BMJ Open

The BDS checklist as measure of illness severity

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2020-042880
Article Type:	Original research
Date Submitted by the Author:	17-Jul-2020
Complete List of Authors:	Petersen, Marie; Aarhus Universitetshospital, The Research Clinic for Functional Disorders and Psychosomatics Rosendal, Marianne; Aarhus University Hospital, Research Clinic for Functional Disorders and Psychosomatics Ørnbøl, Eva Fink, Per Jørgensen, Torben; Frederiksberg Hospital, Dantoft, Thomas; Region Hovedstaden Forskningscenter for forebyggelse og sundhed, Schroeder, Andreas; Aarhus Universitetshospital,
Keywords:	STATISTICS & RESEARCH METHODS, EPIDEMIOLOGY, PSYCHIATRY, GENERAL MEDICINE (see Internal Medicine)

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

The BDS checklist as measure of illness severity

Marie Weinreich Petersen[1], Marianne Rosendal[1,2], Eva Ørnbøl[1], Per Fink[1], Torben Jørgensen[3,4,5], Thomas Meinertz Dantoft[3], Andreas Schröder[1]

Author affiliations:
[1] The Research Clinic for Functional Disorders and Psychosomatics, Aarhus University Hospital, Aarhus C, Denmark
[2] Research Unit for General Practice, Aarhus University, Aarhus C, Denmark
[3] Center for Clinical Research and Prevention, Bispebjerg and Frederiksberg Hospital, The Capital Region of Denmark, Denmark
[4] Department of Public Health, Faculty of Health and Medical Science, University of Copenhagen, Copenhagen C, Denmark
[5] Faculty of Medicine, Aalborg University

Corresponding author:
Marie Weinreich Petersen
The Research Clinic for Functional Disorders and Psychosomatics
Aarhus University Hospital
Nørrebrogade 44, building 2C, second floor
8200 Aarhus C
Denmark
Phone: +45 29467824
Fax: +45 7846 4340
E-mail: mawept@rm.dk

Word count in main text: 3997
Word count abstract: 290
Number of references: 50

Abstract

Objectives

The Bodily Distress Syndrome (BDS) checklist has proven to be useful for diagnostic categorization and screening tool for functional somatic disorders (FSD). This study aims to investigate whether the BDS checklist total sum score (0-100) can be used as measure of physical symptom burden and FSD illness severity.

Design

Cross-sectional.

Participants

Three cohorts of adult individuals; a general population cohort (n=9656), a primary care cohort (n=2480), and a cohort of multi-organ BDS patients from specialized clinical setting (n=492).

Outcome measures

All data were self-reported. Physical symptoms were measured with the 25-items BDS checklist. Overall self-perceived health was measured with one item from the 36-items Short Form Health Survey (SF-36). Physical functioning was measured with an aggregate score of 4 items from the SF-36/SF-12 scales 'physical functioning', 'bodily pain', and 'vitality'. Emotional distress was measured with the mental distress subscale (SCL-8) from the Danish version of the Hopkins Symptom Checklist-90. Illness worry was measured with the 6-items Whiteley Index.

Results

For all cohorts, bi-factor models established that despite some multi-dimensionality, the total sum score of the BDS checklist adequately reflected physical symptom burden and illness severity. The BDS checklist had acceptable convergent validity to measures of overall health, physical functioning, emotional distress, and illness worry. Acceptability was good with low numbers of missing responses to items (<3%). Internal consistency was high ($\alpha \geq 0.879$). BDS score means varied and reflected symptom burden across cohorts. We provide normative data for the Danish general population.

Conclusions

The BDS checklist total sum score can be used as measure of symptom burden and FSD illness severity across settings. These findings establish the usefulness of the BDS checklist in clinic and in research both as a diagnostic screening and as an instrument for assessment of illness severity.

Strengths and limitations of this study

- The study included data from three cohorts and settings: A general population, primary care patients, and patients from a specialized setting
- Well-validated measures were used to determine convergent validity
- All included cohorts had large sample sizes
- Only self-reported measures were included
- Convergent validity was not investigated with other measures of physical symptom burden

Introduction

Persistent physical symptoms (PPS) are common in medical settings and the general population [1-4]. The symptoms present across a continuum from one or a few momentary to numerous symptoms from multiple locations in the body. Having a high number of symptoms has been associated with poor health, poor functional status, and increased health care use [5-9]. Hence, assessment of the burden of persistent physical symptoms is valuable in both clinical care and in research.

For this purpose, self-reported symptom questionnaires are frequently used. They are manageable, non-invasive tools. Several screening questionnaires exist: The Hopkins Symptom Checklist [10], the Patient Health Questionnaire [11], the Somatic Symptom Scale-8 [12 13], the brief form of the Giessen Subjective Complaints List [14], and others [15-17]. However, the existing questionnaire measure PPS without consideration of the well-known aggregation of such symptoms into symptom clusters, and hence, without acknowledgement of the real structure of PPS as it occurs in both the community and in clinical setting[18-21].

When PPS occur in the absence of (other) physical or mental conditions, or when they cause individual suffering and functional limitations beyond what could be expected based on such diseases, they constitute the very core of the disorders captured under the umbrella definition of Functional Somatic Disorders (FSD). FSD cover both specialty-specific syndrome diagnoses such as fibromyalgia, irritable bowel and chronic fatigue, but also their pendants in psychosomatic medicine, somatoform and somatic symptom disorders [22].

In contrast to the above mentioned speciality-specific diagnoses, the proposed research diagnosis bodily distress syndrome (BDS) covers a broader range of functional somatic symptoms ranging from few symptoms with some effect on functioning to severe and disabling functional somatic disorders [18 19 21]. Hence, BDS provides the opportunity to assess and distinguish between conditions persisting as mono- or multi-syndromatic and still within the same framework of diagnostic approach [21 23]. The diagnostic construct was developed in a sample of patients from primary and secondary care, and the 30-items BDS checklist emerged [18]. BDS was confirmed in a new sample of primary care patients where the shortened 25-items BDS checklist was developed [19]. Subsequently, the construct of BDS has been confirmed in general population samples as well[21 24]. BDS presents symptoms grouped in four symptom clusters: Cardiopulmonary (CP), gastrointestinal (GI), musculoskeletal (MS), and general symptoms (GS), and its usefulness and properties used for diagnostic categorisation into no BDS, a single/oligo-organ BDS type and a multi-organ BDS type has been established [19 21 24]. A major strength of the BDS checklist is its usefulness both as a

screening and as diagnostic tool within clinical practise and within epidemiological research [18 19 21 23], but the total BDS sum score has not yet been validated as a measure for the assessment of symptom burden and illness severity.

This study aims to explore whether the BDS checklist can be used as a continuous score to measure symptom burden (i.e. in those individuals that may fall under the diagnostic threshold or what we believe to be clinically relevant) and illness severity (in those individuals fulfilling diagnostic criteria for FSD). In order to elicit the BDS checklist's usability across settings, its structural validity and psychometric properties will be explored in three different populations: the general population, primary care patients, patients in a specialized clinical setting.

Methods

Population

This cross-sectional study included baseline data from three cohorts:

Cohort 1: A general population cohort (DanFunD, n=9656) established with the purpose to investigate and unravel the epidemiology of FSD [25]. The cohort was obtained from the Danish Central Personal Register and drawn as a random sample of the adult Danish background population aged 18-69 years. Participants lived in 10 municipalities in the south-western part of the greater Copenhagen area. All participants were born in Denmark.

Cohort 2: A cohort of primary care patients (KOS, n=2480) established in order to investigate contact and disease patterns in general practice [26]. Participants were included consecutively from 388 general practitioners from the Central Denmark Region. Included participants were 18 years or older and had completed a health-related face-to-face consultation with their general practitioner.

Cohort 3: Data from a specialized clinical setting at the Research Clinic for Functional Disorders and Psychosomatics, Aarhus University Hospital in Denmark (STreSS-3, STreSS-4, STreSS-5, n=492) [27-31]. These cohorts had been part of a group of studies with the shared aim to investigate new treatments for patients with multi-organ BDS aged 20 years or older.

Measures

Self-reported data of physical symptoms, overall health, physical health, mental health, and illness worry was included. The measures and data were not completely consistent across the three included cohorts.

Physical symptoms were assessed with the Danish version of the 25-items BDS checklist (Appendix A) [19 21]. The checklist asks "during the last *(specific time frame)* have you been bothered by" followed by a list of 25 symptoms comprising the four symptom clusters of BDS. The BDS checklist measures symptoms on a five-point rating scale from 0 ('not at all bothersome') to 4 ('a lot bothersome'). We calculated a sum score by adding the single item scores from the 25 items (ranging from 0 to 100). The timeframe covered was 12 months for the general population cohort and four weeks for the other two cohorts.

Overall health was assessed with a single item from the 36-items Short Form Health Survey (SF-36) [32], estimating self-perceived health on a five-point rating scale from 'excellent' to 'poor'. Higher score on this item indicates poorer health. No specific time frame was surveyed in neither of the cohorts.

Physical functioning was measured with an aggregate score of four items from the SF-36 subscales 'physical functioning', 'bodily pain', and 'vitality' [30 32-34]. The aggregate score consisted of four items which are part of the SF-12, addressing limitations in moderate and strenuous activities because of physical health and pain interference. Higher scores indicate better physical health. We tested the correlation of the mean *t*-score of the four item aggregate score against the full SF-36 aggregate score in cohort 3, and correlation was high (*Spearman rho*=0.89, 95% CI: 0.87;0.91). Unfortunately, data on the primary care cohort did not allow us to investigate convergent validity to the aggregate score, while these analyses were only performed in the general population cohort and the cohort from specialized clinical setting. The time frame covered was four weeks for both cohorts.

Emotional distress was measured with the mental distress subscale (SCL-8) from the Danish version of the Hopkins Symptom Checklist (SCL-90) [35 36]. SCL-8 consists of eight items addressing impairment of overall worries, depression, and anxiety. Answers were calculated as mean scores from a scale ranging from 0 ('not at all bothersome') to 4 ('a lot bothersome'). Higher scores indicated higher emotional distress. The time frame covered was one week for the general population cohort and four weeks for the two other cohorts.

Illness worry was measured with the Whiteley Index 6 items version revised (Whiteley-6-R) [37], addressing the respondent's fear of being ill and whether they attribute current bodily sensations to somatic illness[1]. Answers were calculated as mean scores from a scale ranging from 0 ('not at all

[1] In the primary care sample, one of the items in the WI-6 "Do you worry about the possibility that you suffer from an illness you have heard or read about" was expressed as "Do you worry about the possibility that you suffer from an illness".

bothersome') to 4 ('a lot bothersome'). Higher scores indicated higher health anxiety. The time frame covered was 12 months for the general population cohort and four weeks for the two other cohorts.

Validation procedure and statistical analyses

The analyses for the current study were performed according to the Consensus-based Standards for the selection of health Measurement Instrument (COSMIN) framework [38].

All statistical analyses were performed using STATA version 16.0 [39], except for the structural equation modelling which was performed using Mplus version 8.1 [40].

Construct validity was tested by means of structural validity and convergent validity.

Structural validity was tested with confirmatory factor analyses (CFA) with WLSMV (Weighted Least Squares Means and Variance adjusted) estimation due to categorical responses for all items [40]. We wanted to test if it was permissible to model the BDS checklist as unidimensional despite the previous evidence of some multi-dimensionality [18 19 21 24]. Furthermore, we wanted to test if the raw total BDS sum score would be an adequate reliable measure of the general factor (BDS). Therefore, four different CFAs were performed: 1) An one-level one factor model, 2) an one-level four factor model, using factors resembling the four BDS symptom clusters previously reported [19 21], 3) a two-level four factor model, representing a second order common factor (BDS) underlying the four BDS symptom clusters, and 4) a bi-factor CFA, reflecting each symptom to load on a general factor (BDS) and on one of the four specific BDS symptom clusters. Illustrations of the four types of CFAs are displayed in Appendix B.

In all CFAs, model fit were assessed as follows: A Root Mean Square of Approximation (RMSEA) <0.05 indicates very good fit, 0.05-0.08 indicates a good fit, and ≥0.08 indicates a poor fit. Comparative fit Index (CFI) and Tucker-Lewis fit Index (TLI) at 0.90-0.95 indicate an acceptable fit and levels >0.95 indicate a good fit. A Standardized Root Mean square Residual (SRMR) <0.08 indicates good fit [41].

Convergent validity was tested with Spearman's correlations, and associations between the BDS checklist and overall health (one item from SF-36)[32], physical function (an aggregate score of four items from the SF-36) [42], emotional distress (SCL-8) [35], and illness worry (Whiteley-6-R) (Carstensen) were performed. Based on previous literature [12 14 15 17 43], we hypothesized that the BDS checklist would show moderate convergent validity (r=0.40-0.60) with the four measures, and we expected lower correlations in the sample from specialized setting. Expected differences on the

BDS checklist with one unit difference to the SCL-8, the four items aggregate score for physical functioning, and Whiteley-6-R were estimated with linear regression.

BDS checklist item and scale characteristics, i.e. item means (SD), sum score means, score distribution, item-rest correlations, and aspects of acceptability, i.e. percentage of missing items, were examined and computed as descriptive statistics for each of the three samples.

Internal consistency was measured with Cronbach's α coefficients.

Ethical considerations

The current study was carried out in accordance with the relevant guidelines and regulations.

For all three cohorts, written informed consent was obtained from each participant before entering the studies [25-31].

Cohort 1: Approved by the independent ethics committees the Ethical Committee of Copenhagen County (Ethics Committee: KA-2006-0011; H-3-2011-081; H-3-2012-0015) and the Danish Data Protection Agency.

Cohort 2: Approved by the Danish Data Protection Agency and the Health and Medicines Authority. According to Danish law, approval from the health research ethics system was not needed.

Cohort 3: The STreSS-3 trial was registered with ClinicalTrials.gov, number NCT01518634, EudraCT number 2011-004294-87. The STreSS-4 trial was registered with ClinicalTrials.gov, number NCT01518647. The STreSS-5 cohort study was approved by the Danish Data Protection Agency.

It was not appropriate to involve patients or the public in the design, or conduct, or reporting, or dissemination plans of our research.

Results

Sample characteristics

Median age in the general population sample was 54 years (IQR: 44-64), and 53.9% were females.

In the primary care sample, mean age was 54.3 years (SD: 17.5), and 62.5% were females.

In the sample from specialized setting, mean age was 39.4 years (SD: 8.8), and 81.1% were females.

Structural validity

The one-level one factor model showed unacceptable fit indices in all three cohorts (Table 1).

Table 1: Goodness of fit parameters from the CFA models

One-level one factor CFA

	RMSEA	95% CI		CFI	TLI	SRMR	X^2	df	p
General population	0.111	0.110	0.112	0.723	0.697	0.09	32743.1	275	<0.0001
Primary care	0.419	0.147	0.151	0.697	0.670	0.119	15126.5	275	<0.0001
Specialized setting	0.149	0.144	0.153	0.621	0.586	0.115	3261.5	275	<0.0001

One-level four factor CFA

	RMSEA	95% CI		CFI	TLI	SRMR	X^2	df	p
General population	**0.062**	0.061	0.063	**0.914**	**0.905**	**0.052**	10290.96	269	<0.0001
Primary care	0.082	0.080	0.084	**0.91**	**0.90**	**0.067**	4666.85	269	<0.0001
Specialized setting	0.091	0.086	0.096	0.862	0.846	**0.076**	1355.96	269	<0.0001

Two-level four factor CFA

	RMSEA	95% CI		CFI	TLI	SRMR	X^2	df	p
General population	**0.061**	0.060	0.062	**0.917**	**0.908**	**0.052**	9951.02	271	<0.0001
Primary care	0.08	0.078	0.082	**0.914**	**0.905**	**0.068**	4482.39	271	<0.0001
Specialized setting	0.089	0.084	0.093	0.867	0.853	**0.076**	1315.04	271	<0.0001

Bi-factor CFA

	RMSEA	95% CI		CFI	TLI	SRMR	X^2	df	p
General population	**0.048**	0.046	0.049	**0.954**	**0.944**	**0.04**	5680.8	250	<0.0001
Primary care	**0.053**	0.051	0.055	**0.965**	**0.958**	**0.042**	1977.4	250	<0.0001
Specialized setting	**0.059**	0.054	0.065	**0.945**	**0.934**	**0.051**	681.1	250	<0.0001

Abbreviations: CFA=Confirmatory factor analysis; RMSEA=Root Mean Square Error of Approximation; CI=Confidence Interval; CFI=Comparative Fit Index; TLI=Tucker-lewis fit Index; χ^2=Likelyhood Ratio Test; df=degrees of freedom, p=p-value.
Bold: Indicates a good or acceptable fit between the specified model and the observed model in the data.

Fit indices for the one-level four factor model which has been confirmed in previous studies [19 21] and the two-level four factor model showed more acceptable fits. These models revealed correlations between the four BDS symptom clusters and loadings from an underlying BDS factor to the four BDS symptom clusters that may imply a bi-factor model. Good fit indices were seen for the bi-factor model. Hence, a model reflecting a general factor (BDS) and four independent factors (BDS symptom clusters) all explaining the variance of the 25 symptoms in the BDS checklist, was confirmed (Figure 1). Loadings from the general BDS factor were generally higher than loadings from the four symptom clusters; for the population cohort this was the case for 72% of symptoms, in the primary care cohort it accounted for 64% of symptoms, and in the specialized setting it accounted for 52% of symptoms. Loading from the general BDS factor was smaller than loading from the four symptom clusters for six symptoms (*frequent, loose bowel movements; diarrhoea,*

pains in arms and legs; muscular aches or pains; pains in the joints; concentration difficulties) in all three cohorts.

Figure 1 around here

Figure 1: Illustration and factor loadings from the bi-factor model across all three cohorts.

Abbreviations: Gen.=general population; Prim.=primary care; Spec.=specialized clinical setting; CP=cardiopulmonary, GI=gastrointestinal; MS=musculoskeletal; GS=general symptoms; BDS=bodily distress syndrome

Convergent validity

In the general population sample, our hypothesis was met for all measures. The BDS checklist had moderate convergent validity compared to the SF-36 item for overall health (r=0.48, 95% CI: 0.46;0.49, p<0.0001), the four items aggregate score for physical health (r=-0.58, 95% CI: -0.59;-0.56, p<0.0001), the SCL-8 for emotional distress (r=0.52, 95% CI: 0.51;0.54, p<0.0001), and the Whiteley-6-R for illness worry (r=0.53, 95% CI: 0.52;0.55, p<0.0001). Expected difference on the BDS checklist with one unit difference on the four items aggregate score was -0.80 (95% CI: -0.82;-0.78), 12.26 (95% CI: 11.89;12.63) with SCL-8, and 8.93 (95% CI: 8.64;9.21) with WI-6 (Appendix C).

For the primary care sample, our hypothesis was met for all measures as well, however, for some of the measures, the association was stronger than hypothesized. We found moderate convergent validity compared to the SF-36 item for overall health (r=0.58, 95% CI: 0.56;0.61, p<0.0001), the SCL-8 for emotional distress(r=0.62, 95% CI: 0.59;0.64, p<0.0001), and the WI-6 for illness worry (r=0.55, 95% CI: 0.52;0.58, p<0.0001). Expected difference on the BDS checklist with one unit difference on the SCL-8 was 10.60 (95% CI: 10.08;11.12) and 10.01 (95% CI: 9.44;10.59) with Whiteley-6-R (Appendix C).

For the sample from specialized setting, our hypothesis about the correlations being weaker in the specialized setting was met. Moderate convergent validity was seen with emotional distress (r=0.47, 95% CI: 0.40;0.54, p<0.0001) while weaker correlations were seen for overall health (r=0.25, 95% CI: 0.17;0.33, p<0.0001), physical health (r=-0.22, 95% CI: -0.30;-0.12, p<0.0001), and illness worry (r=0.36, 95% CI: 0.28;0.43, p<0.0001). Expected difference on the BDS checklist with one unit difference on the four items aggregate score for physical health was -0.41 (95% CI: -0.56;-0.26), 7.92 (95% CI: 6.65;9.18) with SCL-8, and 5.88 (95% CI: 4.58;7.17) with Whiteley-6-R (Appendix C).

Response distributions and acceptability

BDS checklist item and scale characteristics are displayed in Table 2. Item means varied from 0.15-1.09 in the general population sample, from 0.31-1.53 in the primary care sample, and from 0.81-3.34 in the sample from specialized setting. While the item with the lowest mean varied across samples, the item *'excessive fatigue*' had the highest mean value in all samples. Most item-rest correlations exceeded 0.4.

Table 2: Item and scale characteristics

Item	General population (n=9656)			Primary care (n=2480)			Specialized setting (n=492)		
	Missing %	Mean (SD)	Item-rest correlation	Missing %	Mean (SD)	Item-rest correlation	Missing %	Mean (SD)	Item-rest correlation
Palpations/heart pounding	0.9	0.45 (0.74)	0.427	3.1	0.61 (0.93)	0.544	0.2	1.46 (1.25)	0.480
Precordial discomfort	1.1	0.29 (0.61)	0.410	2.9	0.46 (0.81)	0.517	0.2	1.09 (1.16)	0.391
Breathlessness without exertion	1.0	0.36 (0.71)	0.426	2.9	0.63 (0.99)	0.509	0.2	1.27 (1.25)	0.511
Hyperventilation	1.1	0.15 (0.47)	0.321	3.3	0.37 (0.81)	0.448	0.2	0.93 (1.19)	0.380
Hot and cold sweats	1.2	0.46 (0.80)	0.429	3.2	0.64 (0.96)	0.533	0.2	1.88 (1.35)	0.523
Dry mouth	1.2	0.39 (0.76)	0.432	3.4	0.59 (0.98)	0.501	0.2	1.33 (1.36)	0.442
Frequent loose bowel movements	1.0	0.65 (0.86)	0.403	3.4	0.61 (0.95)	0.391	0.2	1.41 (1.32)	0.357
Abdominal pains	1.4	0.48 (0.76)	0.511	3.5	0.57 (0.90)	0.548	0.2	1.81 (1.22)	0.491
Feeling bloated/full of gas/distended	1.1	0.74 (0.91)	0.524	2.9	0.78 (1.03)	0.532	0.2	2.09 (1.30)	0.484
Diarrhoea	1.1	0.33 (0.63)	0.387	3.5	0.37 (0.80)	0.361	0.2	0.81 (1.13)	0.348
Regurgitations	1.4	0.43 (0.74)	0.392	3.3	0.35 (0.74)	0.456	0.2	1.05 (1.09)	0.430
Nausea	0.9	0.26 (0.57)	0.465	2.9	0.50 (0.87)	0.564	0.2	1.73 (1.27)	0.402
Burning sensation of the upper part of stomach/epigastrium	0.9	0.31 (0.68)	0.441	3.1	0.31 (0.72)	0.483	0.2	1.11 (1.26)	0.512
Pains in arms or legs	1.0	0.87 (1.08)	0.538	3.1	1.21 (1.29)	0.563	0.2	2.68 (1.24)	0.472
Muscular aches or pains	1.3	0.98 (1.01)	0.572	3.2	1.30 (1.23)	0.584	0.2	2.96 (1.10)	0.485
Pains in the joints	1.6	0.96 (1.07)	0.490	3.9	1.20 (1.27)	0.560	0.2	2.57 (1.32)	0.491
Feeling of paresis or localized weakness	1.4	0.16 (0.55)	0.365	4.0	0.33 (0.83)	0.460	0.2	1.22 (1.40)	0.481
Back ache	1.3	1.00 (1.06)	0.492	3.3	1.21 (1.29)	0.542	0.2	2.49 (1.37)	0.377
Pain moving from one place to another	1.4	0.27 (0.71)	0.489	3.8	0.54 (0.99)	0.544	0.2	2.13 (1.48)	0.403
Unpleasant numbness or tingling sensations	1.3	0.25 (0.67)	0.410	4.0	0.34 (0.83)	0.475	0.2	2.00 (1.43)	0.528
Concentration difficulties	0.7	0.60 (0.82)	0.545	2.9	0.85 (1.05)	0.540	0.2	2.53 (1.11)	0.437
Excessive fatigue	0.7	1.09 (1.01)	0.614	2.3	1.53 (1.20)	0.625	0.2	3.34 (0.83)	0.418
Headache	0.8	0.66 (0.89)	0.455	2.8	0.89 (1.08)	0.489	0.2	2.25 (1.22)	0.326
Impairment of memory	0.7	0.60 (0.83)	0.517	2.7	0.80 (1.06)	0.521	0.2	2.29 (1.27)	0.476
Dizziness	0.8	0.34 (0.67)	0.491	2.5	0.58 (0.94)	0.553	0.2	1.75 (1.30)	0.505
Scale									
Total scale missing (%)		0.6			2.7			0.2	
Mean (SD)		13.03 (10.36)			17.33 (13.79)			46.15 (15.91)	
Percentiles									
5%		1			2			22	
10%		3			3			26	
25%		6			7			34	
50%		11			14			45	
75%		18			24			57	
90%		27			37			67	
		34			45			73	

Abbreviations: SD=standard deviation; IQR=interquartile range

Internal consistency was good in all three samples: α=0.887 in the general population sample, α=0.908 in the primary care sample, and α=0.879 in the sample from specialized setting.

BDS score distribution differed across samples (Figure 2) as did total sum score means; it was lowest in the general population (13.03, SD: 10.36) and highest in specialized setting (46.15, SD: 15.91) (Table 2). Acceptability was good, and the numbers of missing responses were generally low in the general population (total 0.6%) and specialized setting (total 0.2%) while it was slightly higher in primary care (total 2.7%).

Figure 2 around here.

Figure 2: Distribution of the BDS total sum score across all three cohorts.

BDS total sum scores were grouped into five categories: 0-20, 21-40, 41-60, 61-80, and 81-100. The vast majority of the general population respondents (96.6%) and primary care patients (90%) scored below 41, while this was only the case for a smaller fraction of the patients from specialized setting (38.7%) (Table 3). Data from each of the three samples and for all three samples pooled together are shown as cumulative percentages across sex and age groups in Tables 1-4 in Appendix D.

Table 3: Grouping of BDS scores across samples

	General population		Primary care		Specialized setting	
Categories of BDS score	n	%	n	%	n	%
0-20	7.762	80.4	1.617	65.2	20	4.1
21-40	1.607	16.6	616	24.8	170	34.6
41-60	208	2.2	156	6.3	204	41.5
61-80	18	0.2	23	0.9	87	17.7
81-100	0	0	2	0.1	10	2.0
Missing	61	0.6	66	2.7	1	0.2

Abbreviations: BDS=Bodily distress syndrome

More detailed information about the normative data from each of the three samples and for all three samples pooled together are shown in Tables 1-4 in Appendix D.

Discussion

Principal findings

This is the first study to establish that, despite some multi-dimensionality, the 25-items BDS checklist can be used as a continuous score to measure symptom burden and illness severity in the general population, in primary care, and in specialized settings. Used as a total sum score with a range from 0-100, the BDS checklist had acceptable convergent validity with measures of overall health, physical health, emotional distress, and illness worry. Internal consistency was good in all three cohorts ($\alpha \geq 0.879$) as was acceptability. Thus, the BDS checklist may work as a simple symptom checklist but also as a diagnostic screening tool for use in clinical work and in research across different settings.

We found the symptom *'excessive fatigue*' to have the highest mean value in all three cohorts. This is in line with a recent German population-based study, finding *'tiredness*' to be one of the leading symptoms [44].

