The Somatic Symptom Disorder – B Criteria Scale (SSD-12) in a Dutch Clinical Sample. A validation study.

C.A.D. Kamp[1,2], L. de Vroege[1,2], J. F. van Eck van der Sluijs[1,2], W.J. Kop [3], & C.M. van der Feltz-Cornelis[1,2]

[1] Department Tranzo, Tilburg School of Social and Behavioral Sciences, Tilburg University, Tilburg, the Netherlands
[2] Clinical Centre of Excellence for Body, Mind and Health, GGz Breburg, Tilburg, the Netherlands
[3] Center of Research on Psychological and Somatic disorders (C*o*RPS), Tilburg University, Tilburg, the Netherlands

Supervisors:
Prof. Christina M. van der Feltz-Cornelis
Prof. Willem J. Kop
Lars de Vroege, MSc
Jonna F. van Eck van der Sluijs, MD, MSc

Name: Remona C.A.D. Kamp
ANR: 792073
Date: 30-06-2017
Word count: 8260
Word count abstract: 299

Abstract

Objectives: The factor structure of the Somatic Symptom Disorder–B Criteria Scale (SSD-12) will be determined in a clinical sample and compared with previous studies. Psychometric properties and cut-off points will be established so that the questionnaire can be optimally used in clinical practice.

Methods: The study sample consisted of 102 outpatients from an Expert Center on Psychosomatic Disorders (mean age 40.64 (*SD* = 14.06) years; 41.2% male). Patients completed the SSD-12 and SSD-12 scores were evaluated against a psychiatric interview-based checklist used as the gold standard for clinical Somatic Symptom and Related Disorders (SSRD). Patients also completed a set of questionnaires to establish convergent and divergent validity.

Results: The SSD-12 displayed a one-factor structure, explaining 69.9% of the common variance. The reliability of the SSD-12 total score was high (Cronbach's α = .91; Guttman's lambda^2 = .92). The SSD-12 displayed high correlations (range *r* = .55-.75) with questionnaires assessing illness anxiety (IAS and WI), reassurance (RQ), anxiety (GAD-7), and depression (PHQ-9). Moderate correlations (range *r* = .43-.48) were found between the SSD-12 and somatic symptoms (PSC-51 and PHQ-15). Patients with SSD had higher SSD-12 scores (27.78 ± 9.80) than patients without SSRD (20.76 ± 12.32; *p* = .015). When using a cut-off value of 15, the SSD-12 is sensitive (94.1%), but not very specific (35.3%) against the SSRD diagnosis according to the diagnostic gold standard.

Conclusions: The SSD-12 is a reliable instrument to measure SSRD in psychosomatic outpatients. Results also supported the convergent and divergent validity of the SSD-12. The questionnaire can be used as a screening tool for SSRD in patients with psychosomatic complaints, but there is a chance of a high number of false-positives. Additional research is needed to examine the sensitivity to change during interventions and the validity of the SSD-12 in populations with a low prevalence of SS(R)D.

Keywords: SSD = somatic symptom disorder, SSD-12 = the Somatic Symptom Disorder–B Criteria Scale, validity, psychometrics, diagnostic instrument.

The Somatic Symptom Disorder – B Criteria Scale (SSD-12) in a Dutch Clinical Sample: A validation study.

Approximately 80% of the general population shows at least one somatic symptom with at least mild impairment during the last week, while roughly one out of five individuals report severe somatization symptoms [1]. In the primary care setting medically unexplained symptoms (MUS) are among the most prevalent symptoms, with 49% of the patients reporting at least one MUS in one year [2]. In the fourth edition of the Diagnostic and Statistical Manual of Mental Disorders (text rev.; DSM-IV-TR) [3] the classification 'Somatoform Disorders' is defined by somatic symptoms in absence of a medical explanation for at least six months causing clinically significant distress or impairment in an important area of functioning. The negative criterion, the absence of a medical explanation, has been frequently criticized [4, 5], because it is often uncertain and difficult to determine whether a somatic symptom can or cannot be medically explained [6,7]. This controversy led to replacing the category Somatoform Disorders with a new category, named Somatic Symptom and Related Disorders (SSRD) in the fifth edition of the DSM (DSM-5) [5, 8]. In Somatic Symptom Disorder (SSD) a focus on dysfunctional coping with the somatic symptoms rather than on the cause of the symptoms (i.e. medically explained or unexplained) is key [5, 6, 9].

The diagnostic criteria for SSD in the DSM-5 consist of criteria A, B, and C [8]. To fulfill criterion A, one or more somatic symptoms have to be present, which have to be distressing or result in significant disruption of functioning [5, 8]. These symptoms can be either medically explained or unexplained [5]. The B criterion for SSD is the presence of abnormal, maladaptive, excessive, and disproportionate thoughts, feelings, and/or behaviors related to the somatic symptoms [5, 8]. At least one of these three psychological symptoms needs to be present to fulfill criterion B [8] and the more of this criterion is fulfilled, the more serious the disorder is. Criterion C includes that the symptoms must be persistent, typically for at least 6 months [8].

The current prevalence of SSD is unknown because of its recent introduction [8]. Prevalence estimates for somatoform disorders in the general population, the occupational health care setting, and the primary health care setting range from 10% to 21% [9-13]. However, the prevalence of SSD is expected to be lower than the prevalence of undifferentiated somatoform disorder [8], because the DSM criteria for an undifferentiated somatoform disorder were defined in such a way that a large proportion of the general population meets the criteria [14]. Though, the prevalence of SSD is expected to be higher than the more restrictive DSM-IV diagnosis somatization disorder (<1%) [8]. The prevalence of SSD in the general population is therefore estimated around 5% to 7% [8].

In clinical practice, diagnostic assessment by psychiatric consultation is considered a gold standard to establish a diagnosis. If no psychiatric consultation is possible, structured diagnostic interviews for DSM-Disorders such as the Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I) [15] and Mini International Neuropsychiatric Interview (MINI) [16] have a high level of consistency due to the standardization. On the other hand, they are inflexible and time consuming [17]. Modules for SSD have been included in the Structured Clinical Interview for DSM-5 (SCID-5), but the psychometric properties are not studied yet [18]. Axelsson et al. developed a structured diagnostic interview, the Health Preoccupation Diagnostic Interview (HPDI), which can be used to reliably diagnose SSD [19]. However, the Dutch translation of both interviews is not yet available [18, 19]. Since SSD is a new classification, establishing a self-report questionnaire to assess criterion B would be very useful in clinical practice to minimalize the risk of mislabeling a person with SSD [9, 20]. Furthermore, such an instrument would be very practical, although it may not be as valid as a structured interview or psychiatric evaluation [6].

There are several reliable and valid self-report questionnaires that effectively measure severity of somatoform disorders according to DSM-IV [17]. Examples of these measures are the Whiteley Index (WI) [21] and the Illness Attitude Scale (IAS) [22] for measuring the subcategory hypochondriasis and the

Physical Symptoms Checklist-51 (PSC-51) [23] and the Patient Health Questionnaire 15-item somatic scale (PHQ-15) [24] for measuring physical symptoms [25]. The PHQ-15 can be used as a screening tool for somatoform disorders, with a sensitivity of 80% and specificity of 59% in the primary care [6] till moderate validity, with reasonable sensitivity but limited efficiency, in the occupational health care setting [11]. The sensitivity and specificity of the WI and IAS against the DSM-IV diagnosis of hypochondriasis are high (≥95%) [17]. These findings indicate that these questionnaires are good predictors of the DSM-IV diagnosis of somatoform disorders. However, their applicability for the revised SSD diagnosis is questionable [9]. Nevertheless, the criterion A of the recently introduced SSD diagnosis (somatic symptoms that are either very distressing or result in significant disruption of functioning) could be adequately reflected by an instrument like the PHQ-15 [9, 24]. Furthermore, criterion C can be answered by the question if the symptoms are present for at least 6 months [27]. However, until recently there are no valid instruments to explicitly cover the B criterion of SSD [28], especially the behavioral aspect [9].

Toussaint et al. (2016) developed a self-report questionnaire called the SSD–B Criteria Scale (SSD-12), to assess criterion B of the SSD diagnosis. This questionnaire, based on the DSM-5 criteria, assesses the patients' perceptions of their symptom-related thoughts, feelings, and behaviors [9]. They validated the SSD-12 in a sample of German patients from a psychosomatic outpatient clinic and concluded that the SSD-12 is a valid tool for clinical and research purposes. In 2016 Kop et al. explored the associations of the Dutch SSD-12 with medical status and illness anxiety in the Dutch general population. They concluded that the SSD-12 is a valid tool in studies examining the general population [29]. The SSD-12 scores were highly correlated with illness-related anxiety measured with the IAS and were elevated in individuals with medical disorders [29].

Rationale

The study of Toussaint et al. (2016) focused on the construct validity and internal consistency of the SSD-12 [6, 9]. The present study aims to take into account other validity measures as well, such as discriminant validity by validating this instrument in Dutch patients from a psychosomatic outpatient clinic and compare them with patients from an outpatient clinic for anxiety and depression. Another novel aspect of the present study is that the SSD-12 score will be compared with a gold standard, to establish the predictive criterion and diagnostic validity. During a psychiatric consultation, a checklist will be used to structure the diagnostic assessment and draw conclusions about the diagnoses of SS(R)D.

Furthermore, this study aims to establish cut-off points for the SSD-12 to discriminate between participants with and without SS(R)D. These can be of great use in clinical practice. With regard to the factor structure, the convergent validity and the internal consistency of the SSD-12, this study aims to replicate the findings of Toussaint et al. (2016) and Kop et al. (2016) in a different sample. Taken together, these objectives contribute to enhanced insight in the structure of the SSD-12, its validity and its usage in a clinical setting.

To explore whether or not the SSD-12 specifically screens for SS(R)D and no other (mental) disorders, we also evaluated its discriminant validity calculating its correlation with questionnaires that do not include SS(R)D-like symptoms such as the Generalized Anxiety Disorder Scale (GAD-7) [30] and Patient Health Questionnaire-9 (PHQ-9) [31]. In addition, we explored the influence of physical health by taking into account chronic physical illness as factor that may influence SSD-12 scores.

Objectives

The first study objective is to explore the factor structure of the SSD-12 and to compare this with the structure found by Toussaint et al. (2016). Toussaint et al. (2016) found a three-factor model that corresponds to the three sub criteria of the B criterion: cognitive, affective, and behavioral aspects.

The second study objective is to establish psychometric properties of the SSD-12 questionnaire among Dutch patients in a Clinical Centre of Excellence. The third study objective is to establish predictive validity as compared with the gold standard, psychiatric evaluation. In addition, cut-off points for the SSD-12 will be established so that the SSD-12 can be used in clinical practice. This objective is explorative because there is no research done yet on cut-off points of the SSD-12.

The following specific hypotheses will be tested:

1. We expect that a three-factor structure will provide an optimal characterization of the SSD-12, consistent with the factor structure reported by Toussaint et al. (2016).

2. We hypothesize that the psychometric properties and validity of the SSD-12 will be good

2a. The internal consistency (reliability) as determined by a Cronbach's α is expected to be > .80.

