Comparing the Patient Health Questionnaire – 15 and the Somatic Symptom Scale – 8 as Measures of Somatic Symptom Burden

Anne Toussaint[1] PhD, Kurt Kroenke[2,3,4] MD, Fitsum Baye[5] MS, Spencer Lourens[5], PhD

[1] Department of Psychosomatic Medicine and Psychotherapy, University Medical Center Hamburg-Eppendorf and Schön Klinik Hamburg Eilbek, Hamburg, Germany

[2] VA HSR&D Center for Health Information and Communication, Roudebush VA Medical Center, Indianapolis, IN, United States

[3] Department of Medicine, Indiana University School of Medicine, Indianapolis, IN, United States

[4] Regenstrief Institute, Inc., Indianapolis, IN, United States

[5] Department of Biostatistics, Indiana University School of Medicine, Indianapolis, IN, United States

Correspondence concerning this article should be addressed to Anne Toussaint, Department of Psychosomatic Medicine and Psychotherapy, University Medical Center Hamburg-Eppendorf, Martinistr. 52, 20246 Hamburg, Germany. E-Mail: a.toussaint@uke.de, Phone: +49 (0)40 7410 52972

e-mail: [1] a.toussaint@uke.de, [2,3,4] kkroenke@regenstrief.org.de, [5] fbaye@iu.edu, [5] slourens@iu.edu

Abstract

Purpose: The Patient Health Questionnaire – 15 (PHQ-15) and the Somatic Symptom Scale – 8 (SSS-8) are self-report measures which assess somatic symptom burden. The present study investigates whether the two measures are comparable in terms of their psychometric properties and estimates of symptom burden.

Method: Item characteristics, reliability, symptom severity and construct validity with regard to other relevant psychological, health-related quality of life and disability measures were compared for the PHQ-15m and the SSS-8 in 294 primary care patients who participated in a randomized comparative effectiveness trial targeting pain and mood symptoms.

Results: The reliabilities of the PHQ-15m and the SSS-8 were $\alpha = 0.66$ and $\alpha = 0.72$, respectively. Both measures were highly correlated ($r = 0.79$). All item characteristics were comparable and both instruments showed the same pattern of correlations with instruments measuring depression, anxiety, pain, quality of life and impairment ($r = 0.25$ to 0.53). A 1-point score increase (worsening of somatic symptoms) on either instrument resulted in a 3.7% to 3.9% increase in the number of disability days reported for the last four weeks. Using the same severity thresholds (5: low, 10: medium, 15: high), both measures identified nearly identical subgroups of patients with regard to health-related quality of life and disability.

Conclusion: The PHQ-15m and the SSS-8 are comparable measures in terms of reliability and validity and severity classifications. These findings are in line with previous results and support the use of the SSS-8 as a valuable and short alternative to the original PHQ-15 in settings with limited assessment time.

Keywords: PHQ-15, SSS-8, somatic symptom burden, psychometrics

Introduction

Somatic symptoms are ubiquitous in the general population; an estimated 80% of individuals will experience one or more somatic symptoms in any given month [1]. Symptoms may include pain as well as digestive, cardiovascular, pulmonary, urological, neurological, or sensory complaints. Many symptoms are neither exclusive correlates of an organic disease (e.g. cancer or coronary heart disease) nor exclusive symptoms of a psychiatric condition (e.g. depression or anxiety disorders) [2-4]. Somatic symptoms which are either part of a functional somatic syndrome, or otherwise unexplained by pathology, are the reason for at least 33% of primary care consultations and between 15 and 54% of specialist referrals across many medical disciplines. Usually, only those individuals who are actually distressed or impaired by their somatic symptoms present to clinical practice. About one fourth of all patients develops persistent symptoms [5].

Persistent somatic symptoms usually represent a substantial burden, they significantly impair patients' quality of life and level of functioning. Psychological factors like depression or anxiety as well as symptom-specific concerns or expectations are important contributors to high levels of health care use in these patients [6]. Repeated investigations and hospital treatment are frequent consequences and lead to high socio-economic costs. The health care burden due to persistent somatic symptoms is comparable to anxiety and depressive disorders, and there is a high co-morbidity between these disorders [7, 8]. Strategies to improve the early recognition and identification of patients with high somatic symptom burden is important to initiate adequate treatment [9].

Standardized patient-reported outcome measures like self-report questionnaires are a good option to assess, quantify, and monitor common conditions in clinical, and especially in general practice. A frequent challenge in this context is to make precise assessments within a limited amount of time. Additionally, self-report represents a complementary source of information by capturing the patients' own perspectives of their symptoms [10].

From a research perspective, it is important to assess the number, type and severity of somatic symptoms as change in symptoms will continue to be a central outcome feature of treatments for patients and physicians/therapists alike [11]. There are several standardized and validated instruments which effectively measure the patients' burden due to specific somatic complaints (e.g. PHQ-15 [12], SSS-8 [13], SCL-90R or BSI [14]).