Previous studies have argued that the best fitting model for the BDS checklist was a one-level four factor model (Appendix B) [19 21 24]. However, the objectives of these studies were to confirm the BDS as case finding instrument in other samples with inspiration from the original studies in which the concept of BDS was developed and initially tested [18 19]. In the current study, we have taken it several steps further and tested various structural equation models in three different populations at the same time. The indicators of a bi-factor model is 1) if inter-correlation between the sub-scales in the CFA exceeds 0.3, 2) if loading on the first order factors on the second order factors exceeds 0.5 [45], and 3) if the ratio between the first and second eigenvalues exceeds 3 [46]. All parameters were fulfilled in the general population cohort and in the primary care cohort. For the cohort from specialized setting the ratio between the first and second eigenvalues was 2.68, but otherwise the parameters were fulfilled. This implies that the results from this study do not disqualify results from previous research, but the presence of some multidimensionality is not strong enough to disqualify the interpretation of the BDS checklist as unidimensional as well.

Correlations between the BDS checklist and self-rated measures of overall health, physical health, emotional distress, and illness worry were generally moderate, especially in the general population and primary care cohort. This was as expected as previous literature has shown the same association between symptom load and reduced function [6 7] . The difference between results on patients in the specialised settings and the two other populations may be caused by the nature of self-reported

measures, where patients in specialized setting still have the opportunity to rate their perceived health as excellent even though they have been referred to specialized medical care because of invalidating physical symptoms. These aspects may produce precision limitations in some settings and may especially be pronounced in smaller samples. Furthermore, the distribution of sex differs across populations which may affect the results on convergent validity.

Strengths/weakness of the study

A major strength of this study is the inclusion of three different populations. To our knowledge, this approach of testing an instrument and using the same methodology in different populations is rare as most other studies concern only one setting at a time [11 12 14 17]. Also, the sample size within each cohort was large. We conducted a thorough validation procedure, using different structural equation models and testing convergent validity to several valid measures.

Weaknesses of the study include: Only self-reported outcomes were used and data measures were not completely consistent across the included cohorts; hence, we chose to apply the intersection of items in order to gain equivalent proxy measures. We did not compare the BDS checklist to other measures of physical symptoms. Finally, as this study had a cross-sectional design, it was not possible to evaluate responsiveness of the BDS checklist.

Diff. in results compared to others

To our knowledge, this is the first study to address the usefulness of the BDS checklist as a measure of physical symptom burden and illness severity. Another symptom checklist which has been widely used within primary care and general population studies for measuring the severity of physical symptoms is the 15-items Patient Health Questionnaire (PHQ-15) [11 17]. It consists of 15 items concerning some of the symptoms from the same four organ systems as the BDS checklist, plus the symptoms 'fainting', 'sleeping problems', 'menstrual problems', and 'sexual pains/problems' not included in the BDS checklist. The PHQ-15 is scored on a three-point rating scale from 'not bothered at all' (0) to 'bothered at lot' (2), whereas the BDS checklist uses a five-point rating scale. In one study, including a sample from the general Swedish population, factor analyses of the structural properties of the PHQ-15 showed a four-factor model, but on the basis on a scree test plot they finally concluded that only one factor should be extracted [47]. Other studies found a bi-factor model to have the best fit to the PHQ-15 [48 49]. Hence, the PHQ-15 may have the same structural

properties as the BDS checklist, but with fewer items to take into account as well as fewer response categories which may make it more prone to floor and ceiling effects. In a shorter version of the PHQ-15, the Somatic Symptoms Scale-8 (SSS-8), the three-point rating scale is replaced with a five-point rating option as in the BDS checklist [12 13 50]. However, neither the PHQ-15 nor the SSS-8 is validated for use as diagnostic categorization of respondents. Other symptom questionnaires resembling the same four factor structure and the same five answer categories as the BDS checklist are the 24-items Giessen Subjective Complaints List and its newer shortened version with 8 items (GBB-8), however, they have only been established and used in German speaking countries[14].
The BDS checklist is, at present, the only symptom checklist providing both diagnostic categorization and a measure of symptom load/illness severity.

Clinical implications

This study provides a self-reported symptom checklist for measuring symptom burden and illness severity which can be used both as a diagnostic screening tool and as a measure of illness severity in large epidemiological studies and also in more selected patient samples and severely ill patients. Regarding FSD, previous research has suggested measures of symptom burden as the primary outcome [33]. However, the current study shows that the BDS checklist shows weaker correlation with measures of overall health, physical health, emotional distress, and illness worry in patients from highly specialized setting than in the general population and primary care. Hence, a simple count of bothersome symptoms may not be adequate when dealing with the more severely ill patients, as symptom burden may not be the only important domain of illness severity – others may be the level of impairment and mental morbidity.
Currently, it is unclear whether the here presented BDS total sum score reliably captures FSD illness severity than the distinction in single vs. multi-organ BDS (e.g. three vs. four clusters fulfilled). Nevertheless, a tool which is also able to measure severity of specific symptom clusters is helpful in specialized settings, as it is possible to elucidate which symptom cluster is experienced most bothersome by the patients.

Future research/perspectives

In this study we suggest the BDS checklist as a prominent tool as it can be used both as a measure of symptom burden and as a diagnostic screening tool for FSD, and we argue for its usefulness in

both epidemiological and clinical research as well as in clinical practice. However, the criterion validity of the self-reported BDS checklist with physician's established diagnoses e.g. specialty-specific syndrome diagnoses and psychiatric diagnoses, is yet to be investigated across settings, and future studies regarding these aspects would be valuable in order to further establish the usefulness of the BDS checklist. Moreover, the additional value of counting the number of symptom clusters fulfilled in the staging of FSD deserves attention. Finally, we need a valid instrument to measure change over time, and the responsiveness of the BDS checklist sum score is worth exploring.

Acknowledgements

The authors would like to thank senior research Grete Moth and data manager Kaare Rud Flarup from the Research Unit of primary care for providing date from the primary care cohort. We are also thankful to the members of the DanFunD steering committee; Senior consultant MD PhD Lene Falgaard Eplov, MSc PhD Allan Linneberg, MSc PhD Susanne Brix Pedersen, and MD PhD Michael Eriksen Benros.

Contributors

MWP contributed to the conception and design of the study and the statistical analyses, interpreted the data, and drafted the article. AS and MR contributed to the conception and design of the study and interpretation of the data and provided general supervision of the work. EØ performed the statistical analyses, and contributed to the conception and design of the study and the interpretation of the data. TJ, TMD and PF contributed to the interpretation of the data. All authors contributed to critically revising the article for important intellectual content, and all authors read and approved the final version of the article.

Funding

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests

The authors declare no competing interests.

Data availability statement

Data are available on reasonable request from the corresponding author.

References

1. Henningsen P, Gundel H, Kop WJ, et al. Persistent Physical Symptoms as Perceptual Dysregulation: A Neuropsychobehavioral Model and Its Clinical Implications. *Psychosomatic medicine* 2018;80(5):422-31. doi: 10.1097/PSY.0000000000000588 [doi]
2. Kroenke K. A practical and evidence-based approach to common symptoms: a narrative review. *Annals of Internal Medicine* 2014;161(8):579-86. doi: 10.7326/M14-0461 [doi]
3. Henningsen P, Zipfel S, Sattel H, et al. Management of Functional Somatic Syndromes and Bodily Distress. *Psychotherapy and psychosomatics* 2018;87(1):12-31. doi: 10.1159/000484413 [doi]
4. Eliasen M, Kreiner S, Ebstrup JF, et al. Somatic Symptoms: Prevalence, Co-Occurrence and Associations with Self-Perceived Health and Limitations Due To Physical Health - A Danish Population-Based Study. *PloS one* 2016;11(3):e0150664. doi: 10.1371/journal.pone.0150664 [doi]
5. Tomenson B, Essau C, Jacobi F, et al. Total somatic symptom score as a predictor of health outcome in somatic symptom disorders. *The British journal of psychiatry : the journal of mental science* 2013;203(5):373-80. doi: 10.1192/bjp.bp.112.114405 [doi]
6. Bruusgaard D, Tschudi-Madsen H, Ihlebaek C, et al. Symptom load and functional status: results from the Ullensaker population study. *BMC Public Health* 2012;12:1085. doi: 10.1186/1471-2458-12-1085 [published Online First: 2012/12/20]
7. Creed FH, Davies I, Jackson J, et al. The epidemiology of multiple somatic symptoms. *Journal of psychosomatic research* 2012;72(4):311-17. doi: 10.1016/j.jpsychores.2012.01.009 [doi]
8. Ladwig KH, Marten-Mittag B, Lacruz ME, et al. Screening for multiple somatic complaints in a population-based survey: does excessive symptom reporting capture the concept of somatic symptom disorders? Findings from the MONICA-KORA Cohort Study. *Journal of psychosomatic research* 2010;68(5):427-37. doi: 10.1016/j.jpsychores.2010.01.009 [doi]
9. Eliasen M, Schröder A, Fink P, et al. A step towards a new delimitation of functional somatic syndromes: A latent class analysis of symptoms in a population-based cohort study. *J Psychosom Res* 2018;108:102-17.

10. Derogatis LR, Lipman RS, Rickels K, et al. The Hopkins Symptom Checklist (HSCL): a self-report symptom inventory. *Behavioral science* 1974;19(1):1-15.
11. 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):258-66.
12. Gierk B, Kohlmann S, Kroenke K, et al. The somatic symptom scale-8 (SSS-8): a brief measure of somatic symptom burden. *JAMA internal medicine* 2014;174(3):399-407. doi: 10.1001/jamainternmed.2013.12179 [doi]
13. Gierk B, Kohlmann S, Toussaint A, et al. Assessing somatic symptom burden: a psychometric comparison of the patient health questionnaire-15 (PHQ-15) and the somatic symptom scale-8 (SSS-8). *Journal of psychosomatic research* 2015;78(4):352-55. doi: 10.1016/j.jpsychores.2014.11.006 [doi]
14. Kliem S, Lohmann A, Klatt T, et al. Brief assessment of subjective health complaints: Development, validation and population norms of a brief form of the Giessen Subjective Complaints List (GBB-8). *J Psychosom Res* 2017;95:33-43. doi: 10.1016/j.jpsychores.2017.02.003 [published Online First: 2017/03/21]
15. Sitnikova K, Dijkstra-Kersten SMA, Mokkink LB, et al. Systematic review of measurement properties of questionnaires measuring somatization in primary care patients. *J Psychosom Res* 2017;103:42-62. doi: 10.1016/j.jpsychores.2017.10.005 [published Online First: 2017/11/24]
16. Hyland ME, Bacon AM, Lanario JW, et al. Symptom frequency and development of a generic functional disorder symptom scale suitable for use in studies of patients with irritable bowel syndrome, fibromyalgia syndrome or chronic fatigue syndrome. *Chronic Dis Transl Med* 2019;5(2):129-38. doi: 10.1016/j.cdtm.2019.05.003 [published Online First: 2019/08/02]
17. Zijlema WL, Stolk RP, Lowe B, et al. How to assess common somatic symptoms in large-scale studies: a systematic review of questionnaires. *Journal of psychosomatic research* 2013;74(6):459-68. doi: 10.1016/j.jpsychores.2013.03.093 [doi]
18. Fink P, Toft T, Hansen MS, et al. Symptoms and syndromes of bodily distress: an exploratory study of 978 internal medical, neurological, and primary care patients. *Psychosomatic medicine* 2007;69(1):30-39. doi: 69/1/30 [pii]
19. Budtz-Lilly A, Fink P, Ornbol E, et al. A new questionnaire to identify bodily distress in primary care: The 'BDS checklist'. *Journal of psychosomatic research* 2015;78(6):536-45. doi: 10.1016/j.jpsychores.2015.03.006 [doi]

20. Rosmalen JG, Tak LM, de Jonge P. Empirical foundations for the diagnosis of somatization: implications for DSM-5. *Psychological medicine* 2011;41(6):1133-42. doi: 10.1017/S0033291710001625 [doi]

21. Petersen MW, Schröder A, Jørgensen T, et al. The unifying diagnostic construct of bodily distress syndrome (BDS) was confirmed in the general population. *Journal of psychosomatic research* 2019;128:109868. doi: S0022-3999(19)30601-4 [pii]

22. Burton C, Fink P, Henningsen P, et al. Functional somatic disorders: discussion paper for a new common classification for research and clinical use. *BMC Med* 2020;18(1):34. doi: 10.1186/s12916-020-1505-4 [published Online First: 2020/03/04]

23. Petersen MW, Schröder A, Jørgensen T, et al. Irritable bowel, chronic widespread pain, chronic fatigue and related syndromes are prevalent and highly overlapping in the general population: DanFunD. *Scientific reports* 2020;10(1):3273-73. doi: 10.1038/s41598-020-60318-6

24. Schmalbach B, Roenneberg C, Hausteiner-Wiehle C, et al. Validation of the German version of the Bodily Distress Syndrome 25 checklist in a representative German population sample. *Journal of Psychosomatic Research* 2020;132:109991. doi: https://doi.org/10.1016/j.jpsychores.2020.109991

25. Dantoft TM, Ebstrup JF, Linneberg A, et al. Cohort description: The Danish Study of Functional Disorders. *Clin Epidemiol* 2017;9:127-39.

26. Moth G, Vedsted P, Olesen F, et al. Kontakt- og sygdomsmønsterundersøgelse KOS 2008: Forskningsenheden for Almen Praksis 2010.

27. Frølund Pedersen H, Holsting A, Frostholm L, et al. "Understand your illness and your needs": Assessment-informed patient education for people with multiple functional somatic syndromes. *Patient Educ Couns* 2019;102(9):1662-71. doi: 10.1016/j.pec.2019.04.016 [published Online First: 2019/04/27]

28. Agger JL, Schröder A, Gormsen LK, et al. Imipramine versus placebo for multiple functional somatic syndromes (STreSS-3): a double-blind, randomised study. *Lancet Psychiatry* 2017;4(5):378-88. doi: 10.1016/S2215-0366(17)30126-8 [published Online First: 2017/04/15]

29. Pedersen HF, Agger JL, Frosthom L, et al. Acceptance and Commitment group Therapy for patients with multiple functional somatic syndromes: a three-armed trial comparing ACT in

a brief and extended version with enhanced care. *Psychological medicine* 2018:1-10. doi: 10.1017/S0033291718001666.

10.1017/S0033291718001666 [doi]

30. Schröder A, Rehfeldt E, Ørnbøl E, et al. Cognitive-behavioural group treatment for a range of functional somatic syndromes: randomised trial. *Br J Psychiatry* 2012;200(6):499-507.
31. Fjorback LO, Arendt M, Ørnbøl E, et al. Mindfulness therapy for somatization disorder and functional somatic syndromes: randomized trial with one-year follow-up. *Journal of psychosomatic research* 2013;74(1):31-40. doi: 10.1016/j.jpsychores.2012.09.006 [doi]
32. Ware JJ, Kosinski, M., Gandek B. SF-36 Health Survey: Manual and Interpretation Guide. *Quality Metric* 1993
33. Rief W, Burton C, Frostholm L, et al. Core Outcome Domains for Clinical Trials on Somatic Symptom Disorder, Bodily Distress Disorder, and Functional Somatic Syndromes: European Network on Somatic Symptom Disorders Recommendations. *Psychosom Med* 2017;79(9):1008-15. doi: 10.1097/psy.0000000000000502 [published Online First: 2017/07/12]
34. Schröder A, Ørnbøl E, Licht RW, et al. Outcome measurement in functional somatic syndromes: SF-36 summary scores and some scales were not valid. *Journal of clinical epidemiology* 2012;65(1):30-41. doi: 10.1016/j.jclinepi.2011.03.012 [doi]
35. Derogatis LR. SCL-90-R, administration, scoring and procedures manual-II for the R(evised) version and other instruments of the Psychopathology Rating Scale Series. Townson: Clinical Psychometric Research 1992.
36. Christensen KS BP, Fink P. Measuring Mental Health by Questionnaires in Primary Care - Unidimensionality. Responsiveness and Compliance. . *European Psychiatric Review* 2010;3:8-12.
37. Carstensen TØ, E. Fink, P. Pedersen, MM. Jørgensen, T. Dantoft, TM. Benros, ME. Frostholm, L. Detection of illness worry in the general population: A specific item on illness rumination improves the Whiteley Index - In review. *Journal of psychosomatic research* 2020
38. Mokkink LB, Terwee CB, Patrick DL, et al. COSMIN Checklist Manual. Amsterdam: EMGO Institute for Health and Care Research 2012.
39. StataCorp. Stata Statistical Software. Release 16: College Station, TX: StataCorp LLC 2019.
40. Muthén LK, Muthén BO. Mplus User's Guide. Eighth Edition. Los Angeles, CA: Muthén & Muthén1998.

41. Hu Lt, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal* 1999;6(1):1-55. doi: 10.1080/10705519909540118

42. Ware J, Jr., Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Medical care* 1996;34(3):220-33.

43. Terluin B, van Marwijk HW, Ader HJ, et al. The Four-Dimensional Symptom Questionnaire (4DSQ): a validation study of a multidimensional self-report questionnaire to assess distress, depression, anxiety and somatization. *BMC Psychiatry* 2006;6:34. doi: 10.1186/1471-244X-6-34 [published Online First: 2006/08/24]

44. Beutel ME, Klein EM, Henning M, et al. Somatic Symptoms in the German General Population from 1975 to 2013. *Sci Rep* 2020;10(1):1595. doi: 10.1038/s41598-020-58602-6 [published Online First: 2020/02/02]

45. MD. HJT. Bifactor analysis in MPlus [Video file]: UK College of Education; 2016 [Available from: https://education.uky.edu/edp/apslab/events/2020.

46. Cho SJ, Wilmer J, Herzmann G, et al. Item response theory analyses of the Cambridge Face Memory Test (CFMT). *Psychol Assess* 2015;27(2):552-66. doi: 10.1037/pas0000068 [published Online First: 2015/02/03]

47. Nordin S, Palmquist E, Nordin M. Psychometric evaluation and normative data for a Swedish version of the Patient Health Questionnaire 15-Item Somatic Symptom Severity Scale. *Scandinavian Journal of Psychology* 2013;54(2):112-17. doi: 10.1111/sjop.12029 [doi]

48. Cano-Garcia FJ, Munoz-Navarro R, Sese Abad A, et al. Latent structure and factor invariance of somatic symptoms in the patient health questionnaire (PHQ-15). *J Affect Disord* 2020;261:21-29. doi: 10.1016/j.jad.2019.09.077 [published Online First: 2019/10/11]

49. Witthoft M, Hiller W, Loch N, et al. The latent structure of medically unexplained symptoms and its relation to functional somatic syndromes. *International Journal of Behavioral Medicine* 2013;20(2):172-83. doi: 10.1007/s12529-012-9237-2 [doi]

50. Toussaint A, Kroenke K, Baye F, et al. Comparing the Patient Health Questionnaire - 15 and the Somatic Symptom Scale - 8 as measures of somatic symptom burden. *J Psychosom Res* 2017;101:44-50. doi: 10.1016/j.jpsychores.2017.08.002 [published Online First: 2017/09/05]

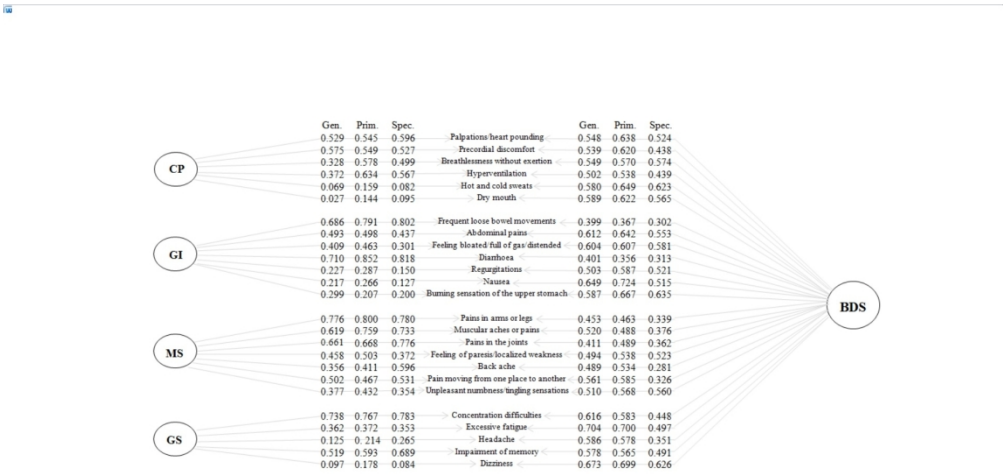


Figure 1: Illustration and factor loadings from the bi-factor model across all three cohorts. Abbreviations: Gen.=general population; Prim.=primary care; Spec.=specialized clinical setting; CP=cardiopulmonary, GI=gastrointestinal; MS=musculoskeletal; GS=general symptoms; BDS=bodily distress syndrome

508x285mm (96 x 96 DPI)

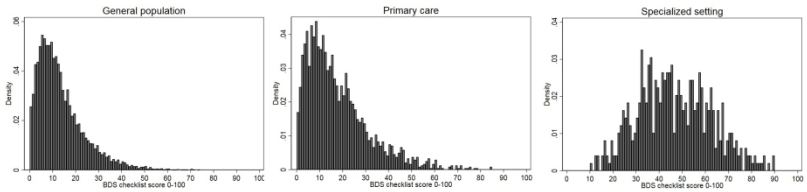


Figure 2: Distribution of the BDS total sum score across all three cohorts.

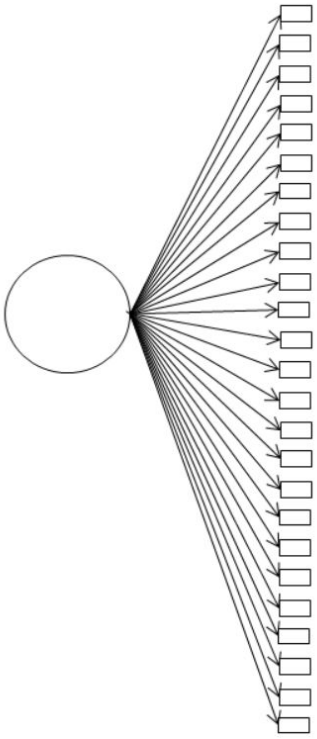
754x311mm (96 x 96 DPI)

Appendix A: The 25-items BDS checklist

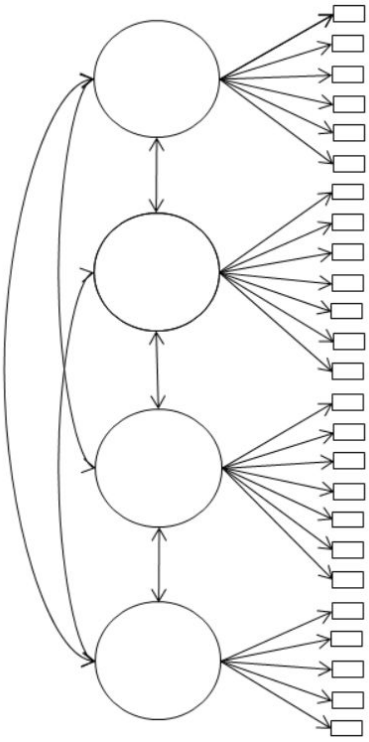
During the last 4 weeks*, have you been bothered by…		Not at all	A bit	Somewhat	Quite a bit	A lot
1	Palpations and heart pounding	□	□	□	□	□
2	Precordial discomfort	□	□	□	□	□
3	Breathlessness without exertion	□	□	□	□	□
4	Hyperventilation	□	□	□	□	□
5	Hot and cold sweats	□	□	□	□	□
6	Dry mouth	□	□	□	□	□
7	Frequent loose bowel movements	□	□	□	□	□
8	Abdominal pains	□	□	□	□	□
9	Feeling bloated/full of gas/distended	□	□	□	□	□
10	Diarrhoea	□	□	□	□	□
11	Regurgitations	□	□	□	□	□
12	Nausea	□	□	□	□	□
13	Burning sensation of the upper part of stomach/epigastrium	□	□	□	□	□
14	Pains in arms or legs	□	□	□	□	□
15	Muscular aches or pains	□	□	□	□	□
16	Pains in the joints	□	□	□	□	□
17	Feeling of paresis or localized weakness	□	□	□	□	□
18	Back ache	□	□	□	□	□
19	Pain moving from one place to another	□	□	□	□	□
20	Unpleasant numbness or tingling sensations	□	□	□	□	□
21	Concentration difficulties	□	□	□	□	□
22	Excessive fatigue	□	□	□	□	□
23	Headache	□	□	□	□	□
24	Impairment of memory	□	□	□	□	□
25	Dizziness	□	□	□	□	□

* This time frame was changed to 12 months in the general population cohort

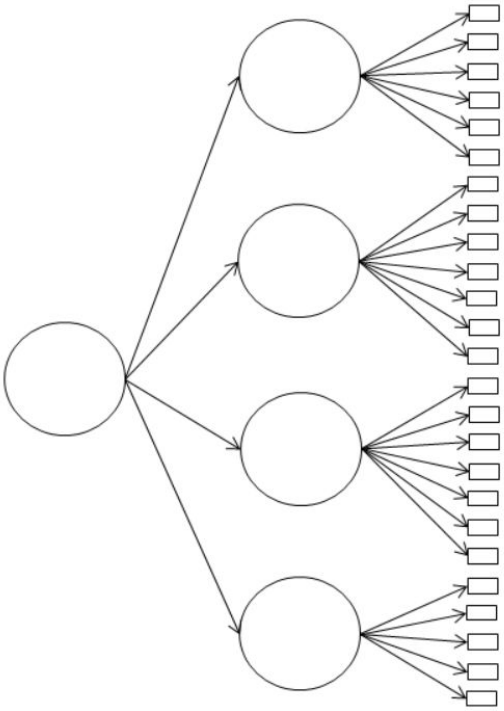
Appendix B: Illustrations of the different models of confirmatory factor analyses



