



# BMJ Open The BDS checklist as measure of illness severity: a cross-sectional cohort study in the Danish general population, primary care and specialised setting

Marie Weinreich Petersen <sup>1</sup>, Marianne Rosendal,<sup>1,2</sup> Eva Ørnboel,<sup>1</sup> Per Fink,<sup>1</sup> Torben Jørgensen,<sup>3,4,5</sup> Thomas Meinertz Dantoft,<sup>3</sup> Andreas Schröder <sup>1</sup>

**To cite:** Petersen MW, Rosendal M, Ørnboel E, *et al.* The BDS checklist as measure of illness severity: a cross-sectional cohort study in the Danish general population, primary care and specialised setting. *BMJ Open* 2020;**10**:e042880. doi:10.1136/bmjopen-2020-042880

► Prepublication history and additional materials for this paper are available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2020-042880>).

Received 17 July 2020

Revised 17 November 2020

Accepted 25 November 2020



© Author(s) (or their employer(s)) 2020. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

## Correspondence to

Marie Weinreich Petersen; [mawept@rm.dk](mailto:mawept@rm.dk)

## ABSTRACT

**Objective** The bodily distress syndrome (BDS) checklist has proven to be useful in the diagnostic categorisation and as 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 a measure of physical symptom burden and FSD illness severity.

**Design** Cross-sectional.

**Setting** Danish general population, primary care and specialised clinical setting.

**Participants** A general population cohort (n=9656), a primary care cohort (n=2480) and a cohort of patients with multiorgan BDS from specialised clinical setting (n=492).

**Outcome measures** All data were self-reported. Physical symptoms were measured with the 25-item BDS checklist. Overall self-perceived health was measured with one item from the 36-item Short-Form Health Survey (SF-36). Physical functioning was measured with an aggregate score of four 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 six-item Whiteley Index.

**Results** For all cohorts, bifactor models established that despite some multidimensionality the total sum score of the BDS checklist adequately reflected physical symptom burden and illness severity. The BDS checklist had acceptable convergent validity with 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 a low number 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 a measure of symptom burden and FSD illness severity across settings. These findings establish the usefulness of the BDS checklist in clinics and in research, both as a diagnostic screening tool and as an instrument to assess illness severity.

## Strengths and limitations of this study

- The study included data from three cohorts and settings: a general population, patients from primary care and patients from a specialised 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 in the general population.<sup>1–4</sup> 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 healthcare use.<sup>5–9</sup> Hence, assessment of the burden of PPS is valuable in both clinical care and in research.

For this purpose, self-reported symptom questionnaires are frequently used, which are manageable, non-invasive tools. Several screening questionnaires exist: the Hopkins Symptom Checklist,<sup>10</sup> the Patient Health Questionnaire-15 (PHQ-15),<sup>11</sup> the Somatic Symptom Scale-8 (SSS-8),<sup>12 13</sup> the brief form of the Giessen Subjective Complaints List (GSCL-8)<sup>14</sup> and others.<sup>15–17</sup> However, the existing questionnaires measure PPS without considering the well-known aggregation of such symptoms into symptom clusters, and hence without acknowledgement of the real structure of PPS as they occur in both the community and in clinical setting.<sup>18–21</sup>

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 covers both specialty-specific syndrome diagnoses such as fibromyalgia, irritable bowel and chronic fatigue, and also their pendants in psychosomatic medicine, somatoform and somatic symptom disorders.<sup>22</sup>

In contrast to the above-mentioned specialty-specific diagnoses, the proposed research diagnosis bodily distress syndrome (BDS) covers a broader range of functional somatic symptoms, ranging from a few symptoms with some effect on functioning to severe and disabling FSD.<sup>18 19 21</sup> Hence, BDS provides the opportunity to assess and distinguish between conditions persisting as mono-syndromatic or multi-syndromatic and still within the same framework of diagnostic approach.<sup>21 23</sup> The diagnostic construct was developed in a sample of patients from primary and secondary care, and the 30-item BDS checklist emerged.<sup>18</sup> The BDS checklist was confirmed in a new sample of primary care patients where the shortened 25-item BDS checklist was developed.<sup>19</sup> Subsequently, the construct of the BDS checklist has been confirmed in general population samples as well.<sup>21 24</sup> The BDS checklist presents symptoms grouped in four clusters—cardiopulmonary, gastrointestinal, musculoskeletal and general symptoms—and its usefulness and the properties used for diagnostic categorisation into no BDS, single/oligo-organ BDS type and multiorgan BDS type have been established.<sup>19 21 24</sup> A major strength of the BDS checklist is its usefulness both as a screening and as a diagnostic tool within clinical practice and within epidemiological research<sup>18 19 21 23</sup>; however, the total BDS sum score has not yet been validated as a measure to assess 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 (ie, in those individuals who may fall under the diagnostic threshold or what we believe to be clinically relevant) and illness severity (in those individuals fulfilling the 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, patients in primary care and patients in a specialised clinical setting.

