“ (Khan et al., 2003; Kroenke, 2003, 2014, 2016). Für Deutschland ermittelten Rief et al. (2001) in einer repräsentativen Allgemeinbevölkerungsstudie, dass bei den etwa zweitausend Befragten im Mittel drei medizinisch nicht erklärbare Körperbeschwerden innerhalb der letzten zwei Jahre auftraten.

Die dichotome Sichtweise von entweder organisch oder psychisch bedingten Körperbeschwerden geriet in den letzten zwei Jahrzehnten immer mehr in die Kritik, da Studien zeigten, dass es sich bei Körperbeschwerden um

ein multikausales Phänomen handelt, bei dem sich wechselseitig bedingende organische, psychische und soziale Einflussgrößen eine Rolle spielen (Dimsdale et al., 2013; Kroenke, 2003; Sharpe, 2013; Voigt et al., 2010, 2012). Das Konzept der organischen Nicht-Erklärbarkeit von Körperbeschwerden war ein kontrovers diskutierter Gegenstand bei der Revision der somatoformen Störungen in der fünften Auflage des Diagnostischen und Statistischen Manuals für psychische Störungen (Dimsdale et al., 2013). Letztendlich wurde die organische Nicht-Erklärbarkeit von Körperbeschwerden zugunsten der Einführung psychologischer Kriterien, die auf Körperbeschwerden bezogen sind, fallengelassen. Eine weitere Auseinandersetzung mit diesem Aspekt findet im Abschnitt „Erfassung des Somatic Symptom Disorder B-Kriteriums“ und in der vierten Studie dieser Dissertation statt.

2.2.3 Komorbidität von Körperbeschwerden, Depressivität und Ängstlichkeit

Kroenke (2003) berichtet, dass das Vorliegen medizinisch nicht erklärbarer Körperbeschwerden im klinischen Kontext häufig mit dem Vorliegen von Angststörungen und depressiver Störungen einher gehe. Nimnuan et al. (2001) berichten ebenfalls eine hohe Prävalenz von Angststörungen und depressiven Störungen, wenn bei Personen eine großen Anzahl von Körperbeschwerden vorliegen. Löwe et al. (2008) bestätigten das häufige gemeinsame Auftreten von Körperbeschwerden, Ängsten und Depressivität im Rahmen einer groß angelegten Studie, bei der etwa zweitausend PatientInnen aus fünfzehn

primärmedizinischen Versorgungseinrichtungen der USA mittels Fragebögen untersucht wurden: Die AutorInnen fanden heraus, dass bei 54% der PatientInnen, die ein hohes Ausmaß an Körperbeschwerden angaben, ebenfalls ein hohes Ausmaß an Ängstlichkeit und / oder Depressivität vorlag. Löwe et al. bezeichneten dies als Somatisierungs-Angst-Depressions-Triade (vgl. Kroenke, Spitzer, Williams, & Löwe, 2010; Löwe et al., 2008).

Kohlmann, Gierk, Hilbert, Brähler und Löwe (2016) untersuchten im Rahmen einer aktuellen repräsentativen Allgemeinbevölkerungsstudie in Deutschland, welchen relativen Einfluss Ängstlichkeit, Depressivität und Körperbeschwerden (unabhängig von ihrer Ätiologie) auf die Inanspruchnahme medizinischer Versorgung haben. Sie fanden heraus, dass depressive oder ängstliche Stimmung allein nur eine geringe Vorhersagekraft für die Inanspruchnahme von Versorgung aufwiesen. Die größte Vorhersagekraft für die Inanspruchnahme von Versorgung fanden die AutorInnen für die Prädiktoren „Körperbeschwerden“ und „Komorbidität aus Ängstlichkeit, Depressivität und Körperbeschwerden“. Dies deutet eine zentrale Rolle von Körperbeschwerden beim Aufsuchen medizinischer Versorgungseinrichtungen an.

2.2.4 Zeitlicher Verlauf von Körperbeschwerden

Zum zeitlichen Verlauf von Körperbeschwerden geben Khan et al. (2003) an, dass bei der überwiegenden Mehrheit (etwa 75%) ihrer aus der Allgemeinmedizin stammenden, naturalistischen Stichprobe Körperbeschwerden zwölf Monate nach der ersten Schilderung nicht mehr vorhanden waren. Nur bei etwa

einem Viertel der in die Studie eingeschlossenen Personen zeigte sich beim 12-Monats-Follow-Up eine Persistenz der Körperbeschwerden. Als Risikofaktor für eine solche Chronifizierung berichten die AutorInnen, dass die Körperbeschwerden bereits bei einem früheren Arztkontakt angegeben wurden. Das Vorliegen organischer Erkrankungen stellte in der Studie ebenfalls einen Risikofaktor für die Chronifizierung der Körperbeschwerden dar. Creed, Davis, Jackson, Littlewood, Chew-Graham, Tomenson et al. (2012) identifizierten weitere Risikofaktoren für die Chronifizierung von Körperbeschwerden. Sie führten dafür eine große, zwölf Monate andauernde Beobachtungsstudie mit Teilnehmenden aus der Allgemeinbevölkerung Englands durch. Als Risikofaktoren identifizierten sie einen niedrigeren Bildungsstatus (weniger als 12 Bildungsjahre), keine Partnerschaft (getrennt, verwitwet, geschieden lebend), psychischer Missbrauch in der Kindheit, das Vorliegen einer organischen Erkrankung sowie hohe Ängstlichkeit und Depressivität. In der Studie ging die Persistenz von Körperbeschwerden mit einer reduzierten gesundheitsbezogenen Lebensqualität und Funktionseinschränkungen einher.

Zu vergleichbaren Zahlen zum Verlauf von Körperbeschwerden kommen auch Jackson und Passamonti (2005). In einer großen Beobachtungsstudie analysierten sie den Verlauf von Körperbeschwerden bei 500 PatientInnen. Sie fanden, dass bei 70% der Stichprobe die Beschwerden nach 14 Tagen gebessert oder sogar abgeklungen waren. Nach drei Monaten waren es bereits 80%, bei denen eine Besserung oder Symptomfreiheit eintrat. Vor dem Hintergrund der häufig schnellen Rückbildung vieler Körperbeschwerden warnt

Kroenke (2014, 2016) vor einer „Übermedikalisierung“ (z.B. wiederholte, invasive diagnostische Prozeduren), sofern bei der Erstkonsultation kein Hinweis auf eine schwerwiegende körperliche Erkrankung gefunden wurde.

2.2.5 Zusammenfassung der Epidemiologie

In der Zusammenschau der oben aufgeführten Studien kann zur Epidemiologie von Körperbeschwerden festgehalten werden, dass Körperbeschwerden häufig und meist zu mehreren auftreten, subjektiv belastend sind, mit beruflichen Einschränkungen assoziiert sind, zu einer Reduzierung von Lebensqualität und Funktionsfähigkeit beitragen, nicht selten mit Ängstlichkeit und Depressivität einher gehen, zur Inanspruchnahme medizinischer Versorgung führen, zu einem großen Teil innerhalb von zwei Wochen remittieren, jedoch bei etwa einem Viertel der PatientInnen einen chronischen Verlauf nehmen. Je nach Studie können etwa die Hälfte bis zwei Drittel der Körperbeschwerden einer organischen Ursache zugeordnet werden. Der Rest wird in der klinischen Praxis häufig als „medizinisch nicht erklärbar“ oder mit anderen Labels gekennzeichnet. Dies weist auf eine dichotome Sichtweise „organischer“ oder „psychischer“ Ursachen hin, die in die Kritik geraten ist. Vielmehr können für das Auftreten und die Persistenz von Körperbeschwerden organische, psychische und soziale Einflussgrößen identifiziert werden.

2.3 Ableitung der Fragestellung

2.3.1 Einsatz von Körperbeschwerdenfragebögen

Damit Körperbeschwerden durch eine Behandlung angemessen adressiert werden können, müssen sie zunächst einmal erkannt und präzise erfasst werden. Dafür stehen neben dem Anamnesegespräch auch Fragebögen zur Verfügung. Diese standardisierten Selbstberichte werden im englischen Sprachraum als „patient reported outcome measures (PROMs)“ bezeichnet (Black, 2013). Dabei handelt es um psychometrische Messinstrumente, die Aufschluss über die Art, Häufigkeit und subjektive Belastung durch Körperbeschwerden geben und diese Informationen numerisch quantifizieren (Kroenke, Monahan, & Kean, 2015; Zijlema et al., 2013).

In der klinischen Versorgung können Fragebögen als Screeninginstrumente eingesetzt werden, das heißt allen PatientInnen vorgelegt werden, um zeitökonomisch die PatientInnen mit einem hohen Ausmaß an Körperbeschwerden zu identifizieren. Diese PatientInnen können dann eingehender untersucht werden. Das Screening mittels Fragebögen hat den Vorteil, dass standardisiert weitverbreitete Körperbeschwerden erfragt werden, die über den Fokus der jeweiligen medizinischen Fachdisziplin hinaus gehen können. Eine Studie von Dowrick und KollegInnen (2009) zeigte, dass der Einsatz von Screeningfragebögen (im Fall ihrer Studie Depressionsfragebögen in Allgemeinarztpraxen) von den PatientInnen positiv bewertet wurde. Die PatientInnen berichteten, dass Fragebögen für sie eine effektive und strukturierte Ergänzung

zum Anamnesegespräch seien. Der Einsatz von Fragebögen hinterließe bei ihnen den Eindruck, dass die BehandlerInnen sie ernst nehmen und sich bemühten, die Gesundheitsprobleme genau zu erfassen.

Werden Fragebögen bei einer Wiederholungsmessung eingesetzt, kann die numerische Veränderung zur Ausgangsmessung Informationen über den Symptom- und Belastungsverlauf bereitstellen (Kroenke et al., 2015). Dies kann beispielsweise Aufschluss über das Ansprechen auf eine Therapie geben, setzt aber voraus, dass die Änderungssensitivität des Messinstrumentes vorher nachgewiesen wurde. Änderungssensitivität bedeutet, dass ein Messinstrument in der Lage ist, eine in der Realität erfolgte Veränderung abzubilden (Stratford & Riddle, 2005). Ein änderungssensitiver Fragebogen sollte eine numerische Differenz anzeigen, wenn PatientInnen diesen ausgefüllt haben, bevor und nachdem sie eine evidenzbasierte Behandlung erhalten haben. Ob ein Fragebogen änderungssensitiv ist, wird im Rahmen einer psychometrischen Studie bei der Etablierung des Fragebogens ermittelt. Ist ein Fragebogen änderungssensitiv und sein minimaler klinisch bedeutsamer Unterschied bekannt, erlaubt dies den BehandlerInnen Rückschlüsse darüber, ob für die PatientInnen ein bedeutsames, positives Behandlungsergebnis erzielt wurde. Der minimale klinisch bedeutsame Unterschied wird ebenfalls in einer psychometrischen Studie ermittelt und bezeichnet einen Schwellenwert, bei dessen Überschreitung von einer für die PatientInnen bedeutsamen, positiven Veränderung ausgegangen werden kann (Copay, Subach, Glassman, Polly, & Schuler, 2007).

Standardisierte Verlaufsmessungen mit Fragebögen sind ebenfalls von großem Interesse bei der Qualitätssicherung in Versorgungseinrichtung. Sie können wertvolle Informationen darüber liefern, ob im Schnitt die erwarteten Therapieeffekte erreicht werden. So können auf Gruppenebene Effektstärken der Behandlung berechnet werden und diese mit Effektstärken aus wissenschaftlichen Studien verglichen werden. Für die BehandlerInnen dienen die Verlaufswerte außerdem als gutes Feedback über die Wirksamkeit ihrer Behandlung, was gerade bei der Aus- und Weiterbildung besonders nützlich sein kann. Die Verlaufswerte können auch bei der Beurteilung helfen, ob eine Optimierung der Versorgungseinrichtung die gewünschte Verbesserung für die PatientInnen erbracht hat.

In der Kommunikation mit PatientInnen können die Fragebogenwerte dazu genutzt werden, den Schweregrad der Belastung allgemeinverständlich mitzuteilen. Über die Prozentrangnormierung kann den PatientInnen mitgeteilt werden, dass sie „in einer Reihe von 100 Menschen, die dem Schwergrad nach aufgereiht sind, an Platz x stehen“. Diese Information lässt sich ebenfalls einfach in Form von Punktdiagrammen darstellen (100 Punkte, von denen x Punkte ausgemalt sind), was zu einer höheren Verständlichkeit führt (Galesic, Garcia-Retamero, & Gigerenzer, 2009) und in einer Studie zur Mitteilung von Depressionsfragebogenergebnissen eingesetzt wurde (Löwe et al., 2016).

Nicht wegzudenken sind Fragebögen zur Erfassung von Körperbeschwerden in wissenschaftlichen Studien (Zijlema et al., 2013). In diesen ist es unumgänglich, die subjektive Symptombelastung zu quantifizieren. Selbst wenn

ökonomisch die Möglichkeit bestünde, mit allen Teilnehmenden ein ausführliches Interview zu Körperbeschwerden durchzuführen, so löst dies nicht die Problematik einer Quantifizierung der Angaben. Da die Belastung durch Körperbeschwerden in der Regel subjektiver Natur ist, zeigt sich hier der deutliche Vorteil von Fragebögen, da diese per sé die subjektive Perspektive der Beantwortenden erfassen (Black, 2013; Kroenke et al., 2015).

2.3.2 Kritische Übersicht über Körperbeschwerdenfragebögen

Eine aktuelle Übersichtsarbeit von Zijlema et al. (2013) führte eine Bestandsaufnahme der Fragebögen zur Erfassung von Körperbeschwerden durch. Die AutorInnen sichteten 99 Studien, in denen Körperbeschwerden mit Fragebögen erfasst wurden. Sie fanden vierzig verschiedene Instrumente, die sich stark hinsichtlich der beinhalteten Symptome, des Antwortformates, der Länge, der verfügbaren Übersetzungen, der Zielpopulationen und der psychometrischen Gütekriterien unterschieden. Sie gingen weiterhin der Frage nach, welche dieser vierzig Fragebögen vor allem weit verbreitete Körperbeschwerden erfassen, ohne dabei einen Bezug zur Ätiologie der Symptome herzustellen, da sie sich so für viele Populationen und Fragestellungen eignen. Die AutorInnen operationalisierten das Konstrukt „weit verbreitete Körperbeschwerden“ durch den Einbezug von Vorstudien. Sie legten fest, dass Fragebögen Items zu kardiopulmonalen, gastrointestinalen, muskuloskelettalen und allgemeinen Symptome enthalten sollten. Außerdem beurteilten sie die Validität, Reliabilität und Testökonomie der Messinstrumente.

Ziejlema et al. (2013) fanden in ihrer Übersichtsarbeit, dass nur zwei der vierzig untersuchten Fragebögen ihren Anforderungen an die Testgütekriterien und die Symptombereiche entsprachen. Dabei handelt es sich um die Somatisierungs-Subskala der Symptom Checklist 90 (SCL-90 SOM) von Derogatis, Lipmann und Covi (1973) und den Patient Health Questionnaire – 15 (PHQ-15) von Kroenke, Spitzer und Williams (2002). Sechs der in beiden Instrumenten enthaltenen Items ähnelten sich stark, wobei die SCL-90 SOM entwicklungsgeschichtlich einen Fokus auf Körperbeschwerden legte, die im Kontext psychischer Belastung auftraten (Derogatis et al., 1973; Zijlema et al., 2013). Demgegenüber beinhaltet der PHQ-15 fünfzehn Symptome, die die größte Häufigkeit in der medizinischen Primärversorgung der USA haben (Kroenke et al., 2002; Kroenke, Spitzer, et al., 2010; Zijlema et al., 2013). Im direkten Vergleich lässt dies den PHQ-15 als vorteilhaft erscheinen, wenn es um die Erfassung häufiger Körperbeschwerden bei weit gefassten Populationen geht (Zijlema et al., 2013). Ein weiterer Vorteil des PHQ-15 gegenüber der SCL-90 SOM besteht darin, dass für den wissenschaftlichen und klinischen, nicht kommerziellen Einsatz keine Nutzungskosten in Form von Lizenzgebühren anfallen. Demgegenüber belaufen sich die Nutzungskosten der aktuellen Testversion SCL-90 R auf 1,30 Euro pro Person und Testung (Quelle: https://www.testzentrale.de/shop/symptom-checklist-90r-standard.html, abgerufen am 20. November 2016). Gerade für Studien mit großen Stichproben oder hoch frequentierte Versorgungseinrichtungen spielt dies eine zentrale Rolle bei der Entscheidung für ein Testverfahren (Kroenke et al., 2015).

2.3.3 Beschreibung des PHQ-15

Der PHQ-15 erfragt, wie stark die Beantwortenden innerhalb der letzten vier Wochen durch fünfzehn weit verbreitete Körperbeschwerden belastet waren. Die beinhalteten Körperbeschwerden decken dabei 90% der Körperbeschwerden ab, die in der primärmedizinischen Versorgung von PatientInnen geäußert werden, sofern Beschwerden der oberen Atemwege (z.B. Husten, Halsschmerzen) außer Acht gelassen werden (Kroenke et al., 2002; Kroenke, Spitzer, et al., 2010). Als Antwortformat liegt eine dreistufige Likertskala (0 „nicht beeinträchtigt“, 1 „wenig beeinträchtigt“, 2 „stark beeinträchtigt“) vor. Die Antworten werden zu einem Summenwert zwischen 0 und 30 Punkten aufaddiert. Cut-off-Werte von 5, 10 und 15 Punkten geben an, ob es sich insgesamt um eine geringe, milde, mittelgradige oder schwere somatische Symptombelastung handelt. Der PHQ-15 weist eine hohe interne Konsistenz von Cronbachs $\alpha = 0{,}80$ (Kroenke et al., 2002) und eine hohe 14-Tage-Test-Retest-Reliabilität von $r_{icc} = 0{,}83$ (van Ravesteijn et al., 2009) auf. Bezüglich der Validität zeigt er Zusammenhänge mit Funktionseinschränkungen, Inanspruchnahme medizinischer Versorgung, Arbeitsunfähigkeitstagen und psychischer Belastung in Form von Depressivität und Ängstlichkeit (Kroenke et al., 2002; Kroenke, Spitzer, et al., 2010). In klinischen Studien erwies sich der PHQ-15 als sensitiv für Veränderungen (Kroenke, Spitzer, et al., 2010), was ihn zu einem geeigneten Instrument für Verlaufsmessungen macht.

Angaben zur Dimensionalität des Messinstrumentes machen Kroenke, Spitzer, DeGruy und Swindle (1998). Sie unterzogen PHQ-15-Daten von eintausend PatientInnen aus der Primärversorgung einer exploratorischen Faktorenanalyse mit schiefwinkeliger Rotation und fanden drei hoch korrelierte latente Dimensionen für 1. kardiopulmonale, 2. gastrointestinale und 3. allgemeine Schmerz- und Erschöpfungssymptome. In einer späteren Studie untersuchten Lee, Ma und Tsang (2011) etwa dreitausend Personen aus der Allgemeinbevölkerung Hong Kongs mit dem PHQ-15. In einer Hauptkomponentenanalyse mit Varimax-Rotation fanden sie vier orthogonale Komponenten, die sie als 1. kardiopulmonal, 2. gastrointestinal, 3. Schmerz und Erschöpfung und 4. neurologisch bezeichneten. Fink, Toft, Hansen, Ørnbøl und Olesen (2007) wendeten eine Hauptkomponentenanalyse mit Varimax-Rotation auf einen Fragebogen an, der ebenfalls acht PHQ-15-Items enthielt. Die Analyse lieferte drei Komponenten, die die AutorInnen als 1. kardiopulmonal, 2. muskuloskelettal und 3. gastrointestinal bezeichneten. Vier latente Dimensionen sowie einen zusätzlichen Generalfaktor fanden Witthöft, Hiller, Loch und Jasper (2013). Sie wendeten konfirmatorische Faktorenanalysen auf PHQ-15-Daten aus der Allgemeinbevölkerung und der medizinischen Primärversorgung an und testeten verschiedene, in der Literatur diskutierte Modelle hinsichtlich ihrer Modellpassung. Die beste Passung erzielte ein Modell mit vier latenten Dimensionen für 1. kardiopulmonale, 2. gastrointestinale, 3. Schmerz- und 4. Erschöpfungssymptome mit einem zusätzlichen Generalfaktor (bi-Faktormodell).

Insgesamt weist der PHQ-15 also eine multidimensionale Struktur auf. Inwiefern dieser multidimensionale Struktur durch einen einfachen Summenwert der fünfzehn Einzelantworten angemessen Rechnung getragen werden kann, ist nicht abschließend geklärt. Beispielsweise berichten Eliasen et al. (2016), dass in der Wechselwirkung zwischen Körperbeschwerden zusätzliche Informationen stecken. Mit Hilfe komplexer statistischer Modelle zeigten sie, dass die statistischen Zusammenhänge einzelner Körperbeschwerden mit den Outcomes „selbstberichteter Gesundheitszustand“ und „gesundheitliche Funktionseinschränkungen“ davon abhängig waren, welche anderen Körperbeschwerden ebenfalls präsent waren (Interaktion). Ein einfaches „Auszählen“ der Körperbeschwerden unterschlage den AutorInnen zufolge wichtige Informationen. In der klinischen und wissenschaftlichen Praxis erscheint die Fragebogenauswertung durch das Bilden eines einfachen Summenwertes jedoch pragmatisch und ist daher weit verbreitet.

Für den internationalen Einsatz des PHQ-15 liegen mittlerweile viele Übersetzungen vor, unter anderem ins Deutsche (Löwe, Spitzer, Zipfel, & Herzog, 2002), Spanische (Ros Montalbán, Comas Vives, & Garcia-Garcia, 2010), Schwedische (Nordin, Palmquist, & Nordin, 2013), Niederländische (Persoons, Luyckx, Desloovere, Vandenberghe, & Fischler, 2003; van Ravesteijn et al., 2009), Chinesische (Lee et al., 2011; Zhang et al., 2016), Koreanische (Han et al., 2009), Ungarische (Szemerszky, Köteles, Lihi, & Bárdos, 2010), Arabische (Daradkeh, Eapen, & Ghubash, 2005) und Griechische (Hyphantis et al., 2014).

2.3.4 Notwendigkeit einer Kurzversion des PHQ-15

Obwohl es sich beim PHQ-15 bereits um ein kurzes Messinstrument handelt, ergab sich im Zuge der Überarbeitung des Diagnostischen und Statistischen Manuals für psychische Störungen (DSM-5) die Notwendigkeit der Entwicklung einer Kurzversion. Anlass war das Etablieren der neuen Diagnose Somatic Symptom Disorder, dessen A-Kriterium das Vorliegen von Körperbeschwerden erfordert (Dimsdale et al., 2013). Zur Etablierung der neuen Diagnose wurden Großstudien („field trials“) durchgeführt, die empirische Daten über die Diagnose erhoben, beispielsweise die Prävalenz (Narrow et al., 2013). Als Messinstrument für Körperbeschwerden wurde durch die Amerikanische Psychiatrische Fachgesellschaft (APA) der PHQ-15 in Betracht gezogen. Da es bei Großstudien ökonomisch notwendig ist, nur wenig Messzeit beanspruchende Erhebungsmethoden einzusetzen, wurde eine Kurzversion des PHQ-15 benötigt. Diese wurde im Jahr 2010 durch den Erstautor des PHQ-15 Kroenke erstellt (Narrow et al., 2013). Er entfernte das genderspezifische Item (Menstruationsbeschwerden) und Items mit relativ niedrigen Prävalenzen, geringer Kommunalität mit anderen Items und geringen Zusammenhängen mit Lebensqualität, Funktionsfähigkeit und der Inanspruchnahme von medizinischer Versorgung (sexuelle Probleme, Ohnmachtsanfälle). Fünf weitere hoch korrelierte Items zu kardiopulmonalen und gastrointestinalen Beschwerden fasste er zu zwei Items zusammen. So entstand eine Kurzversion mit insgesamt

acht Items, die unter dem Arbeitstitel „Patient Health Questionnaire Somatic Symptom Short-Form (PHQ-SSS)“ in den field trials eingesetzt wurde (Narrow et al., 2013).

Zwar weist der PHQ-SSS durch seine stringente Ableitung von dem psychometrisch bereits sehr gut evaluierten PHQ-15 eine hohe Augenscheinvalidität auf, jedoch ersetzt dies nicht eine detaillierte psychometrische Überprüfung (Kroenke et al., 2015). Diese Überprüfung ist Gegenstand der ersten drei Studien dieser kumulativen Dissertation.

2.3.5 Erfassung des Somatic Symptom Disorder B-Kriteriums

Die Erstellung des DSM-5 war nicht nur Anlass für die Entwicklung und psychometrische Überprüfung der PHQ-15-Kurzversion, die in Studie 1-3 beschrieben wird. Die Neukonzeption der somatoformen Störungen als Somatic Symptom Disorder machte es notwendig, dass zusätzlich zu Körperbeschwerden (A-Kriterium) darauf bezogene kognitive, affektive und behaviorale Symptome (B-Kriterium) erhoben werden müssen (Dimsdale et al., 2013). Damit verließ das DSM-5 die vorher geltende Vorstellung, dass der Kern der somatoformen Störungen durch das Vorliegen organisch nicht erklärbarer Körperbeschwerden definiert ist, zugunsten der Vorstellung, dass psychische Störungen durch positiv bestimmbare psychische Symptome definiert werden (Voigt et al., 2010, 2012). Eine Somatic Symptom Disorder kann diagnostiziert werden, wenn über mindestens sechs Monate belastende Körperbeschwerden vorliegen (A-Kriterium), deren Ätiologie nicht relevant ist, und die von

ausgeprägtem gedanklichen Kreisen um die Beschwerden (kognitiv), Krankheitsängsten (affektiv) und exzessivem, Zeit und Energie beanspruchendem Krankheitsverhalten (behavioral) begleitet werden (B-Kriterium) (Dimsdale et al., 2013).

Ein Messinstrument für die Erfassung des B-Kriterium der Somatic Symptom Disorder existierte zum Zeitpunkt des Erscheinens des DSM-5 nicht. Zwar lagen Messinstrumente für Teilaspekte des B-Kriteriums vor, wie etwa der Whiteley Index, der kognitiv-affektive Aspekte von „Hypchondrie“ erfasst (Fink et al., 1999). Jedoch gab es kein Messinstrument, das die auf die Körperbeschwerden bezogenen kognitiven, affektiven und behavioralen Begleitsymptome vollumfänglich erfasst. Daher bildete sich eine Arbeitsgruppe mit dem Ziel, direkt aus der DSM-5-Definition des B-Kriteriums theoriegeleitet ein Selbstbeurteilungsverfahren zu entwickeln. Die Entwicklung und psychometrische Überprüfung dieses Fragebogens ist Gegenstand der vierten Studie dieser Dissertation.