2b. High correlations (r >.50 according to Cohen (1988)) between the SSD-12 and several other questionnaires that measure related constructs (e.g., illness anxiety) are expected, as signs of good convergent validity.

2c. Low to moderate correlations (r <.50 according to Cohen (1988)) between the SSD-12 and questionnaires that measure different disorders or constructs (e.g., anxiety and depression) are expected, as signs of good discriminant validity.

3. The SSD-12 had good criterion validity

3a. A higher mean SSD-12 score is expected in patients with SS(R)D compared with patients without SS(R)D, using psychiatric diagnosis for assessing criterion validity. These analyses will be conducted stratified for individuals with and without general chronic physical illness.

3b. The SSD-12 is a good instrument to identify individuals with a clinician-based diagnosis of SS(R)D with an estimated sensitivity of at least 80% and a specificity of at least 60% in an outpatient sample attending a psychosomatic specialty center.

Method

Study design

A cross-sectional study design using consecutive patients from two outpatient clinics was used in this validation study.

Setting

Patients were recruited at two outpatient clinics that are both part of GGz Breburg, situated in Tilburg, the Netherlands:

- The Clinical Centre of Excellence for Body, Mind and Health (Dutch abbreviation: CLGG) provides treatment to patients referred by a medical specialist, psychiatrist or general practitioner with combined physical (explained and unexplained) and psychological symptoms. Referral criteria for CLGG are somatic symptoms as primary problem in combination with significant psychological suffering. It was expected that many of these patients will suffer from SSRD.
- The outpatient Anxiety and Depression clinic, that provides treatment to patients with mood disorders. These patients are referred by the general practitioner to the outpatient clinic Anxiety and Depression in case of an anxiety and/or depression disorder. It was expected that many of these patients will suffer from anxiety or depressive symptoms that may also be present in the CLGG patients, but in general did not suffer from SSRD.

Participants

Newly registered patients from CLGG and from the outpatient Anxiety and Depression clinic were included in this study. Inclusion started in September 2016 and ended in March 2017.

Eligible patients

Inclusion criteria

In order to be eligible to participate in this study, a participant had to be a newly registered patient from CLGG or from the outpatient clinic Anxiety and Depression. Patients of all ages were included.

Exclusion criteria

The exclusion criteria of this study were an active suicide risk, current or recent (< 3 months) history of alcohol or drug abuse, illiteracy, and insufficient knowledge of Dutch language.

Ethical issues

Every patient had to provide informed consent in order to use their diagnostic and treatment data for scientific research purposes on anonymous basis. Data of patients who did not consent with this during intake were not included in this study. This protocol was evaluated and approved by the scientific committee of GGz Breburg (CWO2016-14).

Measurements

Consecutive newly registered patients from CLGG and from the outpatient Anxiety and Depression clinic took part in Routine Outcome Monitoring (ROM), psycho-diagnostic assessment, anamnesis and psychiatric consultation by psychiatrists trained in the SSD criteria by use of a checklist [33, 34]. According to standard care, participants first fill in the ROM, after which they have several appointments for the anamnesis, the psychiatric evaluation and the psycho-diagnostic assessment. A subset of the questionnaires used in the ROM, psycho-diagnostic assessment, and anamnesis is used for the purposes of this project. Demographic variables such as age, sex, and education level (classified using the method described by Verhage [35]) were obtained during intake. The following questionnaires and checklists assessed in ROM, in the psycho-diagnostic assessment and in the anamnesis are part of the regular intake procedure of CLGG and are used for this study.

The main outcome measure of this study is the SSD-12, developed to quantify the B criterion for SSD [9]. For convergent validity the following constructs were evaluated: illness anxiety (measured by the WI and the IAS), reassurance (measured by the Reassurance Questionnaire; RQ) [36], and physical symptoms (measured by the PSC-51 and the PHQ-15). Discriminant validity was investigated using the following measures: anxiety and depression (measured by the GAD-7 and PHQ-9 respectively).

Somatic Symptom Disorder–B Criteria Scale (SSD-12)

The SSD-12 [9] consists of 12 items. Responders rate how frequently they experienced each cognition, emotion, or behavior on a 5-point Likert scale, ranging from never (0) to very often (4; see Appendix C). The total score ranges from 0 (minimal) to 48 (severe), where higher scores reflect higher levels of the B criterion of SSD. Toussaint et al. (2016) found that the SSD-12 has good item characteristics and excellent reliability (Cronbach's α = .95). A three-factor structure was found, that reflects the three psychological aspects of the B criterion of the DSM-5: cognitive, affective, and behavioral.

Whiteley Index (WI)

Aspects of illness anxiety were assessed by the WI [21]. The Whiteley Index consisting of 14 items is a self-report questionnaire useful in screening for illness anxiety [21]. For the present study, we used the validated Dutch version with binary scoring (yes = 1, no = 0), which can be summed up to an overall score ranging from 0 to 14 [37, 38]. The optimal cut-off point for identifying a diagnosis of severe illness anxiety is 5 [39]. The discriminative power was very high. Reliability (Cronbach's α ranged from .76-.80), stability, concurrent and discriminative validity are adequate [37].

Illness Attitude Scale (IAS)

The IAS [22] is a self-rating questionnaire designed to measure attitudes, fears and beliefs associated with hypochondriasis and abnormal illness behavior [22]. The 29 items are rated on a 5-point Likert scale ranging from no (0) to most of the time (4). The questions 23-25 have divergent response

options, but these are also scored on a 5-point Likert scale (0-4). Item 22 and 26 are not used in the total score, which ranges from 0 to 108. The questionnaire originally consisted of 9 subscales. However, different studies have not consistently supported the suggested nine factors [37-41]. Validation studies found only 2 subscales namely health anxiety and illness behavior, which did not cover all the items [37, 38]. Hadjistavropoulos, Frombach, and Asmundson (1998) compared different models and found the greatest support for a five-factor model, but they did not include a two factor model in their study. According to Sirri, Grandi, and Fava (2008) is the questionnaire developed to be of high clinical value and not to adhere to psychometric criteria of item homogeneity. Based on this background, Hedman et al. (2015) suggested to take the whole questionnaire into account. They found that a cut-off point of 47 on the IAS yielded the best results with a sensitivity of 96% and specificity of 95%. The scale has been shown to have high test-retest reliability (*r* = .89) [41].

Reassurance Questionnaire (RQ)

The RQ [36] is a self-report questionnaire designed to assess the extent to which patients feel reassured by their physician(s). The RQ consists of 10 items, e.g. “Do you think your physician is keeping something from you?”. Answers were scored on a 4-point Likert scale ranging from no (0) to mostly (4). The sum score has a minimum of 0 and a maximum of 40. High scores reflect poor reassurability. For this study, reassurability is defined as a score of 9 or less, based on the Dutch norms. The internal consistency was moderate to high (Cronbach’s α ranging from .66 to .88) and the test–retest reliability was high (*r* = .85) [36]. The convergent validity of the RQ was satisfactory to good [36].

Physical Symptoms Checklist-51 (PSC-51)

Physical symptomatology was assessed using the PSC-51 [23]. The PSC-51 is a checklist consisting of 51 items, representing all the somatic symptoms of the DSM-IV [23]. The question for every symptom is, if the participant has experienced the symptom during the last week. It takes 10-15 minutes to complete the questionnaire. For the present study, we used the version with dichotomous scoring

(regularly/often = 1, never/sometimes = 0). The total score ranges from 0 to 51. It includes five subscales: Urogenital symptoms (9 items; score range 0-9), Gastrointestinal symptoms (13 items; range 0-13), Pain (8 items; range 0-8), Neurological symptoms (11 items; range 0-11), and Autonomous symptoms (10 items; range 0-10). A cut-off score of 11 (men) or 13 (women) will be used [42]. The psychometric properties of the scale have not investigated. However, in a study with primary care patients, the internal consistency was good (Cronbach's α = .88) [43].

Patient Health Questionnaire 15-item somatic scale (PHQ-15)

The PHQ-15 [24] is the somatic symptom subscale derived from the PHQ. The PHQ-15 scale assesses the presence and severity of common somatic symptoms within the last 4 weeks using 15 items [24]. Each item representing one somatic symptom or symptom cluster. Thirteen items are derived from the PHQ somatic symptom module, in which patients are asked to rate the severity of each symptom with response options ranging from not bothered at all (0) to bothered a lot (2). Two additional physical symptoms, feeling tired or having little energy, and trouble sleeping, are derived from the PHQ depression module. In which the question is how often the participant experienced these symptoms in the last 2 weeks, with response options ranging from not at all (0) to nearly every day (2). The total score ranges from 0 to 30, with higher scores indicating a higher self-rated symptom burden. A PHQ-15 score of 9 represents the cut-off point for severe somatic symptoms, which has been used as an indication of a somatoform disorder using DSM-IV criteria [24].

Generalized Anxiety Disorder Scale (GAD-7)

The GAD-7 [30] was used to measure anxiety. The GAD-7 is the short version of the 13-item GAD, consisting of the 7 items with the highest correlation with the 13-item scale score [30]. The self-report questionnaire is used to identify probable cases of GAD and other anxiety disorders. It asks patients how often, during the last 2 weeks, they experienced each of the 7 symptoms. Response options range from not at all (0) to nearly every day (3). The total score ranges from 0 (minimal) to 21

(severe). There is one additional question, about the disruption of functioning. The questionnaire has good reliability and good criterion, construct, factorial, and procedural validity [30]. Cut-off points of 5, 10, and 15 represent mild, moderate, and severe levels of anxiety, respectively [30]. So a cut-off score of 10 can be used to indicate at least moderate anxiety. This cut-off point optimized the sensitivity (89%) and specificity (82%) [30].

Patient Health Questionnaire-9 (PHQ-9)

PHQ-9 [31] was used to assess depressive symptoms. The PHQ is the self-report version of the PRIME-MD diagnostic instrument for common mental disorders [31]. The PHQ-9 is the depression module consisting of 9-items, used as a continuous measure of depression severity. The 9 items represent the 9 DSM-IV criteria for major depression within the last 2 weeks, these criteria correspond with the DSM-5 criteria for depression [9]. Response options are ranging from not at all (0) to nearly every day (3). The total score ranges from 0 (minimal) to 27 (severe). There is one additional question, about the disruption of functioning. Cut-off points of 5, 10, 15, and 20 represent mild, moderate, moderately severe, and severe depression [31]. Therefore, a cut-off score of 10 can be used to indicate at least moderate depression [31].

Central Bureau of Statistics (CBS) list

The CBS list [44] was used to measure chronic physical illness and was part of the anamnesis. The CBS list is a checklist with 28 chronic diseases [44]. It asks patients to mark the chronic diseases like diabetes the patient is suffering from and to fill in when the disease is diagnosed (which month and year). The data of the list is combined with information found in the patients files about physical illness.