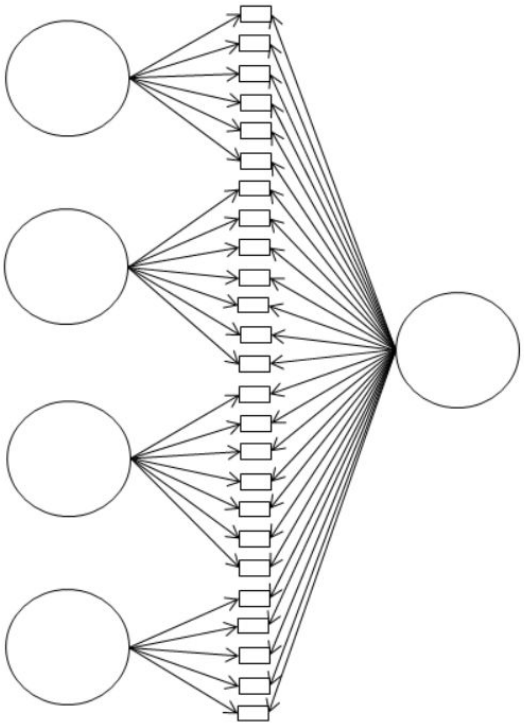
One-level one factor model



One-level four factor model



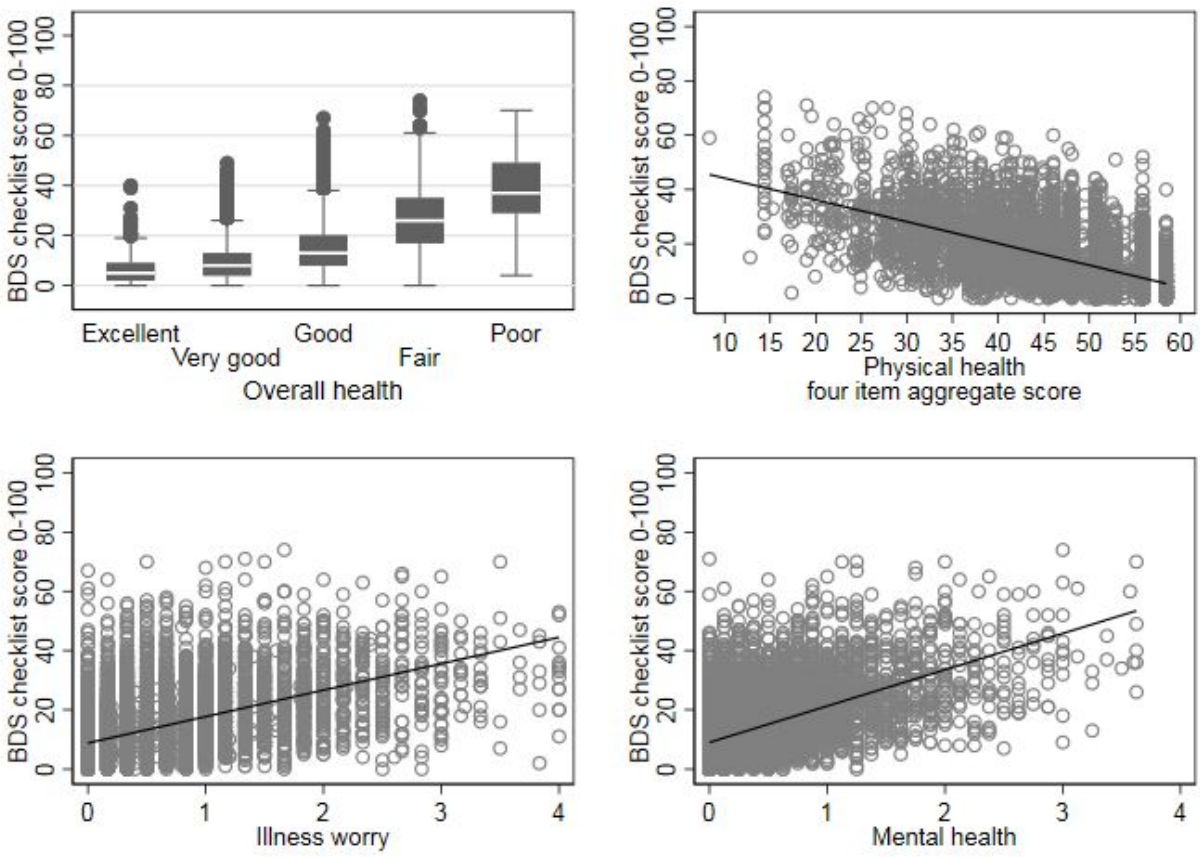
Two-level four factor model



Bi-factor model

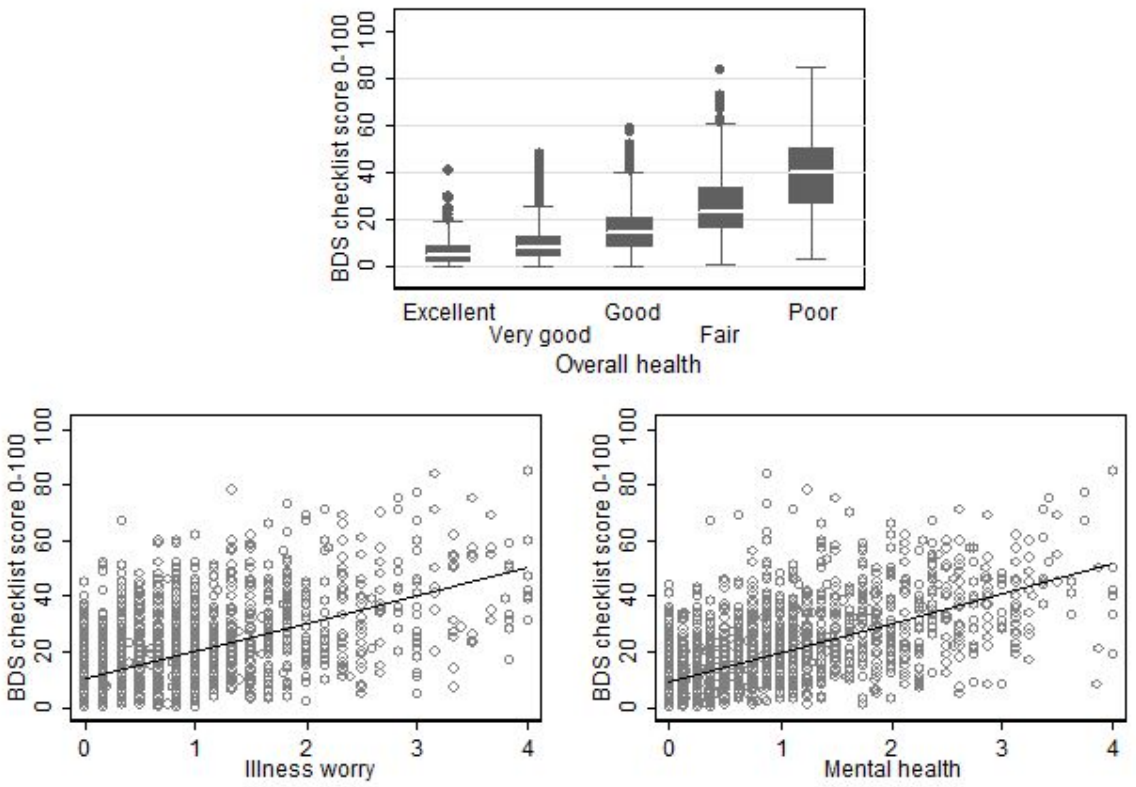
Appendix C: Descriptive plots of associations between the BDS checklist and other measures

General population



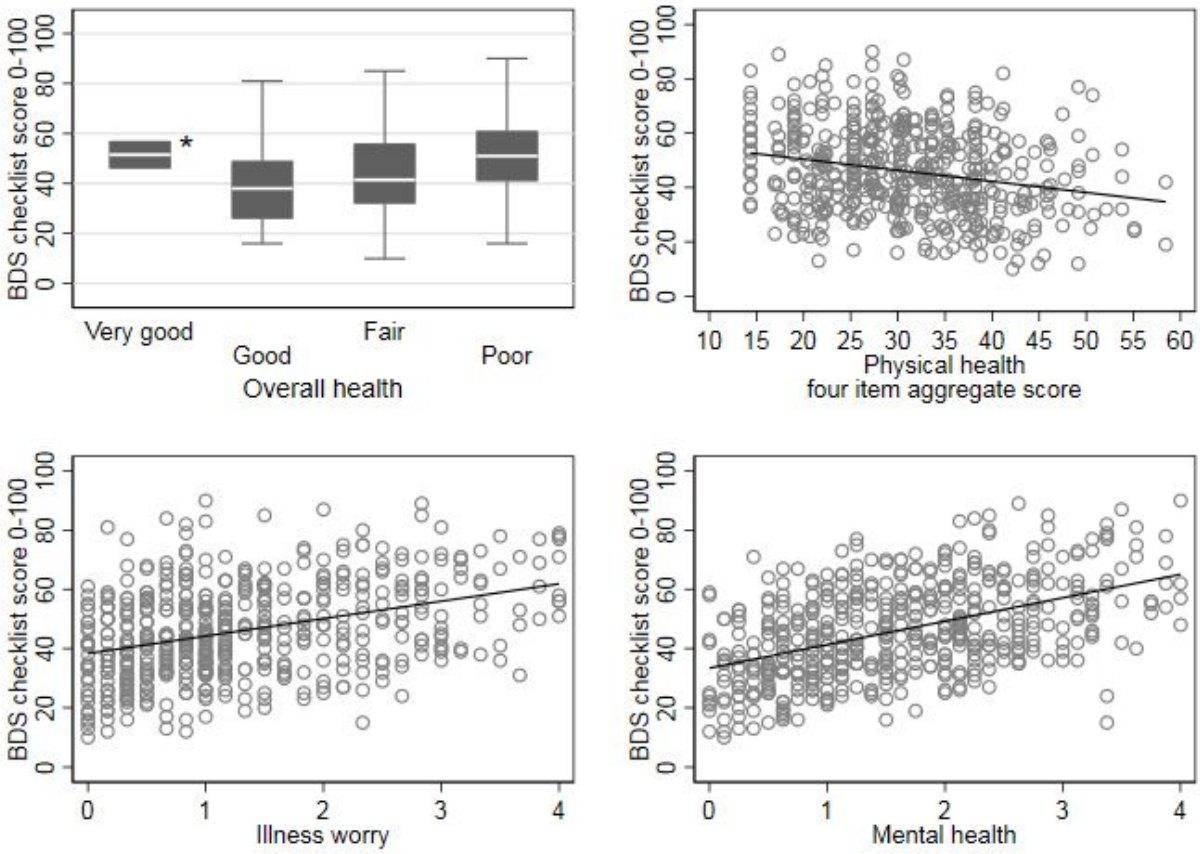
Correlation of the BDS checklist score and measures of overall health, perceived physical health, illness worry, and mental health in the general population cohort.

Primary care



Correlation of the BDS checklist score and measures of overall health, illness worry, and mental health in the primary care cohort.

Specialized setting



Correlation of the BDS checklist score and measures of overall health, perceived physical health, illness worry, and mental health in the cohort from specialized setting.
* Only based on two individuals

Appendix D. Data across sex and age groups

Table 1: Data for the general population cohort (n=9656): Cummulative percentages

	Male				Female			
	18-39	40-49	50-59	60-76	18-39	40-49	50-59	60-76
	N= 719	N=902	N=1.130	N=1.702	N=905	N=1.139	N=1.401	N=1.758
BDS score groups								
0-	21	24	25	28	10	15	10	18
5-	50	55	54	55	34	38	33	42
10-	73	76	71	75	56	59	56	62
15-	85	86	82	86	72	73	71	76
20-	91	92	88	91	83	82	82	85
25-	95	95	93	94	90	89	88	91
30-	97	97	96	97	94	93	93	94
35-	98	97	97	98	96	95	96	96
40-	99	98	98	98	97	97	97	98
45-	99	99	99	99	98	98	98	99
50-	99	99	99	99	98	99	99	99
55-	99	99	99	99	99	99	99	99
60-	99	99	99	99	99	99	100	99
65-	99	99	99	99	99	99	100	99
70-	99	99	99	99	99	99	100	99
75-	99	99	99	99	99	99	100	99
80-	99	99	99	99	99	99	100	99
85-	99	99	99	99	99	99	100	99
90-	99	99	99	99	99	99	100	99
95-	99	99	99	99	99	99	100	99
Missing	100	100	100	100	100	100	100	100

Abbreviations: BDS=Bodily distress syndrome

Table 2: Data for the primary care cohort (n=2480): Cummulative percentages

	Male						Female					
	18-39	40-49	50-59	60-76	70-79	80-	18-39	40-49	50-59	60-69	70-79	80-
	N=146	N=147	N= 172	N=245	N=153	N=67	N=404	N=280	N=271	N=287	N=203	N=105
BDS score groups												
0-	23	20	16	18	14	10	16	14	8	16	14	9
5-	49	37	32	40	35	36	37	31	23	32	31	27
10-	67	60	52	58	54	51	53	47	38	45	45	42
15-	73	70	67	72	65	58	66	59	54	58	59	55
20-	79	80	76	83	72	67	76	75	66	73	72	64
25-	89	86	82	90	79	82	80	80	78	82	79	70
30-	92	89	87	95	82	82	86	84	85	85	85	72
35-	93	91	90	96	88	85	90	88	89	89	88	74
40-	97	93	94	98	91	87	93	90	92	93	90	76
45-	97	94	96	98	94	88	95	94	95	95	91	78
50-	98	95	98	98	95	88	96	96	96	96	93	80
55-	99	98	98	100	95	88	96	97	97	97	94	81
60-	99	99	98	100	95	88	97	98	97	97	95	82
65-	99	99	99	100	95	88	97	98	97	98	96	83
70-	99	99	99	100	95	88	98	98	98	98	96	83
75-	99	99	99	100	95	88	98	98	98	98	96	83
80-	99	99	99	100	95	88	98	98	99	98	96	83
85-	99	99	99	100	95	88	98	99	99	98	96	83
90-	99	99	99	100	95	88	98	99	99	98	96	83
95-	99	99	99	100	95	88	98	99	99	98	96	83
Missing	100	100	100	100	100	100	100	100	100	100	100	100

Abbreviations: BDS=Bodily distress syndrome

Table 3: Data for the cohort from specialized setting (n=492): Cummulative percentages

	Male				Female			
	18-39	40-49	50-59	60-69	18-39	40-49	50-59	60-69
	N=42	N=40	N=11	N=0	N=199	N=162	N=36	N=2
BDS score groups								
0-	0	0	0	-	0	0	0	0
5-	0	0	0	-	0	0	0	0
10-	0	3	0	-	2	1	0	0
15-	5	5	9	-	4	2	3	0
20-	12	15	9	-	8	7	6	0
25-	21	23	9	-	14	13	14	0
30-	36	45	36	-	20	25	17	50
35-	48	50	55	-	36	35	19	50
40-	57	55	73	-	50	48	25	50
45-	74	60	82	-	60	59	36	50
50-	81	70	82	-	68	69	56	50
55-	86	85	82	-	79	75	72	50
60-	90	88	82	-	87	84	83	100
65-	95	93	82	-	92	92	83	100
70-	95	98	91	-	95	96	94	100
75-	95	100	100	-	96	99	97	100
80-	95	100	100	-	99	99	97	100
85-	100	100	100	-	100	100	97	100
90-	100	100	100	-	100	100	97	100
95-	100	100	100	-	100	100	97	100
Missing	100	100	100	-	100	100	100	100

Abbreviations: BDS=Bodily distress syndrome

Table 4: Data for the three pooled cohorts (n=12628): Cummulative percentages

	Male						Female					
	18-39	40-49	50-59	60-69	70-79	80-	18-39	40-49	50-59	60-69	70-79	80-
	N=907	N=1089	N=1313	N=1665	N=435	N=67	N=1508	N=1581	N=1708	N=1785	N=465	N=105
BDS score groups												
0-	20	23	24	26	26	10	10	13	9	18	15	9
5-	47	51	50	52	50	36	31	33	31	41	37	27
10-	69	71	68	71	70	51	48	51	52	59	53	42
15-	80	81	80	83	80	58	61	63	67	73	67	55
20-	86	88	86	90	86	67	71	73	78	83	79	64
25-	91	91	91	94	89	82	77	80	85	89	85	70
30-	94	94	94	97	91	82	82	84	90	93	89	72
35-	95	95	96	98	94	85	86	88	93	95	92	74
40-	97	96	97	99	95	87	90	91	95	97	94	76
45-	97	97	98	99	97	88	92	93	96	98	95	78
50-	98	97	99	99	97	88	94	95	98	99	96	80
55-	98	98	99	99	97	88	95	96	98	99	97	81
60-	99	99	99	99	97	88	97	97	99	99	97	82
65-	99	99	99	99	97	88	98	98	99	99	98	83
70-	99	99	99	99	97	88	98	99	99	99	98	83
75-	99	99	99	99	97	88	99	99	99	99	98	83
80-	99	99	99	99	97	88	99	99	99	99	98	83
85-	99	99	99	99	97	88	99	99	99	99	98	83
90-	99	99	99	99	97	88	99	99	99	99	98	83
95-	99	99	99	99	97	88	99	99	99	99	98	83
Missing	100	100	100	100	100	100	100	100	100	100	100	100

Abbreviations: BDS=Bodily distress syndrome

STROBE 2007 (v4) Statement—Checklist of items that should be included in reports of *cross-sectional studies*

Section/Topic	Item #	Recommendation	Reported on page #
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	2
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4-5
Objectives	3	State specific objectives, including any prespecified hypotheses	5
Methods			
Study design	4	Present key elements of study design early in the paper	5
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	5
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants	5
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	5-8
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	5-8
Bias	9	Describe any efforts to address potential sources of bias	
Study size	10	Explain how the study size was arrived at	5
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	6-8
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	7-8
		(*b*) Describe any methods used to examine subgroups and interactions	
		(*c*) Explain how missing data were addressed	
		(*d*) If applicable, describe analytical methods taking account of sampling strategy	
		(*e*) Describe any sensitivity analyses	
Results			

Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	
		(b) Give reasons for non-participation at each stage	
		(c) Consider use of a flow diagram	
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	5
		(b) Indicate number of participants with missing data for each variable of interest	
Outcome data	15*	Report numbers of outcome events or summary measures	
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	8-13
		(*b*) Report category boundaries when continuous variables were categorized	8-13
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	
Discussion			
Key results	18	Summarise key results with reference to study objectives	14
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	15
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	15-16
Generalisability	21	Discuss the generalisability (external validity) of the study results	14-16
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	17

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

BMJ Open

Original research: The BDS checklist as measure of illness severity: A cross-sectional cohort study in the general population, primary care and specialized setting

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2020-042880.R1
Article Type:	Original research
Date Submitted by the Author:	20-Oct-2020
Complete List of Authors:	Petersen, Marie; Aarhus Universitetshospital, The Research Clinic for Functional Disorders and Psychosomatics Rosendal, Marianne; Aarhus University Hospital, Research Clinic for Functional Disorders and Psychosomatics Ørnbøl, Eva Fink, Per Jørgensen, Torben; Frederiksberg Hospital, Dantoft, Thomas; Region Hovedstaden Forskningscenter for forebyggelse og sundhed, Schroeder, Andreas; Aarhus Universitetshospital,
Primary Subject Heading:	Research methods
Secondary Subject Heading:	Diagnostics, Epidemiology, Mental health
Keywords:	STATISTICS & RESEARCH METHODS, EPIDEMIOLOGY, PSYCHIATRY, GENERAL MEDICINE (see Internal Medicine)

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Original research: The BDS checklist as measure of illness severity: A cross-sectional cohort study in the general population, primary care and specialized setting

Marie Weinreich Petersen[1], Marianne Rosendal[1,2], Eva Ørnbøl[1], Per Fink[1], Torben Jørgensen[3,4,5], Thomas Meinertz Dantoft[3], Andreas Schroeder[1]

Author affiliations:
[1] The Research Clinic for Functional Disorders and Psychosomatics, Aarhus University Hospital, Aarhus C, Denmark
[2] Research Unit for General Practice, Aarhus University, Aarhus C, Denmark
[3] Center for Clinical Research and Prevention, Bispebjerg and Frederiksberg Hospital, The Capital Region of Denmark, Denmark
[4] Department of Public Health, Faculty of Health and Medical Science, University of Copenhagen, Copenhagen C, Denmark
[5] Faculty of Medicine, Aalborg University

Corresponding author:
Marie Weinreich Petersen
The Research Clinic for Functional Disorders and Psychosomatics
Aarhus University Hospital
Nørrebrogade 44, building 2C, second floor
8200 Aarhus C
Denmark
Phone: +45 29467824
Fax: +45 7846 4340
E-mail: mawept@rm.dk

Word count in main text: 4221
Word count abstract: 300
Number of references: 50

Abstract

Objectives

The Bodily Distress Syndrome (BDS) checklist has proven to be useful for diagnostic categorization and screening tool for functional somatic disorders (FSD). This study aims to investigate whether the BDS checklist total sum score (0-100) can be used as measure of physical symptom burden and FSD illness severity.

Design

Cross-sectional.

Setting

Danish general population, primary care, and specialized clinical setting.

Participants

A general population cohort (n=9656), a primary care cohort (n=2480), and a cohort of multi-organ BDS patients from specialized clinical setting (n=492).

Outcome measures

All data were self-reported. Physical symptoms were measured with the 25-items BDS checklist. Overall self-perceived health was measured with one item from the 36-items Short Form Health Survey (SF-36). Physical functioning was measured with an aggregate score of 4 items from the SF-36/SF-12 scales 'physical functioning', 'bodily pain', and 'vitality'. Emotional distress was measured with the mental distress subscale (SCL-8) from the Danish version of the Hopkins Symptom Checklist-90. Illness worry was measured with the 6-items Whiteley Index.

Results

For all cohorts, bi-factor models established that despite some multi-dimensionality, the total sum score of the BDS checklist adequately reflected physical symptom burden and illness severity. The BDS checklist had acceptable convergent validity to measures of overall health (r=0.25-0.58), physical functioning (r=0.22-0.58), emotional distress (r=0.47-0.62), and illness worry (r=0.36-0.55). Acceptability was good with low numbers of missing responses to items (<3%). Internal consistency was high ($\alpha \geq 0.879$). BDS score means varied and reflected symptom burden across cohorts (13.03-46.15). We provide normative data for the Danish general population.

Conclusions

The BDS checklist total sum score can be used as measure of symptom burden and FSD illness severity across settings. These findings establish the usefulness of the BDS checklist in clinic and in research both as a diagnostic screening and as an instrument for assessment of illness severity.

Strengths and limitations of this study

- The study included data from three cohorts and settings: A general population, primary care patients, and patients from a specialized setting
- Well-validated measures were used to determine convergent validity
- All included cohorts had large sample sizes
- Only self-reported measures were included
- Convergent validity was not investigated with other measures of physical symptom burden

Introduction

Persistent physical symptoms (PPS) are common in medical settings and the general population [1-4]. The symptoms present across a continuum from one or a few momentary to numerous symptoms from multiple locations in the body. Having a high number of symptoms has been associated with poor health, poor functional status, and increased health care use [5-9]. Hence, assessment of the burden of persistent physical symptoms is valuable in both clinical care and in research.

For this purpose, self-reported symptom questionnaires are frequently used. They are manageable, non-invasive tools. Several screening questionnaires exist: The Hopkins Symptom Checklist [10], the Patient Health Questionnaire [11], the Somatic Symptom Scale-8 [12 13], the brief form of the Giessen Subjective Complaints List [14], and others [15-17]. However, the existing questionnaire measure PPS without consideration of the well-known aggregation of such symptoms into symptom clusters, and hence, without acknowledgement of the real structure of PPS as it occurs in both the community and in clinical setting[18-21].

When PPS occur in the absence of (other) physical or mental conditions, or when they cause individual suffering and functional limitations beyond what could be expected based on such diseases, they constitute the very core of the disorders captured under the umbrella definition of Functional Somatic Disorders (FSD). FSD cover both specialty-specific syndrome diagnoses such as fibromyalgia, irritable bowel and chronic fatigue, but also their pendants in psychosomatic medicine, somatoform and somatic symptom disorders [22].

In contrast to the above mentioned speciality-specific diagnoses, the proposed research diagnosis bodily distress syndrome (BDS) covers a broader range of functional somatic symptoms ranging from few symptoms with some effect on functioning to severe and disabling functional somatic disorders [18 19 21]. Hence, BDS provides the opportunity to assess and distinguish between conditions persisting as mono- or multi-syndromatic and still within the same framework of diagnostic approach [21 23]. The diagnostic construct was developed in a sample of patients from primary and secondary care, and the 30-items BDS checklist emerged [18]. BDS was confirmed in a new sample of primary care patients where the shortened 25-items BDS checklist was developed [19]. Subsequently, the construct of BDS has been confirmed in general population samples as well[21 24]. BDS presents symptoms grouped in four symptom clusters: Cardiopulmonary (CP), gastrointestinal (GI), musculoskeletal (MS), and general symptoms (GS), and its usefulness and properties used for diagnostic categorisation into no BDS, a single/oligo-organ BDS type and a multi-organ BDS type has been established [19 21 24]. A major strength of the BDS checklist is its usefulness both as a screening and as diagnostic tool within clinical

practise and within epidemiological research [18 19 21 23], but the total BDS sum score has not yet been validated as a measure for the assessment of symptom burden and illness severity.

This study aims to explore whether the BDS checklist can be used as a continuous score to measure symptom burden (i.e. in those individuals that may fall under the diagnostic threshold or what we believe to be clinically relevant) and illness severity (in those individuals fulfilling diagnostic criteria for FSD). In order to elicit the BDS checklist's usability across settings, its structural validity and psychometric properties will be explored in three different populations: the general population, primary care patients, patients in a specialized clinical setting.

Methods

Population

This cross-sectional study included baseline data from three cohorts:

Cohort 1: A general population cohort (DanFunD, n=9656, response rate=33.7%) established with the purpose to investigate and unravel the epidemiology of FSD [25]. The cohort was obtained from the Danish Central Personal Register and drawn as a random sample of the adult Danish background population aged 18-69 years. Participants lived in 10 municipalities in the south-western part of the greater Copenhagen area. All participants were born in Denmark.

Cohort 2: A cohort of primary care patients (KOS, n=2480, response rate=59.5%) established in order to investigate contact and disease patterns in general practice [26]. Participants were included consecutively from 388 general practitioners from the Central Denmark Region. Included participants were 18 years or older and had completed a health-related face-to-face consultation with their general practitioner.

Cohort 3: Data from a specialized clinical setting at the Research Clinic for Functional Disorders and Psychosomatics, Aarhus University Hospital in Denmark (STreSS-3, STreSS-4, STreSS-5, n=492, response rate=100%) [27-31]. These cohorts had been part of a group of studies with the shared aim to investigate new treatments for patients with multi-organ BDS aged 20 years or older.

Measures

Self-reported data of physical symptoms, overall health, physical health, mental health, and illness worry was included. The measures and data were not completely consistent across the three included cohorts.

Physical symptoms were assessed with the Danish version of the 25-items BDS checklist (Appendix A) [19 21]. The checklist asks "during the last *(specific time frame)* have you been bothered by" followed by a list of 25 symptoms comprising the four symptom clusters of BDS. The BDS checklist measures symptoms on a five-point rating scale from 0 ('not at all bothersome') to 4 ('a lot bothersome'). We calculated a sum score by adding the single item scores from the 25 items (ranging from 0 to 100). The timeframe covered was 12 months for the general population cohort and four weeks for the other two cohorts.

Overall health was assessed with a single item from the 36-items Short Form Health Survey (SF-36) [32], estimating self-perceived health on a five-point rating scale from 'excellent' to 'poor'. Higher score on this item indicates poorer health. No specific time frame was surveyed in neither of the cohorts.

Physical functioning was measured with a shortened version of an aggregate score of the SF-36 subscales 'physical functioning', 'bodily pain', and 'vitality [30 32-34]. The shortened version consisted of four items (two items from the 'physical function' subscale, one item from the 'bodily pain' subscale, and one item from the 'vitality' subscale) which are part of the SF-12, addressing limitations in moderate and strenuous activities because of physical health and pain interference. For each item a *z*-score was calculated using mean and standard deviation (SD) from the general Danish population. Mean of the *z*-scores from the three subscales results in an aggregate *z*-score. This is then transformed into a *t*-score (mean=50, SD=10). Higher scores indicate better physical health. We tested the correlation of the *t*-score of the shortened version aggregate score against the full SF-36 aggregate score in cohort 3, and correlation was high (*Spearman rho*=0.89, 95% CI: 0.87;0.91). Unfortunately, it was not possible to investigate convergent validity to the aggregate score in the data on the primary care cohort, because we had limited access to data. These analyses were therefore only performed in the general population cohort and the cohort from specialized clinical setting. The time frame covered was four weeks for both cohorts.

Emotional distress was measured with the mental distress subscale (SCL-8) from the Danish version of the Hopkins Symptom Checklist (SCL-90) [35 36]. SCL-8 consists of eight items addressing impairment of overall worries, depression, and anxiety. Answers were calculated as mean scores from a scale ranging from 0 ('not at all bothersome') to 4 ('a lot bothersome'). Higher scores indicated higher emotional distress. The time frame covered was one week for the general population cohort and four weeks for the two other cohorts.

Illness worry was measured with the Whiteley Index 6 items version revised (Whiteley-6-R) [37], addressing the respondent's fear of being ill and whether they attribute current bodily sensations to

somatic illness[1]. Answers were calculated as mean scores from a scale ranging from 0 ('not at all bothersome') to 4 ('a lot bothersome'). Higher scores indicated higher health anxiety. The time frame covered was 12 months for the general population cohort and four weeks for the two other cohorts.

Validation procedure and statistical analyses

The analyses for the current study were performed according to the Consensus-based Standards for the selection of health Measurement Instrument (COSMIN) framework [38].

All statistical analyses were performed using STATA version 16.0 [39], except for the structural equation modelling which was performed using Mplus version 8.1 [40].

Construct validity was tested by means of structural validity and convergent validity.