The Patient Health Questionnaire PHQ-15 is one of the most frequently used instruments to identify people at risk for somatization. It has well-established psychometric properties, is available in multiple languages and has been recommended for use in large-scale studies [15]. The PHQ-15 assesses the presence and severity of common somatic symptoms in primary care, such as fatigue, gastrointestinal, musculoskeletal, pain, and cardiopulmonary symptoms within the last four weeks using 15 items. Sum-scores range from 0 to 30 and indicate the self-rated symptom burden with higher scores indicating higher burden (0-4 no-minimal; 5-9 low; 10-14 medium; 15-30 high).

The Somatic Symptom Scale-8 (SSS-8) was developed within the process of the DSM-5 field trials as a measure of somatic symptom burden related to the new diagnosis of somatic symptom disorder (former title: Patient Health Questionnaire Somatic Symptom Short-Form, PHQ-SSS) [16]. It is an abbreviated 8-item version of the PHQ-15. The items of the SSS-8 were selected on the basis of symptom prevalence in primary care, association with measures of functioning, and statistical commonalities with the items of the complete scale. Some original items were condensed from two into one, and a few items were deleted. A 5-point response option (0-4) for each SSS-8 item and a 7-day time frame are used. 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. Gender and age specific norms are available [13]. Previous studies demonstrated good item characteristics and excellent reliability, a sound factor structure and significant associations with related constructs like depression, anxiety, quality of life, and health care use [13, 17]. The SSS-8 is available in

English, German, and Japanese [13, 17, 18], and its sensitivity to change has recently been demonstrated [19].

Aims of the study

Gierk and colleagues [17] examined within a sample of outpatients from a psychosomatic clinic in Germany whether both measures were comparable in terms of their psychometric properties and estimates of symptom burden. The correlation between both instruments was high (r = 0.83) and they showed similar results considering reliability and validity. The SSS-8 performed well as a short version of the PHQ-15. Also, analyses suggested that similar cut-points might be used for both measures in grading somatic symptom burden as mild, moderate or severe. The aim of this paper is to replicate and extend the psychometric comparison of both measures using baseline data from a large clinical trial. Specifically, we compare the two measures in terms of item characteristics, reliability, and construct validity of the severity thresholds with regard to health related quality of life, functional impairment and work disability.

Method

Procedure and participants

Data were drawn from the Comprehensive vs. Assisted Management of Mood and Physical Symptoms Study (CAMMPS: https://clinicaltrials.gov/ct2/show/NCT01757301), a randomized comparative effectiveness trial designed to test the relative effectiveness of a lower-resource vs. a higher-resource enhancement of usual primary care in the management of Veterans suffering from pain plus comorbid anxiety and/or depression. The trial enrolled a total of 294 patients between January 2014 and June 2016. Baseline data was used for all analyses in this study. All measures were interviewer-administered. The trial was approved by the Indiana University institutional review board and the Roudebush VAMC research review committee.

Somatic symptom measures

The CAMMPS trial used a modified, 14 item version of the PHQ-15, hereinafter referred to as the PHQ-15m. The item on *sexual pain and problems* was left out for several reasons. First, it is the least commonly endorsed item in multiple epidemiological studies and also the one item that a subset of respondents are most uncomfortable answering. Second, the item has shown the lowest item-total correlation (0.33) of any of the PHQ-15 items, and all correlations with other items of the scale were low (<.20). Third, this item demonstrates among the lowest correlations with multiple domains of quality of life, disability and health care use. Fourth, in factor analysis, this item and the item on menstrual problems had the lowest commonality with the other 13 items and were excluded from the three factors (cardiopulmonary, gastrointestinal, and fatigue/pain) [12]. Fifth, reliability in terms of internal consistency (Cronbach α) for the modified 14 item version (PHQ-15m) ranged from 0.76 - 0.77 in three large clinical trials [20-22], which is similar to the reliability reported for the original PHQ-15 (Cronbach α = .80 [12]). Finally, these trials demonstrated the responsiveness to treatment of the PHQ-15m. The Somatic Symptom Scale-8 (SSS-8) was used in its original form [13].

Other Mental Health, Quality of Life and Disability Measures

Depression was measured with the Patient Health Questionnaire 9-item depression scale (PHQ-9; [23] which assesses the presence of the nine DSM criteria for major depression within the last two weeks. Scores range from 0 to 27 and indicate the severity of depression (high scores reflect high symptom load).

Anxiety was assessed with the7-item Generalized Anxiety Disorder Scale (GAD-7) [24], a self-administered patient questionnaire which is used as a screening tool and severity measure of both generalized anxiety disorder as well as other common anxiety disorders. Scores range from 0 (minimal) to 21 (severe).

The Brief Pain Inventory (BPI) [25] is an 11-item self-administered questionnaire used to evaluate the severity of a patient's pain and the impact of this pain on the patient's daily functioning. Patients are asked to rate their worst, least, average, and current pain intensity, , and the degree that pain interferes with general activity, mood, walking ability, normal work, relations with other persons, sleep, and enjoyment of life on a 0 (none) to 10 (worst) . We used the mean total pain score for our analyses.