Psychiatric evaluation for SSD (Gold Standard)

A checklist was made to assess the diagnostic criteria of SSD following the DSM 5 (see Appendix D). One of three psychiatrists at the outpatient clinic (CFC, JvE, AF) performed the psychiatric evaluation for this study and used this checklist to draw conclusions about the SSD diagnosis. Besides the

psychiatrists filled in on the checklist if the patient was suffering from a physical illness or other SSRD diagnosis as well. The checklist is filled in afterwards.

Statistical Power and sample size calculation

Sample size calculations were done using G*Power 3.19.2 and based on previous findings from the literature. The power analysis was performed for the main hypothesis of interest (hypothesis 2). Therefore, we used G*Power 3.1.9.2 'Correlate bivariate normal model' to calculate the sample size. In the study of Toussaint et al. (2016) correlations of a wide variety of sizes are found. Therefore, the effect size was estimated at a medium level (r = .3). To generate sufficient power (equal to .80) with a medium effect size (r = .3) according to Cohen's guidelines [45] and an alpha level of .05, a total sample size of 84 participants is needed. The expectation is that 5% of the patients will have unusable data which results in a required sample size of 89 patients.

Statistical methods

To analyze the data of this study the statistical software package Statistical Package for the Social Sciences (SPSS) version 22.0 was used [46]. When assumptions of the analysis were violated, non-parametric tests were used. There were some potential confounding variables that had to be taken into account. Therefore, demographic variables (i.e., age, sex, and education) and chronic physical illness, derived from the Central Bureau of Statistics (CBS) list and patient files, were included in the analysis. Missing variables were not imputed and resulted in listwise exclusion of cases in multivariable models. Parallel analysis was conducted by means of the free software program FACTOR version 10.3.01 [47].

Descriptive data

Several categorical and continuous variables were used to describe the sample. The original plan of this project was to compare patients with SSD who were evaluated and treated at the CLGG with patients attending the outpatient Anxiety and Depression clinic, but due to a lack of inclusion of patients of the outpatient Anxiety and Depression clinic the SSD-12 data will be compared for patients with vs.

without SS(R)D based on the gold standard psychiatric evaluation. Education level is obtained following Verhage coding in 7 categories in which the levels 1, 2, 3, and 4 are considered as low education, level 5 is considered as middle education and levels 6 and 7 are considered as high education [35]. The education level was recoded to a low, middle and high education level according to these guidelines. Frequency of the categorical variables (i.e., sex and educational level) were determined. In addition, the mean and standard deviation of continuous variables (i.e., age) were determined. To examine if there are differences between the responders and the non-responders and between patients with and without SSD, Chi-square tests for the categorical variables and independent t-tests for the continuous variables were used.

Examining the hypotheses

Factor structure

The factorial structure of a questionnaire is thought to be stable over different patient groups. Therefore, the factor structure of the SSD-12 was explored in the total study group. Prior research supports a one-factor model [29] and a three-factor model that corresponds to the three sub criteria of the B criterion: cognitive, affective, and behavioral aspects [9]. The factor structure of the SSD-12 was explored using Exploratory factor analysis (EFA) and Confirmative Factor Analysis (CFA) for one and three factors based on the breakpoint of eigenvalues and statistical methods using parallel analysis [48] in combination with minimum rank factor analysis (MRFA) [49, 50].

Parallel analysis is considered to be one of the best methods for determining the number of factors [47]. Parallel analysis compares the percentage of variance explained by the factors with the percentage of variance explained by the same number of factors resulting from randomly generated data. The method generates random correlation matrices by means of permutation of the raw data. In this study 500 random correlation matrices were generated and analyzed by means of MRFA.

MRFA is the only factor method that computes the proportion of variance explained by each factor [50]. The MRFA method divides observed variables into common and unique parts [47]. MRFA maximizes the item communalities given the number of factors and gives a statistically correct reduced correlation matrix [49]. The method minimizes the common variance that is ignored when not all factors are retained in the model, this allows interpretation of the proportion of common variance explained by each of the hypothesized factors [47]. Factors were considered meaningful if the percentage of variance explained by these factors exceeded the percentage of variance explained by the random-data factors. We evaluated the data-model fit using the Root Mean Square Residual (RMSR) and Kelley's criterion [51, 52], the model cannot fit the data adequately if the RMSR is larger than Kelley's criterion (.0995).

Psychometric properties

Psychometric properties were explored for the total group. To explore the psychometric properties of the SSD-12 thoroughly, we calculated reliability and validity estimates in SPSS. Since Cronbach's α is a lower bound of the reliability [53-55] we also provided Guttman's lambda^2, a greater lower bound of the reliability [56] that lies closer to the true reliability than Cronbach's α [55].

To evaluate the construct validity, several correlations were calculated and evaluated. To determine the convergent validity bivariate Pearson (product-moment) correlations between the SSD-12 and several other questionnaires that measure related constructs (WI, IAS, RQ, PSC-51, and PHQ-15) were calculated. To determine the discriminant validity correlations between the SSD-12 and questionnaires that measure different disorders or constructs (GAD-7 and PHQ-9) were evaluated.

Criterion validity and cut-off points

To determine the criterion validity, independent t-tests were used comparing mean scores on the SSD-12 for patients with versus without SSD and patients with and without SSRD based on clinical psychiatric diagnosis. These analyses were conducted stratified for individuals with and without general chronic physical illness. Cohen's *d* was used as a measure for the effect size. We also used a one-way

between-groups analysis of covariance (ANCOVA) to examine differences in the mean scores on the SSD-12 for patients with and without SSD and SSRD when adjusting for sex, age, chronic physical illness, and education level.

For an ideal cut-off point the sensitivity and specificity need to be as high as possible and the false positives and false negatives need to be as low as possible (see Appendix A). Sensitivity, specificity, false negative rate, false positive rate, negative predictive value, positive predictive value, efficiency and Youden's J were calculated using formulas presented in Appendix A. To determine an ideal cut-off point the sensitivity, specificity, predictive values, and efficiency will be calculated for different cut-off points. Youden's J is calculated, the highest Youden's J is the value with an optimal balance between sensitivity and specificity [57]. This value indicates the optimal cut-off point. In addition, receiver operating characteristic (ROC) and an area under the curve (AUC) were also calculated to find the optimal cut-off point. These measures assess the diagnostic performance and validity. The curve plots the sensitivity against the false positive rate (1-specificity). The optimal cut-off point is the value for which the distance between the upper left hand corner and the point on ROC curve is minimized.

Results

Participants

A total of 111 patients were approached to fill out the SSD-12 between September 2016 and March 2017. Of those 111 patients, 103 (92.8%) returned the questionnaire. One patient was not included because of not showing up for the psychiatric interview. The present analyses therefore focuses on the data of 102 patients for whom we obtained both a SSD-12 score and a diagnostic checklist for SSD; this is 91.9% of the patients who were approached. No differences in patients' sex, marital status, age and level of education were found between responders and non-responders. Of the 102 patients 3 patients (2.9%) were patients of the outpatient Anxiety and Depression clinic.

Sample characteristics

Table 1 shows the demographic characteristics of the total sample. The total sample (N = 102) existed of 42 (41.2%) male participants. The mean age was 40.64 years (*SD* = 14.06). Clinical SSD by psychiatric evaluation was present in 64 (62.7%) patients and the remaining 38 (37.3%) patients did not meet diagnostic criteria for SSD (21 of the 38 had another somatic symptom disorder (see below for details) and 17 of the 38 had no SSRD). Data for SSRD were 85 (83.3%) and 17 (16.7%), respectively. A low level education was observed in 30 (29.7%) patients. A medical condition was observed in 61 (59.8%) patients. As shown in Table 1 the SSD group did not differ from the non-SSD group in demographic characteristics. Table E is attached as Appendix E to show the SSRD group also did not differ from the non-SSRD group in demographic characteristics.

Factor analyses and Reliability

The inter-item correlations range from .19 to .79 (see Table 2), only item 7 has weak correlations ($r < .30$) with the other items. Parallel analysis of the EFA suggested a one-factor model. The one-factor model fit the data in an acceptable way (RMSR = .077; Kelley's criterion = .0995). The one-factor model explained 52.2% of the total variance and 69.9% of the common variance (MRFA). Table 3 shows the standardized factor loadings for the one-factor model. All standardized item loadings were greater than .415.

The model fit of the three-factor model was acceptable (RMSR = .028; Kelley's criterion = .0995). The three-factor model explained 69.9% of the total variance and 88.1% of the common variance (MRFA). Table 4 shows the standardized factor loadings for the three-factor model. The loadings for item 10 are not shown, because the highest loading for item 10 was .292 on factor 2 and only loadings of .300 or higher are shown. The items 8, 11, and 12 loaded on the first factor, the items 1, 2, 3, 4, 5, 6, 9, and 11 loaded on the second factor and item 7 loaded on the third factor. These items do not match the items of the three subscales according to the study of Toussaint et al. (2016). The reliability of the

three subscales suggested by Toussaint et al. (2016) was sufficient in the present sample (see Table 5). In the current study the Cronbach's α for the SSD-12 total score was .91. The Guttman's lambda[2] was .92 (see Table 5).

Convergent and discriminant validity

Table 6 shows the mean scores and ranges of the questionnaires. Table 7 shows the correlations between the SSD-12 and the other questionnaires. The SSD-12 has high correlations (r > .50) with the WI, IAS, RQ, GAD-7 and PHQ-9 and moderate correlations (r = .30-.50) with the PSC-51 and PHQ-15.

SS(R)D

Table 8 shows the amount of patients with and without a SSRD diagnosis and the average score on the SSD-12. In the total sample (N = 102), 64 participants fulfilled the classification criteria for SSD according to the checklist (62.7%) while 38 participants did not fulfil diagnostic criteria and were not classified with SSD (37.7%). 21 participants were classified with another SSRD diagnosis (20.6%), in total 85 participants fulfilled the classification criteria for SSRD (83.3%). Illness Anxiety Disorder was the next most frequent SSRD after SSD with 9 participants (8.8%), followed by Conversion Disorder with 7 participants (6.9%) and Unspecified SSRD (n = 5, 4.9%). 17 participants did not fulfil the diagnostic criteria and were not classified with SSRD according to the DSM (16.7%).

Criterion validity

The mean SSD-12 score in the total study sample was 26.92 (SD = 10.67; range = 0–46). The mean scores of the SSD-12 were significantly higher (p = .008) in the SSRD group (M = 28.15, SD = 9.94, range = 12–46) than the no-SSRD group (M = 20.76, SD = 12.32; range = 0–44). The magnitude of the differences in the means was medium (d = 0.71). After adjusting for participants' age, sex, chronic physical illness, and education level, there was still a significant difference between the two groups on

the SSD-12 score (F (1,99) = .527, p = .001, partial eta squared = .10). When we conducted stratified analyses for individuals with and without general chronic physical illness the difference between the SSD-12 means in the SSRD group (M = 29.19, SD = 10.18) and the no-SSRD group (M = 18.71, SD = 12.34) was only significant in the group with chronic physical illness (p = .002, d = 0.98) and not in the group without chronic physical illness (p = .547, d = -0.36). A significant interaction effect was found (p = .046; see Figure 1).