Structural validity was tested with confirmatory factor analyses (CFA) with WLSMV (Weighted Least Squares Means and Variance adjusted) estimation due to categorical responses for all items [40]. We wanted to test if it was permissible to model the BDS checklist as unidimensional despite the previous evidence of some multi-dimensionality [18 19 21 24]. Furthermore, we wanted to test if the raw total BDS sum score would be an adequate reliable measure of the general factor (BDS). Therefore, four different CFAs were performed: 1) An one-level one factor model, 2) an one-level four factor model, using factors resembling the four BDS symptom clusters previously reported [19 21], 3) a two-level four factor model, representing a second order common factor (BDS) underlying the four BDS symptom clusters, and 4) a bi-factor CFA, reflecting each symptom to load on a general factor (BDS) and on one of the four specific BDS symptom clusters. Illustrations of the four types of CFAs are displayed in Appendix B.

In all CFAs, model fit were assessed as follows: A Root Mean Square Error of Approximation (RMSEA) <0.05 indicates very good fit, 0.05-0.08 indicates a good fit, and ≥0.08 indicates a poor fit. Comparative fit Index (CFI) and Tucker-Lewis fit Index (TLI) at 0.90-0.95 indicate an acceptable fit and levels >0.95 indicate a good fit. A Standardized Root Mean square Residual (SRMR) <0.08 indicates good fit [41].

Convergent validity was tested with Spearman's correlations, and associations between the BDS checklist and overall health (one item from SF-36)[32], physical function (an aggregate score of four items from the SF-36) [42], emotional distress (SCL-8) [35], and illness worry (Whiteley-6-R)

[1] In the primary care sample, one of the items in the WI-6 "Do you worry about the possibility that you suffer from an illness you have heard or read about" was expressed as "Do you worry about the possibility that you suffer from an illness".

(Carstensen) were performed. Based on previous literature [12 14 15 17 43], we hypothesized that the BDS checklist would show moderate convergent validity (r=0.40-0.60) with the four measures, and we expected lower correlations in the sample from specialized setting. Expected differences on the BDS checklist with one unit difference to the SCL-8, the four items aggregate score for physical functioning, and Whiteley-6-R were estimated with linear regression.

BDS checklist item and scale characteristics, i.e. item means (SD), sum score means, score distribution, item total correlation, corrected for overlap, and aspects of acceptability, i.e. percentage of missing items, were examined and computed as descriptive statistics for each of the three samples. Internal consistency was measured with Cronbach's α coefficients where values between 0.7 and 0.95 are acceptable [38].

Ethical considerations

The current study was carried out in accordance with the relevant guidelines and regulations.

For all three cohorts, written informed consent was obtained from each participant before entering the studies [25-31].

Cohort 1: Approved by the independent ethics committees the Ethical Committee of Copenhagen County (Ethics Committee: KA-2006-0011; H-3-2011-081; H-3-2012-0015) and the Danish Data Protection Agency.

Cohort 2: Approved by the Danish Data Protection Agency and the Health and Medicines Authority. According to Danish law, approval from the health research ethics system was not needed.

Cohort 3: The STreSS-3 trial was registered with ClinicalTrials.gov, number NCT01518634, EudraCT number 2011-004294-87. The STreSS-4 trial was registered with ClinicalTrials.gov, number NCT01518647. The STreSS-5 cohort study was approved by the Danish Data Protection Agency.

Patient and Public Involvement

It was not appropriate to involve patients or the public in the design, or conduct, or reporting, or dissemination plans of our research.

Results

Sample characteristics

Median age in the general population sample was 54 years (IQR: 44-64), and 53.9% were females.

In the primary care sample, mean age was 54.3 years (SD: 17.5), and 62.5% were females.

In the sample from specialized setting, mean age was 39.4 years (SD: 8.8), and 81.1% were females.

Structural validity

The one-level one factor model showed unacceptable fit indices in all three cohorts (Table 1).

Table 1: Goodness of fit parameters from the CFA models

	RMSEA	95% CI		CFI	TLI	SRMR	X^2	df	p
One-level one factor CFA									
General population	0.111	0.110	0.112	0.723	0.697	0.09	32743.1	275	<0.0001
Primary care	0.419	0.147	0.151	0.697	0.670	0.119	15126.5	275	<0.0001
Specialized setting	0.149	0.144	0.153	0.621	0.586	0.115	3261.5	275	<0.0001
One-level four factor CFA									
General population	**0.062**	0.061	0.063	**0.914**	**0.905**	**0.052**	10290.96	269	<0.0001
Primary care	0.082	0.080	0.084	**0.91**	**0.90**	**0.067**	4666.85	269	<0.0001
Specialized setting	0.091	0.086	0.096	0.862	0.846	**0.076**	1355.96	269	<0.0001
Two-level four factor CFA									
General population	**0.061**	0.060	0.062	**0.917**	**0.908**	**0.052**	9951.02	271	<0.0001
Primary care	0.08	0.078	0.082	**0.914**	**0.905**	**0.068**	4482.39	271	<0.0001
Specialized setting	0.089	0.084	0.093	0.867	0.853	**0.076**	1315.04	271	<0.0001
Bi-factor CFA									
General population	**0.048**	0.046	0.049	**0.954**	**0.944**	**0.04**	5680.8	250	<0.0001
Primary care	**0.053**	0.051	0.055	**0.965**	**0.958**	**0.042**	1977.4	250	<0.0001
Specialized setting	**0.059**	0.054	0.065	**0.945**	**0.934**	**0.051**	681.1	250	<0.0001

Abbreviations: CFA=Confirmatory factor analysis; RMSEA=Root Mean Square Error of Approximation; CI=Confidence Interval; CFI=Comparative Fit Index; TLI=Tucker-lewis fit Index; χ^2=Likelyhood Ratio Test; df=degrees of freedom, p=p-value.
Bold: Indicates a good or acceptable fit between the specified model and the observed model in the data.

Fit indices for the one-level four factor model which has been confirmed in previous studies [19 21] and the two-level four factor model showed more acceptable fits. These models revealed correlations between the four BDS symptom clusters and loadings from an underlying BDS factor to the four BDS symptom clusters that may imply a bi-factor model. Good fit indices were seen for the bi-factor model. Hence, a model reflecting a general factor (BDS) and four independent factors (BDS symptom clusters) all explaining the variance of the 25 symptoms in the BDS checklist, was confirmed (Figure 1). Loadings from the general BDS factor were generally higher than loadings from the four symptom

clusters; for the population cohort this was the case for 72% of symptoms, in the primary care cohort it accounted for 64% of symptoms, and in the specialized setting it accounted for 52% of symptoms. Loading from the general BDS factor was smaller than loading from the four symptom clusters for six symptoms (*frequent, loose bowel movements; diarrhoea, pains in arms and legs; muscular aches or pains; pains in the joints; concentration difficulties*) in all three cohorts.

Figure 1 around here

Figure 1: Illustration and factor loadings from the bi-factor model across all three cohorts.

Abbreviations: Gen.=general population; Prim.=primary care; Spec.=specialized clinical setting; CP=cardiopulmonary, GI=gastrointestinal; MS=musculoskeletal; GS=general symptoms; BDS=bodily distress syndrome

Convergent validity

In the general population sample, our hypothesis was met for all measures. The BDS checklist had moderate convergent validity compared to the SF-36 item for overall health (r=0.48, 95% CI: 0.46;0.49, p<0.0001), the four items aggregate score for physical health (r=-0.58, 95% CI: -0.59;-0.56, p<0.0001), the SCL-8 for emotional distress (r=0.52, 95% CI: 0.51;0.54, p<0.0001), and the Whiteley-6-R for illness worry (r=0.53, 95% CI: 0.52;0.55, p<0.0001). Expected difference on the BDS checklist with one unit difference on the four items aggregate score was -0.80 (95% CI: -0.82;-0.78), 12.26 (95% CI: 11.89;12.63) with SCL-8, and 8.93 (95% CI: 8.64;9.21) with WI-6 (Appendix C).

For the primary care sample, our hypothesis was met for all measures as well, however, for some of the measures, the association was stronger than hypothesized. We found moderate convergent validity compared to the SF-36 item for overall health (r=0.58, 95% CI: 0.56;0.61, p<0.0001), the SCL-8 for emotional distress (r=0.62, 95% CI: 0.59;0.64, p<0.0001), and the WI-6 for illness worry (r=0.55, 95% CI: 0.52;0.58, p<0.0001). Expected difference on the BDS checklist with one unit difference on the SCL-8 was 10.60 (95% CI: 10.08;11.12) and 10.01 (95% CI: 9.44;10.59) with Whiteley-6-R (Appendix C).

For the sample from specialized setting, our hypothesis about the correlations being weaker in the specialized setting was met. Moderate convergent validity was seen with emotional distress (r=0.47, 95% CI: 0.40;0.54, p<0.0001) while weaker correlations were seen for overall health (r=0.25, 95% CI: 0.17;0.33, p<0.0001), physical health (r=-0.22, 95% CI: -0.30;-0.12, p<0.0001), and illness worry (r=0.36, 95% CI: 0.28;0.43, p<0.0001). Expected difference on the BDS checklist with one unit

difference on the four items aggregate score for physical health was -0.41 (95% CI: -0.56;-0.26), 7.92 (95% CI: 6.65;9.18) with SCL-8, and 5.88 (95% CI: 4.58;7.17) with Whiteley-6-R (Appendix C).

Response distributions and acceptability

BDS checklist item and scale characteristics are displayed in Table 2. Item means varied from 0.15-1.09 in the general population sample, from 0.31-1.53 in the primary care sample, and from 0.81-3.34 in the sample from specialized setting. While the item with the lowest mean varied across samples, the item *'excessive fatigue*' had the highest mean value in all samples. Most item total correlations, corrected for overlap, exceeded 0.4.

Table 2: Item and scale characteristics

Item	General population (n=9656)			Primary care (n=2480)			Specialized setting (n=492)		
	Missing %	Mean (SD)	Item total correlation*	Missing %	Mean (SD)	Item total correlation*	Missing %	Mean (SD)	Item total correlation*
Palpations/heart pounding	0.9	0.45 (0.74)	0.427	3.1	0.61 (0.93)	0.544	0.2	1.46 (1.25)	0.480
Precordial discomfort	1.1	0.29 (0.61)	0.410	2.9	0.46 (0.81)	0.517	0.2	1.09 (1.16)	0.391
Breathlessness without exertion	1.0	0.36 (0.71)	0.426	2.9	0.63 (0.99)	0.509	0.2	1.27 (1.25)	0.511
Hyperventilation	1.1	0.15 (0.47)	0.321	3.3	0.37 (0.81)	0.448	0.2	0.93 (1.19)	0.380
Hot and cold sweats	1.2	0.46 (0.80)	0.429	3.2	0.64 (0.96)	0.533	0.2	1.88 (1.35)	0.523
Dry mouth	1.2	0.39 (0.76)	0.432	3.4	0.59 (0.98)	0.501	0.2	1.33 (1.36)	0.442
Frequent loose bowel movements	1.0	0.65 (0.86)	0.403	3.4	0.61 (0.95)	0.391	0.2	1.41 (1.32)	0.357
Abdominal pains	1.4	0.48 (0.76)	0.511	3.5	0.57 (0.90)	0.548	0.2	1.81 (1.22)	0.491
Feeling bloated/full of gas/distended	1.1	0.74 (0.91)	0.524	2.9	0.78 (1.03)	0.532	0.2	2.09 (1.30)	0.484
Diarrhoea	1.1	0.33 (0.63)	0.387	3.5	0.37 (0.80)	0.361	0.2	0.81 (1.13)	0.348
Regurgitations	1.4	0.43 (0.74)	0.392	3.3	0.35 (0.74)	0.456	0.2	1.05 (1.09)	0.430
Nausea	0.9	0.26 (0.57)	0.465	2.9	0.50 (0.87)	0.564	0.2	1.73 (1.27)	0.402
Burning sensation of the upper part of stomach/epigastrium	0.9	0.31 (0.68)	0.441	3.1	0.31 (0.72)	0.483	0.2	1.11 (1.26)	0.512
Pains in arms or legs	1.0	0.87 (1.08)	0.538	3.1	1.21 (1.29)	0.563	0.2	2.68 (1.24)	0.472
Muscular aches or pains	1.3	0.98 (1.01)	0.572	3.2	1.30 (1.23)	0.584	0.2	2.96 (1.10)	0.485
Pains in the joints	1.6	0.96 (1.07)	0.490	3.9	1.20 (1.27)	0.560	0.2	2.57 (1.32)	0.491
Feeling of paresis or localized weakness	1.4	0.16 (0.55)	0.365	4.0	0.33 (0.83)	0.460	0.2	1.22 (1.40)	0.481
Back ache	1.3	1.00 (1.06)	0.492	3.3	1.21 (1.29)	0.542	0.2	2.49 (1.37)	0.377
Pain moving from one place to another	1.4	0.27 (0.71)	0.489	3.8	0.54 (0.99)	0.544	0.2	2.13 (1.48)	0.403
Unpleasant numbness or tingling sensations	1.3	0.25 (0.67)	0.410	4.0	0.34 (0.83)	0.475	0.2	2.00 (1.43)	0.528
Concentration difficulties	0.7	0.60 (0.82)	0.545	2.9	0.85 (1.05)	0.540	0.2	2.53 (1.11)	0.437
Excessive fatigue	0.7	1.09 (1.01)	0.614	2.3	1.53 (1.20)	0.625	0.2	3.34 (0.83)	0.418
Headache	0.8	0.66 (0.89)	0.455	2.8	0.89 (1.08)	0.489	0.2	2.25 (1.22)	0.326
Impairment of memory	0.7	0.60 (0.83)	0.517	2.7	0.80 (1.06)	0.521	0.2	2.29 (1.27)	0.476
Dizziness	0.8	0.34 (0.67)	0.491	2.5	0.58 (0.94)	0.553	0.2	1.75 (1.30)	0.505
Scale									
Total scale missing (%)		0.6			2.7			0.2	
Mean (SD)		13.03 (10.36)			17.33 (13.79)			46.15 (15.91)	
Percentiles									
5%		1			2			22	
10%		3			3			26	
25%		6			7			34	
50% (median)		11			14			45	
75%		18			24			57	
90%		27			37			67	
		34			45			73	

*Item total correlation, corrected for overlap. 25% percentile and 75% percentile=interquartile ranges.**Abbreviations**: SD=standard deviation; IQR=interquartile range

Internal consistency was good in all three samples: α=0.887 in the general population sample, α=0.908 in the primary care sample, and α=0.879 in the sample from specialized setting.

BDS score distribution differed across samples (Figure 2) as did total sum score means; it was lowest in the general population (13.03, SD: 10.36) and highest in specialized setting (46.15, SD: 15.91) (Table 2). Acceptability was good, and the numbers of missing responses were generally low in the general population (total 0.6%) and specialized setting (total 0.2%) while it was slightly higher in primary care (total 2.7%).

Figure 2 around here.

Figure 2: Distribution of the BDS total sum score across all three cohorts.

BDS total sum scores were grouped into five categories: 0-20, 21-40, 41-60, 61-80, and 81-100. The vast majority of the general population respondents (96.6%) and primary care patients (90%) scored below 41, while this was only the case for a smaller fraction of the patients from specialized setting (38.7%) (Table 3). Data from each of the three samples and for all three samples pooled together are shown as cumulative percentages across sex and age groups in Appendix D.

Table 3: Grouping of BDS scores across samples

	General population		Primary care		Specialized setting	
Categories of BDS score	n	%	n	%	n	%
0-20	7.762	80.4	1.617	65.2	20	4.1
21-40	1.607	16.6	616	24.8	170	34.6
41-60	208	2.2	156	6.3	204	41.5
61-80	18	0.2	23	0.9	87	17.7
81-100	0	0	2	0.1	10	2.0
Missing	61	0.6	66	2.7	1	0.2

Abbreviations: BDS=Bodily distress syndrome

More detailed information about the normative data from each of the three samples and for all three samples pooled together are shown in Tables 1-4 in Appendix D.

Discussion

Principal findings

This is the first study to establish that, despite some multi-dimensionality, the 25-items BDS checklist can be used as a continuous score to measure symptom burden and illness severity in the general population, in primary care, and in specialized settings. Used as a total sum score with a range from 0-100, the BDS checklist had acceptable convergent validity with measures of overall health, physical health, emotional distress, and illness worry. Internal consistency was good in all three cohorts ($\alpha \geq 0.879$) as was acceptability. Thus, the BDS checklist may work as a simple symptom checklist but also as a diagnostic screening tool for use in clinical work and in research across different settings.

We found the symptom *'excessive fatigue*' to have the highest mean value in all three cohorts. This is in line with a recent German population-based study, finding *'tiredness*' to be one of the leading symptoms [44].

The three cohorts differed in number of symptoms that had higher loadings on the general BDS factor than on the four-symptom clusters ranging from 72% of symptoms in the general population cohort to 52% in the cohort from specialized clinical setting. The latter group contains patients with longstanding and severe FSD. In this group, the symptom load is high and specific symptom clusters may therefore stand out compared to the less affected participants from the general population with a more scattered symptom picture.

Previous studies have argued that the best fitting model for the BDS checklist was a one-level four factor model (Appendix B) [19 21 24]. However, the objectives of these studies were to confirm the BDS as case finding instrument in other samples with inspiration from the original studies in which the concept of BDS was developed and initially tested [18 19]. In the current study, we have taken it several steps further and tested various structural equation models in three different populations at the same time. The indicators of a bi-factor model is 1) if inter-correlation between the sub-scales in the CFA exceeds 0.3, 2) if loading on the first order factors on the second order factors exceeds 0.5 [45], and 3) if the ratio between the first and second eigenvalues exceeds 3 [46]. All parameters were fulfilled in the general population cohort and in the primary care cohort. For the cohort from specialized setting the ratio between the first and second eigenvalues was 2.68, but otherwise the parameters were fulfilled. This implies that the results from this study do not disqualify results from previous research, but the presence of some multidimensionality is not strong enough to disqualify the interpretation of the BDS checklist as unidimensional as well.

Correlations between the BDS checklist and self-rated measures of overall health, physical health, emotional distress, and illness worry were generally moderate, especially in the general population and primary care cohort. This was as expected as previous literature has shown the same association between symptom load and reduced function [6 7] . The difference between results on patients in the specialised settings and the two other populations may be caused by the nature of self-reported measures, where patients in specialized setting still have the opportunity to rate their perceived health as excellent even though they have been referred to specialized medical care because of invalidating physical symptoms. These aspects may produce precision limitations in some settings and may especially be pronounced in smaller samples. Furthermore, the distribution of sex differs across populations which may affect the results on convergent validity.

Strengths and weaknesses of the study

A major strength of this study is the inclusion of three different populations. To our knowledge, this approach of testing an instrument and using the same methodology in different populations is rare as most other studies concern only one setting at a time [11 12 14 17]. Also, the sample size within each cohort was large. We conducted a thorough validation procedure, using different structural equation models and testing convergent validity to several valid measures.

Weaknesses of the study include: Only self-reported outcomes were used and data measures were not completely consistent across the included cohorts; hence, we chose to apply the intersection of items in order to gain equivalent proxy measures. We did not have the opportunity to compare the BDS checklist to other measures of physical symptoms or – for the primary care cohort and the cohort from specialized clinical setting – to the physician's report. Furthermore, in the linear regression analyses, the assumption of normality of the residuals was not fully met for the primary care cohort and the cohort from specialized clinical care why these results should be interpreted with caution. Finally, as this study had a cross-sectional design, it was not possible to evaluate responsiveness of the BDS checklist.

Difference in results compared to others

To our knowledge, this is the first study to address the usefulness of the BDS checklist as a measure of physical symptom burden and illness severity. Another symptom checklist which has been widely used within primary care and general population studies for measuring the severity of

physical symptoms is the 15-items Patient Health Questionnaire (PHQ-15) [11 17]. It consists of 15 items concerning some of the symptoms from the same four organ systems as the BDS checklist, plus the symptoms 'fainting', 'sleeping problems', 'menstrual problems', and 'sexual pains/problems' not included in the BDS checklist. The PHQ-15 is scored on a three-point rating scale from 'not bothered at all' (0) to 'bothered at lot' (2), whereas the BDS checklist uses a five-point rating scale. In one study, including a sample from the general Swedish population, factor analyses of the structural properties of the PHQ-15 showed a four-factor model, but on the basis on a scree test plot they finally concluded that only one factor should be extracted [47]. Other studies found a bi-factor model to have the best fit to the PHQ-15 [48 49]. Hence, the PHQ-15 may have the same structural properties as the BDS checklist, but with fewer items to take into account as well as fewer response categories which may make it more prone to floor and ceiling effects. In a shorter version of the PHQ-15, the Somatic Symptoms Scale-8 (SSS-8), the three-point rating scale is replaced with a five-point rating option as in the BDS checklist [12 13 50]. However, neither the PHQ-15 nor the SSS-8 is validated for use as diagnostic categorization of respondents. Other symptom questionnaires resembling the same four factor structure and the same five answer categories as the BDS checklist are the 24-items Giessen Subjective Complaints List and its newer shortened version with 8 items (GBB-8), however, they have only been established and used in German speaking countries[14].
The BDS checklist is, at present, the only symptom checklist providing both diagnostic categorization and a measure of symptom load/illness severity.

Clinical implications

This study provides a self-reported symptom checklist for measuring symptom burden and illness severity which can be used both as a diagnostic screening tool and as a measure of illness severity in large epidemiological studies and also in more selected patient samples and severely ill patients. Regarding FSD, previous research has suggested measures of symptom burden as the primary outcome [33]. However, the current study shows that the BDS checklist shows weaker correlation with measures of overall health, physical health, emotional distress, and illness worry in patients from highly specialized setting than in the general population and primary care. Hence, a simple count of bothersome symptoms may not be adequate when dealing with the more severely ill patients, as symptom burden may not be the only important domain of illness severity – others may be the level of impairment and mental morbidity.

Currently, it is unclear whether the here presented BDS total sum score reliably captures FSD illness severity than the distinction in single vs. multi-organ BDS (e.g. three vs. four clusters fulfilled). Nevertheless, a tool which is also able to measure severity of specific symptom clusters is helpful in specialized settings, as it is possible to elucidate which symptom cluster is experienced most bothersome by the patients.

Future research and perspectives

In this study we suggest the BDS checklist as a prominent tool as it can be used both as a measure of symptom burden and as a diagnostic screening tool for FSD, and we argue for its usefulness in both epidemiological and clinical research as well as in clinical practice. However, the criterion validity of the self-reported BDS checklist with physician's established diagnoses e.g. specialty-specific syndrome diagnoses and psychiatric diagnoses, is yet to be investigated across settings, and future studies regarding these aspects would be valuable in order to further establish the usefulness of the BDS checklist. Moreover, the additional value of counting the number of symptom clusters fulfilled in the staging of FSD deserves attention. Finally, we need a valid instrument to measure change over time, and the responsiveness of the BDS checklist sum score is worth exploring.

Acknowledgements

The authors would like to thank senior research Grete Moth and data manager Kaare Rud Flarup from the Research Unit of primary care for providing date from the primary care cohort. We are also thankful to the members of the DanFunD steering committee; Senior consultant MD PhD Lene Falgaard Eplov, MSc PhD Allan Linneberg, MSc PhD Susanne Brix Pedersen, and MD PhD Michael Eriksen Benros.

Contributors

MWP contributed to the conception and design of the study and the statistical analyses, interpreted the data, and drafted the article. AS and MR contributed to the conception and design of the study and interpretation of the data and provided general supervision of the work. EØ performed the statistical analyses, and contributed to the conception and design of the study and the interpretation of the data. TJ, TMD and PF contributed to the interpretation of the data. All authors contributed to critically revising the article for important intellectual content, and all authors read and approved the final version of the article.

Funding

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests

The authors declare no competing interests.

Data availability statement

Data are available on reasonable request from the corresponding author.

References

1. Henningsen P, Gundel H, Kop WJ, et al. Persistent Physical Symptoms as Perceptual Dysregulation: A Neuropsychobehavioral Model and Its Clinical Implications. *Psychosomatic medicine* 2018;80(5):422-31. doi: 10.1097/PSY.0000000000000588 [doi]
2. Kroenke K. A practical and evidence-based approach to common symptoms: a narrative review. *Annals of Internal Medicine* 2014;161(8):579-86. doi: 10.7326/M14-0461 [doi]
3. Henningsen P, Zipfel S, Sattel H, et al. Management of Functional Somatic Syndromes and Bodily Distress. *Psychotherapy and psychosomatics* 2018;87(1):12-31. doi: 10.1159/000484413 [doi]
4. Eliasen M, Kreiner S, Ebstrup JF, et al. Somatic Symptoms: Prevalence, Co-Occurrence and Associations with Self-Perceived Health and Limitations Due To Physical Health - A Danish Population-Based Study. *PloS one* 2016;11(3):e0150664. doi: 10.1371/journal.pone.0150664 [doi]
5. Tomenson B, Essau C, Jacobi F, et al. Total somatic symptom score as a predictor of health outcome in somatic symptom disorders. *The British journal of psychiatry : the journal of mental science* 2013;203(5):373-80. doi: 10.1192/bjp.bp.112.114405 [doi]
6. Bruusgaard D, Tschudi-Madsen H, Ihlebaek C, et al. Symptom load and functional status: results from the Ullensaker population study. *BMC Public Health* 2012;12:1085. doi: 10.1186/1471-2458-12-1085 [published Online First: 2012/12/20]

7. Creed FH, Davies I, Jackson J, et al. The epidemiology of multiple somatic symptoms. *Journal of psychosomatic research* 2012;72(4):311-17. doi: 10.1016/j.jpsychores.2012.01.009 [doi]
8. Ladwig KH, Marten-Mittag B, Lacruz ME, et al. Screening for multiple somatic complaints in a population-based survey: does excessive symptom reporting capture the concept of somatic symptom disorders? Findings from the MONICA-KORA Cohort Study. *Journal of psychosomatic research* 2010;68(5):427-37. doi: 10.1016/j.jpsychores.2010.01.009 [doi]
9. Eliasen M, Schröder A, Fink P, et al. A step towards a new delimitation of functional somatic syndromes: A latent class analysis of symptoms in a population-based cohort study. *J Psychosom Res* 2018;108:102-17.
10. Derogatis LR, Lipman RS, Rickels K, et al. The Hopkins Symptom Checklist (HSCL): a self-report symptom inventory. *Behavioral science* 1974;19(1):1-15.
11. 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):258-66.
12. Gierk B, Kohlmann S, Kroenke K, et al. The somatic symptom scale-8 (SSS-8): a brief measure of somatic symptom burden. *JAMA internal medicine* 2014;174(3):399-407. doi: 10.1001/jamainternmed.2013.12179 [doi]
13. Gierk B, Kohlmann S, Toussaint A, et al. Assessing somatic symptom burden: a psychometric comparison of the patient health questionnaire-15 (PHQ-15) and the somatic symptom scale-8 (SSS-8). *Journal of psychosomatic research* 2015;78(4):352-55. doi: 10.1016/j.jpsychores.2014.11.006 [doi]
14. Kliem S, Lohmann A, Klatt T, et al. Brief assessment of subjective health complaints: Development, validation and population norms of a brief form of the Giessen Subjective Complaints List (GBB-8). *J Psychosom Res* 2017;95:33-43. doi: 10.1016/j.jpsychores.2017.02.003 [published Online First: 2017/03/21]
15. Sitnikova K, Dijkstra-Kersten SMA, Mokkink LB, et al. Systematic review of measurement properties of questionnaires measuring somatization in primary care patients. *J Psychosom Res* 2017;103:42-62. doi: 10.1016/j.jpsychores.2017.10.005 [published Online First: 2017/11/24]
16. Hyland ME, Bacon AM, Lanario JW, et al. Symptom frequency and development of a generic functional disorder symptom scale suitable for use in studies of patients with irritable bowel syndrome, fibromyalgia syndrome or chronic fatigue syndrome. *Chronic Dis Transl Med* 2019;5(2):129-38. doi: 10.1016/j.cdtm.2019.05.003 [published Online First: 2019/08/02]

17. Zijlema WL, Stolk RP, Lowe B, et al. How to assess common somatic symptoms in large-scale studies: a systematic review of questionnaires. *Journal of psychosomatic research* 2013;74(6):459-68. doi: 10.1016/j.jpsychores.2013.03.093 [doi]

18. Fink P, Toft T, Hansen MS, et al. Symptoms and syndromes of bodily distress: an exploratory study of 978 internal medical, neurological, and primary care patients. *Psychosomatic medicine* 2007;69(1):30-39. doi: 69/1/30 [pii]

19. Budtz-Lilly A, Fink P, Ørnbøl E, et al. A new questionnaire to identify bodily distress in primary care: The 'BDS checklist'. *Journal of psychosomatic research* 2015;78(6):536-45. doi: 10.1016/j.jpsychores.2015.03.006 [doi]

20. Rosmalen JG, Tak LM, de Jonge P. Empirical foundations for the diagnosis of somatization: implications for DSM-5. *Psychological medicine* 2011;41(6):1133-42. doi: 10.1017/S0033291710001625 [doi]

21. Petersen MW, Schröder A, Jørgensen T, et al. The unifying diagnostic construct of bodily distress syndrome (BDS) was confirmed in the general population. *Journal of psychosomatic research* 2019;128:109868. doi: S0022-3999(19)30601-4 [pii]

22. Burton C, Fink P, Henningsen P, et al. Functional somatic disorders: discussion paper for a new common classification for research and clinical use. *BMC Med* 2020;18(1):34. doi: 10.1186/s12916-020-1505-4 [published Online First: 2020/03/04]

23. Petersen MW, Schröder A, Jørgensen T, et al. Irritable bowel, chronic widespread pain, chronic fatigue and related syndromes are prevalent and highly overlapping in the general population: DanFunD. *Scientific reports* 2020;10(1):3273-73. doi: 10.1038/s41598-020-60318-6

24. Schmalbach B, Roenneberg C, Hausteiner-Wiehle C, et al. Validation of the German version of the Bodily Distress Syndrome 25 checklist in a representative German population sample. *Journal of Psychosomatic Research* 2020;132:109991. doi: https://doi.org/10.1016/j.jpsychores.2020.109991

25. Dantoft TM, Ebstrup JF, Linneberg A, et al. Cohort description: The Danish Study of Functional Disorders. *Clin Epidemiol* 2017;9:127-39.