## METHODS

### Population

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

Cohort 1 is a general population cohort (DanFunD, n=9656, response rate=33.7%) established with the purpose to investigate and unravel the epidemiology of FSD.<sup>25</sup> The cohort was obtained from the Danish Central Personal Register and drawn as a random sample of 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 is 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.<sup>26</sup> 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 consists data from a specialised 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%).<sup>27–31</sup>

These cohorts had been part of a group of studies with the shared aim of investigating new treatments for patients aged 20 years or older with multiorgan BDS.

### Measures

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

Physical symptoms were assessed with the Danish version of the 25-item BDS checklist (online supplemental appendix A).<sup>19 21</sup> 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 5-point rating scale from 0 ('not at all bothersome') to 4 ('a lot bothersome'). We calculated the sum score by adding the single item scores from the 25 items (ranging from 0 to 100). The time frame covered was 12 months for the general population cohort and 4 weeks for the other two cohorts.

Overall health was assessed with a single item from the 36-item Short-Form Health Survey (SF-36),<sup>32</sup> estimating self-perceived health on a 5-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'.<sup>30 32–34</sup> 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 due to physical health and pain interference. For each item a z-score was calculated using mean and SD from the general Danish population. The 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 the

correlation was high (Spearman's  $r=0.89$ , 95% CI 0.87 to 0.91). Unfortunately, it was not possible to investigate the convergent validity to the aggregate score in the data of the primary care cohort because we had limited access to the data. These analyses were therefore only performed in the general population cohort and the cohort from specialised clinical setting. The time frame covered was 4 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).<sup>35 36</sup> 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 indicate higher emotional distress. The time frame covered was 1 week for the general population cohort and 4 weeks for the two other cohorts.

Illness worry was measured with the revised version of the six-item Whiteley Index (Whiteley-6-R),<sup>37</sup> addressing respondents' fear of being ill and whether they attribute current bodily sensations to somatic illness. (In the primary care sample, one of the items in the Whiteley-6-R '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?') Answers were calculated as mean scores from a scale ranging from 0 ('not at all bothersome') to 4 ('a lot bothersome'). Higher scores indicate higher health anxiety. The time frame covered was 12 months for the general population cohort and 4 weeks for the two other cohorts.

### Validation procedure and statistical analyses

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

All statistical analyses were performed using STATA V.16.0,<sup>39</sup> except for the structural equation modelling which was performed using Mplus V.8.1.<sup>40</sup> Construct validity was tested by means of structural validity and convergent validity. Structural validity was tested with confirmatory factor analysis (CFA) with WLSMV (weighted least squares means and variance adjusted) estimation due to the categorical responses for all items.<sup>40</sup> We wanted to test if it was permissible to model the BDS checklist as unidimensional despite previous evidence of some multidimensionality.<sup>18 19 21 24</sup> 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) a one-level one-factor model; (2) a one-level four-factor model, using factors resembling the four BDS symptom clusters previously reported<sup>19 21</sup>; (3) a two-level four-factor model, representing a second-order common factor (BDS) underlying the four BDS symptom clusters; and (4) a bifactor 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 online supplemental appendix B.

In all CFAs, model fit was assessed as follows: a root mean square error of approximation  $<0.05$  indicates very good fit, 0.05–0.08 indicates a good fit and  $\geq 0.08$  indicates a poor fit. Comparative fit index and Tucker-Lewis fit index at 0.90–0.95 indicate an acceptable fit and levels  $>0.95$  indicate a good fit. A standardised root mean square residual  $<0.08$  indicates good fit.<sup>41</sup>

Convergent validity was tested with Spearman's correlations, and associations between the BDS checklist and overall health (one item from SF-36),<sup>32</sup> physical function (an aggregate score of four items from the SF-36),<sup>42</sup> emotional distress (SCL-8)<sup>35</sup> and illness worry (Whiteley-6-R)<sup>37</sup> were obtained. Based on previous literature,<sup>12 14 15 17 43</sup> we hypothesised 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 specialised setting. Expected differences on the BDS checklist with one-unit difference from SCL-8, the four-item aggregate score for physical functioning and Whiteley-6-R were estimated with linear regression.