3 Darstellung der Einzelstudien

3.1 Studie 1: Psychometrische Überprüfung der SSS-8

Die erste Studie (Gierk et al., 2014) widmet sich der psychometrischen Überprüfung der Kurzversion des PHQ-15, die bereits unter dem früheren Titel Patient Health Questionnaire Somatic Symptom Short-Form (PHQ-SSS) in den DSM-5 field trials eingesetzt wurde (Narrow et al., 2013). In den field trials wurde jedoch als einziger psychometrischer Parameter die Test-Retest-Reliablität von $r_{icc} = 0{,}69$ berichtet (Narrow et al., 2013). Die Beurteilung der internen Konsistenz, der verschiedenen Aspekte der Validität und Dimensionalität sowie die Bestimmung von Normwerten und Schweregradkategorien standen noch aus. Hierfür wurde ein internationales Kooperationsprojekt an den Hochschulen Indianapolis (Kroenke), Leipzig (Brähler, Spangenberg, Zenger) und Hamburg (Gierk, Kohlmann, Löwe) ins Leben gerufen. Die Eigenanteile des Erstautors lagen in der Konzeption der Studie, der statistischen Auswertung und Interpretation der Daten sowie der Erstellung und Überarbeitung des Manuskriptes.

Die Kurzversion, die zur Vereinfachung in Somatic Symptom Scale – 8 (SSS-8) umbenannt wurde, sowie weitere Messinstrumente zu Depressivität, Ängstlichkeit, Wohlbefinden und Inanspruchnahme von Versorgungsleistungen wurden in einer für Deutschland repräsentativen Allgemeinbevölkerungserhebung eingesetzt. Professionelle MitarbeiterInnen eines Marktforschungsunternehmens legten die Messinstrumente rund 2500 Personen als Papier-

Bleistift-Version vor. Die statistische Auswertung der Daten zeigte eine hohe interne Konsistenz der SSS-8 sowie zufriedenstellende Trennschärfekoeffizienten und Itemschwierigkeiten. Der Summenwert als Schätzer für das Ausmaß der Körperbeschwerden zeigte substanzielle Zusammenhänge mit Depressivität, Ängstlichkeit, Wohlbefinden und der Häufigkeit von Arztbesuchen in den letzten 12 Monaten. Dies ist konsistent mit Befunden aus dem Abschnitt zur Epidemiologie von Körperbeschwerden (siehe oben). Konfirmatorische Faktorenanalysen zeigten vier Symptombereiche (Schmerz, kardiopulmonale Beschwerden, gastrointestinale Beschwerden und Erschöpfung), die gemeinsam einen Generalfaktor höherer Ordnung bildeten. Dies ist gut vereinbar mit der oben dargestellten Dimensionalität des PHQ-15. Überdies konnte über die Faktorenanalysen eine faktorielle Invarianz für die Merkmale Geschlecht und Alter gefunden werden, sodass das Messinstrument bei unterschiedlich alten Personen oder Personen unterschiedlichen Geschlechtes keine systematisch unterschiedliche Dimensionalität aufweist.

Für die praktische Anwendung wurden Schweregradkategorien gebildet, die die Beurteilung der Belastung durch Körperbeschwerden ermöglichen (0-3 „keine bis minimal“, 4-7 „niedrig“, 8-11 „mittel“, 12-15 „hoch“, 16-32 „sehr hoch“). Ein Anstieg von einer zur nächsten Schweregradkategorie ging mit einem signifikanten Anstieg von Arztbesuchen einher. Ebenfalls wurden alters- und geschlechtsspezifische Normwerte in Form von Prozenträngen bestimmt.

Die SSS-8 erwies sich in der Überprüfung als insgesamt psychometrisch gut fundiertes Messinstrument, das besonders für den Einsatz in Situationen mit nur wenig Messzeit (z.B. große Studien oder hoch frequentierte medizinische Versorgungseinrichtungen) geeignet ist. Durch die Ableitung von dem frei verfügbaren PHQ-15 kann sie ebenfalls ohne Nutzungsgebühren verwendet werden.

In der klinischen und wissenschaftlichen Praxis wurde die SSS-8 schnell aufgegriffen und rezipiert. Einen ersten Kommentar verfasste Barsky (2014), in dem er Einsatzbereiche und den Nutzen der SSS-8 skizzierte:

> Integration of the SSS-8 into routine, measurement-based care would constitute a substantial advance. An instrument with excellent psychometric properties like those of the SSS-8 is necessary for research to further our understanding of the phenomenology, epidemiology, natural history, and treatment response of somatic symptoms. (S. 408)

Weiterhin nehmen bisher 42 Forschungsarbeiten Bezug auf die SSS-8 (Suchanfrage über Google Scholar und Thomson Reuters Web of Science, Stand 15.02.2017), darunter auch Artikel, die sich speziell an klinisch Tätige richten und die SSS-8 als praktikables Messinstrument zur Erfassung von Körperbeschwerden ausweisen (Croicu, Chwastiak, & Katon, 2014; Kurlansik & Maffei, 2016; Lahmann, Allwang, & Dinkel, 2016). Für den internationalen Einsatz stehen psychometrisch überprüfte englische (Gierk et al., 2014),

deutsche (Gierk et al., 2014) und japanische (Matsudaira et al., 2017) Versionen zur Verfügung (Stand 15.02.2017). Eine Übersetzungsstudie ins Portugiesische ist geplant.

3.2 Studie 2: Empirischer Vergleich von PHQ-15 und SSS-8

Die erste Studie zeigte, dass es sich bei der SSS-8 um einen psychometrisch gut fundierten, kurzen Fragebogen zur Erfassung von Körperbeschwerden handelt. Zwar liegt die Vermutung nahe, dass es sich bei der SSS-8 durch die stringente Ableitung vom PHQ-15 auch um eine psychometrisch adäquate Kurzversion des PHQ-15 handelt, empirisch überprüft wurde dies jedoch nicht. Mögliche Quellen für Abweichungen bestehen zum Beispiel darin, dass die SSS-8 eine fünfstufige anstelle einer dreistufigen Likertskala als Antwortformat verwendet. Außerdem verwendet die SSS-8 ein 7-Tage-Zeitfenster, auf das die Items bezogen sind, wohingegen der PHQ-15 ein 4-Wochen-Zeitfenster vorgibt. Diese Veränderungen der SSS-8 gegenüber dem PHQ-15 erfolgten, um die Items der SSS-8 vergleichbar mit den Items der Messinstrumente der PROMIS-Initiative („Patient-Reported Outcomes Measurement Information System“, http://www.healthmeasures.net) zu machen. Die PROMIS-Initiative entwickelt gezielt Messinstrumente für wichtige medizinische und psychotherapeutische Outcomemaße und greift dabei auch auf neuere messtheoretische Modelle zurück. Ein mit den PROMIS-Messinstrumenten vergleichbares Antwortformat und Zeitfenster wurden angestrebt, um zukünftige Synergie-Effekte nutzen zu können, wie etwa das Eingehen von

SSS-8-Items in eine Itembank für Körperbeschwerden und um Vergleichbarkeit zu den anderen PROMIS-Messinstrumenten in den DSM-5 field trials (Narrow et al., 2013) herzustellen.

Zur Beantwortung der Frage, ob es sich bei der SSS-8 wirklich um eine psychometrisch angemessene Kurzversion des PHQ-15 handelt, entwarfen Gierk et al. (2015) eine zweite Studie, die einen direkten empirischen Vergleich der psychometrischen Gütekriterien beider Instrumente vornimmt. Die Eigenanteile des Erstautors lagen in der Konzeption, Durchführung und Auswertung der Studie sowie in der Erstellung und Überarbeitung der Publikation. Neben dem direkten Vergleich beider Messinstrumente sollte sich die SSS-8 auch in einer klinischen Stichprobe bewähren, da die erste Studie (Gierk et al., 2014) ausschließlich Daten aus der Allgemeinbevölkerung verwendete. Bei Daten aus der Allgemeinbevölkerung kann von einer insgesamt geringeren Beschwerdelast als bei PatientInnen aus einer klinischen Stichprobe ausgegangen werden. Dies merkte auch Barsky (2014) in seinem Kommentar zur SSS-8 kritisch an: „First, the survey was conducted in a community-dwelling sample—not a medical one— and the relationship between their symptoms and those of patients encountered in medical practice requires further investigation, as the authors emphasize. (S. 408)“

Gierk et al. (2015) schlossen 131 PatientInnen in die Studie ein, die sich für einen Erstkontakt in der Psychosomatischen Hochschulambulanz des Universitätsklinikums Hamburg-Eppendorf vorstellten. Die häufigsten Diagnosen in dieser Stichprobe waren nach der ICD-10 Klassifikation

depressive Störungen, somatoforme Störungen und Angststörungen. Während der Wartezeit vor der Konsultation füllten die PatientInnen eine Fragebogenbatterie aus, in der neben Instrumenten zu Demographie, Depressivität, Ängstlichkeit, hypochondrischen Befürchtungen, Inanspruchnahme von Versorgung und gesundheitsbezogener Lebensqualität auch der PHQ-15 und die SSS-8 in getrennter Reihenfolge enthalten waren.

Die Ergebnisse der Studie zeigten, dass beide Fragebögen eine vergleichbar hohe Reliabilität (gemessen über Cronbachs α) aufwiesen, obwohl die SSS-8 (α = 0,76) nur etwa halb so viele Items wie der PHQ-15 (α = 0,80) beinhaltet. Beide Instrumente zeigten außerdem ein sehr ähnliches Muster von Korrelationen mit den Messinstrumenten zu Depressivität, Ängstlichkeit, hypochondrischen Befürchtungen, Inanspruchnahme von Versorgung und gesundheitsbezogener Lebensqualität. Trotz der unterschiedlichen Skalierungen des Antwortformates verfügten beide Fragebögen über eine nahezu identische Prozentrangverteilung der Summenwerte.

Allerdings bestand nur eine geringe Übereinstimmung bei der Schweregradklassifizierung: Die Belastung durch Körperbeschwerden wurde von den beiden Messinstrumenten unterschiedlich bewertet. Dies führen die AutorInnen darauf zurück, dass es für den PHQ-15 drei Cut-off-Werte gibt (5, 10 und 15 Punkte), für die SSS-8 jedoch vier (4, 8, 12 und 16 Punkte). Wurden bei der SSS-8 dieselben Cut-off-Werte wie beim PHQ-15 eingesetzt (5, 10 und 15 Punkte), zeigten beide eine sehr hohe Übereinstimmung bei der Schweregradbeurteilung der PatientInnen. Für den Aspekt der Schweregradklassifikation

merkten die AutorInnen daher kritisch an, dass weitere Untersuchungen auch in anderen Populationen notwenig seien, um zu einer abschließenden Beurteilung zu kommen. Bis dahin sollten bei der Beurteilung des Schweregrades der Belastung die bereits etablierten Cut-off-Werte beider Messinstrumente eingesetzt werden.

In der Zusammenschau der Ergebnisse folgerten die AutorInnen, dass es sich bei der SSS-8 um eine adäquate Kurzversion der PHQ-15 handelt. Überdies erbrachte die Studie den ersten Nachweis, dass sich die SSS-8 auch für den klinischen Einsatz eignet. Für ambulante PatientInnen in der Psychosomatik gibt sie Vergleichswerte in Form von Prozenträngen an.

3.3 Studie 3: Verlaufsmessungen mit der SSS-8

Ein wichtiger Anwendungsbereich für Selbstbeurteilungsfragebögen ist die Verlaufsmessung von Körperbeschwerden (Black, 2013; Kroenke et al., 2015). Verlaufswerte beinhalten wichtige Informationen über das Ansprechen auf Therapie. In randomisierten, kontrollierten Studien (RCT-Studien) werden Verlaufswerte bei PatientInnen verglichen, die per Zufallsziehung unterschiedlichen Interventionen zugeordnet wurden. Auch zur Qualitätssicherung in Versorgungseinrichtungen können Verlaufswerte wichtige Hinweise auf möglichen Optimierungsbedarf liefern.

Damit ein Messinstrument bei Verlaufsmessungen eingesetzt werden kann, ist es notwendig, dass in einer psychometrischen Studie nachgewiesen wird, dass es änderungssensitiv ist. Änderungssensitiv bedeutet, dass das

Messinstrument eine numerische Änderung anzeigt, wenn in der Realität eine Veränderung in dem zu messenden Konstrukt eingetreten ist (Stratford & Riddle, 2005). Auf den klinischen Kontext übertragen heißt dies, dass ein Messinstrument als änderungssensitiv bezeichnet werden kann, wenn es eine Punktwertveränderung (Prä-Post-Differenz) bei PatientInnen anzeigt, die eine Intervention erhalten haben, von der aus Studien bekannt ist, dass sie Beschwerden reduziert.

Ist ein Messinstrument änderungssensitiv, ist ein weiteres wichtiges Maß der minimale klinisch bedeutsame Unterschied (englisch „minimal clinically important difference“, MCID). Der MCID ist ein Schwellenwert der Punktwertveränderung, bei dessen Überschreiten von einer für die PatientInnen klinisch bedeutsamen Reduzierung der Beschwerden ausgegangen werden kann (Copay et al., 2007; Jaeschke, Singer, & Guyatt, 1989). Vereinfacht ausgedrückt gibt der MCID Auskunft darüber, um wie viele Punkte der Summenwert des Fragebogens bei einer Wiederholungsmessung absinken muss, damit von einer „klinisch bedeutsamen“ Reduzierung der Beschwerden ausgegangen werden kann.

Die Studie 3 widmet sich der Bestimmung der Änderungssensitivität und der MCID der SSS-8 (Gierk, Kohlmann, Hagemann-Göbel, Löwe, & Nestoriuc, submitted). Die Eigenanteile des Erstautors bestehen in der Konzeption, Durchführung, Auswertung der Studie sowie in der Manuskripterstellung. Zur Bestimmung der beiden Parameter untersuchten die AutorInnen den Verlauf der SSS-8 Messwerte bei 55 PatientInnen mit psychischen Störungen, die eine

acht- bis zwölfwöchige evidenzbasierte Behandlung in einer psychiatrischen Tagesklinik erhalten haben. Die PatientInnen erhielten zu Beginn und nach Abschluss der Behandlung die SSS-8 sowie weitere Messinstrumente zu Depressivität, Ängstlichkeit und Funktionseinschränkungen. Als Maß der Änderungssensitivität wurde die Effektstärke Cohens d_z berechnet (Lakens, 2013).

Die AutorInnen fanden ein statistisch signifikantes Absinken der SSS-8 Werte, das auf Ebene der Gesamtstichprobe (n = 55) einer mittleren Effektstärke entsprach (vgl. Cohen, 1992). In einer Subgruppenanalyse (n = 11) zeigten die AutorInnen weiterhin, dass PatientInnen mit einem hohen Ausmaß von Körperbeschwerden zu Behandlungsbeginn (Baseline-SSS-8-Werte ≥ 16 Punkte) eine noch ausgeprägtere Reduzierung der Körperbeschwerden erlebten, die einer großen Effektstärke entsprach (vgl. Cohen, 1992). Sie interpretierten dies als erwartungskonformen Beleg dafür, dass die SSS-8 relevant für die Zielgruppe ist, für die sie entwickelt wurde (Personen mit Körperbeschwerden). Weiterhin zeigten die Analysen, dass eine Reduzierung von Körperbeschwerden ebenfalls mit einer Reduzierung von Depressivität, Ängstlichkeit und Funktionseinschränkungen einher ging, was die AutorInnen aufgrund der hohen Assoziation dieser Konstrukte (vgl. Abschnitt zur Epidemiologie von Körperbeschwerden) als Validitätsnachweis deuteten.

Zur Bestimmung der MCID wurden in der Studie messtheoretische Ansätze (Copay et al., 2007) und die Funktionseinschränkung der PatientInnen als Outcomemaß heran gezogen (Jaeschke et al., 1989). Zunächst

berechneten die AutorInnen, welche Punktwertveränderung überhaupt reliabel erfasst werden kann. Hierfür ermittelten sie den Standardmessfehler der SSS-8 und zeigten, dass eine Veränderung um 3 Punkte notwendig ist, um aus dem „statistischen Rauschen“ hervor zu treten. Dann setzen sie die Punktwertveränderungen der SSS-8 in Bezug zur Veränderung der von den PatientInnen berichteten Funktionseinschränkungen. Dabei folgten sie der Annahme, dass eine klinisch bedeutsame Verbesserung der Körperbeschwerden mit einer von den PatientInnen wahrgenommenen Funktionsverbesserung einher gehen müsse. Die Analysen zeigten, dass eine solche Funktionsverbesserung bei einer Reduzierung der SSS-8 – Summenwerte um 4 Punkte gegeben war. Die AutorInnen folgerten, dass eine Reduzierung um 3 Punkte aus dem messtheoretischen Blickwinkel die Schwelle für eine reliabel erfassbare Veränderung darstellt. Für die klinische Anwendung sei jedoch eine Reduzierung um 4 Punkte ein geeigneter Schwellenwert, da dann von einer für die PatientInnen bedeutsamen Funktionsverbesserung ausgegangen werden könne.

Die Studie 3 hat eine große Bedeutung für die praktische Anwendung der SSS-8. Sie liefert den ersten Beleg dafür, dass die SSS-8 psychometrisch geeignet dafür ist, Verläufe von Körperbeschwerden zu beurteilen. Zwar war von ihrer Eignung für die Verlaufsmessung theoretisch auszugehen, da die SSS-8 stringent von dem änderungssensitiven PHQ-15 abgeleitet wurde (Gierk et al., 2015; Kroenke, Spitzer, et al., 2010), ein empirischer Nachweis stand jedoch noch aus. Dieser ist nun mit der Studie 3 erfolgt. Die Studie bietet AnwenderInnen außerdem eine wichtige Orientierung bei der Interpretation der

Verlaufswerte, da sie zeigt, dass erst ab einer Reduzierung von 4 Punkten von einer klinisch bedeutsamen Veränderung ausgegangen werden kann. Dieser Schwellenwert kann zukünftig bei der Beurteilung der Wirksamkeit von Interventionen zur Zieldefinition heran gezogen werden (i.e., eine Intervention sollte mindestens dazu führen, dass der SSS-8 – Summenwert um 4 Punkte absinkt).

3.4 Studie 4: Entwicklung der Somatic Symptom Disorder B-Criteria Scale

Zwar lässt sich das Vorliegen von Körperbeschwerden ohne ätiologischen Bezug (A-Kriterium) durch den PHQ-15 oder die SSS-8 erfassen, jedoch lagen zum Zeitpunkt der vierten Studie noch keine Selbstbeurteilungsinstrumente für die Erfassung der auf die Körperbeschwerden bezogenen kognitiven, affektiven und behavioralen Symptome (B-Kriterium) der Somatic Symptom Disorder vor. Da ein Messinstrument für das B-Kriterium der Somatic Symptom Disorder sowohl für die klinischen AnwenderInnen als auch für die weitere Erforschung der neuen Diagnose von großer Wichtigkeit ist, bildete sich an den Standorten Hamburg (Gierk, Herzog, Löwe, Murray, Toussaint, Voigt), Indianapolis (Kroenke), Marburg (Rief) und München (Henningsen) eine Arbeitsgruppe, mit dem Ziel, theoriegeleitet ein solches Messinstrument zu entwickeln (Toussaint et al., 2016). Die Eigenanteile des Verfassers dieser Dissertation lagen in der Überarbeitung des Itempools, der statischen Analyse der Faktorstruktur des Messinstrumentes sowie der Erstellung und Überarbeitung von Teilen des Manuskriptes.

Ausgehend von der Definition des Somatic Symptom Disorder B-Kriteriums im DSM-5 (Dimsdale et al., 2013) sowie empirischen Vorarbeiten (Herzog et al., 2015) entwickelte die Arbeitsgruppe einen Itempool mit 98 potentiell geeigneten Items. Dieser wurde schrittweise in Fokusgruppendiskussionen mit ExpertInnen und PatientInnen sowie in empirischen Pilotuntersuchungen auf eine Anzahl von zwölf geeigneten Items reduziert. Dabei dienen jeweils vier Items zur Erfassung kognitiver, affektiver und behavioraler Symptome, deren Ausprägung auf einer fünfstufigen Likertskala von den Beantwortenden im Selbstbericht eingeschätzt wird. Die AutorInnen legten die zwölf Items einer großen Stichprobe von 698 PatientInnen der Psychosomatischen Hochschulambulanz am Universitätsklinikum Hamburg-Eppendorf vor. Als Titel für das entstandene Messinstrument wählten sie „Somatic Symptom Disorder - B Criteria Scale“ (SSD-12). Zusätzlich zur SSD-12 bearbeiteten die PatientInnen Messinstrumente zu Körperbeschwerden, Ängstlichkeit, Depressivität, Hypochondrie, Inanspruchnahme medizinischer Versorgung sowie körperlichen und psychischen Funktionseinschränkungen.

Die Analysen der Daten zeigten, dass die Items allesamt hohe bis sehr hohe Trennschärfekoeffizienten und akzeptable Itemmittelwerte und -streuung sowie Schiefe- und Kurtosisparameter aufwiesen. Die Gesamtskala (0 bis 48 Punkte) zeigte eine sehr hohe interne Konsistenz von $\alpha = 0{,}95$. Die internen Konsistenzen der Subskalen (jeweils 0 bis 16 Punkte) lagen zwischen $\alpha = 0{,}72$ und 0,92. Eine konfirmatorische Faktorenanalyse zeigte eine hohe Passung für das aus dem DSM-5 abgeleitete Modell mit drei latenten Dimensionen für

kognitive, affektive und behaviorale Symptome. Die latenten Dimensionen wiesen dabei hohe Interkorrelationen auf. Eine weitere konfirmatorische Faktorenanalyse mit einem Generalfaktor für alle zwölf Items zeigte eine etwas schlechtere aber gerade noch akzeptable Modellpassung. Die AutorInnen folgerten daraus, dass das Messinstrument insgesamt der theoretischen Konzeption des B-Kriteriums der Somatic Symptom Disorder nach DSM-5 entspricht. Sie wiesen aber auch daraufhin, dass alle drei Facetten hoch korreliert sind und daher eine Aggregation der Itemwerte in einem Summenwert gerechtfertigt ist. Der SSD-12-Summenwert wies hohe statistische Zusammenhänge mit körperlicher Funktionseinschränkung, dem Vorliegen von Körperbeschwerden und Krankheitsängsten auf. Moderate bis geringe Zusammenhänge bestanden mit Depressivität, psychischer Funktionseinschränkung und allgemeiner Ängstlichkeit. Der SSD-12 Summenwert zeigte außerdem einen statistisch signifikanten Zusammenhang mit der Inanspruchnahme medizinischer Versorgungsleistungen. Die AutorInnen interpretieren dieses Muster der Zusammenhänge als Nachweis für die Validität der SSD-12.

Insgesamt wurde mit der SSD-12 das erste Selbstbeurteilungsinstrument für das B-Kriterium der Somatic Symptom Disorder nach DSM-5 vorgelegt. Summenwerte für die Gesamtskala und die Subskalen geben an, in welchem Ausmaß bei den Beantwortenden exzessives gedankliches Kreisen um Körperbeschwerden, auf Körperbeschwerden bezogene Krankheitsängste sowie exzessives Krankheitsverhalten vorliegen. Vergleichswerte der ambulanten

PatientInnen aus der Psychosomatischen Hochschulambulanz liegen in Form von Mittelwerten und Standardabweichungen vor (Toussaint et al., 2016). Gemeinsam mit den Fragebögen PHQ-15 oder SSS-8 ist es mit der SSD-12 möglich, die Ausprägung der A- und B-Kriterien der Somatic Symptom Disorder dimensional zu quantifizieren.

Die SSD-12 liegt bisher in englischer und deutscher Sprache vor und wurde durch die AutorInnen als frei verfügbares Messinstrument konzipiert, um eine breite Anwendung ohne Nutzungskosten zu gewährleisten. Bezogen werden kann es über den elektronischen Anhang des Zeitschriftenartikels (Toussaint et al., 2016).

In der wissenschaftlichen Community wurde die SSD-12 beim Erscheinen durch Barsky (2016) kommentiert. Dieser lobte die SSD-12 für ihre strikte Ableitung aus dem DSM-5 und betonte ihre Praxisrelevanz:

> The rigorous, thoughtful, and meticulous development of such a new self-report instrument is described by Toussaint et al. They generated a large item pool; refined it by eliminating items that were redundant, vague, or lacked clarity; used expert consensus and patient input to examine the face validity of the remaining items in assessing each of the three psychological domains they were designed to measure; and then established the psychometric properties of the resulting instrument in a clinical sample. Their sample was large and they obtained good participation rates. The result is a useful and needed tool that can be used clinically and in

> research. [...] the SSD-12 represents a substantial contribution to our current armamentarium of self-report measures for SSD consistent with DSM-5 criteria and this instrument will be helpful in the diagnosis and clinical monitoring of individuals with this common and troubling condition. (S. 3-4)

Weitere Erwähnung fand die SSD-12 in einem an klinisch Tätige adressierten Fortbildungsartikel von Lahmann et al. (2016), der in deutscher Sprache den aktuellen Forschungsstand zur Diagnostik somatoformer Störungen praxisnah zusammenträgt. Dabei wird die neue Diagnose Somatic Symptom Disorder vorgestellt, die in Deutschland durch das in der Versorgung vorgeschriebene Klassifikationssystem ICD-10 formal noch nicht etabliert wurde. Die SSD-12 wird als Messinstrument für das B-Kriterium vorgestellt.

Zur weiteren Etablierung der SSD-12 stehen noch Studien aus, die in Anschlussprojekten durchgeführt werden sollen: So formulierten Toussaint et al. (2016) das Überprüfen der Änderungssensitivität, das Ermitteln von Vergleichswerten aus anderen Populationen und das Etablieren eines Cut-off-Wertes unter Verwendung eines strukturierten klinischen Interviews als notwendige nächste Schritte.

4 Diskussion

4.1 Zusammenfassung der Einzelstudien

Die vorliegende kumulative Dissertation versammelt vier Einzelarbeiten über die psychometrische Erfassung von Körperbeschwerden und deren kognitiven, affektiven und behavioralen Begleitsymptomen. Ausgehend von einer Darstellung der Epidemiologie von Körperbeschwerden beschreibt sie den Nutzen von Selbstbeurteilungsfragebögen für Körperbeschwerden und gibt einen Überblick über die verfügbaren Messinstrumente. Daran anknüpfend greift sie die Entwicklung eines testökonomischen Körperbeschwerdenfragebogens (SSS-8) auf, der im Kontext der Arbeiten am DSM-5 als Kurzversion des bereits etablierten PHQ-15 entstand, jedoch noch nicht psychometrisch überprüft wurde. Dieser Überprüfung widmeten sich die ersten drei Studien. Sie bestimmen wichtige Kennwerte wie Reliabilität, Validität, Faktorstruktur, Cut-off-Werte, Normwerte, Änderungssensitivität und minimale klinisch bedeutsame Veränderung und kommen zu der Schlussfolgerung, dass es sich bei der SSS-8 um ein psychometrisch solides Messinstrument handelt. Sie zeigen außerdem empirisch, dass es sich bei der SSS-8 um eine psychometrisch adäquate Kurzversion des international anerkannten PHQ-15 handelt, was bisher nur eine Annahme war, die aufgrund der stringenten Ableitung aus dem PHQ-15 bestand.