26. Moth G, Vedsted P, Olesen F, et al. Kontakt- og sygdomsmønsterundersøgelse KOS 2008: Forskningsenheden for Almen Praksis 2010.

27. Frølund Pedersen H, Holsting A, Frostholm L, et al. "Understand your illness and your needs": Assessment-informed patient education for people with multiple functional somatic

syndromes. *Patient Educ Couns* 2019;102(9):1662-71. doi: 10.1016/j.pec.2019.04.016 [published Online First: 2019/04/27]

28. Agger JL, Schröder A, Gormsen LK, et al. Imipramine versus placebo for multiple functional somatic syndromes (STreSS-3): a double-blind, randomised study. *Lancet Psychiatry* 2017;4(5):378-88. doi: 10.1016/S2215-0366(17)30126-8 [published Online First: 2017/04/15]
29. Pedersen HF, Agger JL, Frosthom L, et al. Acceptance and Commitment group Therapy for patients with multiple functional somatic syndromes: a three-armed trial comparing ACT in a brief and extended version with enhanced care. *Psychological medicine* 2018:1-10. doi: 10.1017/S0033291718001666.

10.1017/S0033291718001666 [doi]

30. Schröder A, Rehfeldt E, Ørnbøl E, et al. Cognitive-behavioural group treatment for a range of functional somatic syndromes: randomised trial. *Br J Psychiatry* 2012;200(6):499-507.
31. Fjorback LO, Arendt M, Ornbol E, et al. Mindfulness therapy for somatization disorder and functional somatic syndromes: randomized trial with one-year follow-up. *Journal of psychosomatic research* 2013;74(1):31-40. doi: 10.1016/j.jpsychores.2012.09.006 [doi]
32. Ware JJ, Kosinski, M., Gandek B. SF-36 Health Survey: Manual and Interpretation Guide. *Quality Metric* 1993
33. Rief W, Burton C, Frostholm L, et al. Core Outcome Domains for Clinical Trials on Somatic Symptom Disorder, Bodily Distress Disorder, and Functional Somatic Syndromes: European Network on Somatic Symptom Disorders Recommendations. *Psychosom Med* 2017;79(9):1008-15. doi: 10.1097/psy.0000000000000502 [published Online First: 2017/07/12]
34. Schröder A, Ørnbøl E, Licht RW, et al. Outcome measurement in functional somatic syndromes: SF-36 summary scores and some scales were not valid. *Journal of clinical epidemiology* 2012;65(1):30-41. doi: 10.1016/j.jclinepi.2011.03.012 [doi]
35. Derogatis LR. SCL-90-R, administration, scoring and procedures manual-II for the R(evised) version and other instruments of the Psychopathology Rating Scale Series. Townson: Clinical Psychometric Research 1992.
36. Christensen KS BP, Fink P. Measuring Mental Health by Questionnaires in Primary Care - Unidimensionality. Responsiveness and Compliance. . *European Psychiatric Review* 2010;3:8-12.

37. Carstensen TØ, E. Fink, P. Pedersen, MM. Jørgensen, T. Dantoft, TM. Benros, ME. Frostholm, L. Detection of illness worry in the general population: A specific item on illness rumination improves the Whiteley Index - In review. *Journal of psychosomatic research* 2020

38. Mokkink LB, Terwee CB, Patrick DL, et al. COSMIN Checklist Manual. Amsterdam: EMGO Institute for Health and Care Research 2012.

39. StataCorp. Stata Statistical Software. Release 16: College Station, TX: StataCorp LLC 2019.

40. Muthén LK, Muthén BO. Mplus User's Guide. Eighth Edition. Los Angeles, CA: Muthén & Muthén1998.

41. Hu Lt, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal* 1999;6(1):1-55. doi: 10.1080/10705519909540118

42. Ware J, Jr., Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Medical care* 1996;34(3):220-33.

43. Terluin B, van Marwijk HW, Ader HJ, et al. The Four-Dimensional Symptom Questionnaire (4DSQ): a validation study of a multidimensional self-report questionnaire to assess distress, depression, anxiety and somatization. *BMC Psychiatry* 2006;6:34. doi: 10.1186/1471-244X-6-34 [published Online First: 2006/08/24]

44. Beutel ME, Klein EM, Henning M, et al. Somatic Symptoms in the German General Population from 1975 to 2013. *Sci Rep* 2020;10(1):1595. doi: 10.1038/s41598-020-58602-6 [published Online First: 2020/02/02]

45. MD. HJT. Bifactor analysis in MPlus [Video file]: UK College of Education; 2016 [Available from: https://education.uky.edu/edp/apslab/events/2020.

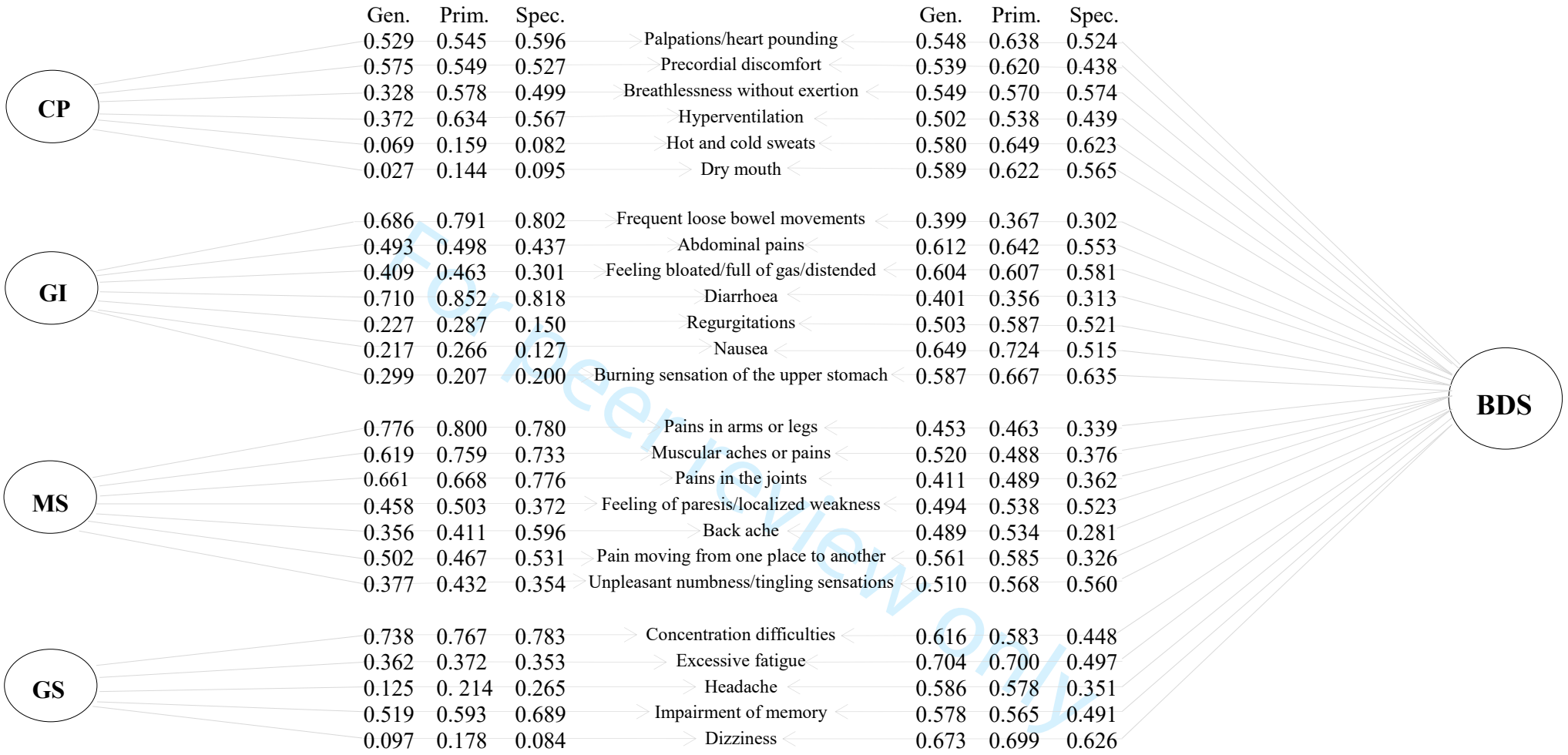
46. Cho SJ, Wilmer J, Herzmann G, et al. Item response theory analyses of the Cambridge Face Memory Test (CFMT). *Psychol Assess* 2015;27(2):552-66. doi: 10.1037/pas0000068 [published Online First: 2015/02/03]

47. Nordin S, Palmquist E, Nordin M. Psychometric evaluation and normative data for a Swedish version of the Patient Health Questionnaire 15-Item Somatic Symptom Severity Scale. *Scandinavian Journal of Psychology* 2013;54(2):112-17. doi: 10.1111/sjop.12029 [doi]

48. Cano-Garcia FJ, Munoz-Navarro R, Sese Abad A, et al. Latent structure and factor invariance of somatic symptoms in the patient health questionnaire (PHQ-15). *J Affect Disord* 2020;261:21-29. doi: 10.1016/j.jad.2019.09.077 [published Online First: 2019/10/11]

49. Witthoft M, Hiller W, Loch N, et al. The latent structure of medically unexplained symptoms and its relation to functional somatic syndromes. *International Journal of Behavioral Medicine* 2013;20(2):172-83. doi: 10.1007/s12529-012-9237-2 [doi]

50. Toussaint A, Kroenke K, Baye F, et al. Comparing the Patient Health Questionnaire - 15 and the Somatic Symptom Scale - 8 as measures of somatic symptom burden. *J Psychosom Res* 2017;101:44-50. doi: 10.1016/j.jpsychores.2017.08.002 [published Online First: 2017/09/05]



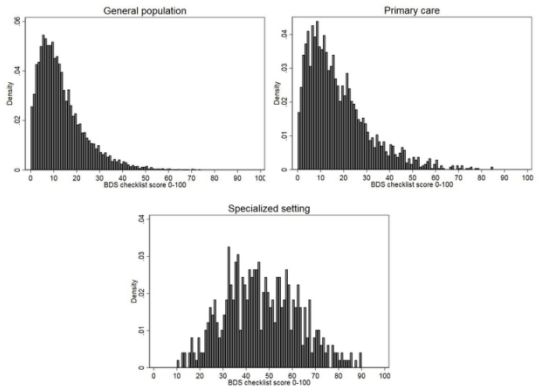


Figure 2: Distribution of the BDS total sum score across all three cohorts.

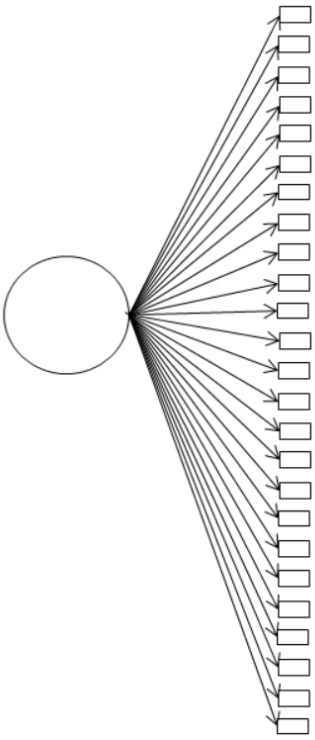
120x49mm (300 x 300 DPI)

Appendix A: The 25-items BDS checklist

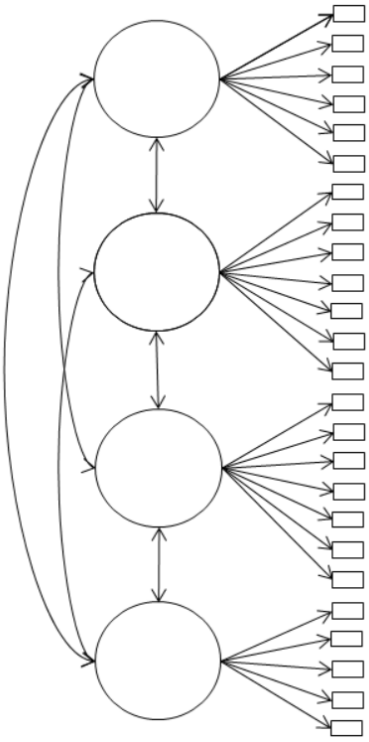
During the last 4 weeks*, have you been bothered by…		Not at all	A bit	Somewhat	Quite a bit	A lot
1	Palpations and heart pounding	□	□	□	□	□
2	Precordial discomfort	□	□	□	□	□
3	Breathlessness without exertion	□	□	□	□	□
4	Hyperventilation	□	□	□	□	□
5	Hot and cold sweats	□	□	□	□	□
6	Dry mouth	□	□	□	□	□
7	Frequent loose bowel movements	□	□	□	□	□
8	Abdominal pains	□	□	□	□	□
9	Feeling bloated/full of gas/distended	□	□	□	□	□
10	Diarrhoea	□	□	□	□	□
11	Regurgitations	□	□	□	□	□
12	Nausea	□	□	□	□	□
13	Burning sensation of the upper part of stomach/epigastrium	□	□	□	□	□
14	Pains in arms or legs	□	□	□	□	□
15	Muscular aches or pains	□	□	□	□	□
16	Pains in the joints	□	□	□	□	□
17	Feeling of paresis or localized weakness	□	□	□	□	□
18	Back ache	□	□	□	□	□
19	Pain moving from one place to another	□	□	□	□	□
20	Unpleasant numbness or tingling sensations	□	□	□	□	□
21	Concentration difficulties	□	□	□	□	□
22	Excessive fatigue	□	□	□	□	□
23	Headache	□	□	□	□	□
24	Impairment of memory	□	□	□	□	□
25	Dizziness	□	□	□	□	□

* This time frame was changed to 12 months in the general population cohort

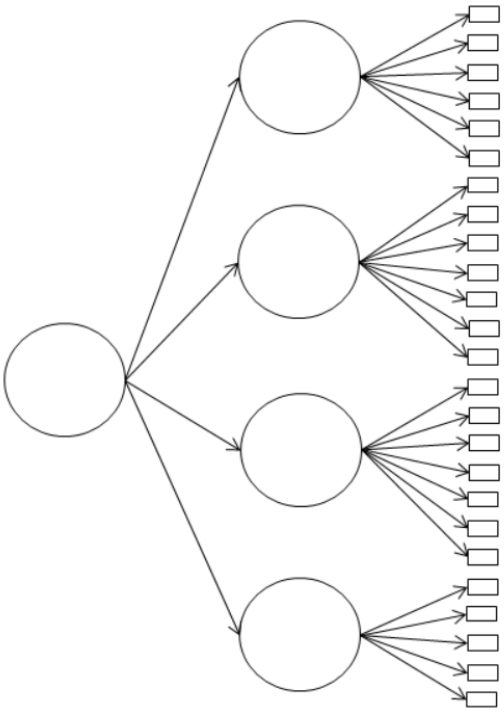
Appendix B: Illustrations of the theoretical models of confirmatory factor analyses



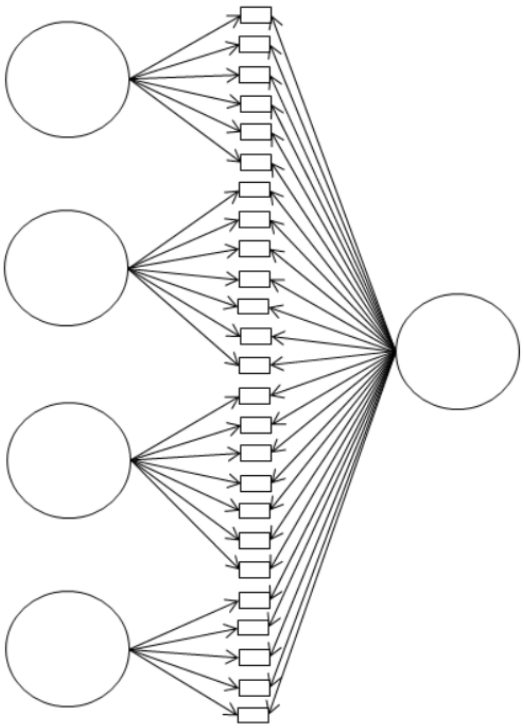
One-level one factor model



One-level four factor model



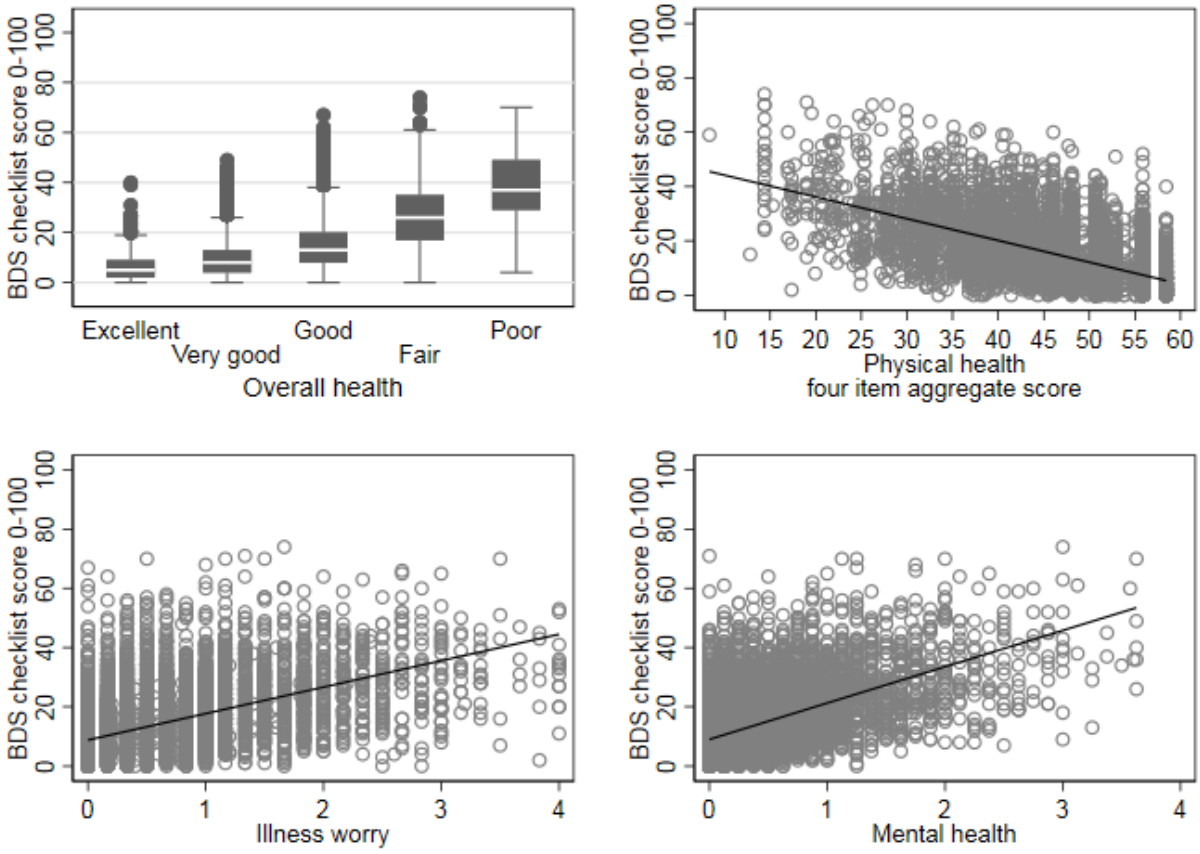
Two-level four factor model



Bi-factor model

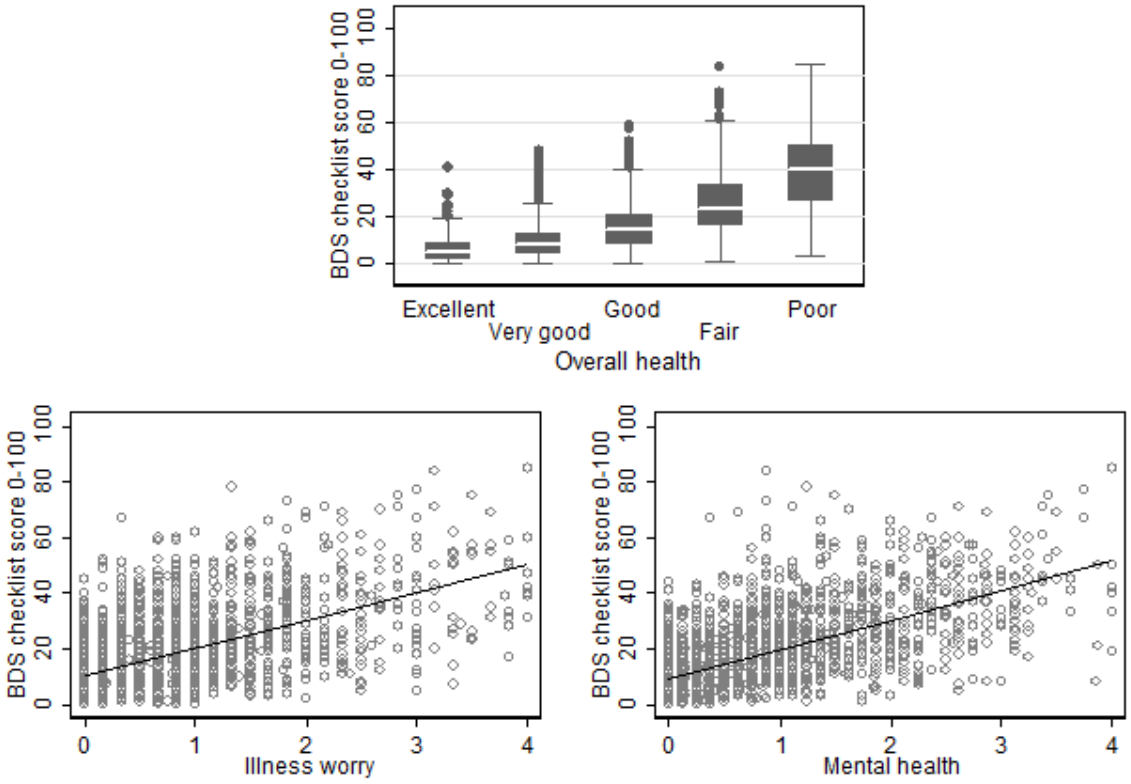
Appendix C: Descriptive plots of associations between the BDS checklist and other measures

General population



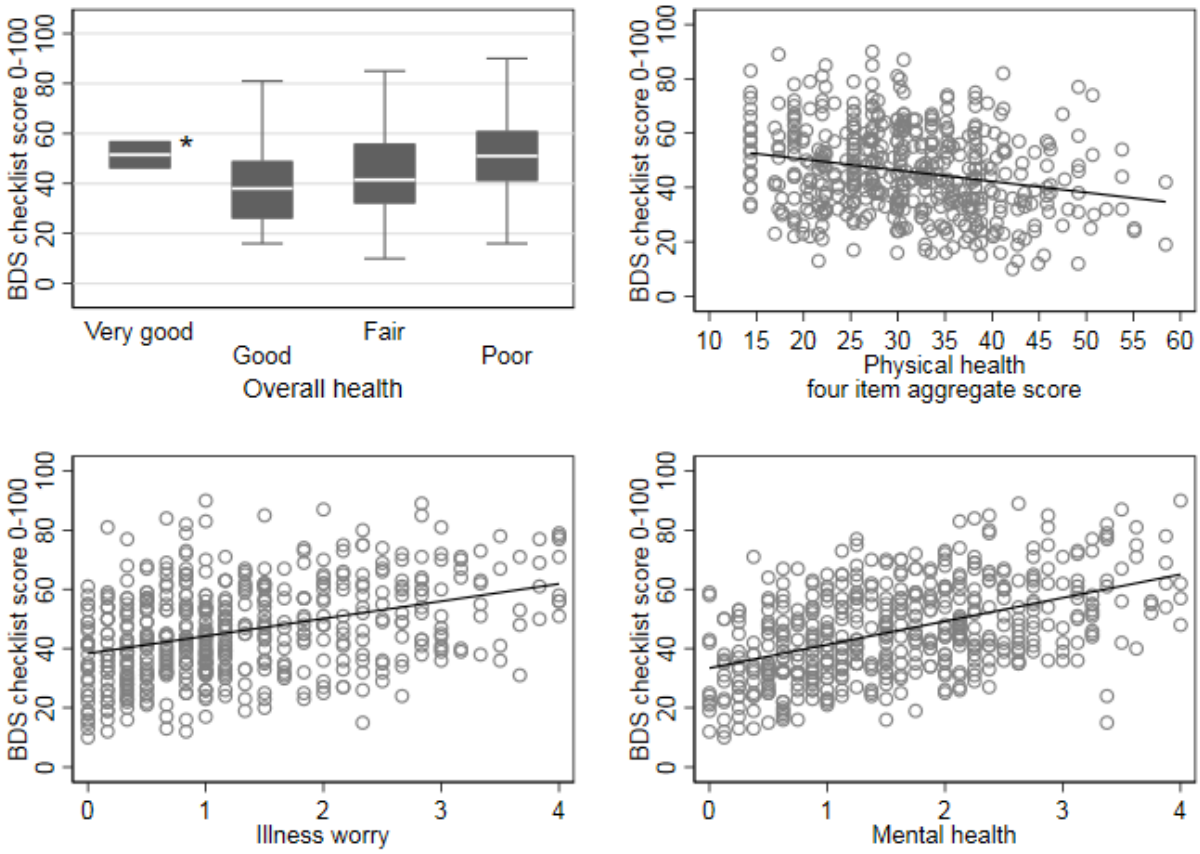
Correlation of the BDS checklist score and measures of overall health, perceived physical health, illness worry, and mental health in the general population cohort.