The BDS checklist item and scale characteristics, that is, item means (SD), sum score means, score distribution and item total correlation, corrected for overlap, and aspects of acceptability, that is, percentage of missing items, were examined and computed as descriptive statistics for each of the three samples. Internal consistency was measured with Cronbach's  $\alpha$  coefficients, where values between 0.7 and 0.95 are acceptable.<sup>38</sup>

For all three cohorts, written informed consent was obtained from each participant before entering the studies.<sup>25–31</sup> The STreSS-3 trial in cohort 3 was registered with ClinicalTrials.gov (NCT01518634) and EudraCT (2011-004294-87), while the STreSS-4 trial was registered with ClinicalTrials.gov (NCT01518647).

### 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

The median age of the general population sample was 54 years (IQR: 44–64) and 53.9% were female. In the primary care sample, the mean age was 54.3 years (SD: 17.5) and 62.5% were female. In the sample from specialised setting, the mean age was 39.4 years (SD: 8.8) and 81.1% were female.

### 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

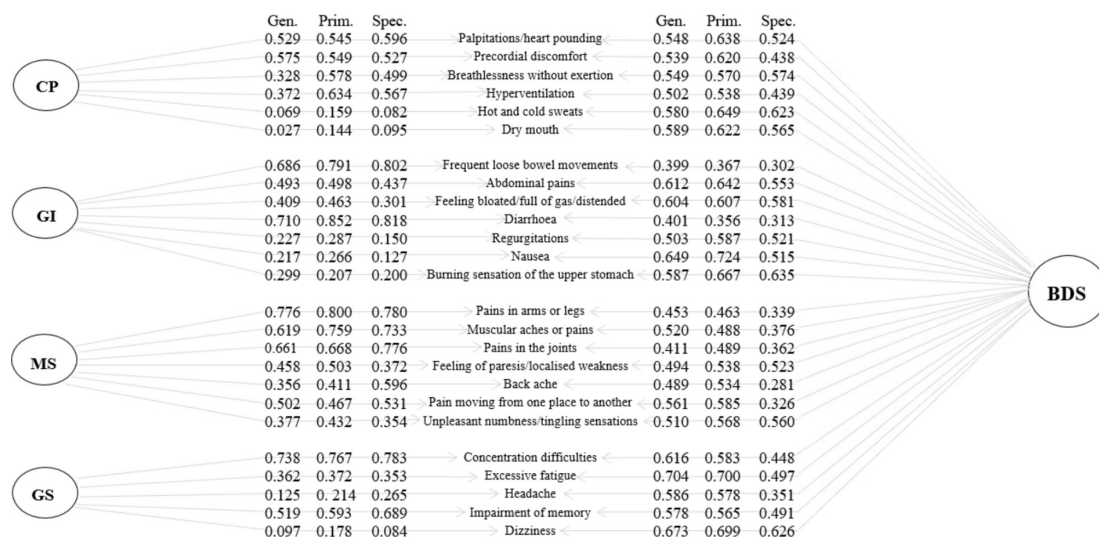
	RMSEA	95% CI	CFI	TLI	SRMR	X <sup>2</sup>	df	P value
<b>One-level one-factor CFA</b>								
General population	0.111	0.110 to 0.112	0.723	0.697	0.09	32 743.1	275	<0.0001
Primary care	0.419	0.147 to 0.151	0.697	0.67	0.119	15 126.5	275	<0.0001
Specialised setting	0.149	0.144 to 0.153	0.621	0.586	0.115	3261.5	275	<0.0001
<b>One-level four-factor CFA</b>								
General population	<b>0.062</b>	0.061 to 0.063	<b>0.914</b>	<b>0.905</b>	<b>0.052</b>	10 290.96	269	<0.0001
Primary care	0.082	0.08 to 0.084	<b>0.91</b>	<b>0.9</b>	<b>0.067</b>	4666.85	269	<0.0001
Specialised setting	0.091	0.086 to 0.096	0.862	0.846	<b>0.076</b>	1355.96	269	<0.0001
<b>Two-level four-factor CFA</b>								
General population	<b>0.061</b>	0.06 to 0.062	<b>0.917</b>	<b>0.908</b>	<b>0.052</b>	9951.02	271	<0.0001
Primary care	0.08	0.078 to 0.082	<b>0.914</b>	<b>0.905</b>	<b>0.068</b>	4482.39	271	<0.0001
Specialised setting	0.089	0.084 to 0.093	0.867	0.853	<b>0.076</b>	1315.04	271	<0.0001
<b>Bifactor CFA</b>								
General population	<b>0.048</b>	0.046 to 0.049	<b>0.954</b>	<b>0.944</b>	<b>0.04</b>	5680.8	250	<0.0001
Primary care	<b>0.053</b>	0.051 to 0.055	<b>0.965</b>	<b>0.958</b>	<b>0.042</b>	1977.4	250	<0.0001
Specialised setting	<b>0.059</b>	0.054 to 0.065	<b>0.945</b>	<b>0.934</b>	<b>0.051</b>	681.1	250	<0.0001

Bold indicates good or acceptable fit between the specified model and the observed model in the data.