The 12-item Short Form Health Survey (SF-12) [26] is a measure of health-related quality of life from which a Physical Component Summary (PCS) score and Mental Component Summary (MCS) score can be derived. Both scores are standardized to a mean of 50 and a standard deviation of 10 to facilitate the comparison of individuals within the general population. Higher scores correspond to better health related quality of life.

The Sheehan Disability Index (SDI) [27] assesses functional impairment in three interrelated domains: work/school, social and family life. Each of its 3 items is scored from 0 (unimpaired) to 10 (highly impaired), with the SDI score being a mean of the 3 items.

In addition, patients reported the number of days in the past four weeks where they had to reduce their usual activities by 50% or more (range: 0 to 28), and provided information on sociodemographic characteristics.

Statistical Analyses

Because the PHQ-15 and SSS-8 were interviewer-administered, the amount of missing data at the item level was very low (1-2%) and comparable for both scales. To provide a conservative estimate of the scale score, a missing item was assigned a value of 0. We calculated means, standard deviations, corrected item-total-correlations and frequency distribution of responses per item for the PHQ-15m and the SSS-8. Reliability was assessed using Cronbach´s α. Pearson correlations between the PHQ-15m, SSS-8, PHQ-9, GAD-7, BPI, MCS, PCS, SDI and disability days were calculated. Percentile distributions of somatic symptom severity for the PHQ-15m and the SSS-8 were calculated. Incidence rate ratios

(IRR) for each 1-point change in PHQ-15m and SSS-8 in association with change in disability days were calculated using parameter estimates from negative binomial regression. In terms of clinically important differences, 1- and 2-standard error of measurement (SEM) changes in PHQ-15m and SSS-8 were calculated (where SEM = standard deviation times the square root of 1-Cronbach's alpha) to quantify the increase in disability days that is associated with change in somatic symptom burden [28]. To study the comparability of the severity thresholds of both scales, we calculated the mean (SD) physical health-related quality of life (PCS), functional impairment (SDI) and disability days for groups of patients that were classified according to the different PHQ-15m and SSS-8 scoring cut-points. We quantified the between-scale differences using the standardized mean difference (Cohen´s d).

Results

Study sample

Table 1 reports sociodemographic and psychopathological characteristics of the study sample. Participants had a mean age of 57.4 and were 87% male and 79% white, characteristic of a Veteran population. Symptom and quality of life scores were in the moderate to moderately severe range as expected for patients with chronic pain and psychiatric comorbidity.

Psychometric characteristics of PHQ-15m and SSS-8

The mean (SD) sum-scores of the PHQ-15m and the SSS-8 were 16.7 (5.8) and 13.5 (4.1), respectively. The correlation between the two scales was 0.79. The internal consistencies were $\alpha = 0.66$ for the PHQ-15m and $\alpha = 0.72$ for the SSS-8. As shown in Table 2, the mean score for specific symptoms varied in both scales, which indicates that some somatic symptoms are reported more frequently and/or at a higher severity level than other symptoms. Pain, energy and sleep items had the highest mean scores, whereas dizziness had the lowest scores. Considering the differing response formats of both scales, frequency distributions of responses are comparable, with the items on chest pain and/or shortness of

breath being the least comparable. All corrected item-total correlations were > 0.50 for the SSS-8; several were lower (< 0.40) for the PHQ-15m.

Association of somatic symptom burden with psychopathology, health-related quality of life, and disability.

The PHQ-15m and the SSS-8 sum-scores were highly correlated (r = 0.79). As shown in **Table 3**, both scales showed a similar pattern of correlations with self-report measures of depression, anxiety, pain (BDI), physical and mental health-related quality of life, functional impairment, and disability days (CUTPH): Higher scores on the PHQ-15m and SSS-8 correspond to higher levels of depression, anxiety, pain, functional impairment and disability days and lower levels of physical and mental health-related quality of life.

Percentile distribution of PHQ-15m and SSS-8 scores

Although the PHQ-15m and the SSS-8 have different numbers of items and a different response format, the sum-scores of both scales have a similar percentile distribution which is shown in **Table 4.** The comparability in the distribution is high (difference ≤ 3 points) up to the 50th percentile, however, it steadily decreases from the 60st percentile (difference = 4 points) to the 99th percentile (difference = 8 points).

Association of somatic symptom scores with disability days

Patient-reported number of disability days in the past four weeks (i.e., number of days where individuals had to reduce their usual activities by 50% or more) had a range of 0 to 28 days with a mean of 15.4 (9.4). Negative binomial regression analysis indicated that the PHQ-15m and the SSS-8 both had significant and similar associations with disability days. A 1-point increase in PHQ-15m (worsening of somatic symptoms) resulted in a 3.7% increment in the number of disability days (IRR = 1.037 [95% CI = 1.015; 1.061], z = 2.23, p <.0001). A 1-point increase in SSS-8 (worsening in somatic symptoms) resulted in a 3.9% increment in the number of disability days (IRR = 1.039 [95% CI = 1.023; 1.055], z = 2.08, p <.0001). Using

the standard error of measurement (SEM) as an approach to estimate the minimal clinically important difference (MCID) for a scale, the SEM was 2.3 for the PHQ-15m 3.1 for the SSS-8. A 1-SEM change in PHQ-15m and SSS-8 resulted in a 9.2% and 12.4% increase in disability days, respectively. A 2-SEM change in PHQ-15m and SSS-8 resulted in a 19.3% and 26.3% increase in disability days, respectively.