Weiterhin nimmt die Dissertation Bezug zu der im DSM-5 etablierten Diagnose Somatic Symptom Disorder. Da mit der SSS-8 nur das A-Kriterium dieser Diagnose erfasst werden kann (A-Kriterium: Vorliegen von Körperbeschwerden), beschreibt die vierte Studie die Entwicklung des SSD-12-Fragebogens, mit dem die Ausprägung des B-Kriterium erfasst werden kann. Dieser Fragebogen besteht aus 12 Items, die auf Körperbeschwerden bezogene kognitive (Grübeln über die Ernsthaftigkeit der Körperbeschwerden), affektive (Krankheitsängste) und behaviorale (exzessives Krankheitsverhalten) Symptome erfassen. Abgeleitet wurden die Items direkt aus der im DSM-5 enthaltenen theoretischen Konzeption des B-Kriteriums. Die vierte Studie berichtet eine hohe Reliabilität der SSD-12 und zeigt erste Hinweise für die Validität des Messinstrumentes. Zur Beurteilung, wie stark das B-Kriterium bei den Beantwortenden ausgeprägt ist, stellt sie Referenzwerte aus einer großen Stichprobe aus einer psychosomatischen Ambulanz zur Verfügung. Mit der SSD-12 liegt das erste Messinstrument vor, mit dem KlinikerInnen und ForscherInnen die Ausprägung des B-Kriteriums bei PatientInnen quantifizieren können.

Insgesamt legt die Doktorarbeit zentrale psychometrische Kennwerte für zwei neue, frei verfügbare und vielversprechende Messinstrumente vor, die in der klinischen und wissenschaftlichen Praxis zur Erfassung von Körperbeschwerden und deren kognitiven, affektiven und behavioralen Begleitsymptomen kostenfrei genutzt werden können.

4.2 Limitationen

Alle vier Studien verwenden hauptsächlich Selbstberichte von PatientInnen oder Personen aus der Allgemeinbevölkerung als Datenquellen. Selbstberichte sind attraktiv, da sie zum einen in besonderem Maße dem subjektiven Charakter vieler Beschwerden Rechnung tragen (Black, 2013; Kroenke et al., 2015) und zum anderen testökonomisch sind, was besonders vorteilhaft ist, wenn große Stichproben erhoben werden sollen. Zwar haben alle eingesetzten Selbstbeurteilungsinstrumente eine Publikationshistorie, die sie als valide und reliable Messinstrumente ausweist, dennoch haben Selbstberichte Grenzen, beispielsweise bei der Erfassung von Krankheitstagen oder Arztbesuchen in den letzten 12 Monaten. Diese können den Befragten nur schwer erinnerlich sein. Zukünftige Studien sollten daher zusätzliche Datenquellen wie etwa Abrechnungsdaten von Krankenversicherungen und ExpertInnen-Fremdbeurteilungen der Funktionsfähigkeit hinzuziehen, um die Validität der beiden neuen Messinstrumente SSS-8 und SSD-12 weiter abzusichern.

Eine weitere Limitation besteht darin, dass alle vier Studien im westlichen Industriestaat Deutschland durchgeführt wurden. Bisher ist lediglich eine internationale psychometrische Studie zum Einsatz der SSS-8 aus einem anderen Kulturkreis erschienen. Hierbei handelt es sich um die japanische Übersetzung der SSS-8 von Matsudaira et al. (2017). Da das Erleben und Präsentieren von Körperbeschwerden und deren kognitiven, affektiven und behavioralen Begleitsymptomen kulturellen Einflüssen unterliegen kann, sind weitere Studien in anderen Kulturkreisen und medizinischen Versorgungs-

systemen notwendig, um zusätzliche Erkenntnisse zur Validität der SSS-8 und SSD-12 zu sammeln. Hierfür sind außerdem weitere Übersetzungsstudien notwendig.

4.3 Schlussfolgerung und Ausblick

Mit der SSS-8 und der SSD-12 liegen zwei neue, frei verfügbare und psychometrisch solide Fragebögen für Körperbeschwerden und darauf bezogene kognitive, affektive und behaviorale Begleitsymptome vor. Die SSS-8 kann zur Quantifizierung (0 bis 32 Punkte) von Körperbeschwerden im Rahmen vieler somatischer Erkrankungen oder psychischer Störungen in Klinik und Forschung eingesetzt werden. Als Screeninginstrument ermöglicht sie in etwa einer Minute Bearbeitungszeit einen raschen Überblick über das Vorliegen häufiger Körperbeschwerden. Der Schweregrad der Gesamtbelastung durch Körperbeschwerden kann durch die Anwendung von Schweregradkategorien und Normwerten aus der Allgemeinbevölkerung beurteilt werden. Durch die nachgewiesene Änderungssensitivität eignet sie sich zur Evaluation von Behandlungen und kann als wichtiges Feedback für klinische Tätige dienen. Die SSS-8 ist besonders geeignet für Situationen, in denen Körperbeschwerden in nur wenig Messzeit erhoben werden müssen. Dies macht sie besonders interessant für große wissenschaftliche Studien und Versorgungseinrichtungen mit einem hohen PatientInnenaufkommen.

Die SSD-12 quantifiziert als erstes Messinstrument dimensional (0 bis 48 Punkte) die Ausprägung des B-Kriteriums der Somatic Symptom Disorder, die im DSM-5 etabliert wurde. Damit bietet sie eine gute Grundlage für die weitere Erforschung dieser neuen Diagnose. Gleichzeitig ist sie für KlinikerInnen ein gutes Hilfsmittel bei der Beurteilung dieses relativ neuen Konstruktes. So können SSD-12 Messwerte den klinischen Eindruck vervollständigen und gezielt die subjektive Sichtweise der PatientInnen einbringen.

Werden die SSS-8 und die SSD-12 gemeinsam bei der klassifikatorischen Diagnostik der Somatic Symptom Disorder eingesetzt, sollte eine zukünftige Kriterium-Standard-Studie unter Zuhilfenahme eines von ExpertInnen geführten strukturierten klinischen Interviews (z.B. SKID-Interview für DSM-5) zunächst noch klären, welche Cut-off-Werte für die Erfüllung des A- und B-Kriteriums anzuwenden sind: Es bedarf einer „Grenzziehung“, ab wann von einer störungswertigen Ausprägung von Körperbeschwerden und den darauf bezogenen kognitiven, affektiven und behavioralen Begleitsymptomen gesprochen werden kann. Hierfür könnte es ebenfalls eine Rolle spielen, ob bei den untersuchten Personen eine schwere somatische Erkrankung vorliegt oder nicht (Dimsdale & Levenson, 2013; Häuser & Wolfe, 2013). Ab welcher Ausprägung beispielsweise das Grübeln über die Ernsthaftigkeit von Körperbeschwerden bei einer an Krebs erkrankten Person als störungswertig bezeichnet werden kann, bedarf weiterer Untersuchungen. Für die Interpretation der Messwerte der SSD-12 steht außerdem noch eine Ermittlung von Normwerten aus der Allgemeinbevölkerung aus. Diese sind wichtig, um beurteilen zu

können, wie stark ausgeprägt die drei Aspekte des B-Kriteriums sein müssen, um über dem „Bevölkerungsdurchschnitt“ zu liegen. Auch muss für die Beurteilung von Therapien, die auf die Veränderung der kognitiven, affektiven und behavioralen Begleitsymptome von Körperbeschwerden abzielen, die Änderungssensitivität der SSD-12 in einer Studie nachgewiesen werden. Diese Studien sind geplant und werden innerhalb der nächsten Jahre erscheinen. Ebenfalls geplant sind weitere Übersetzungen, um die Messinstrumente international zugänglich zu machen. Ob die SSD-12 in der klinischen und wissenschaftlichen Community ebenso gut wie die SSS-8 aufgenommen werden wird, wird in Deutschland nicht zuletzt auch davon abhängen, welche Definition der somatoformen Störungen in der elften Auflage der International Classification of Diseases (ICD-11) erscheinen wird.

5 Literaturverzeichnis

Barsky, A. J. (2014). Assessing Somatic Symptoms in Clinical Practice. *JAMA Internal Medicine*, *174*(3), 407.

Barsky, A. J. (2016). Assessing the New DSM-5 Diagnosis of Somatic Symptom Disorder. *Psychosomatic Medicine*, *78*(1), 2–4.

Black, N. (2013). Patient reported outcome measures could help transform healthcare. *BMJ*, *346*, f167.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159.

Copay, A. G., Subach, B. R., Glassman, S. D., Polly, D. W., & Schuler, T. C. (2007). Understanding the minimum clinically important difference: a review of concepts and methods. *The Spine Journal: Official Journal of the North American Spine Society*, *7*(5), 541–6.

Creed, F. H., Davies, I., Jackson, J., Littlewood, A., Chew-Graham, C., Tomenson, B., … McBeth, J. (2012). The epidemiology of multiple somatic symptoms. *Journal of Psychosomatic Research*, *72*(4), 311–317.

Croicu, C., Chwastiak, L., & Katon, W. (2014). Approach to the Patient with Multiple Somatic Symptoms. *Medical Clinics of North America*, *98*(5), 1079–1095.

Daradkeh, T. K., Eapen, V., & Ghubash, R. (2005). Mental Morbidity in Primary Care in Al Ain (UAE): Application of the Arabic Translation of the PRIME-MD (PHQ) Version. *German Journal of Psychiatry*, *8*, 32–35.

Derogatis, L. R., Lipman, R. S., & Covi, L. (1973). SCL-90: an outpatient psychiatric rating scale--preliminary report. *Psychopharmacology Bulletin*, *9*(1), 13–28.

Dimsdale, J. E., Creed, F., Escobar, J., Sharpe, M., Wulsin, L., Barsky, A. J., … Rief, W. (2013). Somatic Symptom Disorder: An important change in DSM. *Journal of Psychosomatic Research*, *75*(3), 223–228.

Dimsdale, J. E., & Levenson, J. (2013). Diagnosis of somatic symptom disorder requires clinical judgment. *Journal of Psychosomatic Research*.

Dowrick, C., Leydon, G. M., McBride, A., Howe, A., Burgess, H., Clarke, P., … Kendrick, T. (2009). Patients' and doctors' views on depression severity questionnaires incentivised in UK quality and outcomes framework: qualitative study. *British Medical Journal*, *338*, b663.

Eliasen, M., Kreiner, S., Ebstrup, J. F., Poulsen, C. H., Lau, C. J., Skovbjerg, S., … Jørgensen, T. (2016). Somatic Symptoms: Prevalence, Co-Occurrence and Associations with Self-Perceived Health and Limitations Due To Physical Health – A Danish Population-Based Study. *PLOS ONE*, *11*(3), e0150664.

Fink, P., Ewald, H., Jensen, J., Sørensen, L., Engberg, M., Holm, M., & Munk-Jørgensen, P. (1999). Screening for somatization and hypochondriasis in primary care and neurological in-patients: A seven-item scale for hypochondriasis and somatization. *Journal of Psychosomatic Research*, *46*(3), 261–273.

Fink, P., Toft, T., Hansen, M. S., Ørnbøl, E., & Olesen, F. (2007). Symptoms and syndromes of bodily distress: an exploratory study of 978 internal medical, neurological, and primary care patients. *Psychosomatic Medicine*, *69*(1), 30–39.

Galesic, M., Garcia-Retamero, R., & Gigerenzer, G. (2009). Using icon arrays to communicate medical risks: overcoming low numeracy. *Health Psychology*, *28*(2), 210–216.

Gierk, B., Kohlmann, S., Hagemann-Göbel, M., Löwe, B., & Nestoriuc, Y. (2017). Monitoring somatic symptoms in patients with mental disorders: Sensitivity to change and minimally clinical important difference of the Somatic Symptom Scale - 8 (SSS-8). Manuscript submitted for publication.

Gierk, B., Kohlmann, S., Kroenke, K., Spangenberg, L., Zenger, M., Brähler, E., & Löwe, B. (2014). The somatic symptom scale-8 (SSS-8): a brief measure of somatic symptom burden. *JAMA Internal Medicine*, *174*(3), 399–407.

Gierk, B., Kohlmann, S., Toussaint, A., Wahl, I., Brünahl, C. A., Murray, A. M., & Löwe, B. (2015). Assessing somatic symptom burden: A psychometric comparison of the Patient Health Questionnaire-15 (PHQ-15) and the Somatic Symptom Scale-8 (SSS-8). *Journal of Psychosomatic Research*, *78*(4), 352–355.

Han, C., Pae, C.-U., Patkar, A. A., Masand, P. S., Woong Kim, K., Joe, S.-H., & Jung, I. K. (2009). Psychometric Properties of the Patient Health Questionnaire–15 (PHQ–15) for Measuring the Somatic Symptoms of Psychiatric Outpatients. *Psychosomatics*, *50*(6), 580–585.

Häuser, W., & Wolfe, F. (2013). The somatic symptom disorder in DSM 5 risks mislabelling people with major medical diseases as mentally ill. *Journal of Psychosomatic Research*.

Herzog, A., Voigt, K., Meyer, B., Wollburg, E., Weinmann, N., Langs, G., & Löwe, B. (2015). Psychological and interactional characteristics of patients with somatoform disorders: Validation of the Somatic Symptoms Experiences Questionnaire (SSEQ) in a clinical psychosomatic population. *Journal of Psychosomatic Research*, *78*(6), 553–562.

Hyphantis, T., Kroenke, K., Papatheodorou, E., Paika, V., Theocharopoulos, N., Ninou, A., … Guthrie, E. (2014). Validity of the Greek version of the PHQ 15-item Somatic Symptom Severity Scale in patients with chronic medical conditions and correlations with emergency department use and illness perceptions. *Comprehensive Psychiatry*, *55*(8), 1950–1959.

Jackson, J. L., & Passamonti, M. (2005). The outcomes among patients presenting in primary care with a physical symptom at 5 years. *Journal of General Internal Medicine*, *20*(11), 1032–1037.

Jaeschke, R., Singer, J., & Guyatt, G. H. (1989). Measurement of health status. Ascertaining the minimal clinically important difference. *Controlled Clinical Trials*, *10*(4), 407–15.

Katon, W., Lin, E. H. B., & Kroenke, K. (2007). The association of depression and anxiety with medical symptom burden in patients with chronic medical illness. *General Hospital Psychiatry*, *29*(2), 147–155.

Khan, A. A., Khan, A., Harezlak, J., Wanzhu, T., & Kroenke, K. (2003). Somatic symptoms in primary care: etiology and outcome. *Psychosomatics*, *44*(6), 471–8.

Kocalevent, R.-D., Hinz, A., & Brähler, E. (2013). Standardization of a screening instrument (PHQ-15) for somatization syndromes in the general population. *BMC Psychiatry*, *13*(1), 91.

Kohlmann, S., Gierk, B., Hilbert, A., Brähler, E., & Löwe, B. (2016). The overlap of somatic, anxious and depressive syndromes: A population-based analysis. *Journal of Psychosomatic Research*, *90*, 51–56.

Kohlmann, S., Gierk, B., Hümmelgen, M., Blankenberg, S., & Löwe, B. (2013). Somatic Symptoms in Patients With Coronary Heart Disease: Prevalence, Risk Factors, and Quality of Life. *JAMA Internal Medicine*, *173*(15), 1469.

Kroenke, K. (2003). Patients presenting with somatic complaints: epidemiology, psychiatric co-morbidity and management. *International Journal of Methods in Psychiatric Research*, *12*(1), 34–43.

Kroenke, K. (2014). A Practical and Evidence-Based Approach to Common Symptoms. *Annals of Internal Medicine*, *161*(8), 579.

Kroenke, K. (2016). Somatic symptoms deserve our attention. *Families, Systems, & Health*, *34*(4), 330–333.

Kroenke, K., Arrington, M. E., & Mangelsdorff, A. D. (1990). The prevalence of symptoms in medical outpatients and the adequacy of therapy. *Archives of Internal Medicine*, *150*(8), 1685–1689.

Kroenke, K., Jackson, J. L., & Chamberlin, J. (1997). Depressive and anxiety disorders in patients presenting with physical complaints: clinical predictors and outcome. *American Journal of Medicine*, *103*(5), 339–347.

Kroenke, K., & Mangelsdorff, A. D. (1989). Common symptoms in ambulatory care: incidence, evaluation, therapy, and outcome. *The American Journal of Medicine*, *86*(3), 262–6.

Kroenke, K., Monahan, P. O., & Kean, J. (2015). Pragmatic characteristics of patient-reported outcome measures are important for use in clinical practice. *Journal of Clinical Epidemiology*, *68*(9), 1085–92.

Kroenke, K., Spitzer, R. L., DeGruy, F. V., & Swindle, R. (1998). A Symptom Checklist to Screen for Somatoform Disorders in Primary Care. *Psychosomatics*, *39*(3), 263–272.

Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2002). The PHQ-15: Validity of a new measure for evaluating the severity of somatic symptoms. *Psychosomatic Medicine*, *64*(2), 258–266.

Kroenke, K., Spitzer, R. L., Williams, J. B. W., & Löwe, B. (2010). The Patient Health Questionnaire Somatic, Anxiety, and Depressive Symptom Scales: A systematic review. *General Hospital Psychiatry*, *32*(4), 345–359.

Kroenke, K., Zhong, X., Theobald, D., Wu, J., Tu, W., & Carpenter, J. S. (2010). Somatic Symptoms in Patients With Cancer Experiencing Pain or Depression: Prevalence, Disability, and Health Care Use. *Archives of Internal Medicine*, *170*(18), 1686–1694.

Kurlansik, S. L., & Maffei, M. S. (2016). Somatic Symptom Disorder. *American Family Physician*, *93*(1).

Lahmann, C., Allwang, C., & Dinkel, A. (2016). Diagnostik somatoformer Störungen. *PSYCH up2date*, *10*(5), 375–386.

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, *4*, 863.

Lee, S., Ma, Y. L., & Tsang, A. (2011). Psychometric properties of the Chinese 15-item Patient Health Questionnaire in the general population of Hong Kong. *Journal of Psychosomatic Research*, *71*(2), 69–73.

Löwe, B., Blankenberg, S., Wegscheider, K., König, H.-H., Walter, D., Murray, A. M., … Kohlmann, S. (2016). Depression screening with patient-targeted feedback in cardiology: DEPSCREEN-INFO randomised clinical trial. *The British Journal of Psychiatry*, *210*, 132–139.

Löwe, B., Spitzer, R. L., Williams, J. B. W., Mussel, M., Schellberg, D., & Kroenke, K. (2008). Depression, anxiety and somatization in primary care: syndrome overlap and functional impairment. *General Hospital Psychiatry*, *30*(3), 191–9.

Löwe, B., Spitzer, R. L., Zipfel, S., & Herzog, W. (2002). *Gesundheitsfragebogen für Patienten (PHQ D). Komplettversion und Kurzform. Testmappe mit Manual, Fragebögen, Schablonen.* (2nd ed.). Karlsruhe: Pfizer.

Matsudaira, K., Oka, H., Kawaguchi, M., Murakami, M., Fukudo, S., Hashizume, M., & Löwe, B. (2017). Development of a Japanese version of the Somatic Symptom Scale-8: Psychometric validity and internal consistency. *General Hospital Psychiatry*, *45*, 7–11.

Narrow, W. E., Clarke, D., Kuramoto, J. S., Kraemer, H. C., Kupfer, D. J., Greiner, L., & Regier, D. A. (2013). DSM-5 Field Trials in the United States and Canada, Part III: Development and Reliability Testing of a Cross-Cutting Symptom Assessment for DSM-5. *American Journal of Psychiatry*, *170*(1), 71–82.

Nimnuan, C., Hotopf, M., & Wessely, S. (2001). Medically unexplained symptoms: an epidemiological study in seven specialities. *Journal of Psychosomatic Research*, *51*(1), 361–7.

Nordin, S., Palmquist, E., & Nordin, M. (2013). Psychometric evaluation and normative data for a Swedish version of the Patient Health Questionnaire 15-Item Somatic Symptom Severity Scale. *Scandinavian Journal of Psychology*, *54*(2), 112–117.

Persoons, P., Luyckx, K., Desloovere, C., Vandenberghe, J., & Fischler, B. (2003). Anxiety and mood disorders in otorhinolaryngology outpatients presenting with dizziness: validation of the self-administered PRIME-MD Patient Health Questionnaire and epidemiology. *General Hospital Psychiatry*, *25*(5), 316–323.

Rief, W., Hessel, A., & Brähler, E. (2001). Somatization symptoms and hypochondriacal features in the general population. *Psychosomatic Medicine*, *63*(4), 595–602.

Rief, W., Martin, A., Klaiberg, A., & Brähler, E. (2005). Specific effects of depression, panic, and somatic symptoms on illness behavior. *Psychosomatic Medicine*, *67*(4), 596–601.

Roelen, C. A. . M., Koopmans, P. C., & Groothoff, J. W. (2010). Subjective health complaints in relation to sickness absence. *Work*, *37*(1), 15–21.

Ros Montalbán, S., Comas Vives, A., & Garcia-Garcia, M. (2010). Validation of the Spanish version of the PHQ-15 questionnaire for the evaluation of physical symptoms in patients with depression and/or anxiety disorders: DEPRE-SOMA study. *Actas Espanolas de Psiquiatria*, *38*(6), 345–57.

Sharpe, M. (2013). Somatic symptoms: beyond ‘medically unexplained’. *The British Journal of Psychiatry*, *203*(5).

Simon, G. E., von Korff, M., Piccinelli, M., Fullerton, C., & Ormel, J. (1999). An international study of the relation between somatic symptoms and depression. *The New England Journal of Medicine*, *341*(18), 1329–35.

Stratford, P. W., & Riddle, D. L. (2005). Assessing sensitivity to change: choosing the appropriate change coefficient. *Health and Quality of Life Outcomes*, *3*(23).

Szemerszky, R., Köteles, F., Lihi, R., & Bárdos, G. (2010). Polluted places or polluted minds? An experimental sham-exposure study on background psychological factors of symptom formation in 'Idiophatic Environmental Intolerance attributed to electromagnetic fields'. *International Journal of Hygiene and Environmental Health*, *213*(5), 387–394.

Toussaint, A., Murray, A. M., Voigt, K., Herzog, A., Gierk, B., Kroenke, K., … Löwe, B. (2016). Development and Validation of the Somatic Symptom Disorder-B Criteria Scale (SSD-12). *Psychosomatic Medicine*, *78*(1), 5–12.

van Ravesteijn, H., Wittkampf, K., Lucassen, P., van de Lisdonk, E., van den Hoogen, H., van Weert, H., … Speckens, A. (2009). Detecting somatoform disorders in primary care with the PHQ-15. *Annals of Family Medicine*, *7*(3), 232–8.

Voigt, K., Nagel, A., Meyer, B., Langs, G., Braukhaus, C., & Löwe, B. (2010). Towards positive diagnostic criteria: A systematic review of somatoform disorder diagnoses and suggestions for future classification. *Journal of Psychosomatic Research*, *68*(5), 403–414.

Voigt, K., Wollburg, E., Weinmann, N., Herzog, A., Meyer, B., Langs, G., & Löwe, B. (2012). Predictive validity and clinical utility of DSM-5 Somatic Symptom Disorder — Comparison with DSM-IV somatoform disorders and additional criteria for consideration. *Journal of Psychosomatic Research*, *73*(5), 345–350.

Witthöft, M., Hiller, W., Loch, N., & Jasper, F. (2013). The latent structure of medically unexplained symptoms and its relation to functional somatic syndromes. *International Journal of Behavioral Medicine*, *20*(2), 172–183.

Zhang, L., Fritzsche, K., Liu, Y., Wang, J., Huang, M., Wang, Y., … Leonhart, R. (2016). Validation of the Chinese version of the PHQ-15 in a tertiary hospital. *BMC Psychiatry*, *16*(1), 89.

Zijlema, W. L., Stolk, R. P., Löwe, B., Rief, W., White, P. D., & Rosmalen, J. G. M. (2013). How to assess common somatic symptoms in large-scale studies: A systematic review of questionnaires. *Journal of Psychosomatic Research*, *74*(6), 459–468.

6 Anhang

Beigefügt sind die vier Publikationen im Originaltext.

Studie 1

Gierk, B., Kohlmann, S., Kroenke, K., Spangenberg, L., Zenger, M., Brähler, E., & Löwe, B. (2014). The somatic symptom scale-8 (SSS-8): a brief measure of somatic symptom burden. *JAMA Internal Medicine*, *174*(3), 399–407.

Studie 2

Gierk, B., Kohlmann, S., Toussaint, A., Wahl, I., Brünahl, C. A., Murray, A. M., & Löwe, B. (2015). Assessing somatic symptom burden: A psychometric comparison of the Patient Health Questionnaire-15 (PHQ-15) and the Somatic Symptom Scale-8 (SSS-8). *Journal of Psychosomatic Research*, *78*(4), 352–355.

Studie 3

Gierk, B., Kohlmann, S., Hagemann-Göbel, M., Löwe, B., & Nestoriuc, Y. (2017). Monitoring somatic symptoms in patients with mental disorders: Sensitivity to change and minimally clinical important difference of the Somatic Symptom Scale - 8 (SSS-8). Manuscript submitted for publication.

Studie 4

Toussaint, A., Murray, A. M., Voigt, K., Herzog, A., Gierk, B., Kroenke, K., … Löwe, B. (2016). Development and Validation of the Somatic Symptom Disorder-B Criteria Scale (SSD-12). *Psychosomatic Medicine*, *78*(1), 5–12.

Research

Original Investigation

The Somatic Symptom Scale-8 (SSS-8)
A Brief Measure of Somatic Symptom Burden

Benjamin Gierk, MSc; Sebastian Kohlmann, PhD; Kurt Kroenke, MD; Lena Spangenberg, PhD; Markus Zenger, PhD; Elmar Brähler, PhD; Bernd Löwe, MD

IMPORTANCE Somatic symptoms are the core features of many medical diseases, and they are used to evaluate the severity and course of illness. The 8-item Somatic Symptom Scale (SSS-8) was recently developed as a brief, patient-reported outcome measure of somatic symptom burden, but its reliability, validity, and usefulness have not yet been tested.

OBJECTIVE To investigate the reliability, validity, and severity categories as well as the reference scores of the SSS-8.

DESIGN, SETTING, AND PARTICIPANTS A national, representative general-population survey was performed between June 15, 2012, and July 15, 2012, in Germany, including 2510 individuals older than 13 years.

MAIN OUTCOMES AND MEASURES The SSS-8 mean (SD), item-total correlations, Cronbach α, factor structure, associations with measures of construct validity (Patient Health Questionnaire–2 depression scale, Generalized Anxiety Disorder–2 scale, visual analog scale for general health status, 12-month health care use), severity categories, and percentile rank reference scores.

RESULTS The SSS-8 had excellent item characteristics and good reliability (Cronbach α = 0.81). The factor structure reflects gastrointestinal, pain, fatigue, and cardiopulmonary aspects of the general somatic symptom burden. Somatic symptom burden as measured by the SSS-8 was significantly associated with depression (r = 0.57 [95% CI, 0.54 to 0.60]), anxiety (r = 0.55 [95% CI, 0.52 to 0.58]), general health status (r = −0.24 [95% CI, −0.28 to −0.20]), and health care use (incidence rate ratio, 1.12 [95% CI, 1.10 to 1.14]). The SSS-8 severity categories were calculated in accordance with percentile ranks: no to minimal (0-3 points), low (4-7 points), medium (8-11 points), high (12-15 points), and very high (16-32 points) somatic symptom burden. For every SSS-8 severity category increase, there was a 53% (95% CI, 44% to 63%) increase in health care visits.

CONCLUSIONS AND RELEVANCE The SSS-8 is a reliable and valid self-report measure of somatic symptom burden. Cutoff scores identify individuals with low, medium, high, and very high somatic symptom burden.