Results for SSD were similar as those of SSRD, but the effects were not as strong. The difference between the means in the SSD group (M = 27.78, SD = 9.80, range = 12–45) and the no-SSD group (M = 25.47, SD = 12.00; range = 0–46) was not significant (p = .293). The magnitude of the differences in the means was small (d = 0.22). When we conduct stratified analyses for individuals with and without general chronic physical illness the SSD-12 means were significantly higher in the SSD group (M = 29.43, SD = 10.04) than the no-SSD group (M = 21.76, SD = 12.64) in the group with chronic physical illness (p = .012, d = 0.70) and not in the group without chronic physical illness (p = .094, d = -0.54). A significant interaction effect was found (p = .004; see Figure 1). An ANCOVA was conducted to compare the difference between the means of the SSD-12 score in the SSD group versus the no-SSD group. After adjusting for participants' age, sex, chronic physical illness, and education level, there was no significant difference between the two groups on the SSD-12 score (F (1,99) = .185, p = .168, partial eta squared = .02).

When we compare the means of the SSD group (M = 27.78, SD = 9.80, range = 12–45) and the no-SSRD group (M = 20.76, SD = 12.32; range = 0–44) there was a significant difference (p = .015). The magnitude of the differences in the means was medium (d = 0.68). The difference between the means in the Illness Anxiety Disorder group (M = 38.33, SD = 8.17) and the no-SSRD group (M = 20.76, SD = 12.32) was significant (p = .293) as well. The magnitude of the differences in the means was large (d = 1.58). However the difference between the means in the Conversion Disorder group (M = 21.71, SD = 4.99)

and the no-SSRD group (*M* = 20.76, *SD* = 12.32) was not significant (*p* = .847). The magnitude of the differences in the means was very small (*d* = 0.09). The difference between the means in the Unspecified SSRD group (*M* = 23.60, *SD* = 7.60) and the no-SSRD group (*M* = 20.76, *SD* = 12.32) was also not significant (*p* = .634). The magnitude of the differences in the means was small (*d* = 0.25).

Cut-off points

Table 9a shows the sensitivity, the specificity, Youden's J and the predictive values for both positive and negative test results (positive predictive value (PPV), negative predictive value (NPV)) and the efficiency of summed SSD-12 scores when the SSD-12 scores are used to predict if the patient has a SSD diagnosis according to the checklist. Table 9b shows these values when the SSD-12 scores are used to predict if the patient has a SSRD diagnosis according to the checklist. These were calculated for summed scores from 0 to 46.

The highest efficiency (67.6%) of all possible cut-off points, when the SSD-12 scores are used to predict if the patient has a SSD diagnosis according to the checklist, was found at cut-off point 11 and 12. However, at cut-off point 11 and 12 the specificity was unacceptably low with a value of 13.2%. The highest Youden's J of all possible cut-off points when the SSD-12 scores are used to predict if the patient has a SSD diagnosis according to the checklist was .15 at cut-off point 16. At cut-off point 16 the specificity was also unacceptably low with a value of 21.1%.

The highest Youden's J of all possible cut-off points when the SSD-12 scores are used to predict if the patient has a SSRD diagnosis according to the checklist was .29 at cut-off point 11, 12, and 15. The highest efficiency (87.6%) of all possible cut-off points was also found at cut-off point 11 and 12. However, at cut-off point 11 and 12 the specificity was unacceptably low with a value of 29.4%. At cut-off point 15 the specificity was also unacceptably low with a value of 35.3%. A cut-off point of 22

resulted in sensitivity of 65.9% and specificity of 58.8%, PPV of 88.9%, and NPV of 25.6%. At cut-off point 22, efficiency was 64.7% and Youden's J was .25.

The ROC curves, calculated for the SSD-12 summed score vs. the SSRD and the SSD diagnosis are shown in Figure 2. The calculated area under the curve (AUC) of the SSD-12 versus the checklist for SSRD diagnosis was .66 (*SE* = .08; 95% confidence interval (CI) = .51 - .81). These values are significant (*p* = .039). The calculated area under the curve (AUC) of the SSD-12 versus the checklist for SSD diagnosis was 0.54 (*SE* = 0.06; 95% CI = 0.42 - 0.66). These values are not significant.

Discussion

Key results

In this study, the SSD-12 is validated as a screening instrument for SS(R)D in a Dutch psychosomatic outpatient clinic. The results show that a one factor model fits the data. The reliability of the SSD-12 total score is sufficient. The results show that the SSD-12 is sensitive, but not very specific. There was a significant difference between the SSD group and the no-SSRD group on the SSD-12 mean scores.

Factor structure

The results show that the fit of the one factor model is good. We also checked if a three factor model fits the data, because of previous findings [9]. The fit of the three factor model is also good. However, when we look at these items closely we notice that it is hard to explain these factors with constructs. Besides the items of the three-factor model did not match with the items of the subscales following the sub criteria of the B criterion of the DSM-5 according to Toussaint et al. (2016).To conclude, we did not replicate the findings by Toussaint et al. (2016). However, the reliability of the three subscales following the sub criteria according to the study of Toussaint et al. (2016) is sufficient. In contrast to our hypothesis, the results can be explained by the one factor model justifying the use of a general SSD-12 sum score as consistent with the study of Kop et al. (2016).

Psychometric properties

Reliability

Consistent with our hypothesis, the reliability of the SSD-12 total score is sufficient with a good internal consistency. This result is corresponding with results of previous studies [9, 29].

Convergent validity

The SSD-12 has high correlations (r > .50) with the WI, IAS, and RQ indicating a good convergent validity. This is consistent with our hypothesis and with the results of previous studies [9, 29]. In contrast to our hypothesis, the SSD-12 has a moderate (r = .30-.50) correlation with the PSC-51 and PHQ-15. However, in a previous study a moderate correlation between the SSD-12 and PHQ-15 was found as well [9]. These questionnaires measure somatic symptoms. The moderate correlation might be explained by the high prevalence of chronic physical illness in the sample. Patients without SSD but with chronic physical illness score high on these questionnaires as well but will score low on the SSD-12.

Discriminant validity

Low to moderate correlations (r < .50) between the SSD-12 and questionnaires that measure different disorders or constructs (e.g. anxiety and depression) were expected as signs of good discriminant validity, but high correlations (r > .50) are found in this study. Somatization is associated with coexisting psychiatric disorders, especially depressive and anxiety disorders [58]. 58% of somatizing patients and 26% of the patients with a somatoform disorder also meet the criteria for an anxiety or depressive disorder [13, 59]. Furthermore the PHQ-9 might overestimate the amount of patients with a depressive disorder, because the PHQ-9 tends to place individuals in a more severe category [60, 61], this especially holds true for patients with somatic symptoms as found in research with patients with diabetes and traumatic brain injury [62, 63]. This might also apply to patients with somatic symptoms of SS(R)D, which might explain the high correlation. Moreover, the PHQ-9 holds items about somatic

symptoms of depression as well [31], patients with SSD and without depression will score on these items as well. All these factors can contribute to the high correlations found.

However the findings do not coincide with the findings of the study of Toussaint et al. (2016), while the sample of the study of Toussaint et al. (2016) and the sample of this study both consisted of patients of a psychosomatic outpatient clinic. The patients in the study of Toussaint et al. (2016) were diagnosed with a main diagnosis by ICD-10 of depression, anxiety disorder, eating disorder, or somatoform disorder. In this study are, next to the three patients of the outpatient clinic for anxiety and depression, only patients of CLGG included. Patients of CLGG have SSD or a related disorder as the primary reason for seeking treatment [64]. Perhaps our sample consists of patients with more severe complaints and more comorbid disorders given the high complexity and high levels of anxiety and depression in patients of CLGG found in a previous study [64]. On top of that in the study of Toussaint et al. (2016) the German version of the SSD-12 is used, while we used the Dutch version. The translation of the questionnaire from German to Dutch could also have contributed to differences in the findings between this study and the study of Toussaint et al. (2016).

Criterion validity

The mean scores of the SSD-12 did not differ significantly between the SSD group and the no-SSD group. However the means of the SSD group and the no-SSD group did differ in patients with chronic physical illness. We also explored the difference between patients with and without SSRD on the SSD-12 because of the high prevalence of SSRD in the sample and the many similarities between SSD and SSRD [8]. Patients in the SSD and the SSRD group had a significantly higher mean score on the SSD-12 than patients in the no-SSRD group. With regard to our hypothesis, the SSD-12 can differentiate between patients with and without SSRD and patients with SSD and without SSRD. Besides the SSD-12 can differentiate between patients with and without SSD when the patients have chronic physical illness. Somatization occurs in patients with medical disease, somatizing patients even have a higher

level of chronic physical illness than patients who do not somatize [65]. Patients with SSRD also have high scores on the SSD-12, making it difficult to differentiate on the basis of the SSD-12 between SSD and no-SSD in a sample with a high prevalence of SSRD in the no-SSD group.

Cut-off points

Different possible cut-off points are explored. The results show that the SSD-12 is sensitive, which means that patients with SS(R)D are identified. However the SSD-12 is not very specific, resulting in a lot of patients without SS(R)D having a false-positive test result. When the SSD-12 scores are used to predict if the patient has a SSD diagnosis according to the checklist, the highest Youden's J of all possible cut-off points was found at cut-off point 16. However, at these cut-off points the specificity was unacceptably low (21.1%). When the SSD-12 scores are used to predict if the patient has a SSRD diagnosis according to the checklist, the highest Youden's J of all possible cut-off points was found at cut-off point 11, 12, and 15. However, at these cut-off points the specificity was also unacceptably low (range 29.4%-35.3%). Using one of these cut-off points will result in a lot of false-positive test results. When the SSD-12 scores are used to predict if the patient has a SSRD diagnosis according to the checklist, there was a balance between the sensitivity of 65.9% and the specificity of 58.8% found at cut-off point 22. This results in less false-positives but more false-negatives compared with lower cut-off points. The optimal cut-off point clinicians want to select depends heavily on the sample and the setting.

Limitations

We planned to take into account discriminant validity by validating this instrument in Dutch patients from a psychosomatic outpatient specialty clinic and compare them with patients from an outpatient clinic for anxiety and depression. The inclusion of patients from the outpatient clinic Anxiety and Depression did not work out the way we planned. Inclusion of patients from the outpatient clinic

Anxiety and Depression would have resulted in more patients without SSRD and different results. As expected, patients with SSRD score high on the SSD-12.

The translation of the questionnaire from German to Dutch could have contributed to differences in the findings between this study and the study of Toussaint et al. (2016). The German study group of the study of Toussaint et al. (2016) made an English SSD-12 and Kop et al. (2016) translated the English questionnaire to Dutch. Both also used back translation.

Including patients of CLGG can bias the results, because CLGG is a clinical Centre of Excellence which can lead to including patients with more severe symptoms than in an average clinic. Moreover there is a risk of selection bias, because of the exclusion of patients with an active suicide risk, alcohol or drugs abuse, illiteracy, and insufficient knowledge of Dutch language. Therefore, caution should be exercised when generalizing these results to other patient groups. Another shortcoming of this study is the small no SSD-group. The no SSD-group had only 38 participants, while both groups (SSD and no SSD) required 42 participants partly caused by the low prevalence of patients of the outpatient clinic Anxiety and Depression in this sample.