Comparing two ways of classifying severity cut-points on the SSS-8

To examine the comparability of the PHQ-15m and the SSS-8 severity categories, we compared – in accordance with Gierk et al. [17] – the physical health-related quality of life (SF-12 PCS) scores of patients who were grouped by the standard PHQ-15m severity thresholds and two ways of classifying SSS-8 severity thresholds. We additionally compared the groups with regard to their Sheehan Disability Index scores and number of disability days. To quantify the between-scale differences on these three outcomes, we calculated the standardized mean differences (Cohen´s d) for each severity category. The original PHQ-15m (0-4 no – minimal; 5-9 low; 10-14 medium; 15-30 high) and the original SSS-8 thresholds (0-3 no – minimal; 4-7 low; 8-11 medium; 12-15 high; $\geq$ 16 very high) were applied. Additionally the standard PHQ-15m thresholds were applied to the SSS-8 (0-4 no-minimal; 5-9 low; 10-14 medium; $\geq$ 15 high). Only two of the patients from our sample fell within any of the groups of no – minimum symptom burden, so that the results for this threshold group is not displayed. All other results are shown in **Table 5**. For all three outcomes, grouping patients according to the original SSS-8 severity thresholds led to higher group differences between the PHQ-15m and SSS-8 than using the PHQ-15m severity thresholds for the SSS-8. There are, however, two exceptions: in the low category for the SF-12 PCS and in the medium category for the SDI, the SSS-8 original scoring method led to smaller differences.

Discussion

The main aim of the study was to compare the psychometric properties of two instruments assessing somatic symptom burden: the PHQ-15 and the SSS-8. In an earlier

study, Gierk et al. [17] found comparable reliability and construct validity for both measures and therefore suggested the use of the SSS-8 as a more efficient alternative to the original PHQ-15, especially in clinical settings with limited assessment time or in research settings to reduce the measurement burden for study participants. Since the original study was based on a fairly small and selective sample of patients, the findings regarding alternative severity thresholds for the SSS-8 were considered preliminary. We used the PHQ-15m and SSS-8 within a larger sample of 294 patients from a clinical trial involving patients with chronic pain and mood symptoms and compared the psychometric properties of both scales in order to support previous findings. In addition to the SF-12 MCS, we included two additional measures on impairment (SDI and disability days) to further validate severity thresholds.

Cronbach´s α of the SSS-8 (0.72) and PHQ-15m (0.66) were similar, yet lower than in the original study (0.76 and 0.80, respectively). The higher Cronbach's alpha for the SSS-8 may be partly because the SSS-8 has more response options than the PHQ-15 (5 vs. 3). The correlation between the PHQ-15 and SSS-8 scores was higher than the Cronbach's alpha of each scale, which is not surprising. Cronbach's alpha is a measure of the internal reliability of the items constituting a scale. Somatic symptom scales include multiple individual symptoms which originate from different bodily organs or locations. Thus, one might expect greater heterogeneity among items in a somatic symptom scale than among items in other psychological scales (e.g., depression or anxiety). In contrast, each scale score is a single number representing the sum of the individual item scores. Also, the SSS-8 and PHQ-15 include a number of items in common.

The frequency distributions of responses, the mean item severity of specific symptoms, and the standard deviations within the scales were comparable considering the differing response formats of both scales. For both scales, the highest average scores were found for items on pain, energy and sleep which could be expected in a trial on mood and pain symptoms. The lowest average scores were found for dizziness. The items on chest pain

and/or shortness of breath showed the least comparable results which may be due to the fact that they have been combined into a single item in the SSS-8. Both items may be more relevant for patients with burdensome cardiopulmonary symptoms which were not the focus of our study. All item-total correlations were lower for the PHQ-15m which may in general be due to the heterogeneous latent structure of somatic symptom burden [29] which is even more salient to the PHQ-15m which consists of more items.

The correlation between the SSS-8 and the PHQ-15m was high (0.79) and comparable to the findings by Gierk et al (0.83). Both measures showed substantial positive associations with depression and anxiety which is expected since there are high comorbidities between somatoform disorders and depression and anxiety disorders [8]. Associations with the Brief Pain Inventory (BPI) were similar and also substantial which is not surprising given the item overlap with regard to pain. Burdensome somatic symptoms are usually associated with decreased health-related quality of life, reduced daily activities and increased disability across different life domains [30, 31]. With regard to physical and mental health-related quality of life, both the PHQ-15m and SSS-8showed a substantial negative association (i.e. high symptom burden corresponded to poor health-related quality of life). Associations were substantially positive with respect to disability (i.e. high symptom burden corresponded to high impairment). When comparing the correlations of the PHQ-15 and SSS-8 with measures that were used in both Gierk's study and our trial (GAD-7 anxiety, PHQ-9 depression, SF mental and physical health-related quality of life), correlations found in our study were slightly lower, with the SSS-8 showing higher associations than the PHQ-15m with all used.