Invited Commentary page 407

Supplemental content at jamainternalmedicine.com

Author Affiliations: Department of Psychosomatic Medicine and Psychotherapy, University Medical Center Hamburg-Eppendorf, Hamburg, Germany (Gierk, Kohlmann, Löwe); Schön Clinic Hamburg Eilbek, Hamburg, Germany (Gierk, Kohlmann, Löwe); Department of Medicine, Indiana University, Indianapolis (Kroenke); Veterans Affairs Health Services Research and Development Service, Center for Health Information and Communication, Indianapolis, Indiana (Kroenke); Regenstrief Institute, Inc, Indianapolis, Indiana (Kroenke); Department of Medical Psychology and Medical Sociology, University of Leipzig, Leipzig, Germany (Spangenberg, Zenger, Brähler).

Corresponding Author: Benjamin Gierk, MSc, Department of Psychosomatic Medicine and Psychotherapy, University Medical Center Hamburg-Eppendorf, Martinistrasse 52, 20246 Hamburg, Germany (bgierk@uke.de).

JAMA Intern Med. 2014;174(3):399-407. doi:10.1001/jamainternmed.2013.12179
Published online November 25, 2013.

Copyright 2014 American Medical Association. All rights reserved.

Downloaded From: http://archinte.jamanetwork.com/ by a University Medical Center Hamburg - Eppendorf User on 04/10/2014

Somatic symptoms are the common ground of many medical conditions.[1] Presenting symptoms are often the impetus for further diagnostic evaluation, and symptom patterns may point to specific medical diseases. In clinical consultations, somatic symptoms are used as indicators for the severity and course of medical diseases as well as for monitoring treatment progress. They are associated with decreased quality of life, increased psychological distress, and increased use of health care services.[2,3] Somatic symptoms are characteristic of various medical conditions such as cancer or coronary heart disease,[2,3] and they are also present in psychiatric conditions such as depressive disorders, anxiety disorders, or somatoform disorders.[4-6] Previous studies[7-9] suggest that individual somatic symptoms frequently cluster into 4 groups: cardiopulmonary, gastrointestinal, pain, and general.

Although brief self-report instruments are frequently used for screening and severity assessment of depression and anxiety, they are less commonly used for the measurement of somatic symptom burden.[10,11] However, somatic symptom scales might identify symptoms that are not the direct focus of the medical consultation. In addition, they could serve as screening instruments for elevated somatic symptom burden, as indexes for illness severity, and as measures to gauge response to treatment. Moreover, the use of self-report scales to complement medical interviews for symptom-based conditions can increase patient confidence that physicians are correct in their diagnosis and are making systematic efforts to assess and manage their problems.[12]

A recent systematic review[8] indicated that many self-report questionnaires for common somatic symptoms exist but vary considerably with respect to their length, scaling, dimensionality, reliability, validity, and relevance to populations and regions. Among the 40 self-report somatic symptom scales investigated in this review, the Patient Health Questionnaire–15 (PHQ-15)[13] and the 12-item Symptom Checklist–90 somatization scale[14] were identified as the most appropriate measures for large-scale studies because of their well-established psychometric properties, relevance of symptoms, brevity, and availability in multiple languages.[8] Derived from the PHQ-15, the 8-item Somatic Symptom Scale (SSS-8) (**Figure 1**) was recently developed as a brief patient-reported measure of somatic symptom burden. The SSS-8 has been used as a reference measure in the *Diagnostic and Statistical Manual of Mental Disorders* (Fifth Edition) (*DSM-5*)[15] field trials as an instrument to facilitate the diagnosis of somatic symptom disorder. However, the psychometric properties of the instrument have not been investigated, and clinical reference scores are unavailable.

Therefore, the aims of this general-population study were 2-fold: (1) to investigate the reliability, factorial validity, and construct validity of the SSS-8 and (2) to provide reference scores for low, medium, high, and very high overall somatic symptom severity. The ultimate goal was to find a useful scale to evaluate and monitor somatic symptom burden in research and clinical practice. The study was approved by the medical ethics board at Leipzig University, Leipzig, Germany. Oral informed consent was obtained from the participants.

Methods

Design, Selection of Participants, and Procedure

The study was part of a national, representative general-population survey that approached people older than 13 years in Germany. The data were collected in 2 waves between June and July 2012 by professional demographic consultants (Unabhängige Serviceeinrichtung für Umfragen, Methoden und Analysen, Berlin). A random-route sampling procedure with 320 sample points revealed that 4480 households should be contacted as part of the study. Of these, 4436 households were eligible to participate (ie, were not vacant or without individuals who met the inclusion criterion). The households were visited by trained face-to-face interviewers who recorded participants' demographic information; the other information was collected through a paper/pencil self-report. Interviewers made a maximum of 4 contact attempts per household.

Measures

The SSS-8 is an abbreviated 8-item version of the PHQ-15 questionnaire.[13] The PHQ-15 scale assesses the presence and severity of common somatic symptoms and is among the most widely used and best-validated self-report measures of somatic symptom burden.[8,13,16] The SSS-8 was first developed for the *DSM-5* field trials that investigated the newly established somatic symptom disorder.[15] Three PHQ-15 items (menstrual problems, sexual problems, and fainting) were not included in the SSS-8 because of low communalities with other items, low symptom prevalence, and low associations with measures of functioning, quality of life, and health care use.

Figure 1. The Somatic Symptom Scale–8 (SSS-8)

During the past 7 days, how much have you been bothered by any of the following problems?	Not at all	A little bit	Somewhat	Quite a bit	Very much
1. Stomach or bowel problems	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4
2. Back pain	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4
3. Pain in your arms, legs, or joints	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4
4. Headaches	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4
5. Chest pain or shortness of breath	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4
6. Dizziness	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4
7. Feeling tired or having low energy	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4
8. Trouble sleeping	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4
SSS-8 Score ☐ = ___ + ___ + ___ + ___					

Items, response options, and scoring on the SSS-8.

Copyright 2014 American Medical Association. All rights reserved.

Downloaded From: http://archinte.jamanetwork.com/ by a University Medical Center Hamburg - Eppendorf User on 04/10/2014

In addition, 5 PHQ-15 items regarding gastrointestinal and cardiovascular symptoms were collapsed into 2 new items (SSS-8 items 1 and 5). A 5-point response option (0-4) for each SSS-8 item (instead of the 3-point response option of the PHQ-15) and a 7-day time frame (instead of the 4-week time frame of the PHQ-15) were used to mirror the response options and time frame of the Patient Reported Outcomes Measurement Information System measures used in the *DSM-5* field trials. The SSS-8 score can, therefore, range from 0 to 32. Initially labeled as the PHQ-SSS in the *DSM-5* field trials,[15] the scale was renamed to shorten the name and reflect the number of items. The German version of the SSS-8 was developed using several steps of translation and blind back-translation according to state-of-the-art methods of test translation.[17]

To examine the construct validity of the SSS-8, participants also completed the PHQ-2 depression scale,[10,18] the Generalized Anxiety Disorder–2 scale,[10,19] and a visual analog scale to rate their current general health status (0, worst; to 100, best). Health care use was measured as the self-reported numbers of contacts with health care providers in the past 12 months (eg, days in the hospital, general practitioner visits, and emergency treatments). We chose these measures because somatic symptoms often are associated with depression and anxiety,[16] decreased general health status,[13] and increased health care use.[13]

Statistical Analysis

We computed mean (SD) and corrected item-total correlations as descriptive statistics. Reliability was measured using Cronbach α. Factorial validity was analyzed using confirmatory factor analyses that tested 2 models theoretically derived from previous studies[7-9]: a simple general factor model with all items loading on 1 factor and a higher-order general factor model composed of an overarching somatic symptom factor and 4 lower-order symptom clusters (gastrointestinal, pain, cardiopulmonary, and fatigue). Factorial invariance for sex and age was tested using multigroup confirmatory factor analyses with equality constraints.[20] For these analyses, 4 groups based on sex and age were generated by applying a median split: males younger than the median age, males at the median age or older, females younger than the median age, and females at the median age or older. Construct validity was examined using (1) bivariate correlations between SSS-8, PHQ-2, Generalized Anxiety Disorder–2 (GAD-2), and visual analog scale general health and (2) negative binomial regression on the count of health care visits, with SSS-8 score as the predictor and anxiety and depression scores as covariates. The SSS-8 reference scores were presented as percentile ranks. Severity categories were determined in accordance with these percentile ranks. Severity categories were validated using negative binomial regression to predict health care visits at every severity level while adjusting for anxiety and depression scores. Missing values were imputed with a hot deck technique.[21]

Results

Description of the Sample

Within the 4436 eligible households, interviewers had contact with 3736 persons. Of those, 2538 agreed to participate (67.9% cooperation rate[22]). However, 23 individuals were excluded because of severe illness that made participation impossible and 5 individuals were excluded for other reasons. As a result, data from 2510 individuals were included in the analyses (56.6% minimum response rate[22]). Less than 1% of the values in the data set were missing. Missing values showed no systematic pattern and were imputed.[21] **Table 1** reports the demographic characteristics of the study sample.

Item and Scale Characteristics

The SSS-8 had good reliability (Cronbach α = 0.81), and all corrected item-total correlations exceeded 0.40. The item means varied between 0.16 and 0.70, reflecting an acceptable range of item difficulty. In addition, the SDs were comparable among the items (0.50-0.96) and the overall SSS-8 sum score mean was 3.23 (3.96), with an observed range of 0 to 26. Females had slightly higher mean scores compared with males (3.49 vs 2.94; $t_{2508} = 3.50$; $P < .001$; Cohen $d = 0.14$), and age was positively correlated with SSS-8 scores ($r = 0.32$ [95% CI, 0.28-0.35]). Therefore, we calculated age- and sex-specific reference scores (**Table 2**).

Factorial Validity

Confirmatory factor analysis revealed good fit indices for the higher-order general factor model (Tucker-Lewis index, 0.95; comparative fit index, 0.97; root mean square error of approximation, 0.08 [90% CI, 0.08-0.09]).[23] This model, displayed in **Figure 2**, consisted of an overarching somatic symptom factor and 4 lower-order symptom clusters (gastrointestinal, pain, cardiopulmonary, and fatigue). Multigroup confirmatory factor analyses also demonstrated strict factorial invariance for age and sex for this model (Supplement [eTable]). In contrast, the simple general-factor model achieved lower fit indices and was, therefore, rejected (Tucker-Lewis index, 0.91; comparative fit index, 0.94; root mean square error of approximation, 0.11 [90% CI, 0.10-0.12]). Altogether, the results of the confirmatory factor analyses support the organization of individual somatic symptoms in 4 clusters according to a higher-order general-factor model.

Construct Validity

In terms of construct validity, SSS-8 scores were significantly correlated with PHQ-2 depression scores ($r = 0.57$ [95% CI, 0.54 to 0.60]), GAD-2 anxiety scores ($r = 0.55$ [95% CI, 0.52 to 0.58]), and the ratings of general health status ($r = -0.24$ [95% CI, −0.28 to −0.20]). Even after controlling for depression and anxiety, SSS-8 scores significantly predicted health care visits in the past 12 months: each 1-point increase in the SSS-8 score was associated with a 12% increase in health care visits (incidence rate ratio, 1.12 [95% CI, 1.10 to 1.14]).

Severity Categories

We determined severity categories in accordance with the SSS-8 percentile ranks (Table 2). Somatic symptom burden was classified as high if the sum score was between 12 and 15 points (ie, between the 95th and 98th percentile). From 12 points, we went downward in equal 4-point steps and classified 11 to 8 points as medium, 7 to 4 as low, and 3 to 0 as no to minimal

Copyright 2014 American Medical Association. All rights reserved.

Downloaded From: http://archinte.jamanetwork.com/ by a University Medical Center Hamburg - Eppendorf User on 04/10/2014

Table 1. Demographic Characteristics of the Study Sample

Characteristic	No. (%)		
	Total Sample (N = 2510)	Male (n = 1170)	Female (n = 1340)
Age, y			
Mean (SD)	49.4 (18.0)	49.3 (18.0)	49.4 (18.0)
Median (range)	51 (14-91)	51 (14-91)	51 (14-91)
Grouping			
14 to 24	275 (11.0)	129 (11.0)	146 (10.9)
25 to 34	349 (13.9)	166 (14.2)	183 (13.7)
35 to 44	361 (14.4)	166 (14.2)	195 (14.5)
45 to 54	463 (18.4)	209 (18.4)	254 (19.0)
55 to 64	467 (18.6)	215 (18.4)	252 (18.8)
65 to 74	405 (16.1)	209 (17.9)	196 (14.6)
≥75	190 (7.6)	76 (6.5)	114 (8.5)
Living with a partner	1461 (58.2)	738 (63.1)	723 (54.0)
Educational level, y			
0 to 8	42 (1.7)	16 (1.4)	26 (1.9)
9	937 (37.3)	437 (37.4)	500 (37.3)
10 to 11	957 (38.1)	402 (34.4)	555 (41.4)
12	82 (3.3)	46 (3.9)	36 (2.7)
13 to 15	236 (9.4)	123 (10.5)	113 (8.4)
≥16	191 (7.6)	117 (10.0)	74 (5.5)
Current student	65 (2.6)	29 (2.5)	36 (2.7)
Employment			
Full-time	1032 (41.1)	624 (53.3)	408 (30.4)
Part-time	261 (10.4)	24 (2.1)	237 (17.7)
Hourly basis	58 (2.3)	7 (0.6)	51 (3.8)
Civil/military service, childcare	16 (0.6)	1 (0.1)	15 (1.1)
Unemployed	101 (4.0)	52 (4.4)	49 (3.7)
Retired	728 (29.0)	354 (30.3)	374 (27.9)
Stay-at-home	114 (4.5)	10 (0.9)	104 (7.8)
Other	200 (8.0)	98 (8.4)	102 (7.6)
Net household income, per month, €[a]			
<750	72 (2.9)	34 (3.0)	38 (2.9)
750 to 1250	364 (14.9)	107 (9.4)	257 (19.7)
1250 to <2000	751 (30.7)	368 (32.3)	383 (29.3)
≥2000	1259 (51.5)	630 (55.3)	629 (48.1)
Church member	1892 (76.9)	840 (73.5)	1052 (79.9)

[a] Income variable has 64 missing values; these were not imputed because income was not a variable to be included in the statistical models.

burden. Finally, we added a very high category ranging from 16 to 32 points (ie, ≥98th percentile). We chose these cutoffs for 4 reasons. First, in primary care settings, approximately 10% of patients present with high or very high somatic symptom burden.[13] In the present study, however, we surveyed the general population, so the rates of high and very high symptom burden should be lower than those in a typical clinical sample. Second, measures of symptom burden usually follow a left-skewed distribution in a nonclinical sample, ie, most individuals should fall into the categories no to minimal and low. This is also the case for our chosen severity categories. Third, pragmatic cutoffs are needed for clinical applications, and our equal 4-point increments are both valid and memorable (ie, 4, 8, 12, and 16 points representing low, medium, high, and very high levels of somatic symptoms, respectively). Finally, a negative binomial regression analysis revealed that an increase in severity category is significantly associated with an increase in health care use (**Figure 3**): for every SSS-8 severity category increase, there was a 53% increase in health care visits (incidence rate ratio, 1.53 [95% CI, 1.44-1.63]), which provides further evidence of the construct validity of our severity categories.

Discussion

The present study evaluated the SSS-8, a new 8-item self-report measure for the assessment of somatic symptom burden. The results revealed good reliability of the scale and strong evidence of its validity. Cutoff points allow for pragmatic categorization of no to minimal (0-3 points), low (4-7 points), medium (8-11 points), high (12-15 points), and very high (16-32

Copyright 2014 American Medical Association. All rights reserved.

Downloaded From: http://archinte.jamanetwork.com/ by a University Medical Center Hamburg - Eppendorf User on 04/10/2014

Table 2. Normative Data From the General Population for the SSS-8[a]

SSS-8 Severity Category	SSS-8 Sum Score	Total Sample, Aged 14-91 y (N = 2510)	Age Percentile, y						
			14-24	25-34	35-44	45-54	55-64	65-74	≥75
Male									
No. of respondents			129	166	166	209	215	209	76
No to minimal	0	31	59	60	51	31	29	12	13
	1	43	74	73	63	45	40	24	16
	2	56	84	81	77	59	54	37	25
	3	66	90	86	81	69	67	47	34
Low	4	75	93	94	86	79	76	61	41
	5	79	95	95	89	82	83	67	41
	6	84	97	98	93	86	89	76	47
	7	88	98	98	94	90	91	79	66
Medium	8	90	98	98	96	92	93	85	70
	9	92	98	98	96	93	94	89	75
	10	94	98	98	98	95	97	90	83
	11	95	98	99	99	95	98	92	87
High	12	96	98	99	99	96	99	93	92
	13	97	98	99	99	97	99	93	92
	14	97	98	99	99	97	99	93	93
	15	98	98	99	99	98	99	95	96
Very high	16	98	99	99	99	98	99	96	96
	17	99	100	100	99	99	100	97	97
	18	99	100	100	99	99	100	98	99
	19	99	100	100	99	99	100	99	99
	20	100	100	100	100	99	100	100	100
	21	100	100	100	100	99	100	100	100
	22	100	100	100	100	99	100	100	100
	23-32	100	100	100	100	100	100	100	100
Female									
No. of respondents			146	183	195	254	252	196	114
No to minimal	0	31	53	42	30	25	24	15	4
	1	43	66	56	41	37	34	23	8
	2	56	75	74	58	54	45	34	18
	3	66	82	82	67	68	55	47	30
Low	4	75	88	85	76	78	69	58	42
	5	79	91	89	82	81	77	64	48
	6	84	92	92	84	84	84	73	58
	7	88	94	94	88	89	87	80	68
Medium	8	90	94	95	89	91	91	86	76
	9	92	94	96	92	93	93	90	81
	10	94	95	97	94	94	94	92	84
	11	95	95	97	96	96	95	95	85
High	12	96	97	97	97	96	97	97	89
	13	97	97	98	98	98	97	98	90
	14	97	98	99	98	98	97	98	91
	15	98	99	99	99	98	97	98	94

(continued)

points) somatic symptom burden. Simply stated, SSS-8 scores of 4, 8, 12, and 16 represent valid and easy-to-remember cutoff points for low, medium, high, and very high levels of somatic symptoms, respectively. Completion of the scale by the patient takes approximately 1 minute, and the scoring is easily done within another minute. Within this short period of time, health care professionals can obtain a valid score to gauge the total severity of the patients' somatic symptoms and initial information regarding the pattern of the somatic symptoms themselves to inform diagnosis and guide further evaluations.

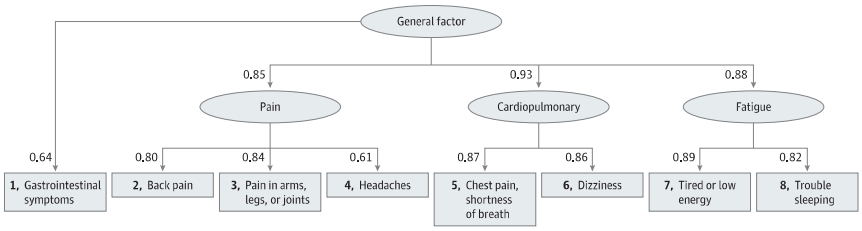
Copyright 2014 American Medical Association. All rights reserved.

Downloaded From: http://archinte.jamanetwork.com/ by a University Medical Center Hamburg - Eppendorf User on 04/10/2014

Table 2. Normative Data From the General Population for the SSS-8[a] (continued)

SSS-8 Severity Category	SSS-8 Sum Score	Total Sample, Aged 14-91 y (N = 2510)	Age Percentile, y						
			14-24	25-34	35-44	45-54	55-64	65-74	≥75
Very high	16	98	99	99	99	98	98	98	96
	17	99	99	99	99	98	98	99	98
	18	99	99	99	99	99	99	100	98
	19	99	99	99	99	99	100	100	98
	20	100	99	99	100	99	100	100	99
	21	100	100	99	100	100	100	100	99
	22	100	100	100	100	100	100	100	100
	23-32	100	100	100	100	100	100	100	100

Figure 2. Confirmatory Factor Analysis of the Higher-Order General-Factor Model



The model was estimated using robust weighted least-squares estimation with mean and variance adjustment based on polychoric correlations. Loadings are standardized. Tucker-Lewis index, 0.95; comparative fit index, 0.97; root mean square error of approximation, 0.08 (90% CI, 0.08-0.09).

The SSS-8 assesses somatic symptom burden as a subjective, patient-reported phenomenon, without making any assumptions concerning its cause. The SSS-8 is potentially useful in many medical settings. First, SSS-8 scores could serve as quantitative markers of somatic symptom burden in patients with chronic diseases, such as cancer[2] or coronary heart disease,[3] who often experience various somatic symptoms that decrease their quality of life and psychological and social functioning. The general-population reference scores allow for comparison with people of the same age and sex. Second, the SSS-8 can facilitate the diagnosis of somatic symptom disorder, which was recently introduced in *DSM-5*.[24] The A criterion of this disorder concerns the presence of burdensome somatic symptoms. Such symptoms can be easily assessed by using the SSS-8, the measure used in the *DSM-5* field trials.[15] Third, serial administration of the SSS-8 can be used to monitor somatic symptom burden and optimize symptom management (eg, reevaluation and modification of treatment if necessary). Fourth, the SSS-8 may detect burdensome and untreated somatic symptoms that are not cardinal symptoms of a particular disease, for example, musculoskeletal pain in patients with coronary heart disease.[3] Fifth, because somatic symptom burden is associated with high health care use and costs,[25,26] completion of the SSS-8 by frequent attenders of outpatient clinics and emergency departments could be helpful in determining whether high somatic symptom burden is a contributing factor to more frequent use of health care services. Furthermore, the scale could be used in managing the patient panels of primary care physicians or in reimbursement for such patients. Sixth, the SSS-8 could be used in risk adjustment of patient populations given that somatic symptom burden is an independent predictor of hospitalization and death, even when controlling for clinical characteristics, chronic medical conditions, self-rated health, and depression.[27]

The SSS-8 was designed to be an abbreviated version of the established PHQ-15 scale.[13,16] Although the SSS-8 is approximately half as long as the PHQ-15, it features comparable construct validity: SSS-8 scores were significantly associated with health care use, anxiety, depression, and general health status.[13,16] Interestingly, the associations between somatic symptom burden, depression, and anxiety were even stronger than the association between somatic symptom burden and perceived general health status. This finding is plausible considering the high comorbidity and partial overlap of somatic, anxiety, and depressive symptoms sometimes referred to as the *somatization-anxiety-depression triad*.[16,28] This triad suggests that somatic, anxiety, and depressive symptoms seldom occur as "pure forms,"[16(p352)] meaning that most patients experience combinations of these symptoms. Nevertheless, studies show that somatic, anxiety, and depressive symptoms "have independent, additive and differential effects on multiple domains of health-related quality of life,

Copyright 2014 American Medical Association. All rights reserved.

Downloaded From: http://archinte.jamanetwork.com/ by a University Medical Center Hamburg - Eppendorf User on 04/10/2014

functional status, disability and health care use," which, therefore, gives support to the differentiation of these 3 symptom domains.[16(p352)28] Given the high comorbidity of somatic symptom burden, anxiety, and depression, we recommend complementing the SSS-8 with either the brief 4-item PHQ-4 anxiety and depression screening tool[10] or the slightly longer GAD-7 anxiety and PHQ-9 depression scales.[16,29,30]

The higher-order general-factor structure revealed by our analyses has been shown in previous studies[9] investigating the latent dimensions of somatic symptoms that have revealed an overarching somatic symptom factor. Similarly, the symptom clusters identified as subcomponents of this general factor in the present study closely mirror symptom clusters reported in earlier studies.[8,9] From a clinical viewpoint, these clusters correspond to 4 common medical syndromes (ie, pain, gastrointestinal symptoms, cardiopulmonary symptoms, and fatigue). Moreover, this factor structure allows for aggregation of the individual item scores into a simple, easily interpretable sum score ranging from 0 to 32 points. The analyses showed that the factor structure is strictly invariant for age and sex. This is important for the use of the scale in diverse populations because it permits the comparison of SSS-8 scores from respondents of different ages and sexes.[9,20]

The 5 severity categories derived from our data were based on the empirical distribution of the SSS-8 sum score. The high and very high symptom burden categories lay in the upper 5% of the severity distribution (ie, ≥95th percentile), which is common in clinical practice. By determining 5 severity categories (no to minimal, low, medium, high, and very high), we follow the taxonomy of frequently used self-report measures, for example, the PHQ-9 depression scale.[16,30] Each increase in severity category is associated with a large increase in health care visits, which emphasizes the validity of the categories. In addition to an interpretation based on the severity category, there is a second way to interpret the SSS-8 score: the sex- and age-specific percentile reference scores can be used to locate an SSS-8 score's relative position within the general population. The reference scores can also be used to aid patient communication because percentage ranks are easy to understand and some patients may feel relieved to know that they are at a certain level of somatic symptom burden (eg, "About X of 100 people experience lower somatic symptom burden than you do").

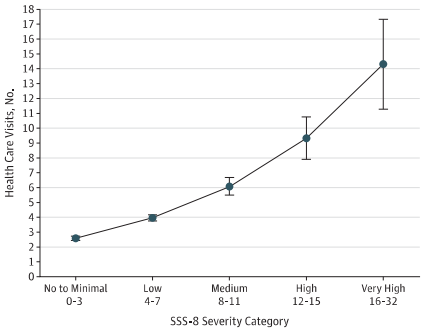
To facilitate the assessment of changes in the severity of somatic symptoms, the SSS-8 refers to recent symptoms (ie, in the past 7 days), whereas the PHQ-15 referred to more chronic symptoms (ie, in the past 4 weeks). Therefore, we believe that the SSS-8 is particularly suitable to assist the close monitoring of somatic symptom burden and to assess response to treatment. Similar to the PHQ-9 depression scale, for which responsiveness to treatment has been demonstrated in longitudinal studies,[31,32] we believe that future studies will demonstrate that the SSS-8 is useful in assessing changes in the severity of somatic symptoms over time.

This study has several strengths, including its large sample size and the representativeness of the study sample. However, there are also several limitations. First, the response rate of 56.6% may be seen as low and could affect the generalizability. However, response rates in population studies are generally lower than those in clinical studies, and our response rate is substantially higher than the average response rates of questionnaire surveys (ranging between 45% and 50%) as reported in 2 meta-analyses.[33,34] Second, because our study was cross-sectional, longitudinal studies are needed to determine the scale's test-retest reliability (ie, the stability of SSS-8 scores over a defined period during which no change in somatic symptom burden occurred), its responsiveness to treatment, and its minimum clinically important difference.[35] The minimum clinically important difference is of special interest for treatment evaluation because it identifies which intraindividual SSS-8 score difference is clinically meaningful. Third, our sample was drawn from the general population; the SSS-8 requires further study in patients receiving clinical care. However, the SSS-8 has been applied successfully in patient samples as part of the *DSM-5* field trials.[15] The *DSM-5* field trials combined with our study findings suggest that the SSS-8 may be applicable to both clinical and general-population samples.

An important question, however, concerns the representativeness of the German study sample and the ability to generalize the study results to other countries. There is good reason to assume that the results regarding validity and normative data of the SSS-8 may be applicable to other Western countries. Studies using the PHQ-9 depression measure (which showed substantial correlations with the SSS-8) revealed similar distributions and cutoff scores in large North American[30] and German[36] samples. Therefore, we do not expect substantial differences in the distribution of SSS-8 scores among Western countries.