Blinding is not guaranteed because clinicians conducting the checklist ratings were not blinded to the SSD-12 and other questionnaire scores. Besides not every patient filled in each questionnaire for varying reasons, resulting in different sample sizes for each questionnaire. For the RQ we only have data for the moiety of the sample, this has probably resulted in bias as well.

Generalisability

The sample, consisting of patients from a Dutch psychosomatic outpatient clinic and only a few patients of the outpatient clinic Anxiety and Depression, is not representative of the general population.

Therefore, the results can only be extrapolated to psychosomatic outpatient clinics in the Netherlands. Whether findings are similar in other populations should be further studied.

Future research

Future research is likely to proceed in larger samples and in samples with patients of different patient groups. Future research should assess the SSD-12 in other medical settings. It might be meaningful to establish different cut off points for different patient groups to simplify the use of the SSD-12 in clinical practice. Furthermore it might be important to explore the use of the SSD-12 as screener in primary care. It might be meaningful to find out if the questionnaire is a predictor of SSRD or if the questionnaire is a specific predictor for SSD.

Conclusion

To conclude, the SSD-12 is a reliable instrument to measure SSRD in a psychosomatic outpatient clinic. The optimal cut-off point depends heavily on the sample and the setting. When few false-negatives are desirable a cut-off point 15 is recommended. However, using cut-off point 15 results in a lot of false-positive test results. When a balance between the sensitivity and specificity is eligible then cut-off point 22 is recommended. Using this cut-off point will result in less false-positives but more false-negatives compared with cut-off point 15. The questionnaire can be used as a screener for SSRD in a psychosomatic outpatient clinic, keeping in mind the high amount of false-positives. Future studies focusing on the SSD-12 should be conducted, especially in populations with a low prevalence of SS(R)D to draw conclusions about the use of the SSD-12 in those populations. Future research should assess the SSD-12 in different patient groups and larger samples to establish different cut off points for each patient group. Furthermore it is important for clinical purposes to study if the SSD-12 is a screener for SSRD or for SSD.

Acknowledgments

We thank the team of CLGG for their contribution to this study, in particular Aziza Foruz (psychiatrist), Adriaan Lemstra (psychiatrist), Keyla Dubero (physician) and Tom Roosen (physician) for filling in the checklists. We would like to thank Iris Koppenol (psychologist) for her help with obtaining data from patients files.

References

1. Hiller W, Rief W, Brahler E. Somatization in the population: from mild bodily misperceptions to disabling symptoms. Social Psychiatry and Psychiatric Epidemiology 2006;41:704–712
2. Haller H, Cramer H, Lauche R, Dobos G. Somatoform disorders and medically unexplained symptoms in primary care. Deutsches Ärzteblatt International 2015;112:279-287.
3. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, 4th text revision ed. (DSM-IV). Washington, DC: American Psychiatric Association;2000. p. 553-557.
4. Sykes R. Somatoform disorder and the DSM-V Workgroup's interim proposals: two central issues. Psychosomatics 2012;53(4):334-338.
5. 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. Annual Review of Clinical Psychology 2014;10(1):339-367.
6. Barsky, AJ. Assessing the New DSM-5 Diagnosis of Somatic Symptom Disorder. Psychosomatic medicine 2016;78(1):2-4.
7. Kirmayer LJ, Groleau D, Looper KJ, Dao MD. Explaining medically unexplained symptoms. The Canadian journal of psychiatry 2004;49(10):663-672.
8. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders (DSM-5®): American Psychiatric Pub;2013. p. 309-327.
9. Toussaint A, Murray AM, Voigt K, Herzog A, Gierk B, Kroenke K, Rief W, Henningsen P, Löwe B. Development and Validation of the Somatic Symptom Disorder–B Criteria Scale (SSD-12). Psychosomatic medicine 2016;78(1):5-12.
10. Hilderink PH, Collard R, Rosmalen JGM, Voshaar RO. Prevalence of somatoform disorders and medically unexplained symptoms in old age populations in comparison with younger age groups: a systematic review. Ageing research reviews 2013;12(1):151-156.

11. Vroege L de, Hoedeman R, Nuyen J, Sijtsma K, Feltz-Cornelis CM van der. Validation of the PHQ-15 for somatoform disorder in the occupational health care setting. Journal of occupational rehabilitation 2012;22(1):51-58.
12. Hoedeman R, Krol B, Blankenstein N, Koopmans PC, Groothoff JW. Severe MUPS in a sick-listed population: a cross-sectional study on prevalence, recognition, psychiatric co-morbidity and impairment. BMC Public Health 2009;9(1):1.
13. Waal MW de, Arnold IA, Eekhof JA, Van Hemert AM. Somatoform disorders in general practice. The British Journal of Psychiatry 2004;184(6):470-476.
14. Dimsdale JE, Creed F, Escobar J, Sharpe M, Wulsin L, Barsky A, Lee S, Irwin MR, Levenson J. Somatic symptom disorder: an important change in DSM. Journal of psychosomatic research 2013;75(3):223-228.
15. First MB, Gibbon M. The Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I) and the Structured Clinical Interview for DSM-IV Axis II Disorders (SCID-II) 2004. Comprehensive handbook of psychological assessment, Vol. 2: Personality assessment , (pp. 134-143). Hoboken, NJ, US: John Wiley & Sons Inc, xvi, 671 pp.
16. Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, Hergueta T, Baker R, Dunbar GC. The Mini-International Neuropsychiatric Interview (MINI): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. Journal of clinical psychiatry 1998;59(20):22-33.
17. Hiller W, Janca A. Assessment of somatoform disorders: a review of strategies and instruments. Acta Neuropsychiatrica 2003;15(4):167-179.
18. American Psychiatric Association (2015). Comparison of diagnostic coverage of the SCID-5-RV vs. SCID-5-CV. Retrieved from https://www.appi.org/File%20Library/Products/AH1522-Comparison-SCID-RV-vs-CV.pdf

19. Axelsson E, Andersson E, Ljótsson B, Wallhed FD, & Hedman E. The Health Preoccupation Diagnostic Interview: inter-rater reliability of a structured interview for diagnostic assessment of DSM-5 somatic symptom disorder and illness anxiety disorder. Cognitive behaviour therapy 2016;45(4):259-269.
20. Frances A. The new somatic symptom disorder in DSM-5 risks mislabeling many people as mentally ill. BMJ 2013;346:f1580.
21. Pilowsky I. Dimensions of hypochondriasis. British Journal of Psychiatry 1967;113:89–93.
22. Kellner R. Manual of the IAS (Illness Attitude Scales). Albuquerque: University of New Mexico; 1981.
23. Hemert AM van. Lichamelijke klachten vragenlijst. Leiden: Leids Universitair Medisch Centrum; 2003.
24. Kroenke K, Spitzer RL, Williams JB. The PHQ-15: validity of a new measure for evaluating the severity of somatic symptoms. Psychosomatic medicine 2002;64(2):258-266.
25. Hemert AM van, Waal MWM de, Rood YR van. Meetinstrumenten bij somatoforme stoornissen. Tijdschrift voor psychiatrie 2004;46(10):693-696.
26. Körber S, Frieser D, Steinbrecher N, Hiller W. Classification characteristics of the Patient Health Questionnaire-15 for screening somatoform disorders in a primary care setting. Journal of Psychosomatic Research 2011;71(3):142–147.
27. Voigt K, Wollburg E, Weinmann N, Herzog A, Meyer B, Langs G, Löwe B. 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 2012;30;73(5):345-350.
28. Klaus K, Mewes R. Assessment of the new DSM-5 diagnosis Somatic Symptom Disorder (300.82). Verhaltenstherapie Verhaltensmedizin 2013;34:399-418.

29. Kop WJ, Mols F, Toussaint A, Loewe B. Somatic symptom disorder in the general population: Associations with medical status and health anxiety using the SSD-12. Psychosomatic Medicine 2016;78(3):A24.
30. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Archives of Internal Medicine 2006;166:1092-1097.
31. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. Journal of general internal medicine 2001;16(9):606-613.
32. Cohen J. Statistical power analysis for the behavioral sciences (2nd ed.). Hillsdale, NJ: Erlbaum; 1988.
33. Metz M, Biemans H, Kessels E, Feltz-Cornelis CM van der. Gebruik van Routine Outcome Monitoring bij Shared Decision Making: Een praktijkvoorbeeld. Het Nederlands Tijdschrift voor Evidence Based Practice 2015;13(5):4-6.
34. Feltz-Cornelis CM van der, Andrea H, Kessels E, Duivenvoorden HJ, Biemans H, Metz M. Shared decision making in combinatie met ROM bij patiënten met gecombineerde lichamelijke en psychische klachten; een klinisch-empirische verkenning. Tijdschrift voor Psychiatrie 2014;56(6):375-384.
35. Verhage F. Intelligentie en leeftijd: Onderzoek bij Nederlanders van twaalf tot zevenenzeventig jaar. Proefschrift. Assen: Van Gorcum; 1964.
36. Speckens AEM, Spinhoven P, Van Hemert AM, Bolk JH. The Reassurance Questionnaire (RQ): psychometric properties of a self-report questionnaire to assess reassurability. Psychological medicine 2000;30(4):841-847.
37. Speckens AE, Spinhoven P, Sloekers PP, Bolk JH, Hemert AM van. A validation study of the Whitely Index, the Illness Attitude Scales, and the Somatosensory Amplification Scale in general medical and general practice patients. Journal of psychosomatic research 1996;40(1):95-104.

38. Speckens AEM, Hemert AM van, Spinhoven P, Bolk JH. The diagnostic and prognostic significance of the Whitely Index, the Illness Attitude Scales and the Somatosensory Amplification Scale. Psychological medicine 1996;26(05):1085-1090.
39. Hadjistavropoulos HD, Frombach IK, Asmundson GJG. Exploratory and confirmatory factor analytic investigations of the Illness Attitudes Scale in a nonclinical sample. Behaviour Research and Therapy 1999;37:671-684.
40. Sirri L, Grandi S, Fava GA. The illness attitude scales. Psychother Psychosom 2008;77:337–50.
41. Hedman E, Lekander M, Ljótsson B, Lindefors N, Rück C, Andersson G, Andersson E. Optimal cut-off points on the health anxiety inventory, illness attitude scales and Whiteley index to identify severe health anxiety. PloS one 2015;10(4):e0123412.
42. Waal MWM de, Hemert AM van. Normscores voor Nederlandse respondenten voor de Lichamelijke klachten vragenlijst (51 items) in de algemene bevolking en in een populatie die de huisarts consulteerde. 2013.
43. Waal MW de, Arnold IA, Spinhoven P, Eekhof JA, Assendelft WJ, Hemert AM van. The role of comorbidity in the detection of psychiatric disorders with checklists for mental and physical symptoms in primary care. Social psychiatry and psychiatric epidemiology 2009;44(1):78-85.
44. Central Bureau of Statistics (CBS). Netharlands Central Bureau of Statistics (CBS) list. CBS vademecum 2005. Den Haag: CBS; 2005.
45. Cohen J. Statistical power analysis for the behavioral sciences. New Jersey: Lawrence Erlbaum; 1988.
46. IBM Corporation. IBM SPSS Statistics for Windows, Version 22.0. Armonk, NY: IBM Corporation; 2013.
47. Lorenzo-Seva U, Ferrando PJ. FACTOR: A computer program to fit the exploratory factor analysis model. Behavior research methods 2006; 38(1): 88-91.