Looking at the percentile distributions, the comparability in the two scales was high up to the 50th percentile (difference ≤ 3 points) and steadily decreased from the 60st percentile on (differences between 4 and 8 points). These differences are greater than the ones found in Gierk et al [17]. Similar to Gierk et al., we found substantial differences in the severity estimation of the PHQ-15m and SSS-8 severity thresholds which could in general be reduced

by scoring the SSS-8 with the PHQ-15 thresholds. This supports a good correspondence between the SSS-8 and the PHQ-15m with data from a larger sample of patients from primary care, not only on the dimension of physical health-related quality of life, but also when measuring physical impairment (SDI) and disability days.

Our study has a few limitations. First and for reasons discussed earlier, we used a modified version of the PHQ-15, the PHQ-15m, which limits the direct comparability of our study results to the ones found in Gierk et al. However, we would assume that the exclusion of the item on sexual pain and problems in the PHQ-15m makes both instruments more comparable since this particular item is not included in the SSS-8. Since it is also one of the items least endorsed, the average PHQ-15 sum-score should not be noticeably influenced by excluding this item. Second, our study uses baseline data only, so that longitudinal questions like the effects of the different time frames of the PHQ-15m (i.e. symptom burden within the last 4 weeks) and the SSS-8 (i.e. symptom burden within the last 7 days) cannot be answered. Due to feasibility reasons, both questionnaires were presented to the patients at baseline as part of an interview-based set of measures. The PHQ-15m was administered first in all patients, so that possible carry-over effects cannot be excluded.

Altogether the SSS-8 performed well as a brief measure of somatic symptom burden in primary care patients with mood and pain symptoms. Its psychometric properties were similar or even slightly better (i.e. higher correlations with other measures) than those of the PHQ-15m. Taking into account all three different impairment measures, the estimates of symptom burden were very similar when the SSS-8 was scored with the PHQ-15 thresholds. Our results therefore support the idea of common scoring thresholds for both instruments. Since there is a growing impetus to incorporate patient-reported outcome measures into clinical practice [32] with the simultaneous need to consider competing demands for treating and managing many different acute and chronic conditions in clinical practice, our findings provide further support for the utility of the SSS-8 as a short measure for somatic symptom burden.

References

[1] Kroenke K, Arrington ME, Mangelsdorff AD. The prevalence of symptoms in medical outpatients and the adequacy of therapy. Arch Intern Med. 1990; 150:1685.9.

[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; 170:1686-94.

[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: 1469-71.

[4] 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; 30:191-9.

[5] Kroenke K. A practical and evidence-based approach to common symptoms. Ann Intern Med 2014; 161:579-586.

[6] Jackson JL, Kroenke K. The effect of unmet expectations among adults presenting with physical symptoms. Ann Intern Med 2001; 134:889-897.

[7] Barsky AJ, Orav EJ, Bates DW. Somatization increases medical utilization and costs independent of psychiatric and medical comorbidity. Archives of General Psychiatry. 2005; 62:903-10.

[8] Kohlmann S, Gierk B, Hilbert A, Brähler E, Löwe B. The overlap of somatic, anxious and depressive syndromes. J Psychosom Res 2016;90:51-6.

[9] Murray A, Toussaint A, Althaus A, Löwe B. The challenge of diagnosing non-specific, functional, and somatoform disorders: A systematic review of barriers to diagnosis in primary care. J Psychosom Res. 2016;80:1-10.

[10] Black N. Patient reported outcome measures could help transform healthcare. BMJ. 2013; 346:167.

[11] Rief W, Burton C, Frostholm L, Henningsen P, Kleinstäuber M, Kop W, Löwe B, Martin A, Malt U, Rosmalen J, Schröder A, Shedden-Mora M, Toussaint A, van der Feltz-Cornelis C. Core outcome domains for clinical trials on somatic symptom disorder, bodily distress disorder and functional somatic syndromes: Euronet-SOMA recommendations. Psychosom Med. 2017: in press.

[12] 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.

[13] 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 Int Med 2014; 174:399-407.

[14] Derogatis, Leonard R.; Savitz, Kathryn L. The SCL-90-R and the Brief Symptom Inventory (BSI) in Primary Care. In Maruish, ME. Handbook of psychological assessment in primary care settings. 2000. Mahwah, NJ: Lawrence Erlbaum Associates. pp. 297–334.

[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; 20:459-68.

[16] 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.

[17] Gierk B, Kohlmann S, Toussaint A, Wahl I, Brünahl C, Murray A, Löwe B. Assessing somatic symptom burden: A psychometric comparison of the Patient Health Questionnaire – 15 (PHQ-15) and the Somatic Symptom Scale – 8 (SSS-8). J Psychosom Res. 2014; 78:352-5.

[18] 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.