Because of the high prevalence and the functional impairment associated with somatic symptoms, their careful consideration remains especially important in patient care. Evidence-based treatments are available for patients with high somatic symptom burden,[37-39] and the SSS-8 may serve as a

Figure 3. Somatic Symptom Scale–8 (SSS-8) Severity Categories and Health Care Use During a 12-Month Period



The number of health care visits was estimated using a negative binomial regression model that adjusted for depression (Patient Health Questionnaire–2) and anxiety (Generalized Anxiety Disorder–2) scores; error bars show 95% CIs.

Copyright 2014 American Medical Association. All rights reserved.

Downloaded From: http://archinte.jamanetwork.com/ by a University Medical Center Hamburg - Eppendorf User on 04/10/2014

patient-reported outcome measure to evaluate these treatments. Given the time constraints and competing demands for busy clinicians,[40] practical instruments for symptom assessment are needed to ease clinicians' workload and facilitate research. The SSS-8 is efficient in that it is brief, reliable, and valid and can be completed entirely by the patient. As with the PHQ-9 depression measure[30] and the GAD-7 anxiety measure,[29] which have had substantial uptake,[16] the SSS-8 somatic symptom measure should have considerable usefulness in busy medical settings and clinical research.

ARTICLE INFORMATION

Accepted for Publication: September 3, 2013.

Published Online: November 25, 2013. doi:10.1001/jamainternmed.2013.12179.

Author Contributions: Mr Gierk and Dr Löwe had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.
Study concept and design: Gierk, Kohlmann, Zenger, Löwe.
Acquisition of data: Spangenberg, Zenger, Brähler.
Analysis and interpretation of data: Gierk, Kohlmann, Kroenke, Zenger, Löwe.
Drafting of the manuscript: Gierk, Kohlmann, Löwe.
Critical revision of the manuscript for important intellectual content: All authors.
Statistical analysis: Gierk, Kohlmann, Zenger, Brähler.
Administrative, technical, or material support: Gierk, Spangenberg, Zenger, Löwe.
Study supervision: Kohlmann, Zenger, Löwe.

Conflict of Interest Disclosures: None reported.

Additional Contributions: Alexandra M. Murray, DPhil, Department of Psychosomatic Medicine and Psychotherapy, University Medical Center Hamburg-Eppendorf, and Schön Clinic Hamburg Eilbek, Germany, provided comments on the manuscript. No financial compensation was given for this service.

Correction: This article was corrected on December 11, 2013, to restore an omitted line of text in the Discussion section.

REFERENCES

1. Kroenke K, Arrington ME, Mangelsdorff AD. The prevalence of symptoms in medical outpatients and the adequacy of therapy. *Arch Intern Med*. 1990;150(8):1685-1689.

2. Kroenke K, Zhong X, Theobald D, Wu J, Tu W, Carpenter JS. Somatic symptoms in patients with cancer experiencing pain or depression: prevalence, disability, and health care use. *Arch Intern Med*. 2010;170(18):1686-1694. doi:10.1001/archinternmed.2010.337.

3. Kohlmann S, Gierk B, Hümmelgen M, Blankenberg S, Löwe B. Somatic symptoms in patients with coronary heart disease: prevalence, risk factors, and quality of life. *JAMA Intern Med*. 2013;173(15):1469-1471. doi:10.1001/jamainternmed.2013.6835.

4. Simon GE, VonKorff M, Piccinelli M, Fullerton C, Ormel J. An international study of the relation between somatic symptoms and depression. *N Engl J Med*. 1999;341(18):1329-1335. doi:10.1056/NEJM199910283411801.

5. Rief W, Martin A, Klaiberg A, Brähler E. Specific effects of depression, panic, and somatic symptoms on illness behavior. *Psychosom Med*. 2005;67(4):596-601. doi:10.1097/01.psy.0000171158.59706.e7.

6. Kroenke K, Jackson JL, Chamberlin J. Depressive and anxiety disorders in patients presenting with physical complaints: clinical predictors and outcome. *Am J Med*. 1997;103(5):339-347.

7. Fink P, Toft T, Hansen MS, Ørnbøl E, Olesen F. Symptoms and syndromes of bodily distress: an exploratory study of 978 internal medical, neurological, and primary care patients. *Psychosom Med*. 2007;69(1):30-39.

8. Zijlema WL, Stolk RP, Löwe B, Rief W, White PD, Rosmalen JGM; BioSHaRE. How to assess common somatic symptoms in large-scale studies: a systematic review of questionnaires. *J Psychosom Res*. 2013;74(6):459-468. doi:10.1016/j.jpsychores.2013.03.093.

9. Witthöft M, Hiller W, Loch N, Jasper F. The latent structure of medically unexplained symptoms and its relation to functional somatic syndromes. *Int J Behav Med*. 2013;20(2):172-183. doi:10.1007/s12529-012-9237-2.

10. Löwe B, Wahl I, Rose M, et al. A 4-item measure of depression and anxiety: validation and standardization of the Patient Health Questionnaire-4 (PHQ-4) in the general population. *J Affect Disord*. 2010;122(1-2):86-95. doi:10.1016/j.jad.2009.06.019.

11. Wahl I, Löwe B, Bjorner JB, et al. Standardization of depression measurement: a common metric was developed for 11 self-report depression scales. *J Clin Epidemiol*. In press.

12. Dowrick C, Leydon GM, McBride A, et al. Patients' and doctors' views on depression severity questionnaires incentivised in UK quality and outcomes framework: qualitative study. *BMJ*. 2009;338:b663. doi:10.1136/bmj.b663.

13. Kroenke K, Spitzer RL, Williams JBW. The PHQ-15: validity of a new measure for evaluating the severity of somatic symptoms. *Psychosom Med*. 2002;64(2):258-266.

14. Derogatis LR. *SCL-90-R—Administration, Scoring, and Procedures Manual*. 3rd ed. Minneapolis, MN: National Computer Systems, Inc; 1994.

15. Narrow WE, Clarke DE, Kuramoto SJ, et al. *DSM-5* field trials in the United States and Canada, part III: development and reliability testing of a cross-cutting symptom assessment for *DSM-5*. *Am J Psychiatry*. 2013;170(1):71-82. doi:10.1176/appi.ajp.2012.12071000.

16. Kroenke K, Spitzer RL, Williams JBW, Löwe B. The Patient Health Questionnaire somatic, anxiety, and depressive symptom scales: a systematic review. *Gen Hosp Psychiatry*. 2010;32(4):345-359. doi:10.1016/j.genhosppsych.2010.03.006.

17. Wild D, Grove A, Martin M, et al; ISPOR Task Force for Translation and Cultural Adaptation. Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: report of the ISPOR Task Force for Translation and Cultural Adaptation. *Value Health*. 2005;8(2):94-104. doi:10.1111/j.1524-4733.2005.04054.x.

18. Löwe B, Kroenke K, Gräfe K. Detecting and monitoring depression with a two-item questionnaire (PHQ-2). *J Psychosom Res*. 2005;58(2):163-171. doi:10.1016/j.jpsychores.2004.09.006.

19. Kroenke K, Spitzer RL, Williams JBW, Monahan PO, Löwe B. Anxiety disorders in primary care: prevalence, impairment, comorbidity, and detection. *Ann Intern Med*. 2007;146(5):317-325.

20. Gregorich SE. Do self-report instruments allow meaningful comparisons across diverse population groups? testing measurement invariance using the confirmatory factor analysis framework. *Med Care*. 2006;44(11)(suppl 3):S78-S94. doi:10.1097/01.mlr.0000245454.12228.8f.

21. Myers TA. Goodbye, listwise deletion: presenting hot deck imputation as an easy and effective tool for handling missing data. *Commun Methods Meas*. 2011;5(4):297-310. doi:10.1080/19312458.2011.624490.

22. American Association for Public Opinion Research. Standard definitions: final dispositions of case codes and outcome rates for surveys. 7th ed. 2011. http://www.aapor.org. Accessed July 19, 2013.

23. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equ Model Multidiscip J*. 1999;6(1):1-55. doi:10.1080/10705519909540118.

24. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 5th ed. Arlington, VA: American Psychiatric Publishing; 2013.

25. Barsky AJ, Orav EJ, Bates DW. Somatization increases medical utilization and costs independent of psychiatric and medical comorbidity. *Arch Gen Psychiatry*. 2005;62(8):903-910. doi:10.1001/archpsyc.62.8.903.

26. Barsky AJ, Orav EJ, Bates DW. Distinctive patterns of medical care utilization in patients who somatize. *Med Care*. 2006;44(9):803-811. doi:10.1097/01.mlr.0000228028.07069.59.

27. Sha MC, Callahan CM, Counsell SR, Westmoreland GR, Stump TE, Kroenke K. Physical symptoms as a predictor of health care use and mortality among older adults. *Am J Med*. 2005;118(3):301-306. doi:10.1016/j.amjmed.2004.12.015.

28. Löwe B, Spitzer RL, Williams JBW, Mussell M, Schellberg D, Kroenke K. Depression, anxiety and somatization in primary care: syndrome overlap and functional impairment. *Gen Hosp Psychiatry*. 2008;30(3):191-199. doi:10.1016/j.genhosppsych.2008.01.001.

29. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med*. 2006;166(10):1092-1097. doi:10.1001/archinte.166.10.1092.

30. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9: validity of a brief depression severity

Copyright 2014 American Medical Association. All rights reserved.

Downloaded From: http://archinte.jamanetwork.com/ by a University Medical Center Hamburg - Eppendorf User on 04/10/2014

measure. *J Gen Intern Med*. 2001;16(9):606-613. doi:10.1046/j.1525-1497.2001.016009606.x.

31. Löwe B, Unützer J, Callahan CM, Perkins AJ, Kroenke K. Monitoring depression treatment outcomes with the Patient Health Questionnaire-9. *Med Care*. 2004;42(12):1194-1201.

32. Löwe B, Kroenke K, Herzog W, Gräfe K. Measuring depression outcome with a brief self-report instrument: sensitivity to change of the Patient Health Questionnaire (PHQ-9). *J Affect Disord*. 2004;81(1):61-66. doi:10.1016/S0165-0327(03)00198-8.

33. Shih T-H, Fan X. Comparing response rates from web and mail surveys: a meta-analysis. *Field Methods*. 2008;20(3):249-271. doi:10.1177/1525822X08317085.

34. Horn PSV, Green KE, Martinussen M. Survey response rates and survey administration in counseling and clinical psychology: a meta-analysis. *Educ Psychol Meas*. 2009;69(3):389-403. doi:10.1177/0013164408324462.

35. Copay AG, Subach BR, Glassman SD, Polly DW Jr, Schuler TC. Understanding the minimum clinically important difference: a review of concepts and methods. *Spine J*. 2007;7(5):541-546. doi:10.1016/j.spinee.2007.01.008.

36. Löwe B, Spitzer RL, Gräfe K, et al. Comparative validity of three screening questionnaires for *DSM-IV* depressive disorders and physicians' diagnoses. *J Affect Disord*. 2004;78(2):131-140.

37. Jackson JL, O'Malley PG, Kroenke K. Antidepressants and cognitive-behavioral therapy for symptom syndromes. *CNS Spectr*. 2006;11(3):212-222.

38. Kroenke K. Efficacy of treatment for somatoform disorders: a review of randomized controlled trials. *Psychosom Med*. 2007;69(9):881-888. doi:10.1097/PSY.0b013e31815b00c4.

39. Henningsen P, Zipfel S, Herzog W. Management of functional somatic syndromes. *Lancet*. 2007;369(9565):946-955. doi:10.1016/S0140-6736(07)60159-7.

40. Klinkman MS. Competing demands in psychosocial care: a model for the identification and treatment of depressive disorders in primary care. *Gen Hosp Psychiatry*. 1997;19(2):98-111. doi:10.1016/S0163-8343(96)00145-4.

Invited Commentary

Assessing Somatic Symptoms in Clinical Practice

Arthur J. Barsky, MD

The interesting study by Gierk et al[1] tests the psychometric properties of a brief, self-report somatic symptom inventory and raises several broader questions about the sources of somatic symptoms and the use of symptom inventories in clinical practice. As the authors point out, routine clinical assessment of somatic symptoms has many functions: although obviously serving as guideposts to aid in diagnosis, they are also important in their own right as indexes of suffering and distress and can be used to assess the effectiveness of treatment over time.

Related article page 399

Somatic symptoms are a complex and multidetermined phenomenon, and they do not bear a fixed, one-to-one relationship with medical morbidity. Their close association with psychiatric disorders, particularly anxiety and depressive disorder, is well established and has been confirmed again in this study. Thus, Gierk and colleagues found that somatic symptoms were more highly correlated with anxiety and depression than with self-reported general health. Therefore, it should not be surprising that there is a substantial body of psychiatric literature on the prevalence, incidence, and nature of somatic symptoms as concomitants of psychiatric disorder. The Somatic Symptom Scale–8 (SSS-8) used in this study is derived from the Patient Health Questionnaire–15, a somatization scale that was developed to assess the somatic symptoms that frequently accompany psychiatric disorders rather than serious medical diseases, that is, to assess somatization rather than medical morbidity. The component items of the SSS-8, therefore, are closely associated with psychiatric distress, medically unexplained symptoms, and diagnosable somatoform disorders. This close association of somatic symptoms with psychiatric disorders is one reason that approximately one-third of the symptoms reported by patients in primary care and specialty medical settings remain medically unexplained after adequate evaluation.[2] This of course does not invalidate their use in the general medical setting but does underscore the fact that they are tapping psychiatric disorder as much as medical disorder.

The lack of a one-to-one relationship between somatic symptoms and medical morbidity is also apparent in the enormous interindividual variability seen in the symptoms of patients with the same severity and extent of medical disease. This has been documented in conditions as diverse as degenerative disk disease, coronary artery disease, asthma, anemia, diabetes mellitus, cholelithiasis, insomnia, and benign prostatic hypertrophy. Even when significant medical disease is present, patients' symptoms may nonetheless be unrelated to their disease, the association being coincidental rather than causal. Symptoms that appear to be disease specific may not in fact be caused by the ostensible disease.[3] When putative disease-specific symptoms (eg, headache and fatigue) have high base rates in the general population and the disease is common (eg, hypertension), the symptoms and disease are likely to occur concomitantly on a statistical basis alone.[4] Somatic symptoms can even result from learning for the first time that one has a medical condition that was previously asymptomatic. Thus, receiving the initial diagnosis of hypertension is associated with the onset of symptoms commonly assumed to be caused by hypertension (eg, headache, dizziness, weakness, and epistaxis), even after controlling for the initiation of antihypertensive therapy.[5]

To add to the complexity of the phenomenology of somatic symptoms, epidemiologic surveys of ostensibly healthy nonpatient populations reveal that bothersome symptoms are prevalent among people with no serious medical and psychiatric disorders.[6] As many as 80% to 90% of adults experience at least 1 symptom every 1 to 3 weeks,[7] and only 19% to 27% of people report having had no symptoms in the previous 3 days.[8] Thus, 14% to 45% of American adults report being bothered

Copyright 2014 American Medical Association. All rights reserved.

Downloaded From: http://archinte.jamanetwork.com/ by a University Medical Center Hamburg - Eppendorf User on 04/10/2014

Journal of Psychosomatic Research 78 (2015) 352–355

Contents lists available at ScienceDirect

Journal of Psychosomatic Research

Assessing somatic symptom burden: A psychometric comparison of the Patient Health Questionnaire—15 (PHQ-15) and the Somatic Symptom Scale—8 (SSS-8)

Benjamin Gierk *, Sebastian Kohlmann, Anne Toussaint, Inka Wahl, Christian A. Brünahl, Alexandra M. Murray, Bernd Löwe

Department of Psychosomatic Medicine and Psychotherapy, University Medical Center Hamburg-Eppendorf, Germany
Schön Clinic Hamburg-Eilbek, Germany

ARTICLE INFO

Article history:
Received 29 July 2014
Received in revised form 4 November 2014
Accepted 9 November 2014

Keywords:
PHQ-15
SSS-8
Somatization
Somatic symptom burden
Psychometrics

ABSTRACT

Objective: The Patient Health Questionnaire—15 (PHQ-15) is a frequently used questionnaire to assess somatic symptom burden. Recently, the Somatic Symptom Scale—8 (SSS-8) has been published as a short version of the PHQ-15. This study examines whether the instruments' psychometric properties and estimates of symptom burden are comparable.
Methods: Psychosomatic outpatients ($N = 131$) completed the PHQ-15, the SSS-8 and other questionnaires (PHQ-9, GAD-7, WI-7, SF-12). Item characteristics and measures of reliability, validity, and symptom severity were determined and compared.
Results: The reliabilities of the PHQ-15 and SSS-8 were $\alpha = 0.80$ and $\alpha = 0.76$, respectively and both scales were highly correlated ($r = 0.83$). The item characteristics were comparable. Both instruments showed the same pattern of correlations with measures of depression, anxiety, health anxiety and health-related quality of life ($r = 0.32$ to 0.61). On both scales a 1-point increase was associated with a 3% increase in health care use. The percentile distributions of the PHQ-15 and the SSS-8 were similar. Using the same thresholds for somatic symptom severity (5, 10, and 15 points), both instruments identified nearly identical subgroups of patients with respect to health related quality of life.
Conclusion: The PHQ-15 and the SSS-8 showed similar reliability and validity but the comparability of severity classifications needs further evaluation in other populations. Until then we recommend the use of the previously established thresholds. Overall, the SSS-8 performed well as a short version of the PHQ-15 which makes it preferable for assessment in time restricted settings.

© 2014 Elsevier Inc. All rights reserved.

Introduction

Somatic symptoms are common in the general population and [1,2]. They are prevalent in different medical conditions such as cancer [3] or coronary heart disease [4] and in the context of mental disorders, e.g. somatoform disorders, anxiety disorders, and depression [5–8]. Because somatic symptoms are related to impaired functioning, decreased health related quality of life, increased health care service use and psychological distress [3,4,7,9,10], the assessment of somatic symptom burden is essential in evidence based patient care and research [11–13].

One of the most frequently used self-report questionnaires to assess somatic symptom severity is the Patient Health Questionnaire—15 (PHQ-15) [14,15]. This questionnaire asks respondents to indicate how much they have been bothered by fifteen common somatic complaints within the past four weeks. The response format is a three point Likert scale and responses can be aggregated to a sum-score that spans between 0 and 30 points. Severity thresholds (0–4 no–minimal, 5–9 low, 10–14 medium, and 15–30 high) and general population normative data are available to aid the interpretation of scores [16]. Studies have demonstrated that the PHQ-15 is a reliable and valid measure of somatic symptom burden [15].

Recently, the Somatic Symptom Scale—8 (SSS-8) has been published as a short version of the PHQ-15 [9]. This eight-item questionnaire was developed for the application in settings with limited measurement time, such as in busy primary care settings [17] or the conduction of epidemiologic surveys. At this time, short scales with good psychometric properties were scarce [15]. The items of the SSS-8 were selected on the basis of empirical studies using the PHQ-15. The item selection criteria were symptom prevalence in primary care, association with measures of functioning, and statistical commonalities with the other items of the scale. The response format was chosen to be a five point

* Corresponding author at: Department of Psychosomatic Medicine and Psychotherapy, University Medical Center Hamburg-Eppendorf, Martinistraße 52, 20246 Hamburg, Germany. Tel.: +49 40 7410 53993; fax: +49 40 7410 54975.
E-mail address: bgierk@uke.de (B. Gierk).

http://dx.doi.org/10.1016/j.jpsychores.2014.11.006
0022-3999/© 2014 Elsevier Inc. All rights reserved.

Likert scale in order to make the SSS-8 compatible with psychometric instruments designed by the US National Institute of Health PROMIS initiative (http://www.nihpromis.org). In order to focus on the most recent symptoms, the response time frame covers the last 7 days. Unlike the PHQ-15, the sum-score of the SSS-8 spans between 0 and 32 points. Severity thresholds and general population normative data are available to facilitate the interpretation of scores. Reliability and validity of the scale could be demonstrated in a first study [9].

To date, no study has empirically compared the measurement characteristics of the PHQ-15 and the SSS-8. The aim of this article is to present data from a psychometric comparison of both instruments using a sample of psychosomatic outpatients. We address the following questions:

1. Are the two instruments comparable with respect to item characteristics, reliability, and validity?
2. Do the severity thresholds of both instruments generate groups with comparable levels of health related quality of life?

Method

Procedure and participants

Between March and April 2014, 136 patients were consecutively approached at their first visit at the psychosomatic outpatient clinic at the University Medical Center Hamburg-Eppendorf, Germany. The outpatient clinic specializes in affective-, anxiety-, somatoform-, eating-, and personality disorders. Of these 136 patients, 131 were included in the study (96% response rate). They completed a set of paper-pencil self-report questionnaires in the waiting room before their consultation with a medical doctor or psychologist. Because the SSS-8 and the PHQ-15 include similar items, both questionnaires were separated by other self-report instruments in order to minimize carry-over-effects. Inclusion criteria were being over 18 years old and providing written informed consent. Exclusion criteria were insufficient German language skills and severe physical illness or mental conditions that made participation impossible. Severe physical illness was defined as illness that required emergency treatment. Mental conditions that made participation impossible were conditions that needed urgent psychiatric treatment or hospitalization (e.g. acute suicidality, severe self-harm, risk of violence against others, dementia, inability to take care of self). The study was approved by the local medical ethics committee (no. PV4222).

Measures

In addition to the PHQ-15 and SSS-8, patients completed four self-report questionnaires:

1. The Patient Health Questionnaire—9 (PHQ-9) [18] as a measure of depression: This questionnaire contains items reflecting the nine DSM-5 criteria of major depression. A sum-score between 0 and 27 points indicates the severity of depression (with a higher score indicative of a higher depression severity).
2. The Generalized Anxiety Disorder Scale—7 (GAD-7) [19,20] as a measure of anxiety: This seven-item instrument was initially designed to measure the severity of generalized anxiety but is also sensitive to posttraumatic stress, panic and social phobia. A sum-score between 0 and 21 points indicates the severity of anxiety (higher scores corresponding to higher levels of anxiety).
3. The Whitley Index—7 (WI-7) [21] as a measure of health anxiety: patients' catastrophizing thoughts and worries about their physical condition is the focus of this seven-item instrument. A sum-score between 0 and 7 indicates the severity of health anxiety (with higher scores corresponding to higher levels of health anxiety).
4. The 12-item Health Survey (SF-12) [22] as a measure of health-related quality of life: The SF-12 comprises of two components: a measure of mental and physical health-related quality of life (i.e. the Mental Component Scale 'SF-12 MCS' and the Physical Component Scale 'SF-12 PCS'). Both measures are standardized to a mean of 50 and a standard deviation of 10 to facilitate the comparison of individuals within the general population. It is important to note that higher scores correspond to higher health-related quality of life.

In addition to these questionnaires, patients reported their number of health care consultations within the last 4 weeks (e.g. visits of general practitioner, days in hospital or in a day clinic) and demographic characteristics. Moreover, the diagnoses made by the consulting clinician according to the 10th edition of the International Classification of Diseases (ICD-10) were recorded.

Statistical analyses

We calculated means, standard deviations, corrected item-total-correlations and the number of missing values per item. To handle missing data, we used all available information (pairwise deletion). Pearson correlations between the PHQ-15, SSS-8, PHQ-9, GAD-7, WI-7, SF-12 PCS, and SF-12 MCS were calculated. Reliability was assessed using Cronbach's α. Percentile distributions of somatic symptom severity for the PHQ-15 and SSS-8 were also calculated. The associations between the PHQ-15 and SSS-8 sum-scores and the number of health care visits were analysed using negative binomial regressions. To study the comparability of the severity thresholds of both scales, we calculated the mean (SD) physical health-related quality of life (SF-12 PCS) for groups of patients which were classified according to the PHQ-15 and the SSS-8. We quantified the differences found using the standardized mean difference (Cohen's d).

Results

Our sample of 131 psychosomatic outpatients was characterized as follows: 66% of the patients were female, the mean age was 41.87 (SD = 15.03), 31% were married, and 50% were living with a partner. The distribution of years of formal education was: 9 years (18%), 10 years (34%), 11–13 years (29%), and more than 13 years (19%). In total, clinicians made 189 diagnoses and multiple diagnoses per patient were possible. The distribution of diagnoses was as follows: 43% depressive disorders (F32-F33), 14% somatoform disorders (F45), 13% anxiety disorders (F40–F41), 12% personality disorders (F60–F61), 5% substance use disorders (F10–F19), 4% eating disorders (F50), 2% adjustment disorders (F43.2), 2% obsessive compulsive disorder (F42), 2% psychological and behavioural factors associated with disorders or diseases classified elsewhere (F54), 1% posttraumatic stress or acute stress reaction (F43.0, F43.1), and 3% other disorders.

The mean (SD) sum-scores of the PHQ-15 and the SSS-8 were 12.12 (5.63) and 12.96 (6.50), respectively. The internal consistencies were $\alpha = 0.80$ for the PHQ-15 and $\alpha = 0.76$ for the SSS-8. The item difficulty varied in both scales, which indicates that some burdensome somatic symptoms are more frequently reported than others. All standard deviations were comparable within the scales and the rates of missing values did not differ markedly (between 4% and 14% per item). Several corrected item-total-correlations were low (r_{it} below 0.40). Table 1 reports the descriptive statistics in detail.

The PHQ-15 and the SSS-8 sum-scores were highly correlated ($r = 0.83$). Both scales showed a similar pattern of correlations with self-report measures of depression (PHQ-9), anxiety (GAD-7), health anxiety (WI-7), and physical and mental health-related quality of life (SF-12 PCS and SF-12 MCS): Higher scores on the SSS-8 corresponded to *higher* levels of depression, anxiety, and health anxiety and *lower* levels of physical and mental health-related quality of life. Correlation coefficients are shown in Table 2.

Although the PHQ-15 and the SSS-8 have different numbers of items and a different response format, the sum-scores of both scales have a comparable percentile distribution which is shown in Table 3. This high similarity in the distribution is present across almost the whole range of PHQ-15 and SSS-8 sum-scores except for the 99th percentile.

Patient-reported counts of health care use over a period of 4-weeks had a range of 0 to 12 health care visits with a median of 2 (1st quartile = 1, 3rd quartile = 4). Negative binomial regressions indicated that the PHQ-15 and the SSS-8 had small but statistically significant associations with health care use. A 1-point increase on the PHQ-15 was associated with a 3% increase in health care use (incidence rate ratio = 1.03 [95% CI = 1.01 to 1.05], $z = 2.64$, $p = 0.008$). The same could be shown for the SSS-8: A 1-point increase on the SSS-8 was associated with a 3% increase in health care use (incidence rate ratio = 1.03 [95% CI = 1.01 to 1.05], $z = 2.56$, $p = 0.01$).