Primary care



Correlation of the BDS checklist score and measures of overall health, illness worry, and mental health in the primary care cohort.

Specialized setting



Correlation of the BDS checklist score and measures of overall health, perceived physical health, illness worry, and mental health in the cohort from specialized setting.
* Only based on two individuals

Appendix D. Data across sex and age groups

Table 1: Data for the general population cohort (n=9656): Cummulative percentages

	Male				Female			
	18-39	40-49	50-59	60-76	18-39	40-49	50-59	60-76
	N= 719	N=902	N=1.130	N=1.702	N=905	N=1.139	N=1.401	N=1.758
BDS score groups								
0-	21	24	25	28	10	15	10	18
5-	50	55	54	55	34	38	33	42
10-	73	76	71	75	56	59	56	62
15-	85	86	82	86	72	73	71	76
20-	91	92	88	91	83	82	82	85
25-	95	95	93	94	90	89	88	91
30-	97	97	96	97	94	93	93	94
35-	98	97	97	98	96	95	96	96
40-	99	98	98	98	97	97	97	98
45-	99	99	99	99	98	98	98	99
50-	99	99	99	99	98	99	99	99
55-	99	99	99	99	99	99	99	99
60-	99	99	99	99	99	99	100	99
65-	99	99	99	99	99	99	100	99
70-	99	99	99	99	99	99	100	99
75-	99	99	99	99	99	99	100	99
80-	99	99	99	99	99	99	100	99
85-	99	99	99	99	99	99	100	99
90-	99	99	99	99	99	99	100	99
95-	99	99	99	99	99	99	100	99
Missing	100	100	100	100	100	100	100	100

Abbreviations: BDS=Bodily distress syndrome

Table 2: Data for the primary care cohort (n=2480): Cummulative percentages

	Male						Female					
	18-39	40-49	50-59	60-76	70-79	80-	18-39	40-49	50-59	60-69	70-79	80-
	N=146	N=147	N= 172	N=245	N=153	N=67	N=404	N=280	N=271	N=287	N=203	N=105
BDS score groups												
0-	23	20	16	18	14	10	16	14	8	16	14	9
5-	49	37	32	40	35	36	37	31	23	32	31	27
10-	67	60	52	58	54	51	53	47	38	45	45	42
15-	73	70	67	72	65	58	66	59	54	58	59	55
20-	79	80	76	83	72	67	76	75	66	73	72	64
25-	89	86	82	90	79	82	80	80	78	82	79	70
30-	92	89	87	95	82	82	86	84	85	85	85	72
35-	93	91	90	96	88	85	90	88	89	89	88	74
40-	97	93	94	98	91	87	93	90	92	93	90	76
45-	97	94	96	98	94	88	95	94	95	95	91	78
50-	98	95	98	98	95	88	96	96	96	96	93	80
55-	99	98	98	100	95	88	96	97	97	97	94	81
60-	99	99	98	100	95	88	97	98	97	97	95	82
65-	99	99	99	100	95	88	97	98	97	98	96	83
70-	99	99	99	100	95	88	98	98	98	98	96	83
75-	99	99	99	100	95	88	98	98	98	98	96	83
80-	99	99	99	100	95	88	98	98	99	98	96	83
85-	99	99	99	100	95	88	98	99	99	98	96	83
90-	99	99	99	100	95	88	98	99	99	98	96	83
95-	99	99	99	100	95	88	98	99	99	98	96	83
Missing	100	100	100	100	100	100	100	100	100	100	100	100

Abbreviations: BDS=Bodily distress syndrome

Table 3: Data for the cohort from specialized setting (n=492): Cummulative percentages

	Male				Female			
	18-39	40-49	50-59	60-69	18-39	40-49	50-59	60-69
	N=42	N=40	N=11	N=0	N=199	N=162	N=36	N=2
BDS score groups								
0-	0	0	0	-	0	0	0	0
5-	0	0	0	-	0	0	0	0
10-	0	3	0	-	2	1	0	0
15-	5	5	9	-	4	2	3	0
20-	12	15	9	-	8	7	6	0
25-	21	23	9	-	14	13	14	0
30-	36	45	36	-	20	25	17	50
35-	48	50	55	-	36	35	19	50
40-	57	55	73	-	50	48	25	50
45-	74	60	82	-	60	59	36	50
50-	81	70	82	-	68	69	56	50
55-	86	85	82	-	79	75	72	50
60-	90	88	82	-	87	84	83	100
65-	95	93	82	-	92	92	83	100
70-	95	98	91	-	95	96	94	100
75-	95	100	100	-	96	99	97	100
80-	95	100	100	-	99	99	97	100
85-	100	100	100	-	100	100	97	100
90-	100	100	100	-	100	100	97	100
95-	100	100	100	-	100	100	97	100
Missing	100	100	100	-	100	100	100	100

Abbreviations: BDS=Bodily distress syndrome

Table 4: Data for the three pooled cohorts (n=12628): Cummulative percentages

	Male						Female					
	18-39	40-49	50-59	60-69	70-79	80-	18-39	40-49	50-59	60-69	70-79	80-
	N=907	N=1089	N=1313	N=1665	N=435	N=67	N=1508	N=1581	N=1708	N=1785	N=465	N=105
BDS score groups												
0-	20	23	24	26	26	10	10	13	9	18	15	9
5-	47	51	50	52	50	36	31	33	31	41	37	27
10-	69	71	68	71	70	51	48	51	52	59	53	42
15-	80	81	80	83	80	58	61	63	67	73	67	55
20-	86	88	86	90	86	67	71	73	78	83	79	64
25-	91	91	91	94	89	82	77	80	85	89	85	70
30-	94	94	94	97	91	82	82	84	90	93	89	72
35-	95	95	96	98	94	85	86	88	93	95	92	74
40-	97	96	97	99	95	87	90	91	95	97	94	76
45-	97	97	98	99	97	88	92	93	96	98	95	78
50-	98	97	99	99	97	88	94	95	98	99	96	80
55-	98	98	99	99	97	88	95	96	98	99	97	81
60-	99	99	99	99	97	88	97	97	99	99	97	82
65-	99	99	99	99	97	88	98	98	99	99	98	83
70-	99	99	99	99	97	88	98	99	99	99	98	83
75-	99	99	99	99	97	88	99	99	99	99	98	83
80-	99	99	99	99	97	88	99	99	99	99	98	83
85-	99	99	99	99	97	88	99	99	99	99	98	83
90-	99	99	99	99	97	88	99	99	99	99	98	83
95-	99	99	99	99	97	88	99	99	99	99	98	83
Missing	100	100	100	100	100	100	100	100	100	100	100	100

Abbreviations: BDS=Bodily distress syndrome

STROBE 2007 (v4) Statement—Checklist of items that should be included in reports of *cross-sectional studies*

Section/Topic	Item #	Recommendation	Reported on page #
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	2
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4-5
Objectives	3	State specific objectives, including any prespecified hypotheses	5
Methods			
Study design	4	Present key elements of study design early in the paper	5
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	5
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants	5
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	5-8
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	5-8
Bias	9	Describe any efforts to address potential sources of bias	
Study size	10	Explain how the study size was arrived at	5
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	6-8
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	7-8
		(*b*) Describe any methods used to examine subgroups and interactions	
		(*c*) Explain how missing data were addressed	
		(*d*) If applicable, describe analytical methods taking account of sampling strategy	
		(*e*) Describe any sensitivity analyses	
Results			

Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	
		(b) Give reasons for non-participation at each stage	
		(c) Consider use of a flow diagram	
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	5
		(b) Indicate number of participants with missing data for each variable of interest	
Outcome data	15*	Report numbers of outcome events or summary measures	
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	8-13
		(*b*) Report category boundaries when continuous variables were categorized	8-13
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	
Discussion			
Key results	18	Summarise key results with reference to study objectives	14
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	15
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	15-16
Generalisability	21	Discuss the generalisability (external validity) of the study results	14-16
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	17

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

BMJ Open

The BDS checklist as measure of illness severity: A cross-sectional cohort study in the Danish general population, primary care and specialized setting

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2020-042880.R2
Article Type:	Original research
Date Submitted by the Author:	17-Nov-2020
Complete List of Authors:	Petersen, Marie; Aarhus Universitetshospital, The Research Clinic for Functional Disorders and Psychosomatics Rosendal, Marianne; Aarhus University Hospital, Research Clinic for Functional Disorders and Psychosomatics Ørnbøl, Eva; The Research Clinic for Functional Disorders and Psychosomatics, Aarhus University Hospital, Aarhus C, Denmark Fink, Per; The Research Clinic for Functional Disorders and Psychosomatics, Aarhus University Hospital, Aarhus C, Denmark Jørgensen, Torben; Frederiksberg Hospital, Dantoft, Thomas; Region Hovedstaden Forskningscenter for forebyggelse og sundhed, Schroeder, Andreas; Aarhus Universitetshospital,
Primary Subject Heading:	Research methods
Secondary Subject Heading:	Diagnostics, Epidemiology, Mental health
Keywords:	STATISTICS & RESEARCH METHODS, EPIDEMIOLOGY, PSYCHIATRY, GENERAL MEDICINE (see Internal Medicine)

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

The BDS checklist as measure of illness severity: A cross-sectional cohort study in the Danish general population, primary care and specialized setting

Marie Weinreich Petersen[1], Marianne Rosendal[1,2], Eva Ørnbøl[1], Per Fink[1], Torben Jørgensen[3,4,5], Thomas Meinertz Dantoft[3], Andreas Schroeder[1]

Author affiliations:
[1] The Research Clinic for Functional Disorders and Psychosomatics, Aarhus University Hospital, Aarhus C, Denmark
[2] Research Unit for General Practice, Aarhus University, Aarhus C, Denmark
[3] Center for Clinical Research and Prevention, Bispebjerg and Frederiksberg Hospital, The Capital Region of Denmark, Denmark
[4] Department of Public Health, Faculty of Health and Medical Science, University of Copenhagen, Copenhagen C, Denmark
[5] Faculty of Medicine, Aalborg University

Corresponding author:
Marie Weinreich Petersen
The Research Clinic for Functional Disorders and Psychosomatics
Aarhus University Hospital
Nørrebrogade 44, building 2C, second floor
8200 Aarhus C
Denmark
Phone: +45 29467824
Fax: +45 7846 4340
E-mail: mawept@rm.dk

Word count in main text: 4221
Word count abstract: 300
Number of references: 50

Abstract

Objectives

The Bodily Distress Syndrome (BDS) checklist has proven to be useful for diagnostic categorization and screening tool for functional somatic disorders (FSD). This study aims to investigate whether the BDS checklist total sum score (0-100) can be used as measure of physical symptom burden and FSD illness severity.

Design

Cross-sectional.

Setting

Danish general population, primary care, and specialized clinical setting.

Participants

A general population cohort (n=9656), a primary care cohort (n=2480), and a cohort of multi-organ BDS patients from specialized clinical setting (n=492).

Outcome measures

All data were self-reported. Physical symptoms were measured with the 25-items BDS checklist. Overall self-perceived health was measured with one item from the 36-items Short Form Health Survey (SF-36). Physical functioning was measured with an aggregate score of 4 items from the SF-36/SF-12 scales 'physical functioning', 'bodily pain', and 'vitality'. Emotional distress was measured with the mental distress subscale (SCL-8) from the Danish version of the Hopkins Symptom Checklist-90. Illness worry was measured with the 6-items Whiteley Index.

Results

For all cohorts, bi-factor models established that despite some multi-dimensionality, the total sum score of the BDS checklist adequately reflected physical symptom burden and illness severity. The BDS checklist had acceptable convergent validity to measures of overall health (r=0.25-0.58), physical functioning (r=0.22-0.58), emotional distress (r=0.47-0.62), and illness worry (r=0.36-0.55). Acceptability was good with low numbers of missing responses to items (<3%). Internal consistency was high ($\alpha \geq 0.879$). BDS score means varied and reflected symptom burden across cohorts (13.03-46.15). We provide normative data for the Danish general population.

Conclusions

The BDS checklist total sum score can be used as measure of symptom burden and FSD illness severity across settings. These findings establish the usefulness of the BDS checklist in clinic and in research both as a diagnostic screening and as an instrument for assessment of illness severity.

Strengths and limitations of this study

- The study included data from three cohorts and settings: A general population, primary care patients, and patients from a specialized setting
- Well-validated measures were used to determine convergent validity
- All included cohorts had large sample sizes
- Only self-reported measures were included
- Convergent validity was not investigated with other measures of physical symptom burden

Introduction

Persistent physical symptoms (PPS) are common in medical settings and the general population [1-4]. The symptoms present across a continuum from one or a few momentary to numerous symptoms from multiple locations in the body. Having a high number of symptoms has been associated with poor health, poor functional status, and increased health care use [5-9]. Hence, assessment of the burden of persistent physical symptoms is valuable in both clinical care and in research.

For this purpose, self-reported symptom questionnaires are frequently used. They are manageable, non-invasive tools. Several screening questionnaires exist: The Hopkins Symptom Checklist [10], the Patient Health Questionnaire [11], the Somatic Symptom Scale-8 [12 13], the brief form of the Giessen Subjective Complaints List [14], and others [15-17]. However, the existing questionnaire measure PPS without consideration of the well-known aggregation of such symptoms into symptom clusters, and hence, without acknowledgement of the real structure of PPS as it occurs in both the community and in clinical setting[18-21].

When PPS occur in the absence of (other) physical or mental conditions, or when they cause individual suffering and functional limitations beyond what could be expected based on such diseases, they constitute the very core of the disorders captured under the umbrella definition of Functional Somatic Disorders (FSD). FSD cover both specialty-specific syndrome diagnoses such as fibromyalgia, irritable bowel and chronic fatigue, but also their pendants in psychosomatic medicine, somatoform and somatic symptom disorders [22].

In contrast to the above mentioned speciality-specific diagnoses, the proposed research diagnosis bodily distress syndrome (BDS) covers a broader range of functional somatic symptoms ranging from few symptoms with some effect on functioning to severe and disabling functional somatic disorders [18 19 21]. Hence, BDS provides the opportunity to assess and distinguish between conditions persisting as mono- or multi-syndromatic and still within the same framework of diagnostic approach [21 23]. The diagnostic construct was developed in a sample of patients from primary and secondary care, and the 30-items BDS checklist emerged [18]. BDS was confirmed in a new sample of primary care patients where the shortened 25-items BDS checklist was developed [19]. Subsequently, the construct of BDS has been confirmed in general population samples as well[21 24]. BDS presents symptoms grouped in four symptom clusters: Cardiopulmonary (CP), gastrointestinal (GI), musculoskeletal (MS), and general symptoms (GS), and its usefulness and properties used for diagnostic categorisation into no BDS, a single/oligo-organ BDS type and a multi-organ BDS type has been established [19 21 24]. A major strength of the BDS checklist is its usefulness both as a screening and as diagnostic tool within clinical

practise and within epidemiological research [18 19 21 23], but the total BDS sum score has not yet been validated as a measure for the assessment of symptom burden and illness severity.

This study aims to explore whether the BDS checklist can be used as a continuous score to measure symptom burden (i.e. in those individuals that may fall under the diagnostic threshold or what we believe to be clinically relevant) and illness severity (in those individuals fulfilling diagnostic criteria for FSD). In order to elicit the BDS checklist's usability across settings, its structural validity and psychometric properties will be explored in three different populations: the general population, primary care patients, patients in a specialized clinical setting.

Methods

Population

This cross-sectional study included baseline data from three cohorts:

Cohort 1: A general population cohort (DanFunD, n=9656, response rate=33.7%) established with the purpose to investigate and unravel the epidemiology of FSD [25]. The cohort was obtained from the Danish Central Personal Register and drawn as a random sample of the adult Danish background population aged 18-69 years. Participants lived in 10 municipalities in the south-western part of the greater Copenhagen area. All participants were born in Denmark.

Cohort 2: A cohort of primary care patients (KOS, n=2480, response rate=59.5%) established in order to investigate contact and disease patterns in general practice [26]. Participants were included consecutively from 388 general practitioners from the Central Denmark Region. Included participants were 18 years or older and had completed a health-related face-to-face consultation with their general practitioner.

Cohort 3: Data from a specialized clinical setting at the Research Clinic for Functional Disorders and Psychosomatics, Aarhus University Hospital in Denmark (STreSS-3, STreSS-4, STreSS-5, n=492, response rate=100%) [27-31]. These cohorts had been part of a group of studies with the shared aim to investigate new treatments for patients with multi-organ BDS aged 20 years or older.

Measures

Self-reported data of physical symptoms, overall health, physical health, mental health, and illness worry was included. The measures and data were not completely consistent across the three included cohorts.

Physical symptoms were assessed with the Danish version of the 25-items BDS checklist (Appendix A) [19 21]. The checklist asks "during the last *(specific time frame)* have you been bothered by" followed by a list of 25 symptoms comprising the four symptom clusters of BDS. The BDS checklist measures symptoms on a five-point rating scale from 0 ('not at all bothersome') to 4 ('a lot bothersome'). We calculated a sum score by adding the single item scores from the 25 items (ranging from 0 to 100). The timeframe covered was 12 months for the general population cohort and four weeks for the other two cohorts.

Overall health was assessed with a single item from the 36-items Short Form Health Survey (SF-36) [32], estimating self-perceived health on a five-point rating scale from 'excellent' to 'poor'. Higher score on this item indicates poorer health. No specific time frame was surveyed in neither of the cohorts.

Physical functioning was measured with a shortened version of an aggregate score of the SF-36 subscales 'physical functioning', 'bodily pain', and 'vitality [30 32-34]. The shortened version consisted of four items (two items from the 'physical function' subscale, one item from the 'bodily pain' subscale, and one item from the 'vitality' subscale) which are part of the SF-12, addressing limitations in moderate and strenuous activities because of physical health and pain interference. For each item a *z*-score was calculated using mean and standard deviation (SD) from the general Danish population. Mean of the *z*-scores from the three subscales results in an aggregate *z*-score. This is then transformed into a *t*-score (mean=50, SD=10). Higher scores indicate better physical health. We tested the correlation of the *t*-score of the shortened version aggregate score against the full SF-36 aggregate score in cohort 3, and correlation was high (*Spearman rho*=0.89, 95% CI: 0.87;0.91). Unfortunately, it was not possible to investigate convergent validity to the aggregate score in the data on the primary care cohort, because we had limited access to data. These analyses were therefore only performed in the general population cohort and the cohort from specialized clinical setting. The time frame covered was four weeks for both cohorts.

Emotional distress was measured with the mental distress subscale (SCL-8) from the Danish version of the Hopkins Symptom Checklist (SCL-90) [35 36]. SCL-8 consists of eight items addressing impairment of overall worries, depression, and anxiety. Answers were calculated as mean scores from a scale ranging from 0 ('not at all bothersome') to 4 ('a lot bothersome'). Higher scores indicated higher emotional distress. The time frame covered was one week for the general population cohort and four weeks for the two other cohorts.

Illness worry was measured with the Whiteley Index 6 items version revised (Whiteley-6-R) [37], addressing the respondent's fear of being ill and whether they attribute current bodily sensations to

somatic illness[1]. Answers were calculated as mean scores from a scale ranging from 0 ('not at all bothersome') to 4 ('a lot bothersome'). Higher scores indicated higher health anxiety. The time frame covered was 12 months for the general population cohort and four weeks for the two other cohorts.

Validation procedure and statistical analyses

The analyses for the current study were performed according to the Consensus-based Standards for the selection of health Measurement Instrument (COSMIN) framework [38].

All statistical analyses were performed using STATA version 16.0 [39], except for the structural equation modelling which was performed using Mplus version 8.1 [40].

Construct validity was tested by means of structural validity and convergent validity.

Structural validity was tested with confirmatory factor analyses (CFA) with WLSMV (Weighted Least Squares Means and Variance adjusted) estimation due to categorical responses for all items [40]. We wanted to test if it was permissible to model the BDS checklist as unidimensional despite the previous evidence of some multi-dimensionality [18 19 21 24]. Furthermore, we wanted to test if the raw total BDS sum score would be an adequate reliable measure of the general factor (BDS). Therefore, four different CFAs were performed: 1) An one-level one factor model, 2) an one-level four factor model, using factors resembling the four BDS symptom clusters previously reported [19 21], 3) a two-level four factor model, representing a second order common factor (BDS) underlying the four BDS symptom clusters, and 4) a bi-factor CFA, reflecting each symptom to load on a general factor (BDS) and on one of the four specific BDS symptom clusters. Illustrations of the four types of CFAs are displayed in Appendix B.

In all CFAs, model fit was assessed as follows: A Root Mean Square Error of Approximation (RMSEA) <0.05 indicates very good fit, 0.05-0.08 indicates a good fit, and ≥0.08 indicates a poor fit. Comparative fit Index (CFI) and Tucker-Lewis fit Index (TLI) at 0.90-0.95 indicate an acceptable fit and levels >0.95 indicate a good fit. A Standardized Root Mean square Residual (SRMR) <0.08 indicates good fit [41].

Convergent validity was tested with Spearman's correlations, and associations between the BDS checklist and overall health (one item from SF-36)[32], physical function (an aggregate score of four items from the SF-36) [42], emotional distress (SCL-8) [35], and illness worry (Whiteley-6-R) [37] were

[1] In the primary care sample, one of the items in the WI-6 "Do you worry about the possibility that you suffer from an illness you have heard or read about" was expressed as "Do you worry about the possibility that you suffer from an illness".

performed. Based on previous literature [12 14 15 17 43], we hypothesized that the BDS checklist would show moderate convergent validity (r=0.40-0.60) with the four measures, and we expected lower correlations in the sample from specialized setting. Expected differences on the BDS checklist with one unit difference to the SCL-8, the four items aggregate score for physical functioning, and Whiteley-6-R were estimated with linear regression.

BDS checklist item and scale characteristics, i.e. item means (SD), sum score means, score distribution, item total correlation, corrected for overlap, and aspects of acceptability, i.e. percentage of missing items, were examined and computed as descriptive statistics for each of the three samples. Internal consistency was measured with Cronbach's α coefficients where values between 0.7 and 0.95 are acceptable [38].

Ethical considerations

The current study was carried out in accordance with the relevant guidelines and regulations.

For all three cohorts, written informed consent was obtained from each participant before entering the studies [25-31].

Cohort 1: Approved by the independent ethics committees the Ethical Committee of Copenhagen County (Ethics Committee: KA-2006-0011; H-3-2011-081; H-3-2012-0015) and the Danish Data Protection Agency.

Cohort 2: Approved by the Danish Data Protection Agency and the Health and Medicines Authority. According to Danish law, approval from the health research ethics system was not needed.

Cohort 3: The STreSS-3 trial was registered with ClinicalTrials.gov, number NCT01518634, EudraCT number 2011-004294-87. The STreSS-4 trial was registered with ClinicalTrials.gov, number NCT01518647. The STreSS-5 cohort study was approved by the Danish Data Protection Agency.

Patient and Public Involvement

It was not appropriate to involve patients or the public in the design, or conduct, or reporting, or dissemination plans of our research.

Results

Sample characteristics

Median age in the general population sample was 54 years (IQR: 44-64), and 53.9% were females.

In the primary care sample, mean age was 54.3 years (SD: 17.5), and 62.5% were females.

In the sample from specialized setting, mean age was 39.4 years (SD: 8.8), and 81.1% were females.

Structural validity

The one-level one factor model showed unacceptable fit indices in all three cohorts (Table 1).

Table 1: Goodness of fit parameters from the CFA models

	RMSEA	95% CI		CFI	TLI	SRMR	X^2	df	p
One-level one factor CFA									
General population	0.111	0.110	0.112	0.723	0.697	0.09	32743.1	275	<0.0001
Primary care	0.419	0.147	0.151	0.697	0.670	0.119	15126.5	275	<0.0001
Specialized setting	0.149	0.144	0.153	0.621	0.586	0.115	3261.5	275	<0.0001
One-level four factor CFA									
General population	**0.062**	0.061	0.063	**0.914**	**0.905**	**0.052**	10290.96	269	<0.0001
Primary care	0.082	0.080	0.084	**0.91**	**0.90**	**0.067**	4666.85	269	<0.0001
Specialized setting	0.091	0.086	0.096	0.862	0.846	**0.076**	1355.96	269	<0.0001
Two-level four factor CFA									
General population	**0.061**	0.060	0.062	**0.917**	**0.908**	**0.052**	9951.02	271	<0.0001
Primary care	0.08	0.078	0.082	**0.914**	**0.905**	**0.068**	4482.39	271	<0.0001
Specialized setting	0.089	0.084	0.093	0.867	0.853	**0.076**	1315.04	271	<0.0001
Bi-factor CFA									
General population	**0.048**	0.046	0.049	**0.954**	**0.944**	**0.04**	5680.8	250	<0.0001
Primary care	**0.053**	0.051	0.055	**0.965**	**0.958**	**0.042**	1977.4	250	<0.0001
Specialized setting	**0.059**	0.054	0.065	**0.945**	**0.934**	**0.051**	681.1	250	<0.0001

Abbreviations: CFA=Confirmatory factor analysis; RMSEA=Root Mean Square Error of Approximation; CI=Confidence Interval; CFI=Comparative Fit Index; TLI=Tucker-lewis fit Index; χ^2=Likelyhood Ratio Test; df=degrees of freedom, p=p-value.
Bold: Indicates a good or acceptable fit between the specified model and the observed model in the data.

Fit indices for the one-level four factor model which has been confirmed in previous studies [19 21] and the two-level four factor model showed more acceptable fits. These models revealed correlations between the four BDS symptom clusters and loadings from an underlying BDS factor to the four BDS symptom clusters that may imply a bi-factor model. Good fit indices were seen for the bi-factor model. Hence, a model reflecting a general factor (BDS) and four independent factors (BDS symptom clusters) all explaining the variance of the 25 symptoms in the BDS checklist, was confirmed (Figure 1). Loadings from the general BDS factor were generally higher than loadings from the four symptom

clusters; for the population cohort this was the case for 72% of symptoms, in the primary care cohort it accounted for 64% of symptoms, and in the specialized setting it accounted for 52% of symptoms. Loading from the general BDS factor was smaller than loading from the four symptom clusters for six symptoms (*frequent, loose bowel movements; diarrhoea, pains in arms and legs; muscular aches or pains; pains in the joints; concentration difficulties*) in all three cohorts.

Figure 1 around here

Figure 1: Illustration and factor loadings from the bi-factor model across all three cohorts.

Abbreviations: Gen.=general population; Prim.=primary care; Spec.=specialized clinical setting; CP=cardiopulmonary, GI=gastrointestinal; MS=musculoskeletal; GS=general symptoms; BDS=bodily distress syndrome

Convergent validity

In the general population sample, our hypothesis was met for all measures. The BDS checklist had moderate convergent validity compared to the SF-36 item for overall health (r=0.48, 95% CI: 0.46;0.49, p<0.0001), the four items aggregate score for physical health (r=-0.58, 95% CI: -0.59;-0.56, p<0.0001), the SCL-8 for emotional distress (r=0.52, 95% CI: 0.51;0.54, p<0.0001), and the Whiteley-6-R for illness worry (r=0.53, 95% CI: 0.52;0.55, p<0.0001). Expected difference on the BDS checklist with one unit difference on the four items aggregate score was -0.80 (95% CI: -0.82;-0.78), 12.26 (95% CI: 11.89;12.63) with SCL-8, and 8.93 (95% CI: 8.64;9.21) with Whiteley-6-R (Appendix C).

For the primary care sample, our hypothesis was met for all measures as well, however, for some of the measures, the association was stronger than hypothesized. We found moderate convergent validity compared to the SF-36 item for overall health (r=0.58, 95% CI: 0.56;0.61, p<0.0001), the SCL-8 for emotional distress (r=0.62, 95% CI: 0.59;0.64, p<0.0001), and the Whiteley-6-R for illness worry (r=0.55, 95% CI: 0.52;0.58, p<0.0001). Expected difference on the BDS checklist with one unit difference on the SCL-8 was 10.60 (95% CI: 10.08;11.12) and 10.01 (95% CI: 9.44;10.59) with Whiteley-6-R (Appendix C).