[19] Gierk B, Kohlmann S, Hagemann-Göbel M, Löwe B, Nestoriuc Y. Monitoring somatic symptoms in patients with mental disorders: Sensitivity to change and minimally clinically important difference of the Somatic Symptom Scale – 8 (SSS-8). In preparation

[20] Kroenke K, Krebs E, Wu J, Bair MJ, Damush T, Chumbler N, York T, Weitlauf S, McCalley S, Evans E, Barnd J, Zhangsheng Y. Stepped care to optimize pain care effectiveness (SCOPE) trial study design and sample characteristics. Contemporary Cinical Trials. 2013; 34:270-281.

[21] Kroenke K, Theobald D, Norton K, Sanders R, Schlundt S, McCalley S, Harvey P, Iseminger K, Morrison G, Carpenter J, Stubbs D, Jacks R, Carney-Doebbeling C, Wu J, Tu W. Indiana cancer pain and depression (INCPAD) trial design of a telecare management intervention for cancer-related symptoms and baseline characteristics of study participants. Gen Hosp Psychiatry. 2009; 31:240-253.

[22] Kroenke K, Bair MJ, Damush TM, Wu J, Hoke S, Sutherland J, Tu W. Optimized antidepressant therapy and pain self-management in primary care patients with depression and musculoskeletal pain: A randomized controlled trial. JAMA. 2009; 301:2099-2110.

[23] Kroenke K, Spitzer RL, Williams JBW. The PHQ-9 - Validity of a brief depression severity measure. J Gen Intern Med. 2001; 16:606-13.

[24] Löwe B, Decker O, Müller S, Brähler E, Schellberg D, Herzog W, York Herzberg P. Validation and standardization of the generalized anxiety disorder screener (GAD-7) in the general population. Med Care. 2008; 46:266-74.

[25] Cleeland CS, Gonin R, Hatfield AK, et al. Pain and its treatment in outpatients with metastatic cancer. N Engl J Med 1994;330:592–596.

[26] Ware JE, Jr., Kosinski M, Keller SD. A 12 Item Short Form Health Survey: Construction of scales and preliminary tests of reliability and validity. Med Care 1996; 34:220-233.

[27] Sheehan DV, Harnett-Sheehan K, Raj BA: The measurement of disability. Int Clin Psychopharmacol 11(suppl 3):89-95, 1996.

[28] 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.

[29] 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.

[30] Henningsen P, Zipfel S, Herzog W. Management of functional somatic syndromes. Lancet. 2007;369: 946-955.

[31] 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; 30:191-9.

[32] Glasgow RE, Riley WT. Pragmatic measures: what they are and why we need them. Am J Prev. 2013; 45:237-43.
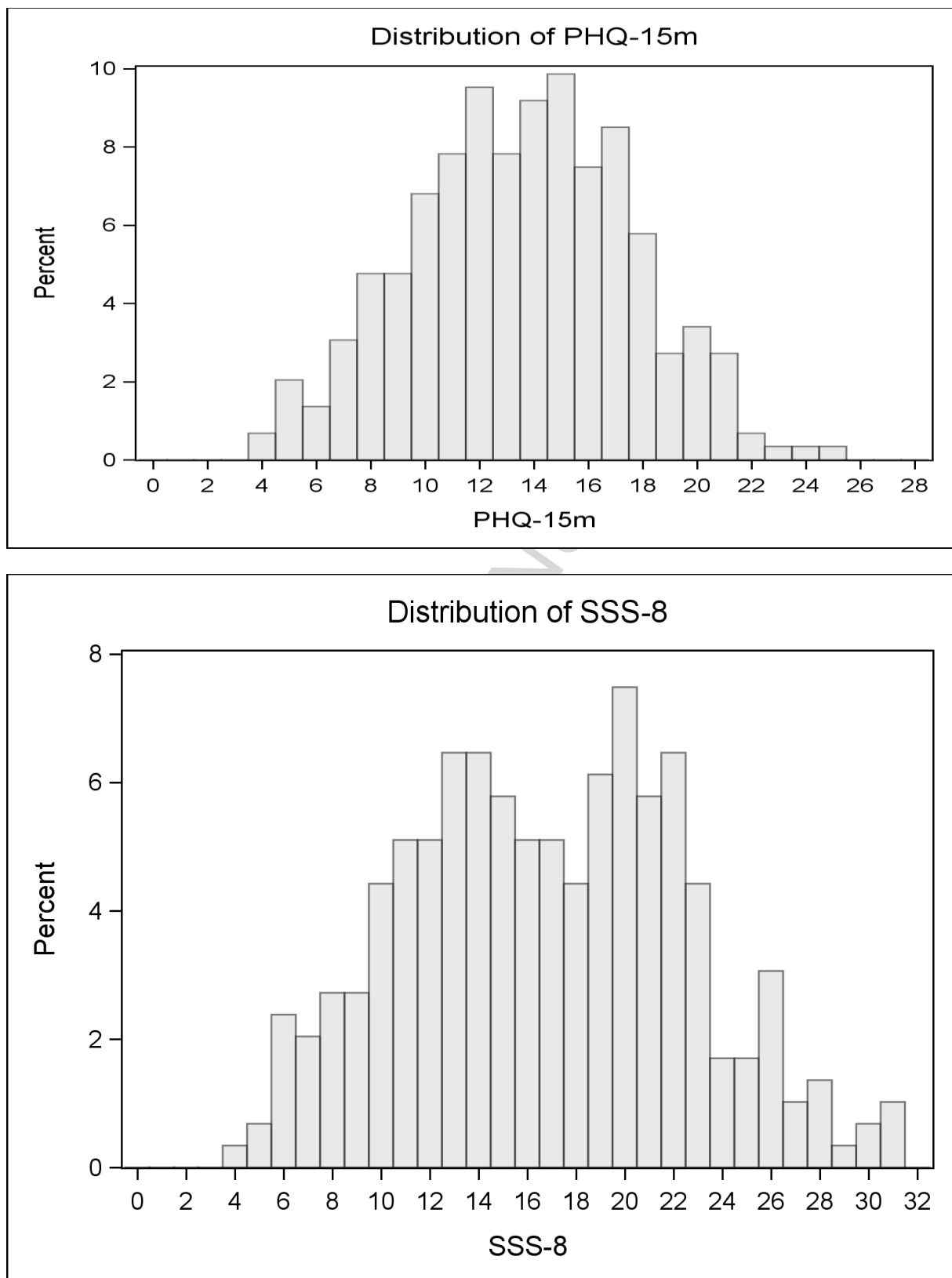
Figure 1