To examine the comparability of the PHQ-15 and the SSS-8 severity estimations, we compared the physical health-related quality of life (SF-12 PCS scores) of patients who were grouped according to the PHQ-15 and SSS-8 severity thresholds. To quantify the differences in severity estimates for each group we calculated the standardized mean difference (Cohen's d). The original PHQ-15 (0–4 no–minimal, 5–9 low, 10–14 medium, and

Table 1
Descriptive statistics of the PHQ-15 and SSS-8 ($n = 131$)

Patient Health Questionnaire—15 (PHQ-15)					Somatic Symptom Scale—8 (SSS-8)				
Item	Missing	M	SD	r_{it}	Item	Missing	M	SD	r_{it}
Stomach pain	8%	0.70	0.80	0.31	Stomach or bowel problems	11%	1.48	1.39	0.33
Back pain	7%	1.02	0.80	0.41	Back pain	8%	1.66	1.41	0.44
Pain in your arms, legs, or joints (knee, hips, etc.)	7%	1.07	0.77	0.37	Pain in your arms, legs, or joints	8%	1.55	1.41	0.54
Menstrual cramps or other problems with your periods [women only]	8%	0.34	0.67	0.24					
Headaches	14%	0.35	0.64	0.26	Headaches	6%	1.51	1.29	0.53
Chest pain	7%	0.93	0.76	0.42	Chest pain or shortness of breath	7%	1.16	1.27	0.58
Dizziness	11%	0.66	0.72	0.58	Dizziness	8%	1.14	1.28	0.40
Fainting spells	7%	0.84	0.76	0.49					
Feeling your heart pound or race	12%	0.14	0.41	0.33					
Shortness of breath	9%	0.86	0.82	0.57					
Pain or problems during sexual intercourse	10%	0.71	0.76	0.48					
Constipation, loose bowels, or diarrhea	10%	0.80	0.78	0.54					
Nausea, gas, or indigestion	10%	0.94	0.81	0.50					
Feeling tired or having low energy	4%	1.60	0.63	0.38	Feeling tired or having low energy	5%	2.54	1.30	0.44
Trouble sleeping	5%	1.44	0.71	0.19	Trouble sleeping	8%	2.11	1.32	0.41
Sum-score		12.12	5.63				12.96	6.50	

Note: Scoring according to the PHQ-15 (0 = not bothered at all, 1 = bothered a little, 2 = bothered a lot) and the SSS-8 (0 = not at all, 1 = a little bit, 2 = somewhat, 3 = quite a bit, 4 = very much), r_{it} = corrected item-total correlation.

15–30 high) and SSS-8 thresholds (0–3 no–minimal, 4–7 low, 8–11 medium, 12–15 high, and 16–32 very high) were applied. Given that our prior analyses suggested that the percentile distributions of the PHQ-15 and the SSS-8 were similar, we also decided to score the SSS-8 using the PHQ-15 thresholds (i.e. 0–4 no–minimal, 5–9 low, 10–14 medium, and 15–32 high) as an additional measure. Results are shown in Table 4.

Grouping patients according to the original severity thresholds of the PHQ-15 and the SSS-8 lead to groups that differed substantially on the SF-12 PCS (Cohen's $d = 0.21$ to 1.02) [23]. However, when the SSS-8 was scored using the PHQ-15 severity thresholds, the differences were remarkably smaller (Cohen's $d = 0.002$ to 0.11). Simply stated, in this sample, the congruence between the PHQ-15 and SSS-8 severity estimations could be increased by scoring the SSS-8 with the PHQ-15 thresholds (i.e. 0–4 no–minimal, 5–9 low, 10–14 medium, and 15–32 high).

Discussion

The present study compared the psychometric properties of the PHQ-15 and the SSS-8 in a sample of psychosomatic outpatients. Altogether, the results show that both instruments have comparable measures of validity and reliability. Hence, we suggest that the SSS-8 is an adequate alternative to the PHQ-15, especially in situations where measurement time is restricted.

Our analyses revealed that the Cronbach's α of the SSS-8 is nearly identical to the PHQ-15, although the SSS-8 contains only about half as many items. The item characteristics are also comparable, specifically: the number of missing values, the variance in the item difficulty, and the standard deviations within the scales. Items with low item-total-correlations can be found in both instruments. This may be due to the multidimensional latent structure of somatic symptom burden which has been previously described [9,24,25].

As expected, the correlation between the PHQ-15 and the SSS-8 was very high. In terms of validity, both measures showed substantial positive associations with depression, anxiety, health anxiety and health care use. This finding is consistent with previous studies which reported high co-morbidity of depressive, anxious and somatic symptoms and suggests that somatic symptom burden is related to health care use [5,26]. With regard to physical and mental health-related quality of life, both measures showed a substantial negative association (i.e. high symptom burden corresponded to poor health-related quality of life). In general, the pattern of correlations with the patient-reported outcomes described above was similar for the PHQ-15 and the SSS-8 which is further evidence of their comparability.

The continuous sum-scores of both PHQ-15 and SSS-8 were also comparable as indicated by similar percentile distributions. However, there were substantial differences in the severity estimation of the PHQ-15 and SSS-8 severity thresholds. These differences could, however, be reduced by scoring the SSS-8 with the PHQ-15 thresholds (i.e. 0–4 no–minimal, 5–9 low, 10–14 medium, and 15–32 high). In doing so, the SSS-8 severity estimations were closer to the PHQ-15 estimations, i.e. satisfactory congruence between the long and the short version could be achieved. However, this finding is based on a small and selective sample of psychosomatic outpatients. Given that both the original PHQ-15 and SSS-8 thresholds performed well in previous studies [9,14,26], it is too early to draw any definite conclusions about these new thresholds. This topic needs further research in large scale studies including several groups of participants and comprehensive measurement of impairment.

Our study has three major limitations: First, our findings are restricted to psychosomatic outpatient settings and cannot be easily generalized to other settings (e.g. items that were not included in the SSS-8 but in the PHQ-15 may be particularly important in other specific settings). Second, the PHQ-15 and the SSS-8 were included in the same

Table 2
Pearson correlations between the PHQ-15, SSS-8 and other included measures ($n = 131$)

	PHQ-15	SSS-8
PHQ-15	1.00	0.83
SSS-8	0.83	1.00
PHQ-9	0.59	0.58
GAD-7	0.51	0.43
WI-7	0.61	0.49
SF-12 PCS	−0.54	-0.53
SF-12 MCS	−0.38	-0.32

Table 3
Percentile distribution of sum-scores ($n = 131$)

Percentile	PHQ-15	SSS-8
1	1	1
5	3	3
10	5	4
20	6	6
30	8	9
40	10	10
50	11	12
60	12	13
70	14	16
80	16	18
90	20	21
95	22	23
99	24	29

Table 4
Analysis of physical health-related quality of life (SF-12 PCS) in patients grouped according to the PHQ-15 (original thresholds), SSS-8 (original thresholds), and the SSS-8 with PHQ-15 scoring thresholds ($n = 131$)

Severity	Descriptives			Standardized differences	
	1. Original PHQ-15 [a]	2. Original SSS-8 [b]	3. SSS-8 with PHQ-15 scoring [c]	1. vs. 2.	1. vs. 3.
No–minimal	50.12 (10.07)	57.56 (2.22)	50.10 (9.00)	1.02	0.002
Low	47.76 (9.19)	49.51 (7.19)	48.77 (8.73)	0.21	0.11
Medium	40.08 (10.69)	44.14 (10.65)	40.32 (11.28)	0.38	0.02
High	33.55 (9.82)	37.54 (11.16)	34.45 (9.43)	0.38	0.09
Very high [d]	–	35.07 (9.59)	–	–	–

Note: Means (SDs) correspond to SF-12 Physical Component Scale scores, Standardized differences are computed as the mean difference divided by the pooled standard deviations (Cohen's *d*).
[a] 0–4 no–minimal, 5–9 low, 10–14 medium, and 15–30 high.
[b] 0–3 no–minimal, 4–7 low, 8–11 medium, 12–15 high, and 16–32 very high.
[c] 0–4 no–minimal, 5–9 low, 10–14 medium, and 15–32 high.
[d] Only available for the SSS-8 original scoring method.

set of self-report questionnaires which was given to the patients. Although the instruments were separated by other questionnaires, carry-over effects could have been present to a certain degree. Still, we used this design instead of randomized presentation because computer-based data collection was not feasible in our setting. Third, our study design was cross-sectional. We cannot assess, therefore, the effect of different time frames of the PHQ-15 (i.e. symptom burden within the last 4 weeks) and SSS-8 (i.e. symptom burden within the last 7 days). A longitudinal study is needed to answer this question.

In this study overall, the SSS-8 performed well as a brief measure of somatic symptom burden in psychosomatic outpatients because psychometric properties are comparable to the PHQ-15. The estimates of symptom burden from both the SSS-8 and PHQ-15 were very similar when the SSS-8 was scored with the PHQ-15 thresholds. However, this finding is based on a small and selective sample of patients and needs further research. Until this time, we recommend that the original severity thresholds should be used [9,14].

Conflict of interest

The study was not externally funded. The authors have no competing interests to report.

Acknowledgments

We thank Aida Akyol for her support in the collection of data.

References

[1] Kroenke K, Arrington ME, Mangelsdorff AD. The prevalence of symptoms in medical outpatients and the adequacy of therapy. Arch Intern Med Aug 1990;150:1685–9.
[2] Van der Windt DAWM, Dunn KM, Spies-Dorgelo MN, Mallen CD, Blankenstein AH, Stalman WAB. Impact of physical symptoms on perceived health in the community. J Psychosom Res Mar 2008;64:265–74.
[3] Kroenke K, Zhong X, Theobald D, Wu J, Tu W, Carpenter JS. Somatic symptoms in patients with cancer experiencing pain or depression: prevalence, disability, and health care use. Arch Intern Med Oct 11 2010;170:1686–94.
[4] Kohlmann S, Gierk B, Hümmelgen M, Blankenberg S, Löwe B. Somatic symptoms in patients with coronary heart disease: prevalence, risk factors, and quality of life. JAMA Intern Med 2013;173:1469–71.
[5] Löwe B, Spitzer RL, Williams JBW, Mussell M, Schellberg D, Kroenke K. Depression, anxiety and somatization in primary care: syndrome overlap and functional impairment. Gen Hosp Psychiatry Jun 2008;30:191–9.
[6] Hanel G, Henningsen P, Herzog W, Sauer N, Schaefert R, Szecsenyi J, et al. Depression, anxiety, and somatoform disorders: vague or distinct categories in primary care? Results from a large cross-sectional study. J Psychosom Res Sep 2009;67:189–97.
[7] Rief W, Martin A, Klaiberg A, Brähler E. Specific effects of depression, panic, and somatic symptoms on illness behavior. Psychosom Med Jan 7 2005;67:596–601.
[8] Simon GE, VonKorff M, Piccinelli M, Fullerton C, Ormel J. An international study of the relation between somatic symptoms and depression. N Engl J Med 1999; 341:1329–35.
[9] Gierk B, Kohlmann S, Kroenke K, Spangenberg L, Zenger M, Brähler E, et al. The Somatic Symptom Scale-8 (SSS-8): a brief measure of somatic symptom burden. JAMA Intern Med Nov 2013;25:E1–9.
[10] Baliatsas C, van Kamp I, Hooiveld M, Yzermans J, Lebret E. Comparing non-specific physical symptoms in environmentally sensitive patients: prevalence, duration, functional status and illness behavior. J Psychosom Res May 2014;76:405–13.
[11] Jackson JL, O'Malley PG, Kroenke K. Antidepressants and cognitive-behavioral therapy for symptom syndromes. CNS Spectr Mar 2006;11:212–22.
[12] Kroenke K. Efficacy of treatment for somatoform disorders: a review of randomized controlled trials. Psychosom Med Dec 2007;69:881–8.
[13] Henningsen P, Zipfel S, Herzog W. Management of functional somatic syndromes. Lancet Mar 17 2007;369:946–55.
[14] Kroenke K, Spitzer RL, Williams JBW. The PHQ-15: validity of a new measure for evaluating the severity of somatic symptoms. Psychosom Med 2002;64:258–66.
[15] Zijlema WL, Stolk RP, Löwe B, Rief W, White PD, Rosmalen JGM. How to assess common somatic symptoms in large-scale studies: a systematic review of questionnaires. J Psychosom Res 2013;74:459–68.
[16] Kocalevent R-D, Hinz A, Brähler E. Standardization of a screening instrument (PHQ-15) for somatization syndromes in the general population. BMC Psychiatry 2013;13:91.
[17] Klinkman MS. Competing demands in psychosocial care: a model for the identification and treatment of depressive disorders in primary care. Gen Hosp Psychiatry Mar 1997;19:98–111.
[18] Kroenke K, Spitzer RL, Williams JBW. The PHQ-9. Validity of a brief depression severity measure. J Gen Intern Med 2001;16:606–13.
[19] Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006;166:1092–7.
[20] Löwe B, Decker O, Müller S, Brähler E, Schellberg D, Herzog W, et al. Validation and standardization of the Generalized Anxiety Disorder Screener (GAD-7) in the general population. Med Care 2008;46:266–74.
[21] Fink P, Ewald H, Jensen J, Sørensen L, Engberg M, Holm M, et al. Screening for somatization and hypochondriasis in primary care and neurological in-patients: a seven-item scale for hypochondriasis and somatization. J Psychosom Res Mar 1999;46:261–73.
[22] Ware JE, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. Med Care 1996;34:220–33.
[23] Cohen J. A power primer. Psychol Bull 1992;112:155–9.
[24] Witthöft M, Hiller W, Loch N, Jasper F. The latent structure of medically unexplained symptoms and its relation to functional somatic syndromes. Int J Behav Med 2012; 1–12.
[25] Fink P, Toft T, Hansen MS, Ørnbøl E, Olesen F. Symptoms and syndromes of bodily distress: an exploratory study of 978 internal medical, neurological, and primary care patients. Psychosom Med 2007;69:30–9.
[26] Kroenke K, Spitzer RL, Williams JBW, Löwe B. The Patient Health Questionnaire Somatic, Anxiety, and Depressive Symptom Scales: a systematic review. Gen Hosp Psychiatry 2010;32:345–59.

Monitoring somatic symptoms in patients with mental disorders:

Sensitivity to change and minimal clinically important difference of the

Somatic Symptom Scale – 8 (SSS-8)

Benjamin Gierk[1], Sebastian Kohlmann[2,3], Marion Hagemann-Göbel[1],

Bernd Löwe[2,3], & Yvonne Nestoriuc[2,3]

1. Department of Psychiatry and Psychotherapy, Asklepios Clinic North, Hamburg, Germany

2. University Medical Centre Hamburg-Eppendorf, Hamburg, Germany

3. Schön Clinic Hamburg Eilbek, Hamburg, Germany

Running tile: Monitoring somatic symptoms using the SSS-8
Word count main text (excl. tables, figures and references): 3.627
Word count abstract: 199
Tables: 4
Figures: 0

Correspondence:
Benjamin Gierk, M.Sc.
Department of Psychiatry and Psychotherapy
Asklepios Clinic North
Day clinic am Ochspark
Langenhorner Chaussee 560
22419 Hamburg
Germany
Phone: +49 40 18 18 87 52 49
Fax: +49 40 18 18 87 22 66
Email: benjamin.gierk@gmx.de

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Abstract

Objective: The Somatic Symptom Scale – 8 (SSS–8) is a brief questionnaire for the assessment of somatic symptom burden. This study examines its sensitivity to change and the minimal clinically important difference (MCID).

Method: 55 outpatients with mental disorders completed the SSS-8 and measures of anxiety, depression, and disability before and after receiving treatment. Effect sizes and correlations between the change scores were calculated. The MCID of SSS-8 was estimated using a one standard error of measurement threshold and the change in disability as an external criterion.

Results: There was a medium decline in somatic symptom symptom burden for the complete sample from pre- to post-assessment ($n = 55$, $d_z = 0.53$) and a large decline in a subgroup with very high somatic symptom burden at baseline ($n = 11$, $d_z = 0.94$). Decreases in somatic symptom burden were associated with decreases in anxiety ($r = 0.68$), depression ($r = 0.62$) and disability ($r = 0.51$). The MCID was estimated as a 4-point decrease on the SSS-8.

Conclusion: The SSS-8 is sensitive to change. A 4-point decrease reflects a clinically important improvement. Due to is brevity and sound psychometric properties, the SSS-8 is useful for monitoring somatic symptom burden.

Keywords: SSS-8, sensitivity to change, minimal clinically important difference, somatic symptom burden

1. Introduction

Burdensome somatic symptoms are one of the main reasons for health care consultations.[1] Patients with a wide range of medical conditions including cancer, coronary heart disease, depression, anxiety, and somatic symptom disorder suffer from burdensome somatic symptoms, which are related to functional impairment, absence from work, and reduced quality of life.[2–8] Therefore, identification of individuals suffering from high somatic symptom burden is important to initiate adequate treatment.[1,9]

Standardized patient-reported outcome measures like self-report questionnaires are a good option to assess, quantify, and monitor frequent conditions in primary care. A frequent challenge in this context is to make precise assessments in a limited quantity of time.[10] The Somatic Symptom Scale – 8 (SSS–8) is a brief self-report measure of somatic symptom burden that was especially designed for situations with restricted measurement time.[11] The SSS-8 is a brief version of the well validated and widely used Patient Health Questionnaire – 15 (PHQ-15).[12] It can be completed within less than one minute which makes it suitable for the assessment in health care units with limited time and resources. The SSS-8 consists of eight items reflecting very common somatic symptoms in primary care, such as gastrointestinal symptoms, musculoskeletal symptoms, pain symptoms, cardiopulmonary symptoms, and fatigue. On a five-point Likert scale, respondents indicate how much they were bothered by these symptoms within the last seven days. Answers are added up to a sum-score ranging between 0 and 32 points. Cut-off-scores indicate whether a patient suffers from minimal (0-3 points), low (4-7), medium (8-11), high (12-15), or very high (16-32) somatic symptom burden. Furthermore, gender and age specific general population norms are available.[11] Studies demonstrated good internal consistencies between $\alpha = 0.76$ and 0.86, a sound factor structure and substantial associations with related constructs such as depression, anxiety, wellbeing, and health care use.[11,13] The correlation with the long version PHQ-15 is very high ($r = 0.83$).[13]

Since its development there has been high resonance to the SSS-8 in the scientific community and clinical context. For example, it was used in the DSM-5 field trials as a measure of somatic symptom burden related to the new diagnosis of somatic symptom disorder (Note: In the DSM-5 field trials the SSS-8 was included with its former title Patient Health Questionnaire Somatic Symptom Short-Form, PHQ-SSS).[2,14] To date, the SSS-8 is available in English, German, and Japanese.[11,13,15] Despite of its fast application in the clinical and scientific context, two important psychometric parameters of the SSS-8 remain unknown: it's sensitivity to change and the minimal clinically important difference.

Sensitivity to change means a "measure's ability to detect change".[16] A questionnaire is sensitive to change if it shows a decline in the severity score when patients are repeatedly assessed before and after receiving evidence-based treatment. Given that a self-report questionnaire is sensitive to change, it is also important to examine the degree of change necessary for being considered clinically meaningful. Copay et al. defines the minimal clinically important difference (MCID) as follows: "A clinically important difference represents a change that would be considered meaningful and worthwhile by the patient such that he / she would consider repeating the intervention if it were his / her choice to make again. The minimum clinically important difference (MCID) is a threshold value for such a change." [17]

This study aims to add to the literature and specifically the usefulness of the scale by providing, for the first time, measures of sensitivity to change and MCID for the SSS-8. We tested whether the SSS-8 is able to detect change in somatic symptom severity in patients with mental disorders who receive evidence-based treatment. Moreover, we calculated the MCID using statistical criteria as well as the improvement in disability as an external criterion of clinical significance.

2. Method

2.1 Sample

We analysed consecutive data from patients who were treated at a psychiatric day clinic (Day clinic am Ochspark, Department of Psychiatry and Psychotherapy, Asklepios Clinic North, Hamburg, Germany). The day clinic uses standardized patient-reported outcomes measures at the beginning and at the end of treatment to monitor treatment efficacy as part of quality management. Written informed consent was obtained at the first day of treatment. Patients completed questionnaires on a voluntarily basis, and data were fully anonymised. The study was conducted in compliance with the Declaration of Helsinki. Harm to patients due to analysis of fully anonymised data could be excluded. An ethical approval from the local medical ethics committee for patient assessment using the SSS-8 was obtained in Gierk et al. (2015).[13]

We included patients aged 18 years or older, who had at least one diagnosis of a mental disorder, received treatment and took part in pre and post assessment. Exclusion criteria were severe visual, motor, or cognitive deficits, life-threatening conditions and insufficient language skills. Sample size was calculated for a paired t-test examining the pre-post difference of the SSS-8 sum-score. A recent meta-analysis showed an averaged pre-post effect size of Cohen's $d = 0.80$ for somatic symptom burden in patients with somatoform disorders who received psychotherapy.[18] We expected a more conservative pre-post effect size of Cohen's $d_z = 0.60$ [19] because we included patients from the whole spectrum of mental disorders (i.e., we included patients not only showing up for predominantly somatic symptom burden). A two sided significance level of 5% and a power of 95% were specified. Power analysis was calculated using the G*Power software, 3rd Version [20] and resulted in a minimum sample size of 39 patients.

2.2 Procedure

Patients took part in a consecutive assessment at the beginning and at the end of treatment. The treatment program includes 8 to 12 weeks (Monday to Friday from 8 a.m. to 4 p.m.). Within this period, patients received evidence-based treatment matched to their individual psychopathology. The following treatment components were conducted: Weekly individual psychotherapy session (cognitive behavioural therapy or psychodynamic therapy), group sessions of cognitive behavioural therapy (depression group, anxiety group, skills training, resources based group) or interactional psychodynamic therapy, group training of social competence, group sessions of psychoeducation about schizophrenia, family based interventions (e.g., psychoeducation, debating family conflicts), occupational therapy (e.g., handcraft, cooking, gardening, creative painting), physical activity (e.g., nordic walking, dancing, ball games, table tennis), support by a social worker (e.g., housing, finances, formalities with public authorities, employment), acupuncture, progressive muscle relaxation, mindfulness training, medical consultations and pharmacological treatment (e.g., antidepressants, antipsychotics, mood stabilizer, stimulants, benzodiazepines). The day clinic is part of the German public health system financed by health insurance companies available to the general population.

2.3 Measures

Besides the SSS-8 patients completed the Generalized Anxiety Disorder – 7 (GAD-7) scale, the Patient Health Questionnaire – 9 (PHQ-9) depression scale, and the WHO Disability Assessment Schedule (WHOADAS-12). The GAD-7 is a seven item self-report questionnaire sensitive to generalized anxiety, social anxiety, panic and posttraumatic stress.[21–23] On a four-point Likert scale patients rated how often they were bothered by anxious symptoms within the last two weeks. A sum score ranging between 0 and 21 points quantifies the severity of anxiety, and 5, 10, and 15 points are cut-off-scores for mild, moderate and severe anxiety. The PHQ-9 is a nine item self report questionnaire that quantifies the severity of depression within the last two weeks.

[22,24] On a four-point Likert scale patients indicate how often they experienced depressive symptoms within this time frame. The total amount of depression is quantified by a sum score ranging between 0 and 27 points, and 5, 10, 15, and 20 points are thresholds for mild, moderate, moderately severe and severe depression. Sensitivity to change of the PHQ-9 has been established in several studies.[25–27]

Disability was measured using the 12-item World Health Organisation Disability Assessment Schedule (WHODAS-12, second edition).[28] Patients indicated the degree of disability they experienced within the last 30 days in cognition, mobility, self-care, getting along with others, life activities, and social participation on a five-point Likert scale. A sum-score ranging between 0 and 48 points quantifies the total amount of disability. Reference values from the Australian general population show that 4, 7, 9, and 14 points correspond to the 75th, 85th, 90th, and 95th percentiles of disability.[29] In addition to the self-report questionnaires, we extracted age, gender, and diagnoses of mental disorders from the patients' health records. Diagnoses were made by the treating clinicians using the 10th edition of the International Classification of Diseases system (ICD-10).

2.4 Statistics

Nominal data are presented as absolute and relative numbers and continuous data are summarized in means and standard deviations. Distributions are described using percentile ranks. Sensitivity to change was examined using two approaches: First, the pre-post difference of the SSS-8 was calculated. A two sided paired t-test tested for statistical significance of the pre-post difference. Statistical significance was evaluated on a 5% α level. The magnitude of effect was quantified by the Cohen's d_z effect size measure (formula: mean of pre-post differences divided by the standard deviation of the pre-post differences; this can be calculated directly from the t-statistic: $d_z = t$ divided by the square root of the sample size).[19] Pre-post differences were examined for the complete sample and for a subgroup with very high somatic symptom burden at baseline. These patients were identified by applying the 'very high' cut-off-score of the SSS-8 ($\geq$ 16 points). We

analysed this specific subgroup because we expected that patients with very high somatic symptom burden at baseline assessment are most likely to show a decline on the SSS-8. If this subgroup does not show a change after receiving complex evidence-based treatment, it is very likely that the SSS-8 is not sensitive to change. Second, Pearson correlations were calculated between the change scores of the SSS-8, GAD-7, PHQ-9, and WHODAS-12 to examine whether a decline in somatic symptom burden was associated with a decline anxiety, depression and disability.

We defined the minimal clinical important difference (MCID) using a statistical distribution based method and an anchor based method that relates the change score to meaningful patient-reported external criterion.[17,30] As a statistical criterion we used the standard error of measurement (SEM) as a threshold value for the MCID. The SEM is a marker of statistical noise due to unreliability of a measurement instrument. A 'true' change score should exceed the SEM. Change scores below the SEM are likely a result of measurement error.[17] Therefore, one SEM can be seen as a benchmark of a reliably detectable change. Still, it is a statistical criterion and it is unclear whether an improvement of one SEM in somatic symptom burden is meaningful to the patients. Therefore, as second step we applied the anchor method proposed by Jaeschke et al. [30] and linked the SSS-8 pre-post change to the change in patient reported disability, i.e. the WHODAS-12 change score. We expected that a meaningful change in somatic symptom burden is related to a substantial change in disability. Missing values in the symptom severity measures were imputed by substituting the missing values with the median score of the available values. Median imputation was conducted if not more than three values per questionnaire were missing. Analysis were performed in SPSS 24 (IBM Inc.).

3. Results

3.1 Sample

160 patients were treated at the day clinic between December 2015 and November 2016. None of the patients had to be excluded due to severe visual, motor, or cognitive deficits, life-

threatening conditions, or insufficient language skills. 117 patients volunteered to take part in the assessment (73% response rate): 55 patients completed the pre and post questionnaire, 51 patients completed only the pre questionnaire, and 11 patients completed only the post questionnaire. In accordance with our inclusion criteria, 55 patients entered into the analysis. We favoured complete case analysis over multiple imputation, because we were mainly interested in the psychometric properties of the SSS-8 (i.e., we did not intend to precisely estimate treatment effects by including plausible substitute data for patients who missed one assessment). Moreover, a sample size of 55 patients exceed the minimum sample size calculated in the a-priori power analysis (n =39) and therefore was sufficient for the primary analysis. Rates of missing values in the questionnaires provided by the 55 patients who entered into the analyses were very low (0.98%) and handled by median imputation.

The mean age of the patients was 44 (SD = 11) years and 78% were female. Overall, 107 diagnoses of mental disorders were made with a median number of 2 diagnoses per patient and a range of 1 to 4 diagnoses per patient. Most common diagnoses were depressive disorders, personality disorders, and substance use disorders, followed by anxiety disorders and psychotic disorders. At baseline, the mean scores for somatic symptom burden, anxiety and depression exceeded the moderate / medium level, showing substantial symptom load. The mean disability score exceed the 95th percentile of the general population, which underlines that we examined a clinical sample of patients. Table 1 shows the characteristics of the sample.