For the sample from specialized setting, our hypothesis about the correlations being weaker in the specialized setting was met. Moderate convergent validity was seen with emotional distress (r=0.47, 95% CI: 0.40;0.54, p<0.0001) while weaker correlations were seen for overall health (r=0.25, 95% CI: 0.17;0.33, p<0.0001), physical health (r=-0.22, 95% CI: -0.30;-0.12, p<0.0001), and illness worry (r=0.36, 95% CI: 0.28;0.43, p<0.0001). Expected difference on the BDS checklist with one unit

difference on the four items aggregate score for physical health was -0.41 (95% CI: -0.56;-0.26), 7.92 (95% CI: 6.65;9.18) with SCL-8, and 5.88 (95% CI: 4.58;7.17) with Whiteley-6-R (Appendix C).

Response distributions and acceptability

BDS checklist item and scale characteristics are displayed in Table 2. Item means varied from 0.15-1.09 in the general population sample, from 0.31-1.53 in the primary care sample, and from 0.81-3.34 in the sample from specialized setting. While the item with the lowest mean varied across samples, the item *'excessive fatigue*' had the highest mean value in all samples. Most item total correlations, corrected for overlap, exceeded 0.4.

Table 2: Item and scale characteristics

Item	General population (n=9656)			Primary care (n=2480)			Specialized setting (n=492)		
	Missing %	Mean (SD)	Item total correlation*	Missing %	Mean (SD)	Item total correlation*	Missing %	Mean (SD)	Item total correlation*
Palpations/heart pounding	0.9	0.45 (0.74)	0.427	3.1	0.61 (0.93)	0.544	0.2	1.46 (1.25)	0.480
Precordial discomfort	1.1	0.29 (0.61)	0.410	2.9	0.46 (0.81)	0.517	0.2	1.09 (1.16)	0.391
Breathlessness without exertion	1.0	0.36 (0.71)	0.426	2.9	0.63 (0.99)	0.509	0.2	1.27 (1.25)	0.511
Hyperventilation	1.1	0.15 (0.47)	0.321	3.3	0.37 (0.81)	0.448	0.2	0.93 (1.19)	0.380
Hot and cold sweats	1.2	0.46 (0.80)	0.429	3.2	0.64 (0.96)	0.533	0.2	1.88 (1.35)	0.523
Dry mouth	1.2	0.39 (0.76)	0.432	3.4	0.59 (0.98)	0.501	0.2	1.33 (1.36)	0.442
Frequent loose bowel movements	1.0	0.65 (0.86)	0.403	3.4	0.61 (0.95)	0.391	0.2	1.41 (1.32)	0.357
Abdominal pains	1.4	0.48 (0.76)	0.511	3.5	0.57 (0.90)	0.548	0.2	1.81 (1.22)	0.491
Feeling bloated/full of gas/distended	1.1	0.74 (0.91)	0.524	2.9	0.78 (1.03)	0.532	0.2	2.09 (1.30)	0.484
Diarrhoea	1.1	0.33 (0.63)	0.387	3.5	0.37 (0.80)	0.361	0.2	0.81 (1.13)	0.348
Regurgitations	1.4	0.43 (0.74)	0.392	3.3	0.35 (0.74)	0.456	0.2	1.05 (1.09)	0.430
Nausea	0.9	0.26 (0.57)	0.465	2.9	0.50 (0.87)	0.564	0.2	1.73 (1.27)	0.402
Burning sensation of the upper part of stomach/epigastrium	0.9	0.31 (0.68)	0.441	3.1	0.31 (0.72)	0.483	0.2	1.11 (1.26)	0.512
Pains in arms or legs	1.0	0.87 (1.08)	0.538	3.1	1.21 (1.29)	0.563	0.2	2.68 (1.24)	0.472
Muscular aches or pains	1.3	0.98 (1.01)	0.572	3.2	1.30 (1.23)	0.584	0.2	2.96 (1.10)	0.485
Pains in the joints	1.6	0.96 (1.07)	0.490	3.9	1.20 (1.27)	0.560	0.2	2.57 (1.32)	0.491
Feeling of paresis or localized weakness	1.4	0.16 (0.55)	0.365	4.0	0.33 (0.83)	0.460	0.2	1.22 (1.40)	0.481
Back ache	1.3	1.00 (1.06)	0.492	3.3	1.21 (1.29)	0.542	0.2	2.49 (1.37)	0.377
Pain moving from one place to another	1.4	0.27 (0.71)	0.489	3.8	0.54 (0.99)	0.544	0.2	2.13 (1.48)	0.403
Unpleasant numbness or tingling sensations	1.3	0.25 (0.67)	0.410	4.0	0.34 (0.83)	0.475	0.2	2.00 (1.43)	0.528
Concentration difficulties	0.7	0.60 (0.82)	0.545	2.9	0.85 (1.05)	0.540	0.2	2.53 (1.11)	0.437
Excessive fatigue	0.7	1.09 (1.01)	0.614	2.3	1.53 (1.20)	0.625	0.2	3.34 (0.83)	0.418
Headache	0.8	0.66 (0.89)	0.455	2.8	0.89 (1.08)	0.489	0.2	2.25 (1.22)	0.326
Impairment of memory	0.7	0.60 (0.83)	0.517	2.7	0.80 (1.06)	0.521	0.2	2.29 (1.27)	0.476
Dizziness	0.8	0.34 (0.67)	0.491	2.5	0.58 (0.94)	0.553	0.2	1.75 (1.30)	0.505
Scale									
Total scale missing (%)		0.6			2.7			0.2	
Mean (SD)		13.03 (10.36)			17.33 (13.79)			46.15 (15.91)	
Percentiles									
5%		1			2			22	
10%		3			3			26	
25%		6			7			34	
50% (median)		11			14			45	
75%		18			24			57	
90%		27			37			67	
		34			45			73	

*Item total correlation, corrected for overlap. 25% percentile and 75% percentile=interquartile ranges.**Abbreviations**: SD=standard deviation; IQR=interquartile range

Internal consistency was good in all three samples: α=0.887 in the general population sample, α=0.908 in the primary care sample, and α=0.879 in the sample from specialized setting.

BDS score distribution differed across samples (Figure 2) as did total sum score means; it was lowest in the general population (13.03, SD: 10.36) and highest in specialized setting (46.15, SD: 15.91) (Table 2). Acceptability was good, and the numbers of missing responses were generally low in the general population (total 0.6%) and specialized setting (total 0.2%) while it was slightly higher in primary care (total 2.7%).

Figure 2 around here.

Figure 2: Distribution of the BDS total sum score across all three cohorts.

BDS total sum scores were grouped into five categories: 0-20, 21-40, 41-60, 61-80, and 81-100. The vast majority of the general population respondents (96.6%) and primary care patients (90%) scored below 41, while this was only the case for a smaller fraction of the patients from specialized setting (38.7%) (Table 3). Data from each of the three samples and for all three samples pooled together are shown as cumulative percentages across sex and age groups in Appendix D.

Table 3: Grouping of BDS scores across samples

	General population		Primary care		Specialized setting	
Categories of BDS score	n	%	n	%	n	%
0-20	7.762	80.4	1.617	65.2	20	4.1
21-40	1.607	16.6	616	24.8	170	34.6
41-60	208	2.2	156	6.3	204	41.5
61-80	18	0.2	23	0.9	87	17.7
81-100	0	0	2	0.1	10	2.0
Missing	61	0.6	66	2.7	1	0.2

Abbreviations: BDS=Bodily distress syndrome

More detailed information about the normative data from each of the three samples and for all three samples pooled together are shown in Tables 1-4 in Appendix D.

Discussion

Principal findings

This is the first study to establish that, despite some multi-dimensionality, the 25-items BDS checklist can be used as a continuous score to measure symptom burden and illness severity in the general population, in primary care, and in specialized settings. Used as a total sum score with a range from 0-100, the BDS checklist had acceptable convergent validity with measures of overall health, physical health, emotional distress, and illness worry. Internal consistency was good in all three cohorts ($\alpha \geq 0.879$) as was acceptability. Thus, the BDS checklist may work as a simple symptom checklist but also as a diagnostic screening tool for use in clinical work and in research across different settings.

We found the symptom *'excessive fatigue*' to have the highest mean value in all three cohorts. This is in line with a recent German population-based study, finding *'tiredness*' to be one of the leading symptoms [44].

The three cohorts differed in number of symptoms that had higher loadings on the general BDS factor than on the four-symptom clusters ranging from 72% of symptoms in the general population cohort to 52% in the cohort from specialized clinical setting. The latter group contains patients with longstanding and severe FSD. In this group, the symptom load is high and specific symptom clusters may therefore stand out compared to the less affected participants from the general population with a more scattered symptom picture.

Previous studies have argued that the best fitting model for the BDS checklist was a one-level four factor model (Appendix B) [19 21 24]. However, the objectives of these studies were to confirm the BDS as case finding instrument in other samples with inspiration from the original studies in which the concept of BDS was developed and initially tested [18 19]. In the current study, we have taken it several steps further and tested various structural equation models in three different populations at the same time. The indicators of a bi-factor model is 1) if inter-correlation between the sub-scales in the CFA exceeds 0.3, 2) if loading on the first order factors on the second order factors exceeds 0.5 [45], and 3) if the ratio between the first and second eigenvalues exceeds 3 [46]. All parameters were fulfilled in the general population cohort and in the primary care cohort. For the cohort from specialized setting the ratio between the first and second eigenvalues was 2.68, but otherwise the parameters were fulfilled. This implies that the results from this study do not disqualify results from previous research, but the presence of some multidimensionality is not strong enough to disqualify the interpretation of the BDS checklist as unidimensional as well.

Correlations between the BDS checklist and self-rated measures of overall health, physical health, emotional distress, and illness worry were generally moderate, especially in the general population and primary care cohort. This was as expected as previous literature has shown the same association between symptom load and reduced function [6 7] . The difference between results on patients in the specialised settings and the two other populations may be caused by the nature of self-reported measures, where patients in specialized setting still have the opportunity to rate their perceived health as excellent even though they have been referred to specialized medical care because of invalidating physical symptoms. These aspects may produce precision limitations in some settings and may especially be pronounced in smaller samples. Furthermore, the distribution of sex differs across populations which may affect the results on convergent validity.

Strengths and weaknesses of the study

A major strength of this study is the inclusion of three different populations. To our knowledge, this approach of testing an instrument and using the same methodology in different populations is rare as most other studies concern only one setting at a time [11 12 14 17]. Also, the sample size within each cohort was large. We conducted a thorough validation procedure, using different structural equation models and testing convergent validity to several valid measures.

Weaknesses of the study include: Only self-reported outcomes were used and data measures were not completely consistent across the included cohorts; hence, we chose to apply the intersection of items in order to gain equivalent proxy measures. We did not have the opportunity to compare the BDS checklist to other measures of physical symptoms or – for the primary care cohort and the cohort from specialized clinical setting – to the physician's report. Furthermore, in the linear regression analyses, the assumption of normality of the residuals was not fully met for the primary care cohort and the cohort from specialized clinical care why these results should be interpreted with caution. Finally, as this study had a cross-sectional design, it was not possible to evaluate responsiveness of the BDS checklist.

Difference in results compared to others

To our knowledge, this is the first study to address the usefulness of the BDS checklist as a measure of physical symptom burden and illness severity. Another symptom checklist which has been widely used within primary care and general population studies for measuring the severity of

physical symptoms is the 15-items Patient Health Questionnaire (PHQ-15) [11 17]. It consists of 15 items concerning some of the symptoms from the same four organ systems as the BDS checklist, plus the symptoms 'fainting', 'sleeping problems', 'menstrual problems', and 'sexual pains/problems' not included in the BDS checklist. The PHQ-15 is scored on a three-point rating scale from 'not bothered at all' (0) to 'bothered at lot' (2), whereas the BDS checklist uses a five-point rating scale. In one study, including a sample from the general Swedish population, factor analyses of the structural properties of the PHQ-15 showed a four-factor model, but on the basis on a scree test plot they finally concluded that only one factor should be extracted [47]. Other studies found a bi-factor model to have the best fit to the PHQ-15 [48 49]. Hence, the PHQ-15 may have the same structural properties as the BDS checklist, but with fewer items to take into account as well as fewer response categories which may make it more prone to floor and ceiling effects. In a shorter version of the PHQ-15, the Somatic Symptoms Scale-8 (SSS-8), the three-point rating scale is replaced with a five-point rating option as in the BDS checklist [12 13 50]. However, neither the PHQ-15 nor the SSS-8 is validated for use as diagnostic categorization of respondents. Other symptom questionnaires resembling the same four factor structure and the same five answer categories as the BDS checklist are the 24-items Giessen Subjective Complaints List and its newer shortened version with 8 items (GBB-8), however, they have only been established and used in German speaking countries[14].
The BDS checklist is, at present, the only symptom checklist providing both diagnostic categorization and a measure of symptom load/illness severity.

Clinical implications

This study provides a self-reported symptom checklist for measuring symptom burden and illness severity which can be used both as a diagnostic screening tool and as a measure of illness severity in large epidemiological studies and also in more selected patient samples and severely ill patients. Regarding FSD, previous research has suggested measures of symptom burden as the primary outcome [33]. However, the current study shows that the BDS checklist shows weaker correlation with measures of overall health, physical health, emotional distress, and illness worry in patients from highly specialized setting than in the general population and primary care. Hence, a simple count of bothersome symptoms may not be adequate when dealing with the more severely ill patients, as symptom burden may not be the only important domain of illness severity – others may be the level of impairment and mental morbidity.

Currently, it is unclear whether the here presented BDS total sum score reliably captures FSD illness severity than the distinction in single vs. multi-organ BDS (e.g. three vs. four clusters fulfilled). Nevertheless, a tool which is also able to measure severity of specific symptom clusters is helpful in specialized settings, as it is possible to elucidate which symptom cluster is experienced most bothersome by the patients.

Future research and perspectives

In this study we suggest the BDS checklist as a prominent tool as it can be used both as a measure of symptom burden and as a diagnostic screening tool for FSD, and we argue for its usefulness in both epidemiological and clinical research as well as in clinical practice. However, the criterion validity of the self-reported BDS checklist with physician's established diagnoses e.g. specialty-specific syndrome diagnoses and psychiatric diagnoses, is yet to be investigated across settings, and future studies regarding these aspects would be valuable in order to further establish the usefulness of the BDS checklist. Moreover, the additional value of counting the number of symptom clusters fulfilled in the staging of FSD deserves attention. Finally, we need a valid instrument to measure change over time, and the responsiveness of the BDS checklist sum score is worth exploring.

Acknowledgements

The authors would like to thank senior research Grete Moth and data manager Kaare Rud Flarup from the Research Unit of primary care for providing date from the primary care cohort. We are also thankful to the members of the DanFunD steering committee; Senior consultant MD PhD Lene Falgaard Eplov, MSc PhD Allan Linneberg, MSc PhD Susanne Brix Pedersen, and MD PhD Michael Eriksen Benros.

Contributors

MWP contributed to the conception and design of the study and the statistical analyses, interpreted the data, and drafted the article. AS and MR contributed to the conception and design of the study and interpretation of the data and provided general supervision of the work. EØ performed the statistical analyses, and contributed to the conception and design of the study and the interpretation of the data. TJ, TMD and PF contributed to the interpretation of the data. All authors contributed to critically revising the article for important intellectual content, and all authors read and approved the final version of the article.

Funding

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests

The authors declare no competing interests.

Data availability statement

Data are available on reasonable request from the corresponding author.

References

1. Henningsen P, Gundel H, Kop WJ, et al. Persistent Physical Symptoms as Perceptual Dysregulation: A Neuropsychobehavioral Model and Its Clinical Implications. *Psychosomatic medicine* 2018;80(5):422-31. doi: 10.1097/PSY.0000000000000588 [doi]
2. Kroenke K. A practical and evidence-based approach to common symptoms: a narrative review. *Annals of Internal Medicine* 2014;161(8):579-86. doi: 10.7326/M14-0461 [doi]
3. Henningsen P, Zipfel S, Sattel H, et al. Management of Functional Somatic Syndromes and Bodily Distress. *Psychotherapy and psychosomatics* 2018;87(1):12-31. doi: 10.1159/000484413 [doi]
4. Eliasen M, Kreiner S, Ebstrup JF, et al. Somatic Symptoms: Prevalence, Co-Occurrence and Associations with Self-Perceived Health and Limitations Due To Physical Health - A Danish Population-Based Study. *PloS one* 2016;11(3):e0150664. doi: 10.1371/journal.pone.0150664 [doi]
5. Tomenson B, Essau C, Jacobi F, et al. Total somatic symptom score as a predictor of health outcome in somatic symptom disorders. *The British journal of psychiatry : the journal of mental science* 2013;203(5):373-80. doi: 10.1192/bjp.bp.112.114405 [doi]
6. Bruusgaard D, Tschudi-Madsen H, Ihlebaek C, et al. Symptom load and functional status: results from the Ullensaker population study. *BMC Public Health* 2012;12:1085. doi: 10.1186/1471-2458-12-1085 [published Online First: 2012/12/20]

7. Creed FH, Davies I, Jackson J, et al. The epidemiology of multiple somatic symptoms. *Journal of psychosomatic research* 2012;72(4):311-17. doi: 10.1016/j.jpsychores.2012.01.009 [doi]
8. Ladwig KH, Marten-Mittag B, Lacruz ME, et al. Screening for multiple somatic complaints in a population-based survey: does excessive symptom reporting capture the concept of somatic symptom disorders? Findings from the MONICA-KORA Cohort Study. *Journal of psychosomatic research* 2010;68(5):427-37. doi: 10.1016/j.jpsychores.2010.01.009 [doi]
9. Eliasen M, Schröder A, Fink P, et al. A step towards a new delimitation of functional somatic syndromes: A latent class analysis of symptoms in a population-based cohort study. *J Psychosom Res* 2018;108:102-17.
10. Derogatis LR, Lipman RS, Rickels K, et al. The Hopkins Symptom Checklist (HSCL): a self-report symptom inventory. *Behavioral science* 1974;19(1):1-15.
11. 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):258-66.
12. Gierk B, Kohlmann S, Kroenke K, et al. The somatic symptom scale-8 (SSS-8): a brief measure of somatic symptom burden. *JAMA internal medicine* 2014;174(3):399-407. doi: 10.1001/jamainternmed.2013.12179 [doi]
13. Gierk B, Kohlmann S, Toussaint A, et al. Assessing somatic symptom burden: a psychometric comparison of the patient health questionnaire-15 (PHQ-15) and the somatic symptom scale-8 (SSS-8). *Journal of psychosomatic research* 2015;78(4):352-55. doi: 10.1016/j.jpsychores.2014.11.006 [doi]
14. Kliem S, Lohmann A, Klatt T, et al. Brief assessment of subjective health complaints: Development, validation and population norms of a brief form of the Giessen Subjective Complaints List (GBB-8). *J Psychosom Res* 2017;95:33-43. doi: 10.1016/j.jpsychores.2017.02.003 [published Online First: 2017/03/21]
15. Sitnikova K, Dijkstra-Kersten SMA, Mokkink LB, et al. Systematic review of measurement properties of questionnaires measuring somatization in primary care patients. *J Psychosom Res* 2017;103:42-62. doi: 10.1016/j.jpsychores.2017.10.005 [published Online First: 2017/11/24]
16. Hyland ME, Bacon AM, Lanario JW, et al. Symptom frequency and development of a generic functional disorder symptom scale suitable for use in studies of patients with irritable bowel syndrome, fibromyalgia syndrome or chronic fatigue syndrome. *Chronic Dis Transl Med* 2019;5(2):129-38. doi: 10.1016/j.cdtm.2019.05.003 [published Online First: 2019/08/02]

17. Zijlema WL, Stolk RP, Lowe B, et al. How to assess common somatic symptoms in large-scale studies: a systematic review of questionnaires. *Journal of psychosomatic research* 2013;74(6):459-68. doi: 10.1016/j.jpsychores.2013.03.093 [doi]

18. Fink P, Toft T, Hansen MS, et al. Symptoms and syndromes of bodily distress: an exploratory study of 978 internal medical, neurological, and primary care patients. *Psychosomatic medicine* 2007;69(1):30-39. doi: 69/1/30 [pii]

19. Budtz-Lilly A, Fink P, Ørnbøl E, et al. A new questionnaire to identify bodily distress in primary care: The 'BDS checklist'. *Journal of psychosomatic research* 2015;78(6):536-45. doi: 10.1016/j.jpsychores.2015.03.006 [doi]

20. Rosmalen JG, Tak LM, de Jonge P. Empirical foundations for the diagnosis of somatization: implications for DSM-5. *Psychological medicine* 2011;41(6):1133-42. doi: 10.1017/S0033291710001625 [doi]

21. Petersen MW, Schröder A, Jørgensen T, et al. The unifying diagnostic construct of bodily distress syndrome (BDS) was confirmed in the general population. *Journal of psychosomatic research* 2019;128:109868. doi: S0022-3999(19)30601-4 [pii]

22. Burton C, Fink P, Henningsen P, et al. Functional somatic disorders: discussion paper for a new common classification for research and clinical use. *BMC Med* 2020;18(1):34. doi: 10.1186/s12916-020-1505-4 [published Online First: 2020/03/04]

23. Petersen MW, Schröder A, Jørgensen T, et al. Irritable bowel, chronic widespread pain, chronic fatigue and related syndromes are prevalent and highly overlapping in the general population: DanFunD. *Scientific reports* 2020;10(1):3273-73. doi: 10.1038/s41598-020-60318-6

24. Schmalbach B, Roenneberg C, Hausteiner-Wiehle C, et al. Validation of the German version of the Bodily Distress Syndrome 25 checklist in a representative German population sample. *Journal of Psychosomatic Research* 2020;132:109991. doi: https://doi.org/10.1016/j.jpsychores.2020.109991

25. Dantoft TM, Ebstrup JF, Linneberg A, et al. Cohort description: The Danish Study of Functional Disorders. *Clin Epidemiol* 2017;9:127-39.

26. Moth G, Vedsted P, Olesen F, et al. Kontakt- og sygdomsmønsterundersøgelse KOS 2008: Forskningsenheden for Almen Praksis 2010.

27. Frolund Pedersen H, Holsting A, Frostholm L, et al. "Understand your illness and your needs": Assessment-informed patient education for people with multiple functional somatic

syndromes. *Patient Educ Couns* 2019;102(9):1662-71. doi: 10.1016/j.pec.2019.04.016 [published Online First: 2019/04/27]

28. Agger JL, Schröder A, Gormsen LK, et al. Imipramine versus placebo for multiple functional somatic syndromes (STreSS-3): a double-blind, randomised study. *Lancet Psychiatry* 2017;4(5):378-88. doi: 10.1016/S2215-0366(17)30126-8 [published Online First: 2017/04/15]
29. Pedersen HF, Agger JL, Frosthom L, et al. Acceptance and Commitment group Therapy for patients with multiple functional somatic syndromes: a three-armed trial comparing ACT in a brief and extended version with enhanced care. *Psychological medicine* 2018:1-10. doi: 10.1017/S0033291718001666.

10.1017/S0033291718001666 [doi]

30. Schröder A, Rehfeldt E, Ørnbøl E, et al. Cognitive-behavioural group treatment for a range of functional somatic syndromes: randomised trial. *Br J Psychiatry* 2012;200(6):499-507.
31. Fjorback LO, Arendt M, Ornbol E, et al. Mindfulness therapy for somatization disorder and functional somatic syndromes: randomized trial with one-year follow-up. *Journal of psychosomatic research* 2013;74(1):31-40. doi: 10.1016/j.jpsychores.2012.09.006 [doi]
32. Ware JJ, Kosinski, M., Gandek B. SF-36 Health Survey: Manual and Interpretation Guide. *Quality Metric* 1993
33. Rief W, Burton C, Frostholm L, et al. Core Outcome Domains for Clinical Trials on Somatic Symptom Disorder, Bodily Distress Disorder, and Functional Somatic Syndromes: European Network on Somatic Symptom Disorders Recommendations. *Psychosom Med* 2017;79(9):1008-15. doi: 10.1097/psy.0000000000000502 [published Online First: 2017/07/12]
34. Schröder A, Ørnbøl E, Licht RW, et al. Outcome measurement in functional somatic syndromes: SF-36 summary scores and some scales were not valid. *Journal of clinical epidemiology* 2012;65(1):30-41. doi: 10.1016/j.jclinepi.2011.03.012 [doi]
35. Derogatis LR. SCL-90-R, administration, scoring and procedures manual-II for the R(evised) version and other instruments of the Psychopathology Rating Scale Series. Townson: Clinical Psychometric Research 1992.
36. Christensen KS BP, Fink P. Measuring Mental Health by Questionnaires in Primary Care - Unidimensionality. Responsiveness and Compliance. . *European Psychiatric Review* 2010;3:8-12.

37. Carstensen TBW, Ørnbøl E, Fink P, et al. Detection of illness worry in the general population: A specific item on illness rumination improves the Whiteley Index. *J Psychosom Res* 2020;138:110245. doi: 10.1016/j.jpsychores.2020.110245 [published Online First: 2020/09/21]

38. Mokkink LB, Terwee CB, Patrick DL, et al. COSMIN Checklist Manual. Amsterdam: EMGO Institute for Health and Care Research 2012.

39. StataCorp. Stata Statistical Software. Release 16: College Station, TX: StataCorp LLC 2019.

40. Muthén LK, Muthén BO. Mplus User's Guide. Eighth Edition. Los Angeles, CA: Muthén & Muthén1998.

41. Hu Lt, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal* 1999;6(1):1-55. doi: 10.1080/10705519909540118

42. Ware J, Jr., Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Medical care* 1996;34(3):220-33.

43. Terluin B, van Marwijk HW, Ader HJ, et al. The Four-Dimensional Symptom Questionnaire (4DSQ): a validation study of a multidimensional self-report questionnaire to assess distress, depression, anxiety and somatization. *BMC Psychiatry* 2006;6:34. doi: 10.1186/1471-244X-6-34 [published Online First: 2006/08/24]

44. Beutel ME, Klein EM, Henning M, et al. Somatic Symptoms in the German General Population from 1975 to 2013. *Sci Rep* 2020;10(1):1595. doi: 10.1038/s41598-020-58602-6 [published Online First: 2020/02/02]

45. MD. HJT. Bifactor analysis in MPlus [Video file]: UK College of Education; 2016 [Available from: https://education.uky.edu/edp/apslab/events/2020.