Distribution of PHQ-15m

Distribution of SSS-8

Table 1: Baseline data of the study sample (n = 294)

Variable	Total (n=294) N (%) or Mean (SD)
Demographic	
Age	57.4 (12.1)
Male sex	257 (87.4%)
Race	
White	233 (79.3%)
Black	45 (15.3%)
Other	16 (5.4%)
Married	167 (56.8%)
Education:	
≤ high school	76 (25.9%)
Some college or trade school	156 (53.1%)
College	62 (21.1%)
Employment:	
Employed	82 (27.9%)
Unemployed	112 (38.1%)
Retired	100 (34.0%)
Scale Scores	
Somatic symptom burden (PHQ-15m)	13.5 (4.1)
Somatic symptom burden (SSS-8)	16.7 (5.8)
Depression severity (PHQ-9)	14.0 (5.2)
Anxiety severity (GAD-7)	11.2 (5.1)
Total Pain Score (BPI)	5.9 (1.7)
Health-related quality of life / physical (SF-12 PCS)	33.2 (8.5)
Health-related quality of life / mental (SF-12 MCS)	37.8 (10.7)
Sheehan disability (SDI)	5.7 (2.3)
Disability days past 4 weeks	15.4 (9.4)

Table 2: Frequency distribution of responses (%), means (SD), and item-total correlations for the items of the PHQ-15m and SSS-8

Somatic Symptom (Scale)	Not at all (both scales)	A little bit (PHQ-15m) A little bit or Somewhat (SSS-8)	A lot (PHQ-15m) Quite a bit or Very much (SSS-8)	Mean (SD)* Item Score	Item-Total Correlation
Headache (PHQ-15m)	34.7	36.7	28.6	0.9 (0.8)	0.49
Headache (SSS-8)	40.8	34.3	24.8	1.4 (1.4)	0.62
Dizziness (PHQ-15m)	41.5	43.5	15.0	0.7 (0.7)	0.54
Dizziness (SSS-8)	51.7	37.1	11.2	0.9 (1.1)	0.52
Back pain (PHQ-15m)	5.8	16.7	77.6	1.7 (0.6)	0.32
Back pain (SSS-8)	8.5	17.6	73.8	2.9 (1.3)	0.65
Feeling tired/having low energy (PHQ-15m)	4.4	16.7	78.9	1.7 (0.5)	0.32
Feeling tired/ having low energy (SSS-8)	2.4	27.9	69.7	2.8 (1.1)	0.65
Trouble sleeping (PHQ-15m)	7.1	13.3	79.6	1.7 (0.6)	0.37
Trouble sleeping (SSS-8)	11.2	20.0	68.7	2.8 (1.4)	0.57
Limb/joint pain (PHQ-15m)	1.7	11.2	87.1	1.9 (0.4)	0.17
Limb/joint pain (SSS-8)	2.0	20.7	77.2	3.2 (1.0)	0.51
Chest pain (PHQ-15m)	64.3	27.9	7.8	0.4 (0.6)	0.46
Shortness of breath (PHQ-15m)	41.8	36.7	21.4	0.8 (0.8)	0.53
Chest pain or shortness of breath (SSS-8)	41.2	43.5	15.3	1.1 (1.2)	0.53
Stomach pain (PHQ-15m)	39.1	37.4	23.5	0.8 (0.8)	0.58
Nausea/gas/indigestion (PHQ-15m)	33.3	31.3	35.4	1.0 (0.8)	0.65
Constipation/loose stools (PHQ-15m)	37.1	29.6	33.3	1.0 (0.8)	0.59
Stomach or bowel problems (SSS-8)	30.6	35.5	34.0	1.7 (1.5)	0.60
Fainting spells (PHQ-15m)	94.9	4.4	0.7	0.1 (0.3)	0.20
Menstrual pain/problems (PHQ-15m)	97.6	2.0	0.3	0.03 (0.2)	0.12
Feeling heart pound/race (PHQ-15m)	43.9	43.2	12.9	0.7 (0.7)	0.51`

*The PHQ-15m asks how much a symptom bothered a patient in the past month; each item is scored from 0 (not at all) to 2 (bothered a lot).