Table 1. Characteristics of the sample

Sample size (no.)	55
Female gender (no., %)	43 (78.18)
Age (M, SD)	43.93 (11.69)
Diagnoses (no., %)	107 (100.00)
Depressive disorder	45 (42.06)
Personality disorder	15 (14.02)
Substance use disorder	10 (9.35)
Anxiety disorder	8 (7.48)
Psychotic disorder	7 (6.54)
Obsessive compulsive disorder	4 (3.74)
Posttraumatic stress disorder	4 (3.74)
Eating disorder	3 (2.80)
Organic brain disorder	3 (2.80)
Bipolar affective disorder	2 (1.87)
Somatoform disorder	2 (1.87)
Other	4 (3.74)
Baseline symptom severity (M, SD)	
Somatic symptom burden (SSS-8)	11.29 (5.74)
Anxiety severity (GAD-7)	11.11 (4.87)
Depression severity (PHQ-9)	13.64 (5.57)
Disability (WHODAS-12)	18.58 (8.29)

Notice: Number (no.), mean (M), standard deviation (SD), multiple diagnoses were possible, Somatic Symptom Scale – 8 (SSS-8), Generalized Anxiety Disorder Scale – 7 (GAD-7), Patient Health Questionnaire – 9 depression scale (PHQ-9), WHO Disability Assessment Schedule (WHODAS-12)

3.2 Sensitivity to change

From pre to post assessment the proportion of patients in the medium, high, and very high somatic symptom severity categories of the SSS-8 declined, whereas the proportion of patients in the no / minimal and low somatic symptom severity categories increased. This indicates an overall reduction in somatic symptom severity (see table 2).

Table 2. Somatic symptom severity at pre and post assessment

	Assessment	
SSS-8 severity category	Pre	Post
No / minimal (0-3)	3 (5.45%)	11 (20.00%)
Low (4-7)	11 (20.00%)	15 (27.27%)
Medium (8-11)	16 (29.09%)	13 (23.64%)
High (12-15)	14 (25.45%)	10 (18.18%)
Very high (16-32)	11 (20.00%)	6 (10.91%)

Notice: Absolute and relative numbers of patients

No change in somatic symptom burden was observed for 7% of patients. In 24% of the cases somatic symptom burden increased from pre to post assessment. In contrast, 69% showed a decrease in somatic symptom burden. Table 3. shows the distribution of the SSS-8 change scores.

Table 3. Distribution of Somatic Symptom Scale 8 change scores

Change score	**n**	**%**	**∑n**	**∑%**
-10	1	1.82	1	1.82
-9	1	1.82	2	3.64
-5	1	1.82	3	5.45
-4	4	7.27	7	12.73
-3	1	1.82	8	14.55
-2	3	5.45	11	20.00
-1	2	3.64	13	23.64
0	4	7.27	17	30.91
1	4	7.27	21	38.18
2	6	10.91	27	49.09
3	6	10.91	33	60.00
4	4	7.27	37	67.27
5	3	5.45	40	72.73
6	2	3.64	42	76.36
7	4	7.27	46	83.64
8	3	5.45	49	89.09
9	1	1.82	50	90.01
10	2	3.64	52	94.55
11	2	3.64	54	98.18
13	1	1.82	55	100.00

Notice: SSS-8 change scores and corresponding absolute and relative numbers of patients

On the complete sample level (n = 55), the mean SSS-8 sum score showed a 2.58 (95% CI 1.26 to 3.91) points decline from pre (M = 11.29, SD = 5.74) to post (M = 8.71, SD = 5.96) assessment, $t = 3.90$, $p < 0.001$, effect size $d_z = 0.53$. This effect was large for the subgroup of patients with very high somatic symptom burden at pre assessment (SSS-8 score ≥ 16 points, n = 11). These patients showed a 4.73 (95% CI 1.31 to 8.14) points decline from pre (M = 19.82, SD = 4.05) to post (M = 15.09, SD = 5.72) assessment, $t = 3.09$, $p = 0.01$, effect size $d_z = 0.94$.

Severity declines were also observed for the GAD-7 anxiety score (4.85 points decline, 95% CI 3.52 to 6.19, $t = 7.29$, $p < 0.001$, effect size $d_z = 0.98$), the PHQ-9 depression score (6.22 points decline, 95% CI 4.77 to 7.66, $t = 8.36$, $p < 0.001$, effect size $d_z = 1.16$) and the WHODAS 12 disability score (6.13 points decline, 95% CI 4.10 to 8.16, $t = 6.05$, $p < 0.001$, effect size $d_z = 0.81$). The SSS-8 change score correlated with the GAD-7 change score ($r = 0.68$, 95% CI 0.51 to 0.80, $p < 0.001$), the PHQ-9 change score ($r = 0.62$, 95% CI 0.42 to 0.76, $p < 0.001$) and the WHODAS-12 change score ($r = 0.51$, 95% CI 0.28 to 0.68, $p < 0.001$) showing that a decrease in somatic symptom burden is associated with a decrease in anxiety, depression and disability.

3.3 Minimal clinically important difference

First, we used a distribution based method to estimate the MCID. We calculated the SEM as a threshold value for a true change that stands out from noise caused by the unreliability of the instrument. The SSS-8 showed a reliability of 0.75 at pre assessment and 0.82 at post assessment (Cronbach's α). This leads to a SEM of 2.90 and 2.51 points, respectively. Therefore, a true change requires at least approximately 3 points pre-post difference. Improvements smaller than 3 points are likely a result of measurement error "rather than a true observed change".[17]

However, even if the pre-post difference equals 3 points, we do not know whether this change is really meaningful to the patients. Therefore we used the anchor based method as a second approach to the MCID and linked the SSS-8's change score to the WHODAS-12's change score. We created a group of patients who did not experience a change in disability or even experienced an

increase (n = 11). The remaining sample (i.e., n = 44 patients who experienced a decrease in disability) was divided into four quartiles according to the amount of decrease in disability they reported. For every group, we calculated the SSS-8 mean change score and its standard deviation. For comparison we also computed mean (SD) change scores of the GAD-7 anxiety scale and PHQ-9 depression scale. The analysis is shown in table 3.

Table 3. SSS-8, GAD-7, and PHQ-9 change scores broken down by WHODAS-12 change scores

	WHODAS-12 disability pre-post change score				
	no change or worse (n = 11)	1st quartile (n = 13)	2nd quartile (n = 11)	3rd quartile (n = 10)	4th quartile (n = 10)
Measure	≤ 0 points	1 to 4 points	5 to 8 points	9 to 12 points	≥ 13 points
SSS-8	0.09 (5.61)	1.15 (5.35)	1.91 (2.91)	4.20 (3.82)	6.30 (4.22)
GAD-7	0.82 (5.49)	3.54 (3.26)	5.09 (3.42)	7.30 (4.37)	8.30 (4.95)
PHQ-9	1.73 (4.96)	5.00 (3.44)	5.73 (4.17)	9.20 (3.99)	10.30 (6.11)

Notice: Cells show means and standard deviations in parenthesis, WHO Disability Assessment Schedule (WHODAS-12), Somatic Symptom Scale – 8 (SSS-8), Generalized Anxiety Disorder Scale – 7 (GAD-7), Patient Health Questionnaire – 9 depression scale (PHQ-9), group size (n).

A 3-point change on SSS-8 revealed by the distribution based method corresponded to a change on the WHODAS-12 that lays between the second and third quartile groups. The mean SSS-8 change score in the second quartile group of the WHODAS-12 was 1.91 and the mean change score in the third quartile group was 4.20. In order to choose a MCID that is really important to the patients we decided to use the more conservative option and defined the MCID as a 4-point decrease on the SSS-8 (i.e., we located the MCID more close to the upper limit of the interval). Using a 4-point decrease on the SSS-8 as a benchmark of the MCID, approximately 40% of the sample achieved at least a minimal clinically important improvement in somatic symptom burden (see table 3). In addition, location of the MCID between the second and third quartile improvement group of the WHODAS-12 leads to a GAD-7 anxiety change score between 5.09 and 7.30, and a PHQ-9 depression change score between 5.73 and 9.20. Both indicate a substantial improvement in anxiety and depression. This underlines the validity of locating the MCID between the second and the third quartile of improvement on the WHODAS-12.

4. Discussion

This study examined the clinical utility and changeability of the SSS-8 as a suitable measure for monitoring somatic symptom burden. We addressed the question, whether the SSS-8 is sensitive to change and determined the minimal clinically important difference. For this purpose, we monitored the course of symptoms in patients with mental disorders who received 8 to 12 weeks evidence-based treatment in a day clinic.

The analyses showed a reduction in somatic symptom burden from pre to post assessment with a medium effect size for the complete sample and a large effect size for a subgroup of patients with very high somatic symptom burden at baseline.[31] Effect sizes were comparable to a meta-analytic effect sizes quantifying the efficacy of psychotherapy in patients with somatoform disorders suffering from physical symptoms.[18] Patients also showed a medium to large decline in anxiety, depression and disability indicating the efficacy of the treatment program. Changes in somatic symptom burden were substantially associated changes in depression, anxiety, and disability. We interpret these findings as evidence of the SSS-8's sensitivity to change: First, as expected for an evidence-based treatment [32,33], patients showed significant decline in somatic symptom burden from pre to post assessment. Second, the associations between the change in somatic symptom burden and the change in depression, anxiety and disability are plausible. As formerly shown, these syndromes are highly interrelated and often referred to as the somatization-anxiety-depression triad.[34,35]

Our study combined two methods to calculate the MCID of the SSS-8 in order to consider both statistical and clinical criteria of significance. The first analysis showed that a 3-point decline on the SSS-8 constitutes a reliably detectable improvement, as it exceeds one standard error of measurement and therefore stands out from noise caused by measurement unreliability. However, when it comes to clinical significance, our second analysis showed that a 4-point decrease is related to a substantial improvement in patient reported disability. A decline of 4 points on the SSS-8 was

achieved by 40% of the patients which is plausible and well in line with the reported effects of evidence-based treatment on somatic symptoms in patients with mental disorders.[18,32,33]

Some limitations have to be considered. First, our study included an unselected sample of patients with mental disorders. Only 1.87% had a diagnosis of a somatoform disorder according to ICD-10. Yet, we decided to include patients from the whole spectrum of varying mental diagnoses to reflect real world conditions of outpatient settings (i.e., to assure high ecological validity). In addition, prior studies showed high amounts of somatic symptoms in subjects diagnosed with depression and anxiety. [1,6,34] Depression or anxiety disorders were present in approximately 50% of our patients. Furthermore, a subgroup of patients with very high somatic symptom burden at baseline showed a large pre-post effect size, which was comparable to the effect size reported by a meta-analysis of psychotherapy studies in patients with somatoform disorders.[18] Second, our study used only patient-reported outcome measures. No expert rating of symptom severity was conducted. However, all of the instruments have a tradition in clinical studies and were well validated.[22,28,29] We believe that an expert rating is likely to show the same results because there is no special reason why patients should dissimulate symptom severity at post assessment (i.e., assessment was on a voluntarily basis, post assessment scores had no clinical implications for the patients). Third, we examined a relatively small sample of patients. Yet, we performed a-priori power analysis to guide our primary analysis. Furthermore, all the reported effects were medium to large which makes it very unlikely that our findings can be attributed to a statistical artefact.[31]

In total, the study showed that the SSS-8 is sensitive to change. A 3-point change can be considered a statistically significant difference, and that a 4-point decrease in symptom severity reflects a clinically important improvement. Taking into account its brevity, the sound psychometric properties, the availability of reference scores, and the straightforward interpretation we recommend the SSS-8 for monitoring somatic symptom burden in patients with mental disorders.

5. References

[1] Kroenke K. Patients presenting with somatic complaints: epidemiology, psychiatric co-morbidity and management. International Journal of Methods in Psychiatric Research 2003;12:34–43.

[2] Dimsdale JE, Creed F, Escobar J, Sharpe M, Wulsin L, Barsky AJ, et al. Somatic Symptom Disorder: An important change in DSM. Journal of Psychosomatic Research 2013;75:223–8.

[3] Katon W, Lin EHB, Kroenke K. The association of depression and anxiety with medical symptom burden in patients with chronic medical illness. General Hospital Psychiatry 2007;29:147–55.

[4] Khan AA, Khan A, Harezlak J, Wanzhu T, Kroenke K. Somatic symptoms in primary care: etiology and outcome. Psychosomatics 2003;44:471–8.

[5] Kohlmann S, Gierk B, Hümmelgen M, Blankenberg S, Löwe B. Somatic Symptoms in Patients With Coronary Heart Disease: Prevalence, Risk Factors, and Quality of Life. JAMA Internal Medicine 2013;173:1469.

[6] Kroenke K, Jackson JL, Chamberlin J. Depressive and anxiety disorders in patients presenting with physical complaints: clinical predictors and outcome. American Journal of Medicine 1997;103:339–47.

[7] Kroenke K, Zhong X, Theobald D, Wu J, Tu W, Carpenter JS. Somatic Symptoms in Patients With Cancer Experiencing Pain or Depression: Prevalence, Disability, and Health Care Use. Archives of Internal Medicine 2010;170:1686–94.

[8] Nimnuan C, Hotopf M, Wessely S. Medically unexplained symptoms: an epidemiological study in seven specialities. Journal of Psychosomatic Research 2001;51:361–7.

[9] Kroenke K, Arrington ME, Mangelsdorff AD. The prevalence of symptoms in medical outpatients and the adequacy of therapy. Archives of Internal Medicine 1990;150:1685–9.

[10]Black N. Patient reported outcome measures could help transform healthcare. BMJ 2013;346:f167.

[11]Gierk B, Kohlmann S, Kroenke K, Spangenberg L, Zenger M, Brähler E, et al. The somatic symptom scale-8 (SSS-8): a brief measure of somatic symptom burden. JAMA Internal Medicine 2014;174:399–407.

[12]Kroenke K, Spitzer RL, Williams JBW. The PHQ-15: Validity of a new measure for evaluating the severity of somatic symptoms. Psychosomatic Medicine 2002;64:258–66.

[13]Gierk B, Kohlmann S, Toussaint A, Wahl I, Brünahl CA, Murray AM, et al. Assessing somatic symptom burden: A psychometric comparison of the Patient Health Questionnaire-15 (PHQ-15) and the Somatic Symptom Scale-8 (SSS-8). Journal of Psychosomatic Research 2015;78:352–5.

[14]Narrow WE, Clarke D, Kuramoto JS, Kraemer HC, Kupfer DJ, Greiner L, et al. DSM-5 Field Trials in the United States and Canada, Part III: Development and Reliability Testing of a Cross-Cutting Symptom Assessment for DSM-5. American Journal of Psychiatry 2013;170:71–82.

[15]Matsudaira K, Oka H, Kawaguchi M, Murakami M, Fukudo S, Hashizume M, et al. Development of a Japanese version of the Somatic Symptom Scale-8: Psychometric validity and internal consistency. General Hospital Psychiatry 2017;45:7–11.

[16]Stratford PW, Riddle DL. Assessing sensitivity to change: choosing the appropriate change coefficient. Health and Quality of Life Outcomes 2005;3.

[17]Copay AG, Subach BR, Glassman SD, Polly DW, Schuler TC. Understanding the minimum clinically important difference: a review of concepts and methods. The Spine Journal: Official Journal of the North American Spine Society 2007;7:541–6.

[18]Koelen JA, Houtveen JH, Abbass A, Luyten P, Eurelings-Bontekoe EHM, Van Broeckhuysen-Kloth SAM, et al. Effectiveness of psychotherapy for severe somatoform disorder: Meta-analysis. British Journal of Psychiatry 2014;204:12–9.

[19]Lakens D. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. Frontiers in Psychology 2013;4:863.

[20]Faul F, Erdfelder E, Lang A-GG, Buchner A. G * Power 3 : A flexible statistical power analysis program for the social , behavioral , and biomedical sciences. Behavior Research Methods 2007;39:175–91.

[21]Spitzer RL, Kroenke K, Williams JBW, Löwe B. A Brief Measure for Assessing Generalized Anxiety Disorder - The GAD-7. Archives of Internal Medicine 2006;166:1092–7.

[22]Kroenke K, Spitzer RL, Williams JBW, Löwe B. The Patient Health Questionnaire Somatic, Anxiety, and Depressive Symptom Scales: A systematic review. General Hospital Psychiatry 2010;32:345–59.

[23]Löwe B, Decker O, Müller S, Brähler E, Schellberg D, Herzog W, et al. Validation and standardization of the Generalized Anxiety Disorder Screener (GAD-7) in the general population. Medical Care 2008;46:266–74.

[24]Kroenke K, Spitzer RL, Williams JBW. The PHQ-9 - Validity of a Brief Depression Severity Measure. Journal of General Internal Medicine 2001;16:606–13.

[25]Löwe B, Unützer J, Callahan CM, Perkins AJ, Kroenke K. Monitoring depression treatment outcomes with the patient health questionnaire-9. Medical Care 2004;42:1194–201.

[26]Löwe B, Kroenke K, Herzog W, Gräfe K. Measuring depression outcome with a brief self-report instrument: Sensitivity to change of the Patient Health Questionnaire (PHQ-9). Journal of Affective Disorders 2004;81:61–6.

[27]Löwe B, Schenkel I, Carney-Doebbeling C, Göbel C. Responsiveness of the PHQ-9 to Psychopharmacological Depression Treatment. Psychosomatics 2006;47:62–7.

[28]Üstün TB, Kostanjsek S, Chatterji S, Rehm J, editors. Measuring Health and Disability: Manual for WHO Disability Assessment Schedule (WHODAS 2.0). World Health Organisation; 2010.

[29]Andrews G, Kemp A, Sunderland M, von Korff M, Üstün TB. Normative data for the 12 item WHO disability assessment schedule 2.0. PLoS ONE 2009;4:1–6.

[30]Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. Controlled Clinical Trials 1989;10:407–15.

[31]Cohen J. A power primer. Psychological Bulletin 1992;112:155–9.

[32]Kroenke K. Efficacy of treatment for somatoform disorders: a review of randomized controlled trials. Psychosomatic Medicine 2007;69:881–8.

[33]Sumathipala A. What is the evidence for the efficacy of treatments for somatoform disorders? A critical review of previous intervention studies. Psychosomatic Medicine 2007;69:889–900.

[34]Löwe B, Spitzer RL, Williams JBW, Mussel M, Schellberg D, Kroenke K. Depression, anxiety and somatization in primary care: syndrome overlap and functional impairment. General Hospital Psychiatry 2008;30:191–9.

[35]Kohlmann S, Gierk B, Hilbert A, Brähler E, Löwe B. The overlap of somatic, anxious and depressive syndromes: A population-based analysis. Journal of Psychosomatic Research 2016;90:51–6.

ORIGINAL ARTICLE

Development and Validation of the Somatic Symptom Disorder–B Criteria Scale (SSD-12)

Anne Toussaint, PhD, Alexandra M. Murray, DPhil, Katharina Voigt, PhD, Annabel Herzog, MSc, Benjamin Gierk, MSc, Kurt Kroenke, MD, Winfried Rief, PhD, Peter Henningsen, MD, and Bernd Löwe, MD

ABSTRACT

Objective: To develop and validate a new self-report questionnaire for the assessment of the psychological features of Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition somatic symptom disorder.

Methods: The Somatic Symptom Disorder–B Criteria Scale (SSD-12) was developed in several steps from an initial pool of 98 items. The SSD-12 is composed of 12 items; each of the three psychological subcriteria is measured by four items. In a cross-sectional study, the SSD-12 was administered to 698 patients (65.8% female, mean [standard deviation] age = 38.79 [14.15] years) from a psychosomatic outpatient clinic. Item and scale characteristics as well as measures of reliability and validity were determined.

Results: The SSD-12 has good item characteristics and excellent reliability (Cronbach α = .95). Confirmatory factor analyses suggested that a three-factorial structure that reflects the three psychological criteria interpreted as cognitive, affective, and behavioral aspects (n = 663, Comparative Fit Index > 0.99, Tucker-Lewis Index > 0.99, Root Mean Square Error of Approximation = 0.06, 90% confidence interval = 0.01-0.08). SSD-12 total sum score was significantly associated with somatic symptom burden (r = 0.47, p < .001) and health anxiety (r = 0.71, p < .001), and moderately associated with general anxiety (r = 0.35, p < .001) and depressive symptoms (r = 0.22, p < .001). Patients with a higher SSD-12 psychological symptom burden reported higher general physical and mental health impairment and significantly higher health care use.

Conclusions: The SSD-12 is the first self-report questionnaire that operationalizes the new psychological characteristics of Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition somatic symptom disorder. Initial assessment indicates that the SSD-12 has sufficient reliability and validity to warrant further testing in both research and clinical settings.

Key words: somatic symptom disorder, self-report, psychometrics, diagnostic instruments, SSD-12, somatoform disorders.

INTRODUCTION

The new classification criteria for somatic symptom and related disorders in Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) emphasize psychological features like abnormal and excessive thoughts, feelings, and behaviors associated with the bothersome somatic symptoms rather than their medical explicability (1,2). These central changes lead to several diagnostic challenges which call the applicability of current assessment instruments into question (3). Because it is a great challenge for physicians and other health care providers to take these new psychological aspects into account when considering somatic symptom disorder (SSD), it is crucial to establish a well-founded diagnostic strategy. The changes in the criteria also imply a need to review and eventually adapt existing supportive diagnostic material because its future feasibility in clinical practice and research settings can no longer be guaranteed (3).

Typically, DSM disorders are assessed using the gold-standard Structured Clinical Interview for DSM-Disorders (SCID) (4). However, given that self-report questionnaires capture the patients' own perspectives of their symptoms, they

➢ Related editorial on pages 2–4

DSM-5 = Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, **SSD** = somatic symptom disorder

SDC Supplemental Content

From the Department of Psychosomatic Medicine and Psychotherapy (Toussaint, Murray, Voigt, Herzog, Gierk, Löwe), University Medical Center Hamburg-Eppendorf & Schön Clinic Hamburg-Eilbek, Hamburg, Germany; VA HSR&D Center for Health Information and Communication (Kroenke), Indiana University School of Medicine, and Regenstrief Institute, Indianapolis, Indiana; Department of Clinical Psychology (Rief), University of Marburg, Marburg, Germany; and Department of Psychosomatic Medicine and Psychotherapy (Henningsen), Technical University of Munich, Munich, Germany.

Address correspondence and reprint requests to Anne Toussaint, PhD, Department of Psychosomatic Medicine and Psychotherapy, University Medical Center Hamburg-Eppendorf, Martinistraße 52, 20246 Hamburg, Germany. E-mail: a.toussaint@uke.de

Received for publication February 10, 2015; revision received July 31, 2015.

DOI: 10.1097/PSY.0000000000000240

Copyright © 2015 by the American Psychosomatic Society

Copyright © 2015 by the American Psychosomatic Society. Unauthorized reproduction of this article is prohibited.

represent an additional and valuable source of information and they can save considerable time within the diagnostic process. There are several standardized and validated instruments that effectively measure the burden of specific somatic complaints (5) or take into account psychological factors associated with somatoform disorders according to DSM-IV and International Classification of Diseases, Tenth Revision (ICD-10; e.g., Whiteley Index [WI-7] (6), Symptom Checklist-90-Revised (7), Scale for the Assessment of Illness Behavior (8), Questionnaire on Pain Regulation [Fragebogen zur Erfassung der Schmerzverarbeitung in German] (9)). Given that most of these questionnaires are oriented toward the former diagnostic criteria of DSM-IV and ICD-10, their applicability for the revised SSD diagnosis is questionable. Although criterion A (somatic symptoms that are either very distressing or result in significant disruption of functioning) could, with minor modifications, be adequately reflected by current instruments like the Patient Health Questionnaire 15-item somatic scale (PHQ-15) (10), Somatic Symptom Scale-8 (11), or Screening for Somatoform Disorders-7 (12), Criterion B (psychological features) can only be partially assessed by questionnaires like the Scale for the Assessment of Illness Behavior or WI-7, which mostly cover maladaptive thoughts and anxieties like catastrophizing thoughts or health anxiety. There are no valid instruments to explicitly cover the behavioral aspects of Criterion B, and most notably, no instruments that allow the direct measurement of all three new B criteria of SSD (3,5).

The objective of the present project was to develop and validate a new self-rating questionnaire based on the reconceptualized SSD criteria. The Somatic Symptom Disorder–B Criteria Scale (SSD-12) assesses patients' perceptions of their symptom-related thoughts, feelings, and behaviors using questions directly based on the DSM-5 criteria. The current data details the development of the SSD-12, its item, and scale characteristics as well as validation analyses using a sample of patients from a psychosomatic outpatient clinic.

METHOD

Development of the SSD-12

A preliminary self-report questionnaire, the Somatic Symptoms Experiences Questionnaire (13,14) was previously developed by our group to assess the psychological characteristics of SSDs. This questionnaire was based on early drafts of the proposed changes to the former somatoform disorders as posted by the DSM-5 task force. It covers four dimensions with 15 items: health worries, illness experience, difficulties in interacting with doctors, and the impact of illness. Studies using this questionnaire provided important information about the psychological dimension of somatoform-type complaints and served as a foundation for further enhancements that ended in the development of the SSD-12. The item content of the new scale was primarily based on the draft SSD criteria published on the DSM-5 development Web site which have been confirmed with the release of DSM-5 (1). The initial item pool and questionnaire design were based on the cognitions, emotions, and behaviors which we interpreted as representing the three B criteria of SSD.

Three authors independently developed a preliminary pool of 98 items. References to specific situational contexts and reliance on academic language were avoided, as recommended by previous research (e.g., (15)). A discussion among the group members led to the immediate elimination of any items which were unclear, too vague, too specific, or repetitive. This process resulted in an initial pool of 40 items. It was decided that responders should rate how frequently they experienced each cognition, emotion, or behavior on a 5-point Likert scale: "never" (0), "rarely" (1), "sometimes" (2), "often" (3), and "very often" (4).

To ensure that the item content captured the three diagnostic subcriteria of the SSD B criterion, a focus group of 14 researchers and clinicians who were familiar with somatoform type disorders was formed. Each participant categorized each item as being associated with the first, second, or third subcriterion and rated how well the item fit the criteria. All items were judged to be understandable. Based on the focus group ratings, some items were excluded, condensed, or made to be more specific to the appropriate subcriterion. This modification process resulted in a set of 26 items.

The remaining 26 items were then evaluated by a focus group of six patients receiving inpatient care for somatoform disorders with predominant pain (as diagnosed by ICD-10) in the Schön Clinic Eilbek, Hamburg, Germany. Spontaneous comments, impressions, and questions were collected during the discussion of the items. The wording of some items was slightly modified afterward, but in general, all items and the response format were understandable. Subsequently, the next version of the questionnaire contained 18 items.

The 18-item version was then tested in a consecutive outpatient sample (n = 156) from the Department of Psychosomatic Medicine and Psychotherapy, University Medical Center Hamburg-Eppendorf. The best four items from each of the three subscales were chosen based on their psychometric item content and item properties. All corrected item-total correlations of the remaining items exceeded 0.30, and the item means varied between 0.20 and 0.70, reflecting an acceptable range of item difficulty. In addition, we changed the order of the 12 remaining items so that they alternated between measuring the three subcriteria (Subcriteria 1, 2, 3, 1, 2, 3… etc). In the pilot study, our final German scale SSD-12 had high total reliability (Cronbach α = .95). Item characteristics and reliabilities of the three subscales as measured with only four items were acceptable (not reported here). The German version of this final 12-item questionnaire and an English version that was simultaneously developed can be found in Supplemental Digital Content 1 (http://links.lww.com/PSYMED/A235).

Validation Sample

The final 12-item SSD-12 version was tested in a larger consecutive outpatient sample from the Department of Psychosomatic Medicine and Psychotherapy, University Medical Center Hamburg-Eppendorf. Inclusion criteria were being 18 years or older and providing written informed consent. Exclusion criteria were having a current psychotic disorder, indications of substance abuse, organic brain disease, active suicidality, personality disorder, and insufficient German language skills. All patients were diagnosed with a main diagnosis by ICD-10 of either depression or an anxiety disorder, eating disorder, or somatoform disorder.