48. Horn JL. A rationale and test for the number of factors in factor analysis. Psychometrika 1965; 30(2): 179-185.
49. Ten Berge JM, Kiers HA. A numerical approach to the approximate and the exact minimum rank of a covariance matrix. Psychometrika 1991; 56(2): 309-315.
50. Timmerman ME, Lorenzo-Seva U. Dimensionality assessment of ordered polytomous items with parallel analysis. Psychological Methods 2011; 16(2): 209–220.
51. Kelley TL. Essential Traits of Mental Life. Harvard Studies in Education. Vol. 26. Cambridge: Harvard University Press; 1935.
52. Harman HH. Modern factor analysis. Chicago: University of Chicago Press; 1976.
53. Cronbach LJ. Coefficient alpha and the internal structure of tests. Psychometrika 1951;16(3):297-334.
54. Novick MR, Lewis C. Coefficient alpha and the reliability of composite measurements. Psychometrika 1967;32(1):1-13.
55. Sijtsma K. Over misverstanden rond Cronbach's alfa en de wenselijkheid van alternatieven. Psycholoog 2009;44(11):561.
56. Guttman L. A basis for analyzing test-retest reliability. Psychometrika 1945;10(4):255-282.
57. Youden WJ. Index for rating diagnostic tests. Cancer 1950;3.1:32-35.
58. Kroenke K. Patients presenting with somatic complaints: epidemiology, psychiatric co-morbidity and management. International journal of methods in psychiatric research 2003;12(1):34-43.
59. Sullivan M, Katon WJ, Russo J, Dobie R. Somatization, co-morbidity, and the quality of life: measuring the effect of depression upon chronic medical illness. Psychiatric Medicine 1992;10(3):61-76.
60. Cameron IM, Cardy A, Crawford JR, du Toit SW, Hay S, Lawton K, Mitchell K, Sharma S, Shivaprasad S, Winning S, Reid IC. Measuring depression severity in general practice:

discriminatory performance of the PHQ-9, HADS-D, and BDI-II. British Journal of General Practice 2011;61(588):e419-e426.

61. Hansson M, Chotai J, Nordstöm A, Bodlund O. Comparison of two self-rating scales to detect depression: HADS and PHQ-9. British Journal of General Practice 2009;59(566):e283-e288.
62. Reddy P, Philpot B, Ford D, Dunbar JA. Identification of depression in diabetes: the efficacy of PHQ-9 and HADS-D. British Journal of General Practice 2010;60(575):e239-e245.
63. Cook KF, Bombardier CH, Bamer AM, Choi SW, Kroenke K, Fann JR. Do somatic and cognitive symptoms of traumatic brain injury confound depression screening? Archives of physical medicine and rehabilitation 2011;92(5):818-823.
64. Vroege L de, Manen AS van, Cees A, & Feltz-Cornelis CM van der. Complexity Assessed by the INTERMED in Patients With Somatic Symptom Disorder Visiting a Specialized Outpatient Mental Health Care Setting: A Cross-sectional Study. Psychosomatics 2017.
65. Barsky AJ, Orav EJ, Bates DW. Somatization increases medical utilization and costs independent of psychiatric and medical comorbidity. Archives of general psychiatry 2005;62(8):903-910.

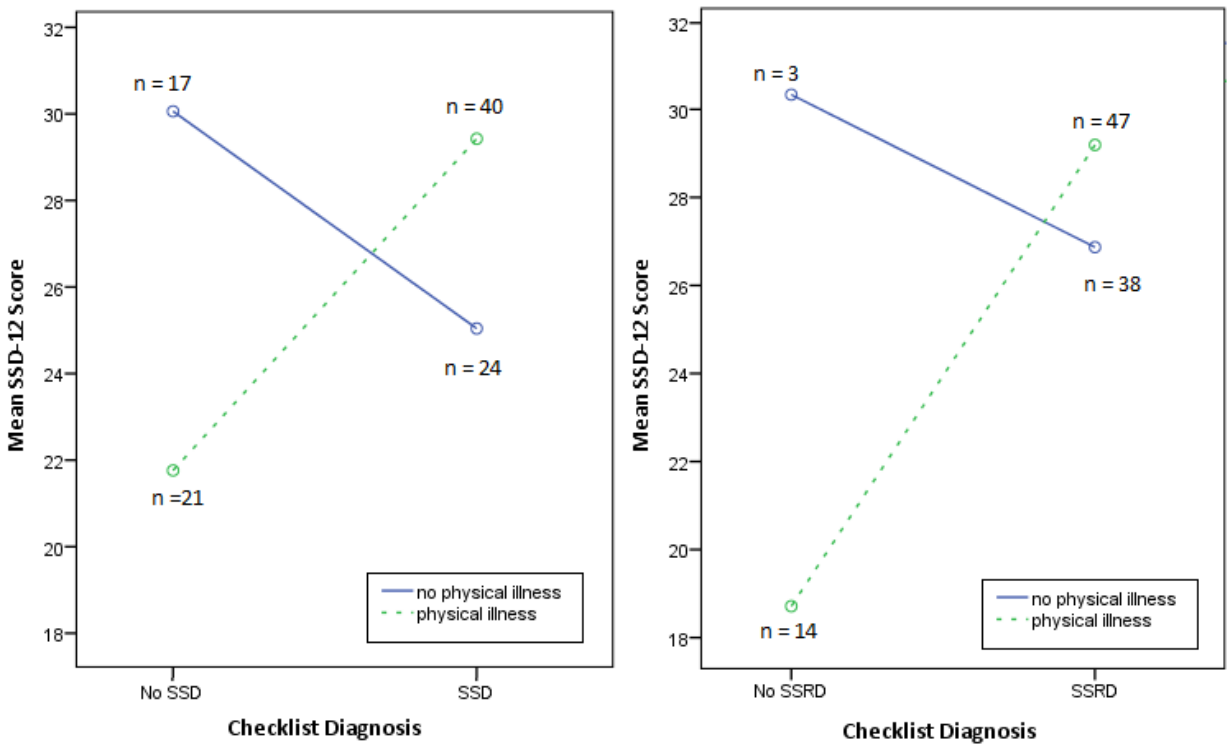

Figure 1. Interaction effect for physical illness on the relation between the SSD-12 and the checklist for the SSD diagnosis vs. the SSRD diagnosis.

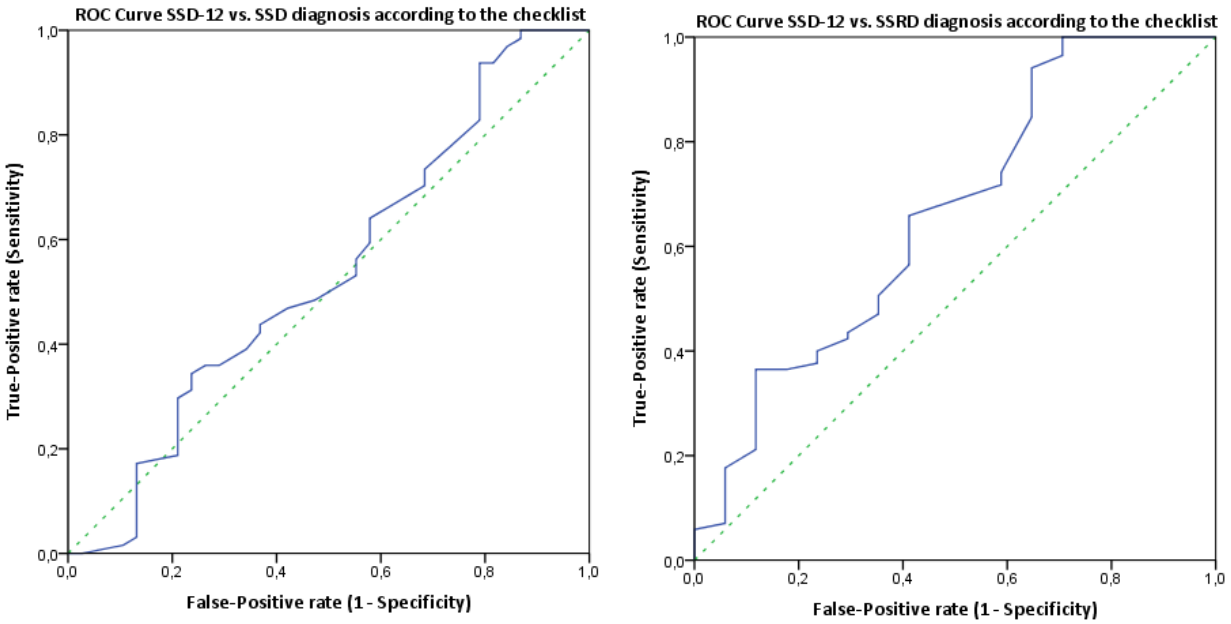

Figure 2. ROC curves for SSD-12 vs. checklist for SSD diagnosis and SSD-12 vs. checklist for SSRD diagnosis. (*Note:* the dotted line is the reference line)

Table 1

Characteristics of the Total Sample (N = 102) and Stratified for SSD

		SSD			
	Total	SSD	No SSD		
	(*N* = 102)	(*n* = 64)	(*n* = 38)	*ES*	*p*
Sample Characteristic	*M (SD)* / *n* (%)	*M (SD) / n (%)*	*M (SD) / n (%)*		
Age	40.6 (14.1)	40.5 (15.1)	41.0 (12.4)	-0.04*	.865
Seks				.14**	.136
Male	42 (41.2%)	23 (35.9%)	19 (50.0%)		
Marital status				.12**	.678
Married	34 (34.0%)	20 (31.7%)	14 (37.8%)		
Living together	22 (22.0%)	14 (22.2%)	8 (21.6%)		
Single	34 (34.0%)	21 (33.3%)	13 (35.1%)		
Living at home	10 (10.0%)	8 (12.7%)	2 (5.4%)		
Missing	2	1	1		
Educational Level***				.11**	.560
Low education	30 (29.7%)	20 (31.7%)	10 (26.3%)		
Middle education	41 (40.6%)	23 (36.5%)	18 (47.4%)		
High education	30 (29.7%)	20 (31.7%)	10 (26.3%)		
Missing	1	1	0		
Chronic Physical Illness				.07**	.304
Yes	61 (59.8%)	40 (62.5%)	21 (55.3%)		
No	41 (40.2%)	24 (37.5%)	17 (44.7%)		

Note. Means (*M*) and standard deviations (*SD*) are presented for the continuous variables and the number of patients (*n*) and percentage (%) are presented for the categorical variables. Cramer's *V* (Chi-square tests) and Cohen's *d* (t-tests) will be used to determine the effect size (ES). Differences between patients with and without Somatic Symptom Disorder (SSD) are tested on significance level. *Cohen's *d*, **Cramer's *V*. ***Following Verhage grading system [36].