The SSS-8 asks how much a symptom bothered the patient in the past week; each item is scored from 0 (not at all) to 4 (very much).

Table 3: Pearson correlations of the PHQ-15m and SSS-8 sum-score with other scales

Scale	PHQ-15m	P value	SSS-8	P value
PHQ-9 depression	0.51	<.0001	0.53	<.0001
GAD-7 anxiety score	0.35	<.0001	0.37	<.0001
Brief Pain Inventory	0.28	<.0001	0.36	<.0001
SF-12 PCS (physical)	- 0.32	<.0001	- 0.40	<.0001
SF-12 MCS (mental)	- 0.32	<.0001	- 0.37	<.0001
Sheehan Disability Index	0.41	<.0001	0.46	<.0001
Disability days in past 4 weeks	0.25	<.0001	0.35	<.0001

Note: PHQ-15m = Patient Health Questionnaire – 15, modified version. SSS-8 = Somatic Symptom Scale-8. PHQ-9 = Patient Health Questionnaire Depression Scale-9. GAD-7 = Generalized Anxiety Disorder Scale-7. SF-12 PCS = Physical Component Summary score of the SF-12 health-related quality of life scale. SF-12 = Mental Component Summary score of the SF-12.

Table 4: Percentile distribution of PHQ-15m and SSS-8 sum-scores (n = 294)

Percentile	PHQ-15m Score	SSS-8 Score
1	5	5
5	7	7
10	8	9
20	10	11
30	11	13
40	12	15
50	14	17
60	15	19
70	16	20
80	17	22
90	19	24
95	20	26
99	23	31

Table 5 Construct validity of PHQ-15m and SSS-8 severity categories (n = 294)

Severity	SF-12 Physical Component (PCS), mean (SD)			Standardized differences*	
	1. PHQ-15 original [a]	2. SSS-8 original [b]	3. SSS-8 with PHQ-15 scoring [c]	1. vs. 2.	1. vs. 3.
Low	38.2 (10.0)	37.9 (9.4)	37.1 (8.9)	0.03	0.12
Medium	34.2 (8.0)	38.5 (9.0)	37.2 (8.1)	0.51	0.37
High	30.2 (7.1)	34.7 (7.6)	30.6 (7.3)	0.61	0.06
Very high [d]		30.7 (7.6)			
Severity	Sheehan Disability Index, mean (SD)			Standardized differences*	
	1. PHQ-15 original [a]	2. SSS-8 original [b]	3. SSS-8 with PHQ-15 scoring [c]	1. vs. 2.	1. vs. 3.
Low	4.4 (2.6)	3.5 (2.6)	4.5 (2.1)	0.35	0.04
Medium	5.1 (2.2)	4.8 (1.9)	4.6 (2.3)	0.15	0.22
High	6.8 (1.9)	4.9 (2.4)	6.4 (2.1)	0.88	0.20
Very high [d]		6.5 (2.0)			
Severity	Disability Days in Past 4 Weeks, mean (SD)			Standardized differences*	
	1. PHQ-15 original [a]	2. SSS-8 original [b]	3. SSS-8 with PHQ-15 scoring [c]	1. vs. 2.	1. vs. 3.
Low	12.7 (9.5)	9.3 (8.3)	9.8 (8.3)	0.38	0.33
Medium	14.3 (9.2)	11.9 (10.0)	13.3 (9.6)	0.25	0.11
High	17.8 (9.1)	13.6 (8.8)	17.4 (8.9)	0.47	0.04
Very high [d]		17.8 (8.9)			

* Standardized differences are computed as the mean difference divided by the pooled standard deviations (Cohen´s d).

[a] 0-4 no-minimal (n = 2)[e], 5-9 low (n = 47), 10-14 medium (n = 121), and 15-30 high (n = 124)

[b] 0-3 no-minimal (n = 0)[e], 4-7 low (n = 16) , 8-11 medium (n = 44), 12-15 high (n = 70), and 16-32 very high (n = 164)

[c] 0-4 no-minimal (n = 1)[e], 5-9 low (n = 31), 10-14 medium (n = 81), and 15-32 very high (n = 181)

[d] only available for the SSS-8 original scoring method

[e] not reported in the table, because of n ≤ 2

Highlights

- The PHQ-15m and the SSS-8 are comparable measures
- The SSS-8 is a short alternative to PHQ-15 in settings with limited assessment time
- Patient-reported outcome measures should be incorporated into clinical practice