Patients were recruited between March and October 2013. The study was approved by the medical ethics board of the Hamburg Medical Chamber. Written informed consent was obtained from all participants.

Other Measures

In addition to the SSD-12, every patient answered a battery of questionnaires as part of standard routine care. We included some measures to examine the construct validity of the SSD-12, with the understanding that somatoform disorders as diagnosed by DSM-IV or ICD-10 are often associated with depression and anxiety (16), decreased general health status, and increased health care use (10). Because it can be assumed that the same applies to SSDs, the following instruments were chosen.

Copyright © 2015 by the American Psychosomatic Society. Unauthorized reproduction of this article is prohibited.

The PHQ-15 scale (10) assesses the presence and severity of common somatic symptoms within the last 4 weeks using 15 items. Sum scores range from 0 to 30 and indicate the self-rated symptom burden, with higher scores indicating a higher burden.

Depression was measured with the Patient Health Questionnaire nine-item depression scale (PHQ-9) (16,17), which assesses the presence of the nine DSM criteria for major depression within the last 2 weeks. Scores range from 0 to 27 and indicate the severity of depression (high scores reflect high symptom load).

The seven-item Generalized Anxiety Disorder Scale (GAD-7) (18) is a self-administered patient questionnaire that is used as a screening tool and severity measure of both generalized anxiety disorder as well as other common anxiety disorders (19). Scores range from 0 (minimal) to 21 (severe).

Correlations with the German version of the WI-7 (6) as a measure of health anxiety using seven dichotomous items were also performed.

Patients also rated their current general physical status and mental health status in the past 2 weeks on two items using a 5-point Likert scale (0 = not at all impaired, 4 = seriously impaired).

Health care use (HCU) was measured as the self-reported number of contacts with health care providers in the past 4 weeks (0 = zero, 1 = once, 2 = twice, 3 = more than twice).

Statistical Analysis

We computed mean (standard deviation), corrected item-total correlations, skewness, and kurtosis as descriptive statistics. In cases of missing data, mean responses were imputed. Reliability of the full scale and subscales was measured using Cronbach α. To investigate the factorial structure of the questionnaire, we conducted two Confirmatory Factor Analyses (CFAs) using the Lavaan package from the statistical computing software R (20). Because of the nonnormal distribution of our data, robust weighted least squares estimation with mean and variance adjustment was used to fit the model and missing cases were excluded listwise. The Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), and the Root Mean Square Error of Approximation (RMSEA) were used to test global model fits. A good fit for continuous data is indicated for RMSEA < 0.60, TLI > 0.95, CFI > 0.95, and SRMR < 0.08 (21,22). Given that the three subcriteria of Criterion B in DSM-5 should form the basis of the severity specification of SSD, the ability to assess each subcriterion independently is important to properly reflect the structure of the criteria. We tested, therefore, a first model (A) to investigate whether three latent dimensions corresponding to the three subcriteria could explain the data.

In contrast, we conducted a second analysis (Model B) to test whether it is appropriate to measure the severity of Criterion B symptoms as a single construct. Specifically, we tested whether all 12 items assessed a general factor "g."

Construct validity was examined using bivariate correlations between SSD-12, PHQ-15, GAD-7, and WI-7. For the two items measuring the general physical and mental health status and the four categories of HCU, 1-way analyses of variance were calculated to test for linear trends in the group means.

RESULTS

Data from 856 eligible patients were recruited during the recruitment period. From those, we analyzed data from 698 (81.5 %) participants who had answered a total of at least 9 of the 12 SSD-12 items. Characteristics of the sample are shown in Table 1.

Overall Item Characteristics and Reliability

Generally speaking, the participants responded well to the questionnaire. There was no indication that particular items were skipped or neglected in a systematic way. Individual item difficulty, skewness, and kurtosis were all acceptable and responses for every item covered the full range of response categories (as shown in Table 2). The 12-item-questionnaire had high reliability in this sample (Cronbach α = .94). Reliability of the subscales was α = .71 for Subcriterion I, α = .92 for Subcriterion II, and α = .92 for Subcriterion III. Overall item and subscale characteristics are shown in Table 2.

Factor Analysis

Model A, displayed in Figure 1, included the three latent dimensions as proposed by the theoretical conceptualization of DSM-5 SSD. The global fit of this model was good (n = 663, CFI > 0.99, TLI > 0.99, RMSEA = 0.06, 90% confidence interval = 0.01-0.08). All standardized item loadings were greater than 0.50. Given that there seems to be substantial overlap in content between the three subcriteria, the moderate to high correlations between the three subcriteria (Subcriteria I and II: r = 0.66; Subcriteria I and III: r = 0.61; Subcriteria II and III: r = 0.74) were in accordance with prior expectations. The general-factor Model B fit the data in a conventionally acceptable way (n = 663, CFI = 0.99, TLI = 0.98, RMSEA = 0.09, 90% confidence interval = 0.07-0.12) (22). All standardized item loadings were greater than 0.49. Altogether, the results of the confirmatory factor analyses support the organization of the different aspects of the B criterion into three subcriteria. However, the satisfactory fit of the general-factor Model B also justifies the use of a general SSD-12 sum score.

Construct Validity

For our further analyses, we calculated an overall sum score of the SSD-12. We then conducted a preliminary investigation of the construct validity of the scale by using bivariate correlations of the SSD-12 sum score and single subcriterion scores with other well-established scales, that is, the PHQ-15 (somatic symptom burden), the WI-7 (symptoms of hypochondriasis), the PHQ-9 (depression), and the GAD-7 (anxiety). Mean responses were imputed for missing data if participants had at least answered 12 items of the PHQ-15, 7 of the PHQ-9, and 5 of the WI-7 and GAD-7. If more than these numbers of responses were missing, the data for that instrument were not included in the analyses. The series of correlational analyses is shown in Table 3.

The two items measuring general physical health status and general mental health status (n = 656) significantly correlated with SSD-12 sum scores (r = 0.63 and r = 0.24, respectively). One-way analyses of variance were performed to assess the linearity of the relationship between these two variables and for the four categories of HCU and SSD-12

Copyright © 2015 by the American Psychosomatic Society. Unauthorized reproduction of this article is prohibited.

TABLE 1. Characteristics of the Psychosomatic Outpatient Sample (n = 698) Used to Test the Final Version of the SSD-12

Sample Characteristics	
Sex, female, n (%)	459 (65.8)
Age, M (SD), y	38.79 (14.15)
Age range, y	18–89
Nationality, n (%)	
German	643 (92.1)
Other	55 (7.9)
Education level, n (%)	
Still in school	6 (9)
<10 y of formal education	118 (16.9)
10 y of formal education	224 (32.1)
12–13 y of formal education	178 (25.5)
University degree/certificate	146 (20.9)
Other	26 (3.8)

SSD-12 = Somatic Symptom Disorder–B Criteria Scale; M = mean; SD = standard deviation.

Overlap between the highest group of formal education (12–13 years) and university degrees is possible.

sum score (n = 669). All trends were significant. The results are shown in Figure 2.

DISCUSSION

In this study, we present the development and initial validation of the first self-report questionnaire that operationalizes the psychological characteristics of DSM-5 SSD: the SSD-12. To our knowledge, there are no other instruments available that directly assess these new and important psychological features.

The SSD-12 assesses the three psychological subcriteria using a total of 12 items based directly on the DSM-5 wording (four items for each subcriterion). Its brevity enables the scale to be completed in approximately 2 minutes which facilitates its use in clinical practice as, for example, a screening instrument. Results from a large outpatient sample (n = 698) indicate that the SSD-12 sum scale has high reliability as measured by Cronbach α and also reliability measures of the subscales are moderate (for Subcriterion I) to high (for Subcriteria II and III). Confirmatory factor analyses performed on a general-factor model fit the data in an acceptable way which justifies the use of a SSD-12 total sum score, which we used for our initial validation analyses: increases in total item scores represent increases in the construct of "SSD psychological dimension," which we here interpreted as a general psychological symptom burden. However, for theoretical reasons, we give preference to the "three latent dimensions" model, which is supported by the good-fit indices revealed by the confirmatory factor analysis. The calculation of three subscores is more in line with the structure of the DSM-5 criteria, and this model might, therefore, be of greater importance for clinical purposes. However, we would like to point out that the three subcriteria are demonstrably highly correlated, a fact that should be discussed further in the future. At this stage, it remains unclear whether these suggested psychological symptoms are the most valid, sensitive, and specific empirically founded features (23).

TABLE 2. Overall Item Characteristics (n = 698) for SSD-12

Subcriterion		M (SD)	Corrected Item Total[a] Correlation	Skewness (SE)	Kurtosis (SE)
I (cognitive aspects)	Item 1	1.60 (1.21)	0.68	0.32 (0.09)	−0.83 (0.19)
	Item 4	1.85 (1.32)	0.66	0.07 (0.09)	−1.12 (0.19)
	Item 7	1.15 (1.20)	0.41	0.61 (0.09)	−0.79 (0.19)
	Item 10	0.85 (1.17)	0.50	1.18 (0.09)	0.29 (0.19)
II (affective aspects)	Item 2	2.08 (1.32)	0.80	−0.03 (0.09)	−1.10 (0.19)
	Item 5	1.93 (1.39)	0.83	0.08 (0.09)	−1.21 (0.19)
	Item 8	1.95 (1.49)	0.81	0.03 (0.09)	−1.39 (0.19)
	Item 11	2.04 (1.44)	0.84	−0.09 (0.09)	−1.30 (0.19)
III (behavioral aspects)	Item 3	2.12 (1.47)	0.80	−0.14 (0.09)	−1.33 (0.19)
	Item 6	1.45 (1.41)	0.81	0.51 (0.09)	−1.09 (0.19)
	Item 9	1.65 (1.39)	0.82	0.29 (0.09)	−1.19 (0.19)
	Item 12	1.74 (1.37)	0.80	0.20 (0.09)	−1.17 (0.19)

SSD-12 = Somatic Symptom Disorder–B Criteria Scale; M = mean; SD = standard deviation; SE = standard error.

Range for all items = 0–4 (with higher scores representing greater severity); range for complete 12-item scale = 0–48.

[a] "Total" refers to the complete 12-item scale.

Copyright © 2015 by the American Psychosomatic Society. Unauthorized reproduction of this article is prohibited.

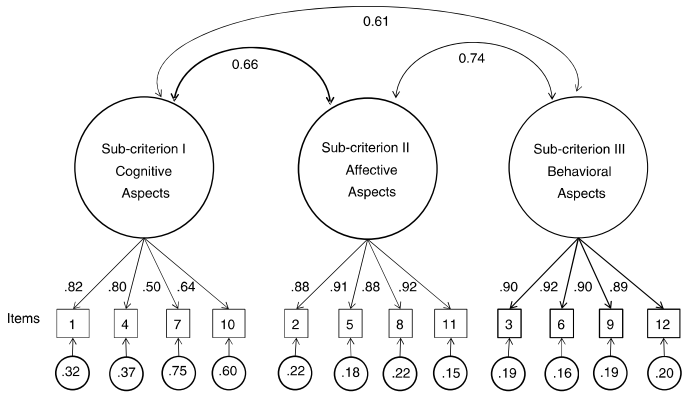


FIGURE 1. Path diagram illustrating the final SSD-12 Model A estimates (n = 663). The 12 items of SSD-12 measure correlated cognitive (Items 1, 4, 7, 10), affective (Items 2, 5, 8, 12), and behavioral (Items 3, 6, 9, 11) latent factors (large circles). Item error/uniquenesses are shown as small circles. The model was estimated using robust WLSMV. Missing cases were excluded listwise. Loadings are standardized. SSD-12 = Somatic Symptom Disorder–B Criteria Scale; WLSMV = weighted least squares estimation with mean and variance adjustments.

Because the psychological dimensions of SSD should be, by definition, highly associated with somatic symptom burden and illness anxiety, we conducted bivariate correlations of SSD-12, PHQ-15, and WI-7 sum scores to examine the construct validity of SSD-12. We found a reasonable association between the sum scores of SSD-12 and PHQ-15 (r = 0.47), which suggests that patients with a higher physical symptom load tend to report higher psychological symptom burden. The high correlation of r = 0.71 with the WI-7 suggests that the dimensions of SSD-12 are highly related to health anxiety. This would be expected, given the emphasis on health-related cognitions in the SSD criteria, especially in subcriterion II. Using these interpretations, the correlational analyses suggested that the SSD-12 has good construct validity.

There are high comorbidities between somatoform disorders and depression and anxiety disorders. Nevertheless, all three are considered to be distinct entities, so the correlations between SSD-12 and PHQ-9 and GAD-7 sum scores were expected to be moderate at best. The correlation with the GAD-7—as a measure of anxiety—was r = 0.35, and that with the PHQ-9—as a measure of depression—was r = 0.22. These psychometric properties are of great importance for the differential diagnosis of SSD, depression, and anxiety disorders (24), and we interpreted them in terms of divergent validity. A similar pattern in the direction and size of correlations was found for the three subcriterion scales.

Correlations with SSD-12 sum scores and the two items measuring the general physical and mental health status were significantly higher for the physical dimension (r = 0.63) than for the mental one (r = 0.24). Because a specific characteristic of somatoform patients is to focus on their physical sensations and symptoms, this particular result is in line with a priori expectations. The highly burdened patients (in terms of higher SSD-12 scores) also reported significantly higher HCU, another cardinal feature of somatoform type disorders.

Our study has several limitations. First, it is important to note that it was conducted in a specific sample of patients from a psychosomatic outpatient clinic in Germany, so the representativeness and generalizability of the findings are limited. In the near future, we will therefore try to replicate the findings in a general population sample and in international settings. Second, Subcriterion I was considered especially problematic by both the members of the research and patient focus groups. It specifies the need for "disproportionate" thoughts about the seriousness of one's symptoms. This disproportion is extremely difficult for patients to judge themselves but also for clinicians, especially because SSD can now also accompany known diagnoses of physical disorders. We decided, therefore, to include two items that reflect the opinions of others (e.g., doctors) at the expense of a higher heterogeneity and lower reliability of that specific subcriterion (α = .71). However, the two items containing the views of others (Items 7 and 10) were judged to be most difficult by our outpatient sample (see Table 2). Because these two items are also the only ones containing a negation, it is unclear which factor is responsible for the higher difficulties, so this should be investigated in the future. A third study limitation is that the diagnostic process did not include "gold-standard" instruments because they were not considered sufficient to

Copyright © 2015 by the American Psychosomatic Society. Unauthorized reproduction of this article is prohibited.

TABLE 3. Descriptive Characteristics and Pearson Correlations Between SSD-12 Sum Score/Subscales and Other Scales

	SSD-12 Total	SSD-12 Subscale I	SSD-12 Subscale II	SSD-12 Subscale III	WI-7	PHQ-15	GAD-7	PHQ-9
Descriptive characteristics								
n	698	698	698	698	674	618	681	685
M (SD)	20.4 (12.7)	5.5 (3.6)	8.0 (5.0)	7.0 (5.1)	3.1 (2.4)	10.6 (5.7)	12.3 (5.4)	14.0 (6.2)
Range	0–48	0–16	0–16	0–16	0–17	0–28	0–21	0–27
Correlation with SSD-12 total score								
Correlation coefficient (*r*)		0.86	0.96	0.94	0.71	0.47	0.35	0.22
p		<.001	<.001	<.001	<.001	<.001	<.001	<.001
Correlation with SSD-12 Subscale I (cognitive aspects)								
Correlation coefficient (*r*)			0.66	0.61	0.66	0.39	0.26	0.14
p			<.001	<.001	<.001	<.001	<.001	<.001
Correlation with SSD-12 Subscale II (affective aspects)								
Correlation coefficient (*r*)				0.74	0.71	0.44	0.32	0.20
p				<.001	<.001	<.001	<.001	<.001
Correlation with SSD-12 Subscale III (behavioral aspects)								
Correlation coefficient (*r*)					0.62	0.46	0.38	0.26
p					<.001	<.001	<.001	<.001

SSD-12 = Somatic Symptom Disorder–B Criteria Scale; WI-7 = Whitely Index-7; PHQ-15 = Patient Health Questionnaire-15; GAD-7 = Generalized Anxiety Disorder-7; PHQ-9 = Patient Health Questionnaire-9; M = mean; SD = standard deviation.

be the bases of sensitivity and specificity analyses of the SSD-12. Even if the current gold-standard SCID had been used, the results could hardly be considered as an external validation criterion because the current SCID version is based on the DSM-IV somatoform disorders criteria. The psychological criteria measured by our instrument are not included in these criteria and are also, therefore, not captured by the current version of SCID. Further testing will be needed once a SCID version for DSM-5 is available.

A further limitation of the study is that neither the SSD-12 nor any other currently available measure provides a complete psychometric assessment of the SSD criteria. The PHQ-15 and the even shorter eight-item version Somatic Symptom Scale-8 can easily be combined

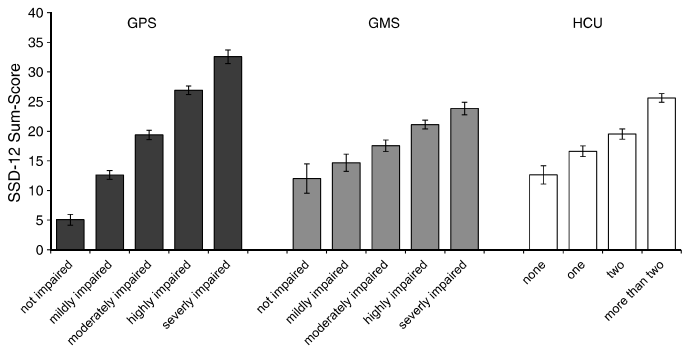


FIGURE 2. SSD-12 sum scores and GPS ($n = 656$), GMS ($n = 656$), and HCU ($n = 669$). Note: SSD-12 range = 0–48. SSD-12 = Somatic Symptom Disorder–B Criteria Scale; GPS = general physical health status (last 2 weeks); GMS = general mental health status (last 2 weeks); HCU = health care use (last 4 weeks): none = 0 medical visits, one = 1 medical visit, two = 2 medical visits, more than two = more than 2 medical visits. Error bars show standard error.

Copyright © 2015 by the American Psychosomatic Society. Unauthorized reproduction of this article is prohibited.

with SSD-12 to assess somatic symptom severity as required by Criterion A of SSD, but DSM-5 also requires a third "chronicity" criterion (C) defined by persistent symptoms (typically >6 months). In addition, DSM-5 allows for the coding of several specifiers including the presence of predominant pain, a persistent course with severe symptoms or marked impairment, and current severity of SSD ranging from mild to moderate to severe (1). Given that the validity of these specifiers currently remains unclear (23), additional items should be developed in the future to capture these aspects. Algorithms about how to combine the measures of Criteria A, B, and C will be developed within a general population sample. We thereby hope to be able to shed some more light on the understanding of how the new diagnosis is conceptualized.

Future research should assess the responsiveness (i.e., sensitivity to change) of the SSD-12 in patients with SSD before and after treatment. Also, because the new concept of SSD may now be very heterogeneous (in terms of patients' diagnostic and symptomatology profile), it might be important to calculate separate norms for different patient groups (e.g., patients with or without accompanying serious physical disease, with monosymptomatic or polysymptomatic disorder, and with pain or illness anxiety disorder in particular).

DSM-5 creates a great opportunity to address the numerous problems of the former diagnostic criteria regarding somatoform disorders. Given the emphasis on the psychological aspects of SSD, it is possible that the disorder may be more widely applied, especially because there is no longer the highly debated requirement of having the lack of a "medical explanation" for symptoms. The development of the SSD-12 was based on the idea to help improve the underdiagnosis of somatoform disorders in the past, so that earlier detection as well as more effective and quality of life-related interventions can be offered to patients with SSD in the future. The SSD-12 is a reliable and valid self-administered questionnaire that can be used to assess patients' psychological symptom burden as described in DSM-5 SSD and to monitor psychological symptom burden in clinical practice and research.

We thank all patients from our somatoform inpatient and outpatient clinics who participated in the study. We also thank all those who gave their valuable contribution and feedback during the development process of SSD-12.

Source of Funding and Conflicts of Interest: The authors report no conflicts of interest and no source of funding.

REFERENCES

1. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders (DSM-5). 5th ed. Arlington, VA: American Psychiatric Association; 2013.
2. Dimsdale JE, Levenson J. What's next for somatic symptom disorder? Am J Psychiatry 2013;170:1393–5.
3. Klaus KM, Mewes R. Assessment of the new DSM-5 diagnosis Somatic Symptom Disorder (300.82). Verhaltensther Verhaltensm 2013;34:399–418.
4. First MB, Spitzer RL, Gibbon M, Williams JBW. Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Patient Edition (SCID-I/P). New York: Biometrics Research, New York State Psychiatric Institute; 2002.
5. Zijlema WL, Stolk RP, Löwe B, Rief W, BioSHaRE, White PD, Rosmalen JG. How to assess common somatic symptoms in large-scale studies: a systematic review of questionnaires. J Psychosom Res 2013;74:459–68.
6. Pilowsky I. Dimensions of hypochondriasis. Br J Psychiatry 1967;113:89–93.
7. Franke GH. SCL-90-R. Die Symptomcheckliste von Derogatis–Deutsche Version–Manual. Vol. 2 vollst. überarb. und neu norm. Aufl. Beltz: Göttingen; 2002.
8. Rief W, Ihle D, Pilger F. A new approach to assess illness behaviour. J Psychosom Res 2003;54:405–14.
9. Geissner E. Fragebogen zur Erfassung der Schmerzverarbeitung (FESV). Göttingen: Hogrefe; 2001.
10. Kroenke K, Spitzer RL, Williams JB. The PHQ-15: validity of a new measure for evaluating the severity of somatic symptoms. Psychosom Med 2002;64:258–66.
11. Gierk B, Kohlmann S, Kroenke K, Spangenberg L, Zenger M, Brähler E, Löwe B. The Somatic Symptom Scale-8 (SSS-8): a brief measure of somatic symptom burden. JAMA Intern Med 2014;174:399–407.
12. Rief W, Hiller W. SOMS—Screening für Somatoforme Störungen Manual. Bern: Huber; 2008.
13. Herzog A, Voigt K, Meyer B, Rief W, Henningsen P, Hausteiner-Wiehle C, Löwe B. The Somatic Symptoms Experiences Questionnaire (SSEQ): a new self-report instrument for the assessment of psychological characteristics of patients with somatoform disorder. Psychother Psychosom Med Psychol 2013;63:112–21.
14. Herzog A, Voigt K, Meyer B, Wollburg E, Weinmann N, Langs G, Löwe B. Psychological and interactional characteristics of patients with somatoform disorders: validation of the Somatic Symptoms Experiences Questionnaire (SSEQ) in a clinical psychosomatic population. J Psychosom Res 2015; 78:553–62.
15. Angleitner A, John OP, Löhr FJ. It's what you ask and how you ask it: an item metric analysis of personality questionnaires. In: Angleitner A, Wiggins JS, editors. Personality Assessment Via Questionnaires. Current Issues in Theory and Measurement. Berlin: Springer; 1986:61–108.
16. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001; 16:606–13.
17. Rief W, Nanke A, Klaiberg A, Braehler E. Base rates for panic and depression according to the Brief Patient Health Questionnaire: a population-based study. J Affect Disord 2004;82:271–6.
18. Löwe B, Decker O, Müller S, Brähler E, Schellberg D, Herzog W, Herzberg PY. Validation and standardization of the Generalized Anxiety Disorder Screener (GAD-7) in the general population. Med Care 2008;46: 266–74.
19. Kroenke K, Spitzer RL, Williams JB, Monahan PO, Löwe B. Anxiety disorders in primary care: prevalence, impairment, comorbidity, and detection. Ann Intern Med 2007; 146:317–25.

Copyright © 2015 by the American Psychosomatic Society. Unauthorized reproduction of this article is prohibited.

ORIGINAL ARTICLE

20. Rosseel Y. *Lavaan*: An R package for structural equation modeling. J Stat Softw 2012;48:1–36.
21. Hu L, Bentler PM. Cutoff criteria for fit indices in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Model 1999;6:1–55.
22. Schreiber JB, Nora A, Stage FK, Barlow EA, King J. Reporting structural equation modeling and confirmatory factor analysis results: a review. J Educ Res 2006;99:323–37.
23. Rief W, Martin A. How to use the new DSM-5 somatic symptom disorder diagnosis in research and practice: a critical evaluation and a proposal for modifications. Annu Rev Clin Psychol 2014;10:339–67.
24. Kroenke K, Spitzer RL, Williams JB, Löwe B. The Patient Health Questionnaire somatic, anxiety, and depressive symptom scales: a systematic review. Gen Hosp Psychiatry 2010; 32:345–59.

Copyright © 2015 by the American Psychosomatic Society. Unauthorized reproduction of this article is prohibited.

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Fakultät für
Psychologie und
Bewegungswissenschaft

Institut für Bewegungswissenschaft
Institut für Psychologie

Erklärung gemäß *(bitte Zutreffendes ankreuzen)*

☐ **§ 4 (1c) der Promotionsordnung**

des Instituts für Bewegungswissenschaft der Universität Hamburg vom 18.08.2010

☐ **§ 5 (4d) der Promotionsordnung**

des Instituts für Psychologie der Universität Hamburg vom 20.08.2003

Hiermit erkläre ich,

_____ (Vorname, Nachname),

dass ich mich an einer anderen Universität oder Fakultät noch keiner Doktorprüfung unterzogen oder mich um Zulassung zu einer Doktorprüfung bemüht habe.

_____ _____

Ort, Datum Unterschrift

Studien- und Prüfungsbüro Bewegungswissenschaft • Fakultät PB • Universität Hamburg • Mollerstraße 10 • 20148 Hamburg
Studien- und Prüfungsbüro Psychologie • Fakultät PB • Universität Hamburg • Von-Melle-Park 5 • 20146 Hamburg

www.pb.uni-hamburg.de

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Fakultät für
Psychologie und
Bewegungswissenschaft

Institut für Bewegungswissenschaft
Institut für Psychologie

Eidesstattliche Erklärung nach *(bitte Zutreffendes ankreuzen)*

☐ **§ 7 (4) der Promotionsordnung des Instituts für Bewegungswissenschaft der Universität Hamburg vom 18.08.2010**

☐ **§ 9 (1c und 1d) der Promotionsordnung des Instituts für Psychologie der Universität Hamburg vom 20.08.2003**

Hiermit erkläre ich an Eides statt,

1. dass die von mir vorgelegte Dissertation nicht Gegenstand eines anderen Prüfungsverfahrens gewesen oder in einem solchen Verfahren als ungenügend beurteilt worden ist.
2. dass ich die von mir vorgelegte Dissertation selbst verfasst, keine anderen als die angegebenen Quellen und Hilfsmittel benutzt und keine kommerzielle Promotionsberatung in Anspruch genommen habe. Die wörtlich oder inhaltlich übernommenen Stellen habe ich als solche kenntlich gemacht.

Ort, Datum | Unterschrift

Studien- und Prüfungsbüro Bewegungswissenschaft • Fakultät PB • Universität Hamburg • Mollerstraße 10 • 20148 Hamburg
Studien- und Prüfungsbüro Psychologie • Fakultät PB • Universität Hamburg • Von-Melle-Park 5 • 20146 Hamburg

· www.pb.uni-hamburg.de