Table 2

Inter-Item Correlations

Item	1	2	3	4	5	6	7	8	9	10	11	12
1												
2	.587											
3	.602	.728										
4	.525	.586	.544									
5	.617	.626	.686	.526								
6	.471	.554	.657	.553	.568							
7	.351	.312	.284	.187	.298	.267						
8	.367	.480	.508	.412	.486	.575	.344					
9	.521	.631	.685	.465	.605	.662	.228	.545				
10	.335	.333	.350	.304	.333	.301	.328	.455	.306			
11	.470	.492	.520	.440	.485	.572	.251	.593	.576	.337		
12	.299	.348	.454	.373	.439	.488	.326	.793	.474	.385	.561	

Table 3

Loading Matrix of the One- Factor Model

Item	Factor 1
1	.697
2	.791
3	.826
4	.675
5	.760
6	.792
7	.415
8	.755
9	.774
10	.484
11	.706
12	.695

Table 4

Loading Matrix of the Three-Factor Model

Item	Factor 1	Factor 2	Factor 3
1		.816	
2		.889	
3		.867	
4		.720	
5		.760	
6		.614	
7			.987
8	.933		
9		.673	
10			
11	.452	.354	
12	.996		

Note. Only factor loadings of .300 or higher are shown

Table 5

Reliability of the Three Subscales According to the Study of Toussaint et al. (2016)

	Cronbach's α	Guttman's lambda[2]
SSD-12 total score	.91	.92
SSD-12 subscale I	.68	.68
SSD-12 subscale II	.81	.81
SSD-12 subscale III	.84	.85

Table 6

Descriptives of the Questionnaires

Questionnaire	Descriptive statistics		
	n	*M (SD)*	Range
SSD-12	*102*	26.92 (10.67)	0-46
WI	74	6.55 (3.38)	0-15
IAS	63	44.32 (19.41)	16-96
RQ	51	12.43 (6.29)	1-27
PSC-51	100	15.28 (8.19)	0-41
PHQ-15	99	14.76 (5.44)	1-27
GAD-7	101	10.75 (6.08)	0-31
PHQ-9	100	13.73 (6.05)	0-27

Note. Means (*M*), standard deviations (*SD*) and ranges of the total score and the number of patients (*n*) are presented for each questionnaire.
Abbreviations: SSD-12, Somatic Symptom Disorder–B Criteria Scale; WI, the Whiteley Index; IAS, the Illness Attitude Scale; RQ, the Reassurance Questionnaire; PSC-51, the Physical Symptoms Checklist-51; PHQ-15, the Patient Health Questionnaire-15; GAD-7, the Generalized Anxiety Disorder-7; PHQ-9, Patient Health Questionnaire-9.

Table 7

Pearson Correlations between the SSD-12 Total Score and Subscales and the Other Questionnaires

	WI	IAS	RQ	PSC-51	PHQ-15	GAD-7	PHQ-9
SSD-12 total score	.66**	.75**	.59**	.48**	.43**	.63**	.55**

Abbreviations: SSD-12, Somatic Symptom Disorder–B Criteria Scale; WI, the Whiteley Index; IAS, the Illness Attitude Scale; RQ, the Reassurance Questionnaire; PSC-51, the Physical Symptoms Checklist-51; PHQ-15, the Patient Health Questionnaire-15; GAD-7, the Generalized Anxiety Disorder-7; PHQ-9, Patient Health Questionnaire-9.

Table 8

Amount of Patients with and without SSRD Diagnosis and Average Score on the SSD-12

	n (%)	SSD-12 score *M (SD)*
Total	102 (100%)	26.92 (10.67)
No Somatic Symptom and Related Disorders	17 (16.7%)	20.76 (12.32)
Somatic Symptom and Related Disorders	85 (83.3%)	28.15 (9.94)
Somatic Symptom Disorder	64 (62.7%)	27.78 (9.80)
Somatic Symptom Disorder and Conversion Disorder	4 (3.9%)	25.25 (9.00)
Somatic Symptom Disorder and no Conversion Disorder	60 (58.8%)	27.95 (9.90)
Another Somatic Symptom and Related Disorders	21 (20.6%)	29.29 (10.52)
Illness Anxiety disorder	9 (8.8%)	38.33 (8.17)
Conversion disorder	7 (6.9%)	21.71 (4.99)
Factitious Disorder	0 (0%)	-
Other Specified Somatic Symptom and Related Disorders	0 (0%)	-
Unspecified Somatic Symptom and Related Disorders	5 (4.9%)	23.60 (7.60)
No Somatic Symptom Disorder	38 (37.3%)	25.47 (12.00)

Note. Means (*M*) and standard deviations (*SD*) of the SSD-12 score and the number of patients (*n*) are presented for each Somatic Symptom and Related Disorders (SSRD) category.

Table 9a

Sensitivity, Specificity, Youden's J, PPV, NPV, Efficiency, False Positives and Negatives of the SSD-12 Scores for SSD diagnosis

Score SSD-12	*N*	SSD (*n* =)	Sensitivity	Specificity	Youden's J	PPV	NPV	Efficiency	False pos.	False neg.
0	1	0	1.00	.00	.00	.55	1.00	.55	46	0
1	0	0	1.00	.03	.03	.63	1.00	.64	37	0
2	0	0	1.00	.03	.03	.63	1.00	.64	37	0
3	0	0	1.00	.03	.03	.63	1.00	.64	37	0
4	0	0	1.00	.03	.03	.63	1.00	.64	37	0
5	1	0	1.00	.03	.03	.63	1.00	.64	37	0
6	1	0	1.00	.05	.05	.64	1.00	.65	36	0
7	0	0	1.00	.08	.08	.65	1.00	.66	35	0
8	0	0	1.00	.08	.08	.65	1.00	.66	35	0
9	1	0	1.00	.08	.08	.65	1.00	.66	35	0
10	1	0	1.00	.11	.11	.65	1.00	.67	34	0
11	0	0	1.00	.13	.13	.66	1.00	.68	33	0
12	1	1	1.00	.13	.13	.66	1.00	.68	33	0
13	2	1	.98	.13	.11	.66	.83	.67	33	1
14	3	2	.97	.16	.13	.66	.75	.67	32	2
15	1	0	.94	.18	.12	.66	.64	.66	31	4
16	3	3	.94	.21	.15	.67	.67	.67	30	4
17	1	1	.89	.21	.10	.66	.53	.64	30	7
18	3	3	.88	.21	.09	.65	.50	.63	30	8
19	10	6	.83	.21	.04	.64	.42	.60	30	11
20	2	2	.73	.32	.05	.64	.41	.58	26	17
21	8	4	.70	.32	.02	.64	.39	.56	26	19
22	3	3	.64	.42	.06	.65	.41	.56	22	23
23	3	2	.59	.42	.01	.63	.38	.53	22	26
24	2	2	.56	.45	.01	.63	.38	.52	21	28
25	6	3	.53	.45	-.02	.62	.36	.50	21	30
26	3	1	.48	.53	.01	.63	.38	.50	18	33
27	4	2	.47	.58	.05	.63	.39	.51	16	34
28	1	1	.44	.63	.07	.67	.40	.51	14	36
29	3	2	.42	.63	.05	.66	.39	.50	14	37
30	2	1	.39	.66	.05	.66	.39	.49	13	39
31	2	1	.38	.68	.06	.67	.39	.49	12	40
32	1	0	.36	.71	.07	.68	.40	.49	11	41
33	2	1	.36	.74	.10	.70	.41	.50	10	41
34	2	2	.34	.76	.10	.71	.41	.50	9	42
35	2	1	.31	.76	.07	.69	.40	.48	9	44
36	1	1	.30	.79	.09	.70	.40	.48	8	45
37	3	3	.28	.79	.07	.69	.39	.47	8	46
38	3	3	.23	.79	.02	.65	.38	.44	8	49
39	4	1	.19	.79	-.02	.60	.37	.41	8	52
40	1	1	.17	.87	.04	.69	.38	.43	5	53
41	1	1	.16	.87	.03	.67	.38	.42	5	54
42	1	1	.14	.87	.01	.64	.38	.41	5	55
43	6	6	.13	.87	.00	.62	.37	.40	5	56
44	2	1	.03	.87	-.10	.29	.35	.34	5	62
45	4	1	.02	.89	.09	.20	.35	.34	4	63
46	1	0	.00	.97	-.03	.00	.37	.36	1	64

Abbreviations: SSD-12, Somatic Symptom Disorder–B Criteria Scale; SSD, Somatic Symptom Disorder; PPV, Positive Predictive Value; NPV, Negative Predictive Value; false pos., false positives; false neg., false negatives.

Table 9b

Sensitivity, Specificity, Youden's J, PPV, NPV, Efficiency, False Positives and Negatives of the SSD-12 Scores for SSRD diagnosis

Score SSD-12	*N*	SSRD (*n* =)	Sensitivity	Specificity	Youden's J	PPV	NPV	Efficiency	False pos.	False neg.
0	1	0	1.00	.00	.00	.83	1.00	.83	17	0
1	0	0	1.00	.06	.06	.84	1.00	.84	16	0
2	0	0	1.00	.06	.06	.84	1.00	.84	16	0
3	0	0	1.00	.06	.06	.84	1.00	.84	16	0
4	0	0	1.00	.06	.06	.84	1.00	.84	16	0
5	1	0	1.00	.06	.06	.84	1.00	.84	16	0
6	1	0	1.00	.12	.12	.85	1.00	.85	15	0
7	0	0	1.00	.18	.18	.86	1.00	.86	14	0
8	0	0	1.00	.18	.18	.86	1.00	.86	14	0
9	1	0	1.00	.18	.18	.86	1.00	.86	14	0
10	1	0	1.00	.24	.24	.87	1.00	.87	13	0
11	0	0	1.00	.29	.29	.88	1.00	.88	12	0
12	1	1	1.00	.29	.29	.88	1.00	.88	12	0
13	2	2	.99	.29	.28	.88	.83	.87	12	1
14	3	2	.96	.29	.25	.87	.63	.85	12	3
15	1	1	.94	.35	.29	.88	.55	.84	11	5
16	3	3	.93	.35	.28	.88	.50	.83	11	6
17	1	1	.89	.35	.24	.87	.40	.80	11	9
18	3	3	.88	.35	.23	.87	.38	.79	11	10
19	10	9	.84	.35	.19	.87	.32	.76	11	13
20	2	2	.74	.41	.15	.86	.24	.69	10	22
21	8	5	.72	.41	.13	.86	.23	.67	10	24
22	3	3	.66	.59	.25	.89	.26	.64	7	29
23	3	3	.62	.59	.21	.88	.24	.62	7	32
24	2	2	.59	.59	.18	.88	.22	.59	7	35
25	6	5	.56	.59	.15	.87	.21	.57	7	37
26	3	3	.51	.65	.16	.88	.21	.53	6	42
27	4	3	.47	.65	.12	.87	.20	.50	6	45
28	1	1	.44	.71	.15	.88	.20	.48	5	48
29	3	2	.42	.71	.13	.88	.20	.47	5	49
30	2	2	.40	.76	.16	.89	.20	.46	4	51
31	2	1	.38	.76	.14	.89	.20	.44	4	53
32	1	0	.36	.82	.18	.91	.21	.44	3	54
33	2	2	.36	.88	.24	.94	.22	.45	2	54
34	2	2	.34	.88	.22	.94	.21	.43	2	56
35	2	2	.32	.88	.20	.93	.21	.41	2	58
36	1	1	.29	.88	.17	.93	.20	.39	2	60
37	3	3	.28	.88	.16	.92	.20	.38	2	61
38	3	3	.25	.88	.13	.91	.19	.35	2	64
39	4	3	.21	.88	.09	.90	.18	.32	2	67
40	1	1	.18	.94	.12	.94	.19	.30	1	70
41	1	1	.16	.94	.10	.94	.18	.29	1	71
42	1	1	.15	.94	.09	.94	.18	.28	1	72
43	6	6	.14	.94	.08	.92	.18	.27	1	73
44	2	1	.07	.94	.01	.86	.17	.22	1	79
45	4	4	.06	1.00	.06	1.00	.18	.22	0	80
46	1	1	.01	1.00	.01	1.00	.17	.18	0	84

Abbreviations: SSD-12, Somatic Symptom Disorder–B Criteria Scale; SSRD, Somatic Symptom and Related Disorders; PPV, Positive Predictive Value; NPV, Negative Predictive Value; false pos., false positives; false neg., false negatives.