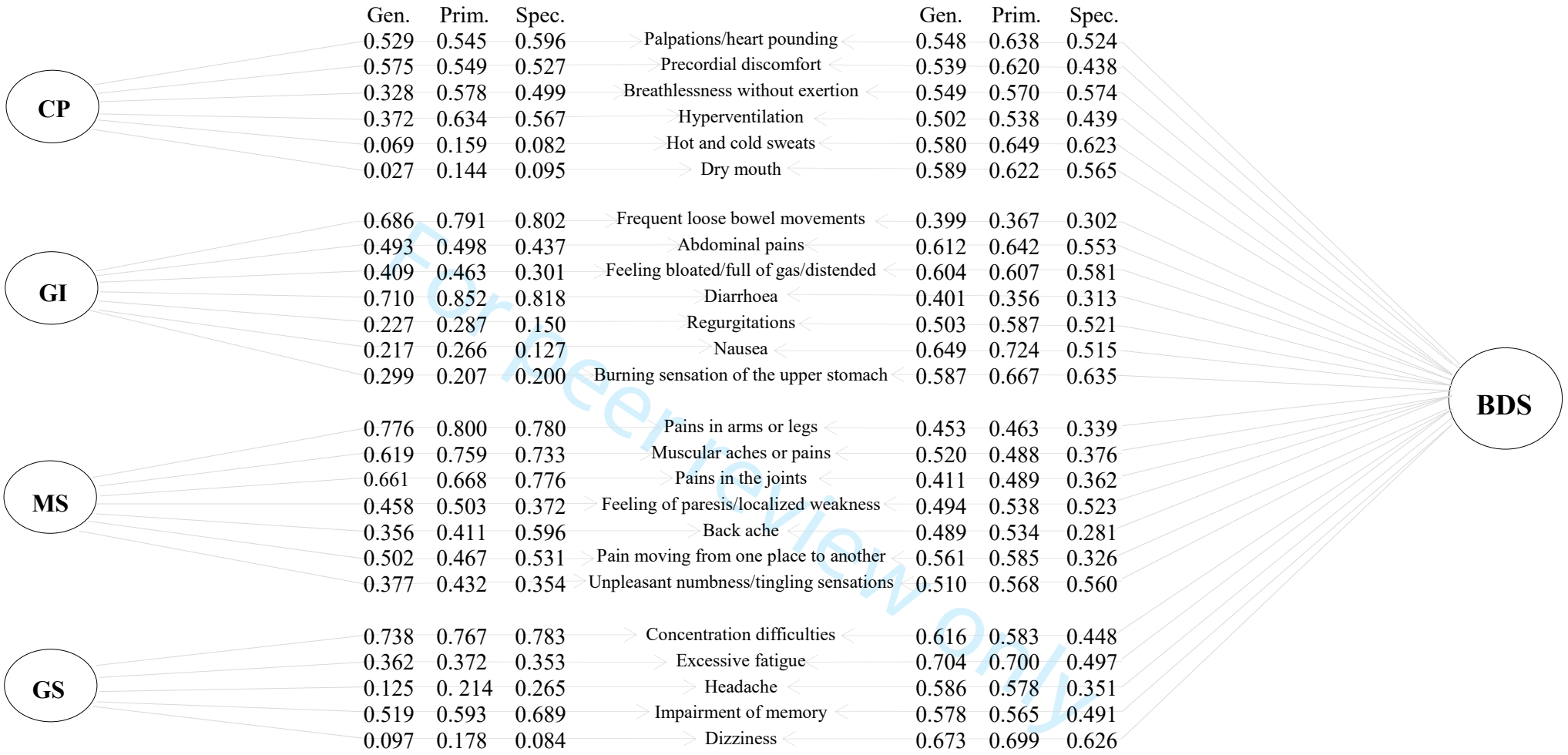
46. Cho SJ, Wilmer J, Herzmann G, et al. Item response theory analyses of the Cambridge Face Memory Test (CFMT). *Psychol Assess* 2015;27(2):552-66. doi: 10.1037/pas0000068 [published Online First: 2015/02/03]

47. Nordin S, Palmquist E, Nordin M. Psychometric evaluation and normative data for a Swedish version of the Patient Health Questionnaire 15-Item Somatic Symptom Severity Scale. *Scandinavian Journal of Psychology* 2013;54(2):112-17. doi: 10.1111/sjop.12029 [doi]

48. Cano-Garcia FJ, Munoz-Navarro R, Sese Abad A, et al. Latent structure and factor invariance of somatic symptoms in the patient health questionnaire (PHQ-15). *J Affect Disord* 2020;261:21-29. doi: 10.1016/j.jad.2019.09.077 [published Online First: 2019/10/11]

49. Witthoft M, Hiller W, Loch N, et al. The latent structure of medically unexplained symptoms and its relation to functional somatic syndromes. *International Journal of Behavioral Medicine* 2013;20(2):172-83. doi: 10.1007/s12529-012-9237-2 [doi]

50. Toussaint A, Kroenke K, Baye F, et al. Comparing the Patient Health Questionnaire - 15 and the Somatic Symptom Scale - 8 as measures of somatic symptom burden. *J Psychosom Res* 2017;101:44-50. doi: 10.1016/j.jpsychores.2017.08.002 [published Online First: 2017/09/05]



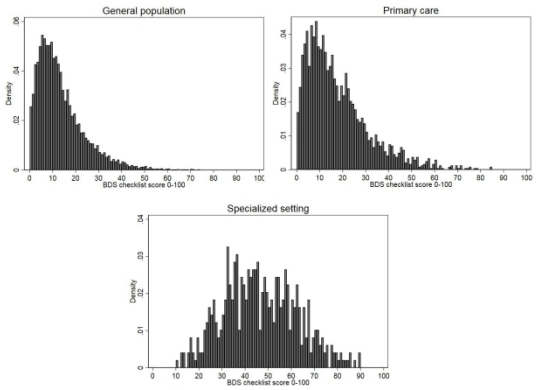


Figure 2: Distribution of the BDS total sum score across all three cohorts.

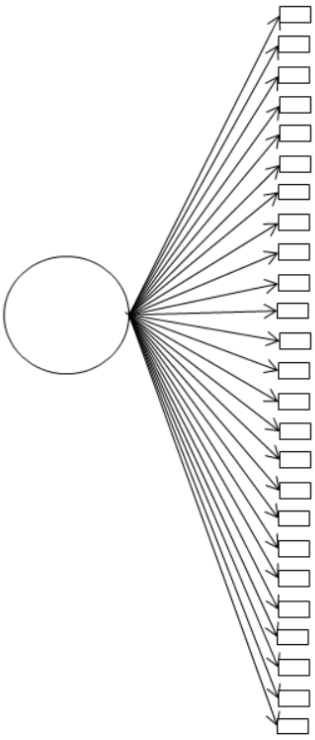
120x49mm (600 x 600 DPI)

Appendix A: The 25-items BDS checklist

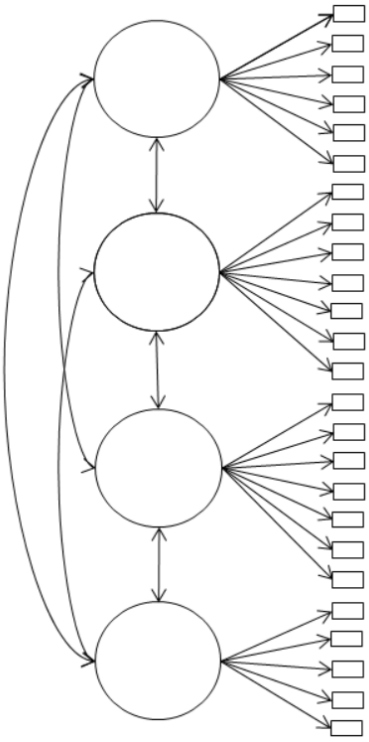
During the last 4 weeks*, have you been bothered by…		Not at all	A bit	Somewhat	Quite a bit	A lot
1	Palpations and heart pounding	□	□	□	□	□
2	Precordial discomfort	□	□	□	□	□
3	Breathlessness without exertion	□	□	□	□	□
4	Hyperventilation	□	□	□	□	□
5	Hot and cold sweats	□	□	□	□	□
6	Dry mouth	□	□	□	□	□
7	Frequent loose bowel movements	□	□	□	□	□
8	Abdominal pains	□	□	□	□	□
9	Feeling bloated/full of gas/distended	□	□	□	□	□
10	Diarrhoea	□	□	□	□	□
11	Regurgitations	□	□	□	□	□
12	Nausea	□	□	□	□	□
13	Burning sensation of the upper part of stomach/epigastrium	□	□	□	□	□
14	Pains in arms or legs	□	□	□	□	□
15	Muscular aches or pains	□	□	□	□	□
16	Pains in the joints	□	□	□	□	□
17	Feeling of paresis or localized weakness	□	□	□	□	□
18	Back ache	□	□	□	□	□
19	Pain moving from one place to another	□	□	□	□	□
20	Unpleasant numbness or tingling sensations	□	□	□	□	□
21	Concentration difficulties	□	□	□	□	□
22	Excessive fatigue	□	□	□	□	□
23	Headache	□	□	□	□	□
24	Impairment of memory	□	□	□	□	□
25	Dizziness	□	□	□	□	□

* This time frame was changed to 12 months in the general population cohort

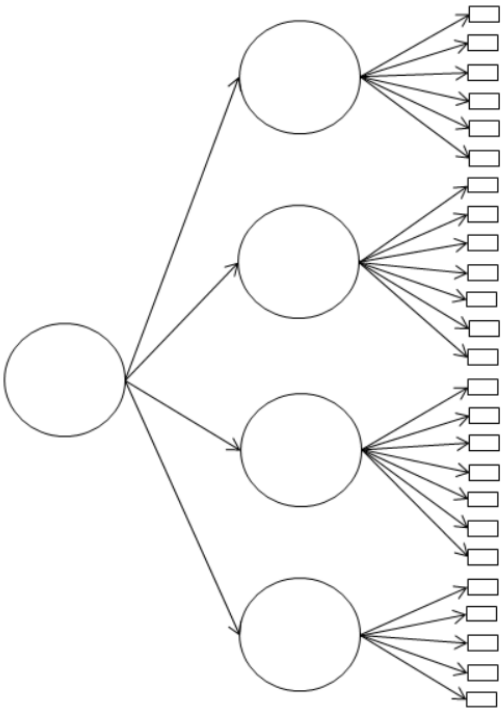
Appendix B: Illustrations of the theoretical models of confirmatory factor analyses



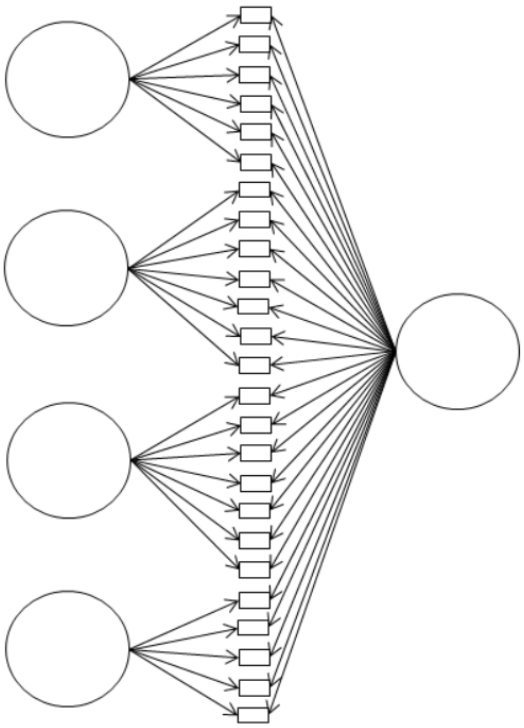
One-level one factor model



One-level four factor model



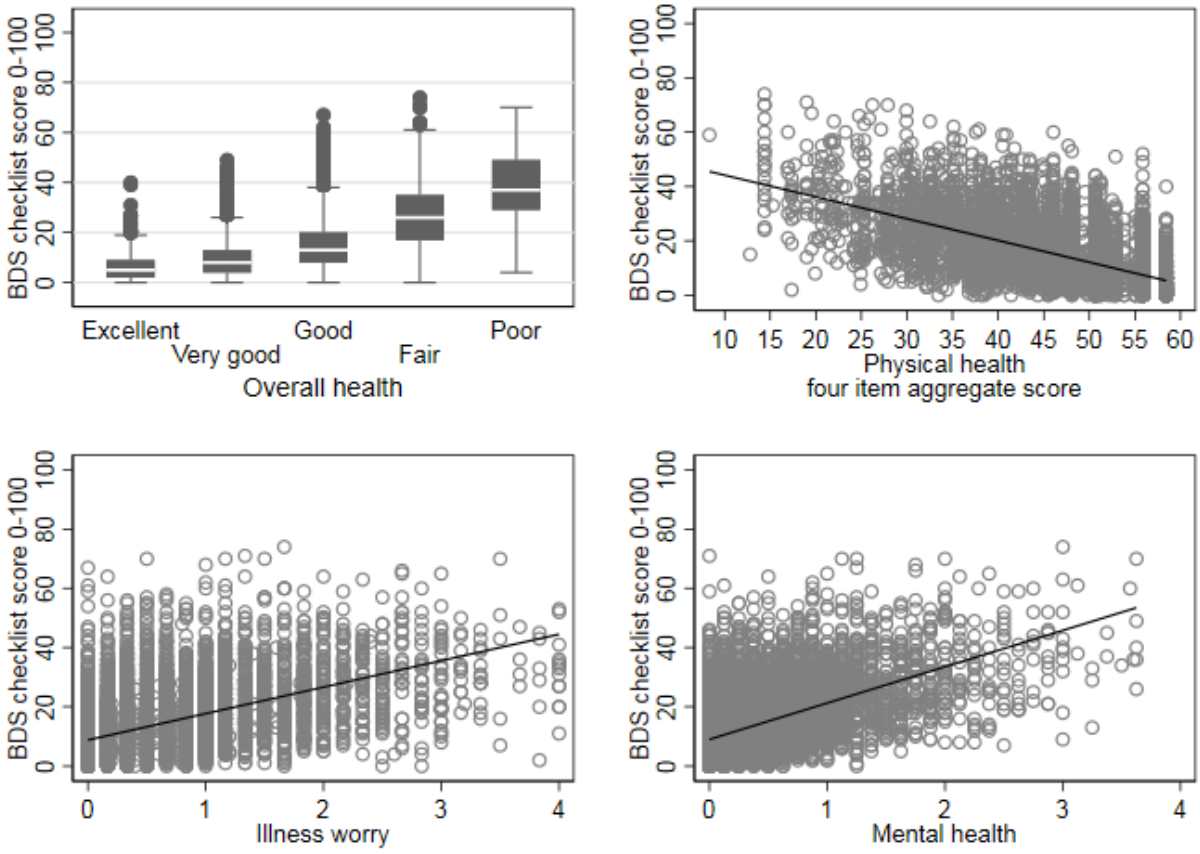
Two-level four factor model



Bi-factor model

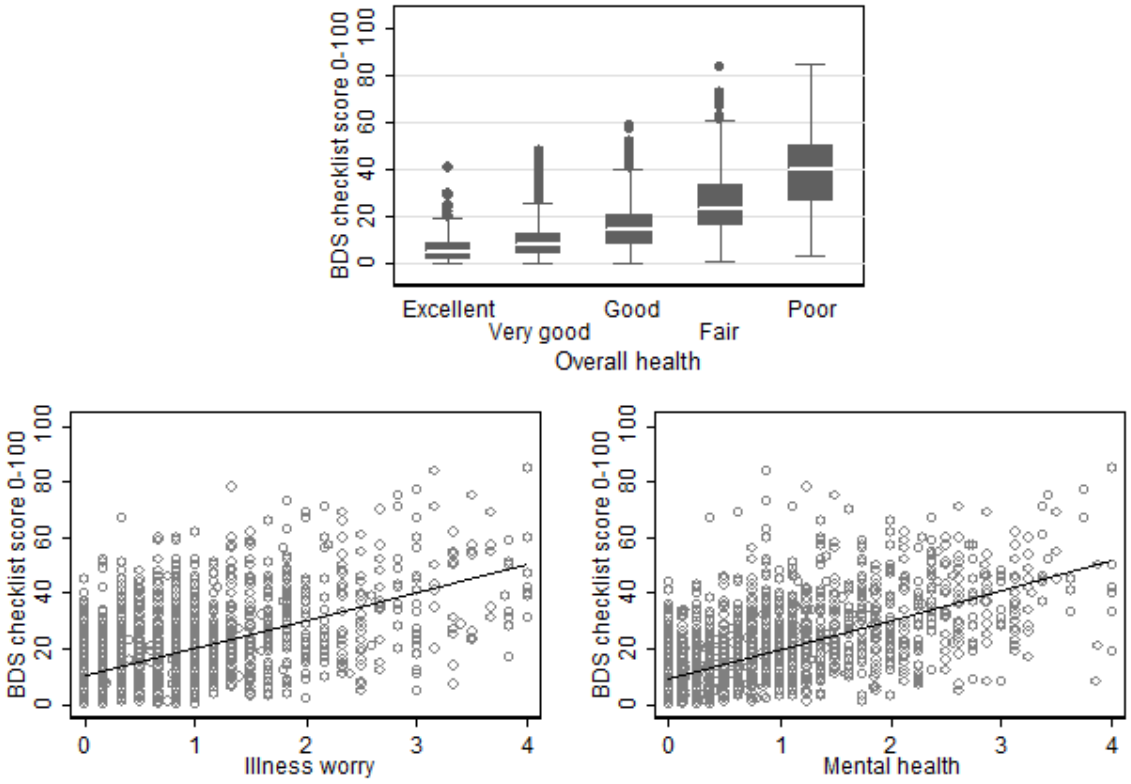
Appendix C: Descriptive plots of associations between the BDS checklist and other measures

General population



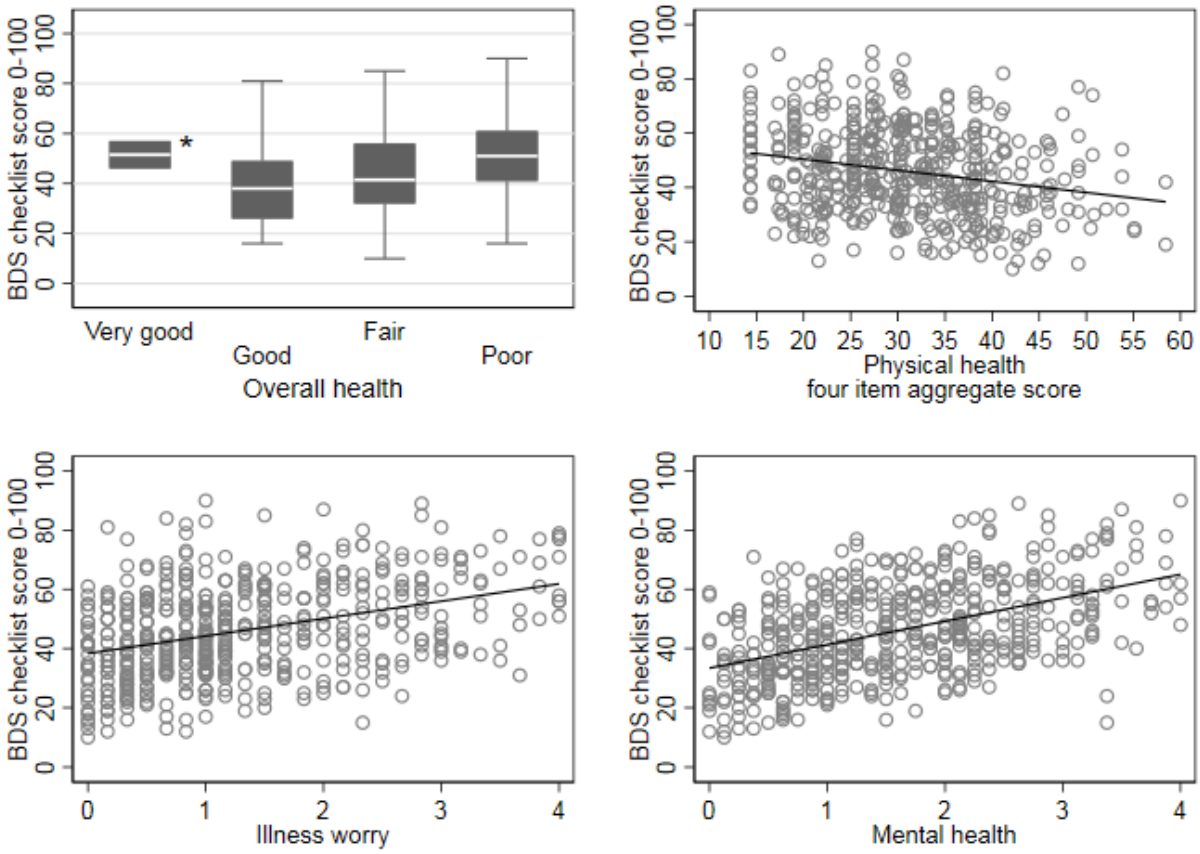
Correlation of the BDS checklist score and measures of overall health, perceived physical health, illness worry, and mental health in the general population cohort.

Primary care



Correlation of the BDS checklist score and measures of overall health, illness worry, and mental health in the primary care cohort.

Specialized setting



Correlation of the BDS checklist score and measures of overall health, perceived physical health, illness worry, and mental health in the cohort from specialized setting.
* Only based on two individuals

Appendix D. Data across sex and age groups

Table 1: Data for the general population cohort (n=9656): Cummulative percentages

	Male				Female			
	18-39	40-49	50-59	60-76	18-39	40-49	50-59	60-76
	N= 719	N=902	N=1.130	N=1.702	N=905	N=1.139	N=1.401	N=1.758
BDS score groups								
0-	21	24	25	28	10	15	10	18
5-	50	55	54	55	34	38	33	42
10-	73	76	71	75	56	59	56	62
15-	85	86	82	86	72	73	71	76
20-	91	92	88	91	83	82	82	85
25-	95	95	93	94	90	89	88	91
30-	97	97	96	97	94	93	93	94
35-	98	97	97	98	96	95	96	96
40-	99	98	98	98	97	97	97	98
45-	99	99	99	99	98	98	98	99
50-	99	99	99	99	98	99	99	99
55-	99	99	99	99	99	99	99	99
60-	99	99	99	99	99	99	100	99
65-	99	99	99	99	99	99	100	99
70-	99	99	99	99	99	99	100	99
75-	99	99	99	99	99	99	100	99
80-	99	99	99	99	99	99	100	99
85-	99	99	99	99	99	99	100	99
90-	99	99	99	99	99	99	100	99
95-	99	99	99	99	99	99	100	99
Missing	100	100	100	100	100	100	100	100

Abbreviations: BDS=Bodily distress syndrome

Table 2: Data for the primary care cohort (n=2480): Cummulative percentages

	Male						Female					
	18-39	40-49	50-59	60-76	70-79	80-	18-39	40-49	50-59	60-69	70-79	80-
	N=146	N=147	N= 172	N=245	N=153	N=67	N=404	N=280	N=271	N=287	N=203	N=105
BDS score groups												
0-	23	20	16	18	14	10	16	14	8	16	14	9
5-	49	37	32	40	35	36	37	31	23	32	31	27
10-	67	60	52	58	54	51	53	47	38	45	45	42
15-	73	70	67	72	65	58	66	59	54	58	59	55
20-	79	80	76	83	72	67	76	75	66	73	72	64
25-	89	86	82	90	79	82	80	80	78	82	79	70
30-	92	89	87	95	82	82	86	84	85	85	85	72
35-	93	91	90	96	88	85	90	88	89	89	88	74
40-	97	93	94	98	91	87	93	90	92	93	90	76
45-	97	94	96	98	94	88	95	94	95	95	91	78
50-	98	95	98	98	95	88	96	96	96	96	93	80
55-	99	98	98	100	95	88	96	97	97	97	94	81
60-	99	99	98	100	95	88	97	98	97	97	95	82
65-	99	99	99	100	95	88	97	98	97	98	96	83
70-	99	99	99	100	95	88	98	98	98	98	96	83
75-	99	99	99	100	95	88	98	98	98	98	96	83
80-	99	99	99	100	95	88	98	98	99	98	96	83
85-	99	99	99	100	95	88	98	99	99	98	96	83
90-	99	99	99	100	95	88	98	99	99	98	96	83
95-	99	99	99	100	95	88	98	99	99	98	96	83
Missing	100	100	100	100	100	100	100	100	100	100	100	100

Abbreviations: BDS=Bodily distress syndrome

Table 3: Data for the cohort from specialized setting (n=492): Cummulative percentages

	Male				Female			
	18-39	40-49	50-59	60-69	18-39	40-49	50-59	60-69
	N=42	N=40	N=11	N=0	N=199	N=162	N=36	N=2
BDS score groups								
0-	0	0	0	-	0	0	0	0
5-	0	0	0	-	0	0	0	0
10-	0	3	0	-	2	1	0	0
15-	5	5	9	-	4	2	3	0
20-	12	15	9	-	8	7	6	0
25-	21	23	9	-	14	13	14	0
30-	36	45	36	-	20	25	17	50
35-	48	50	55	-	36	35	19	50
40-	57	55	73	-	50	48	25	50
45-	74	60	82	-	60	59	36	50
50-	81	70	82	-	68	69	56	50
55-	86	85	82	-	79	75	72	50
60-	90	88	82	-	87	84	83	100
65-	95	93	82	-	92	92	83	100
70-	95	98	91	-	95	96	94	100
75-	95	100	100	-	96	99	97	100
80-	95	100	100	-	99	99	97	100
85-	100	100	100	-	100	100	97	100
90-	100	100	100	-	100	100	97	100
95-	100	100	100	-	100	100	97	100
Missing	100	100	100	-	100	100	100	100

Abbreviations: BDS=Bodily distress syndrome

Table 4: Data for the three pooled cohorts (n=12628): Cummulative percentages

	Male						Female					
	18-39	40-49	50-59	60-69	70-79	80-	18-39	40-49	50-59	60-69	70-79	80-
	N=907	N=1089	N=1313	N=1665	N=435	N=67	N=1508	N=1581	N=1708	N=1785	N=465	N=105
BDS score groups												
0-	20	23	24	26	26	10	10	13	9	18	15	9
5-	47	51	50	52	50	36	31	33	31	41	37	27
10-	69	71	68	71	70	51	48	51	52	59	53	42
15-	80	81	80	83	80	58	61	63	67	73	67	55
20-	86	88	86	90	86	67	71	73	78	83	79	64
25-	91	91	91	94	89	82	77	80	85	89	85	70
30-	94	94	94	97	91	82	82	84	90	93	89	72
35-	95	95	96	98	94	85	86	88	93	95	92	74
40-	97	96	97	99	95	87	90	91	95	97	94	76
45-	97	97	98	99	97	88	92	93	96	98	95	78
50-	98	97	99	99	97	88	94	95	98	99	96	80
55-	98	98	99	99	97	88	95	96	98	99	97	81
60-	99	99	99	99	97	88	97	97	99	99	97	82
65-	99	99	99	99	97	88	98	98	99	99	98	83
70-	99	99	99	99	97	88	98	99	99	99	98	83
75-	99	99	99	99	97	88	99	99	99	99	98	83
80-	99	99	99	99	97	88	99	99	99	99	98	83
85-	99	99	99	99	97	88	99	99	99	99	98	83
90-	99	99	99	99	97	88	99	99	99	99	98	83
95-	99	99	99	99	97	88	99	99	99	99	98	83
Missing	100	100	100	100	100	100	100	100	100	100	100	100

Abbreviations: BDS=Bodily distress syndrome

STROBE 2007 (v4) Statement—Checklist of items that should be included in reports of *cross-sectional studies*

Section/Topic	Item #	Recommendation	Reported on page #
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1+2
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	1+2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4-5
Objectives	3	State specific objectives, including any prespecified hypotheses	5+8
Methods			
Study design	4	Present key elements of study design early in the paper	5
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	5
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants	5
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	5-8
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	5-8
Bias	9	Describe any efforts to address potential sources of bias	N/A
Study size	10	Explain how the study size was arrived at	5
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	5~~6~~-8
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	5~~7~~-8
		(*b*) Describe any methods used to examine subgroups and interactions	N/A
		(*c*) Explain how missing data were addressed	N/A
		(*d*) If applicable, describe analytical methods taking account of sampling strategy	N/A
		(*e*) Describe any sensitivity analyses	N/A
Results			

Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	N/A
		(b) Give reasons for non-participation at each stage	N/A
		(c) Consider use of a flow diagram	N/A
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	8-9~~5~~
		(b) Indicate number of participants with missing data for each variable of interest	N/A
Outcome data	15*	Report numbers of outcome events or summary measures	N/A
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	9-14~~8-13~~
		(*b*) Report category boundaries when continuous variables were categorized	9-14~~8-13~~
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	N/A
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	N/A
Discussion			
Key results	18	Summarise key results with reference to study objectives	15~~4~~
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	16~~5~~
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	~~15-16~~15-17
Generalisability	21	Discuss the generalisability (external validity) of the study results	~~14-16~~17-18
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	18~~7~~

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.