Appendix

Appendix A

Formulae for Test Performance Indices

Table A1

Accuracy of the SSD-12

	Positive SSD checklist	Negative SSD checklist	Total
Positive SSD-12 somscale	a *True Positives*	b *False Positives*	a + b
Negative SSD-12 somscale	c *False Negatives*	d *True Negatives*	c + d
Total	a + c	b + d	a+ b + c + d

Abbreviations: SSD, Somatic Symptom Disorder; SSD-12, Somatic Symptom Disorder–B Criteria Scale.

Table A2

Summary Indices of Test Performance

Variable	Formula	
Sensitivity (True Positive Rate)	a/(a + c)	True Positives/(True Positives + False Negatives)
Specificity (True Negative Rate)	d/(b + d)	True Negatives/(False Positives + True Negatives)
False Negative Rate (1-Sensitivity)	c/(a + c)	False Negatives/(True Positives + False Negatives)
False Positive Rate (1-specificity)	b/(b + d)	False Positives/(True Negatives + False Positives)
Negative Predictive Value (NPV)	d/(c + d)	True Negatives/(False Negatives + True Negatives)
Positive Predictive Value (PPV)	a/(a + b)	True Positives/(True Positives + False Positives)
Efficiency	(a + d)/ (a + b + c + d)	
Youden's J	(a/(a + c)) + (d/(b + d)) - 1	Sensitivity + Specificity – 1

Appendix B

Measures

Measurement	Measuring instrument
Presence of SSD	A checklist was made, based on the criteria of the SSD diagnosis of the DSM 5, to assess the diagnostic criteria of SSD following the DSM 5 (see Appendix D). Three psychiatrists at the outpatient clinic (CFC, JvE, AF) performed the psychiatric evaluations for this study and used this checklist to draw conclusions about the SSD diagnosis.
Presence of anxiety during the past 2 weeks.	Indicated by a score of 10 or higher on the GAD-7.
Presence of depression during the past 2 weeks.	Indicated by a score of 10 or higher on the PHQ-9.
Presence of somatic symptoms.	During the last week, indicated by a score of 11 for men and 13 for women or higher on the PSC-51. During the past 4 weeks, indicated by a score of 9 or higher on the PHQ-15.
Presence of health anxiety.	Indicated by a score of 5 or higher on the WI and a score of 47 or higher on the IAS.
Presence of reassurability.	Indicated by a score of 9 or lower on the RQ.
Presence of chronic physical illness.	Indicated by at least one chronic physical illness on the CBS checklist.

Appendix C
SSD-12

SSD-12

Bijna iedereen heeft wel eens last van lichamelijke klachten, zoals hoofdpijn, rugpijn, misselijkheid, of hartkloppingen. De volgende vragen gaan over uw gedachten en gevoelens bij deze lichamelijke klachten en hoe u met dit soort klachten omgaat. Gelieve bij elk van deze uitspraken een ***cirkeltje*** te plaatsen rond het ***cijfer*** dat het meest op u van toepassing is. Er zijn geen goede of foute antwoorden; we zijn geïnteresseerd hoe u zich voelt

		0 = Nooit **1 = Zelden** **2 = Soms** **3 = Vaak** **4 = Meestal of altijd**				
1.	Ik denk dat mijn lichamelijke klachten een teken zijn van een ernstige ziekte.	0	1	2	3	4
2.	Ik ben erg bezorgd over mijn gezondheid.	0	1	2	3	4
3.	Mijn zorgen over mijn gezondheid hinderen me in het dagelijkse leven.	0	1	2	3	4
4.	Ik ben ervan overtuigd dat mijn klachten ernstig zijn.	0	1	2	3	4
5.	Mijn lichamelijke klachten maken me bang.	0	1	2	3	4
6.	Mijn lichamelijke klachten houden me het merendeel van de dag bezig.	0	1	2	3	4
7.	Andere mensen vertellen me dat mijn lichamelijke klachten niet ernstig zijn.	0	1	2	3	4
8.	Ik maak me zorgen dat mijn klachten nooit zullen verdwijnen.	0	1	2	3	4
9.	Mijn zorgen over mijn gezondheid kosten mij energie.	0	1	2	3	4
10.	Ik denk dat artsen mijn lichamelijke klachten niet serieus nemen.	0	1	2	3	4
11.	Door mijn lichamelijke klachten kan ik me slecht op andere dingen concentreren.	0	1	2	3	4
12.	Ik maak me zorgen dat mijn lichamelijke klachten ook in de toekomst zullen blijven bestaan.	0	1	2	3	4

Appendix D
The psychiatric checklist

Checklist Somatisch-symptoomstoornis tijdens psychiatrisch onderzoek:

Naam: Geboortedatum: .. - .. - .... Cliëntnummer: Geslacht: m/v

A Patiënt(e) heeft de volgende lichamelijke hoofdklacht(en):
……
Deze klacht past bij een bekende somatische aandoening:
□ Nee
□ Ja, namelijk:……
□ Er loopt nog onderzoek naar een chronische lichamelijke aandoening, namelijk……

B Deze lichamelijke klachten gaan gepaard met [vink aan welke van toepassing zijn]:
□ disproportionele en persisterende gedachten over de ernst van de klachten
□ een persisterende hoge mate van ongerustheid over de gezondheid of de klachten
□ het excessief veel tijd en energie besteden aan deze klachten of aan de zorgen over de gezondheid
□ Geen van bovenstaande

C Het hebben van lichamelijke klachten is van langere duur: [NB niet elke afzonderlijke klacht hoeft steeds aanwezig te zijn, maar het hebben van klachten wel (meestal >6 maanden). Dit is dus NIET hetzelfde als persisterend zoals verderop beschreven bij specificatie]
□ Nee
□ Ja, namelijk:…… [specificeer duur]

Voldoet patiënt(e) aan criterium A (de lichamelijke klacht mag zowel verklaard als onverklaard zijn), B (minstens 1 van de punten) en C, en is er dus sprake van een somatisch-symptoomstoornis?
□ Nee
□ Ja → specificeer hieronder:

Specificeer indien:
□ Met voornamelijk pijn [voorheen pijnstoornis, deze specificatie geldt voor mensen van wie de lichamelijke klachten vooral bestaan uit pijn]

Specificeer actuele ernst:
□ Licht [Slechts één van de in criterium B genoemde symptomen is aanwezig]
□ Matig [Er zijn twee of meer van de in criterium B genoemde symptomen aanwezig]
□ Ernstig [Er zijn twee of meer van de in criterium B genoemde symptomen aanwezig en er zijn multipele lichamelijke klachten (of één zeer ernstige lichamelijke klacht)]

Specificeer indien:
□ Persisterend [Een persisterend beloop wordt gekenmerkt door ernstige klachten, duidelijke beperkingen in het functioneren en een lange duur (langer dan zes maanden)]

CONCLUSIE

Omschrijf op basis hiervan [vink aan welke van toepassing zijn]:

- ☐ Deze patiënt heeft een lichte/matige/ernstige [doorhalen wat niet van toepassing is] somatisch-symptoomstoornis, [toevoegen indien van toepassing:] met pijn EN/OF persisterend
- ☐ Deze patiënt heeft tevens een chronische lichamelijke aandoening, namelijk……
- ☐ Er loopt nog onderzoek naar een chronische lichamelijke aandoening, namelijk……
- ☐ Deze patiënt heeft geen Somatisch Symptoom Stoornis
- ☐ Deze patiënt heeft een andere SSRD diagnose, namelijk:
  - ☐ Een Conversiestoornis
  - ☐ Een Nagebootste stoornis
  - ☐ Een Ziekteangststoornis
  - ☐ Een Ongespecificeerde Somatisch Symptoom Stoornis of verwante stoornis

Appendix E
Characteristics Stratified for SSRD

Table E1
Characteristics of the Total Sample (N = 102) and Stratified for SSRD

		SSRD			
	Total	SSRD	No SSRD		
	(*N* = 102)	(*n* = 85)	(*n* = 17)	*ES*	*p*
Sample Characteristic	*M (SD) / n* (%)	*M (SD) / n (%)*	*M (SD) / n (%)*		
Age	40.6 (14.1)	39.5 (14.3)	46.3 (11.4)	0.49*	.069
Sex				.16**	.105
Male	42 (41.2%)	32 (37.7%)	10 (58.8%)		
Marital status				.15**	.501
Married	34 (34.0%)	26 (31.3%)	8 (47.0%)		
Living together	22 (22.0%)	20 (24.1%)	2 (11.8%)		
Single	34 (34.0%)	28 (33.7%)	6 (35.3%)		
Living at home	10 (10.0%)	9 (10.8%)	1 (5.9%)		
Missing	2	2	0		
Educational Level***				.12**	.490
Low education	30 (29.7%)	24 (28.6%)	6 (35.3%)		
Middle education	41 (40.6%)	33 (39.3%)	8 (47.1%)		
High education	30 (29.7%)	27 (32.1%)	3 (17.6%)		
Missing	1	1	0		
Chronic Physical Illness				.04**	.206
Yes	61 (59.8%)	47 (55.3%)	14 (82.4%)		
No	41 (40.2%)	38 (44.7%)	3 (17.6%)		

Note. Means (*M*) and standard deviations (*SD*) are presented for the continuous variables and the number of patients (*n*) and percentage (%) are presented for the categorical variables. Cramer's V (Chi-square tests) and Cohen's d (t-tests) will be used to determine the effect size (ES). Differences between patients with and without Somatic Symptom and Related Disorders (SSRD) are tested on significance level. *Cohen's d, **Cramer's V. ***Following Verhage grading system [36].