χ<sup>2</sup>, likelihood ratio test; CFA, confirmatory factor analysis; CFI, comparative fit index; RMSEA, root mean square error of approximation; SRMR, standardised root mean square residual; TLI, Tucker-Lewis fit index.

Fit indices for the one-level four-factor model which has been confirmed in previous studies<sup>19 21</sup> 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 bifactor model. Good fit indices were seen for the bifactor 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 general population cohort this was the case for 72% of symptoms, in the primary care cohort it accounted for 64% of symptoms, and in the specialised setting it accounted for 52% of symptoms. Loading from the general BDS factor was smaller than loading for six symptoms from the four symptom clusters (*frequent loose bowel movements, diarrhoea, pain in the arms and legs,*



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

*muscular aches or pain, pain in the joints and concentration difficulties*) in all three cohorts.

### Convergent validity

In the general population sample, our hypothesis was met for all measures. The BDS checklist had moderate convergent validity compared with the SF-36 item for overall health ( $r=0.48$ , 95% CI 0.46 to 0.49,  $p<0.0001$ ), the four-item aggregate score for physical health ( $r=-0.58$ , 95% CI  $-0.59$  to  $-0.56$ ,  $p<0.0001$ ), the SCL-8 for emotional distress ( $r=0.52$ , 95% CI 0.51 to 0.54,  $p<0.0001$ ) and the Whiteley-6-R for illness worry ( $r=0.53$ , 95% CI 0.52 to 0.55,  $p<0.0001$ ). The expected difference on the BDS checklist with one-unit difference from the four-item aggregate score was  $-0.80$  (95% CI  $-0.82$  to  $-0.78$ ), 12.26 (95% CI 11.89 to 12.63) with SCL-8, and 8.93 (95% CI 8.64 to 9.21) with Whiteley-6-R (online supplemental 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 hypothesised. We found moderate convergent validity compared with the SF-36 item for overall health ( $r=0.58$ , 95% CI 0.56 to 0.61,  $p<0.0001$ ), the SCL-8 for emotional distress ( $r=0.62$ , 95% CI 0.59 to 0.64,  $p<0.0001$ ) and the Whiteley-6-R for illness worry ( $r=0.55$ , 95% CI 0.52 to 0.58,  $p<0.0001$ ). The expected difference on the BDS checklist with one-unit difference was 10.60 (95% CI 10.08 to 11.12) with SCL-8 and 10.01 (95% CI 9.44 to 10.59) with Whiteley-6-R (online supplemental appendix C).

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

### Response distributions and acceptability

The BDS checklist item and scale characteristics are displayed in table 2. Item means varied from 0.15 to 1.09 in the general population sample, from 0.31 to 1.53 in the primary care sample, and from 0.81 to 3.34 in the sample from specialised 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.

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

The 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 specialised setting (46.15, SD: 15.91) (table 2). Acceptability was good, and the number of missing responses was generally low in the general population (total 0.6%) and specialised setting (total 0.2%), while it was slightly higher in primary care (total 2.7%).

The 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 patients from specialised 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 online supplemental appendix D.

## DISCUSSION

### Principal findings

This is the first study to establish that, despite some multidimensionality, the 25-item 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 specialised settings. Used as a total sum score ranging from 0 to 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.<sup>44</sup>

The three cohorts differed in the 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 specialised clinical setting. The latter group contains patients with long-standing and severe FSD. In this group, the symptom load is high and specific symptom clusters may therefore stand out compared with 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 (online supplemental appendix B).<sup>19 21 24</sup> However, the objectives of these studies were to confirm the BDS checklist as a case-finding instrument in other samples, with inspiration from the original studies in which the concept of the BDS checklist was developed and initially tested.<sup>18 19</sup> 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 following are the indicators of a bifactor model: (1) if the intercorrelation between the subscales in the CFA

**Table 2** Item and scale characteristics

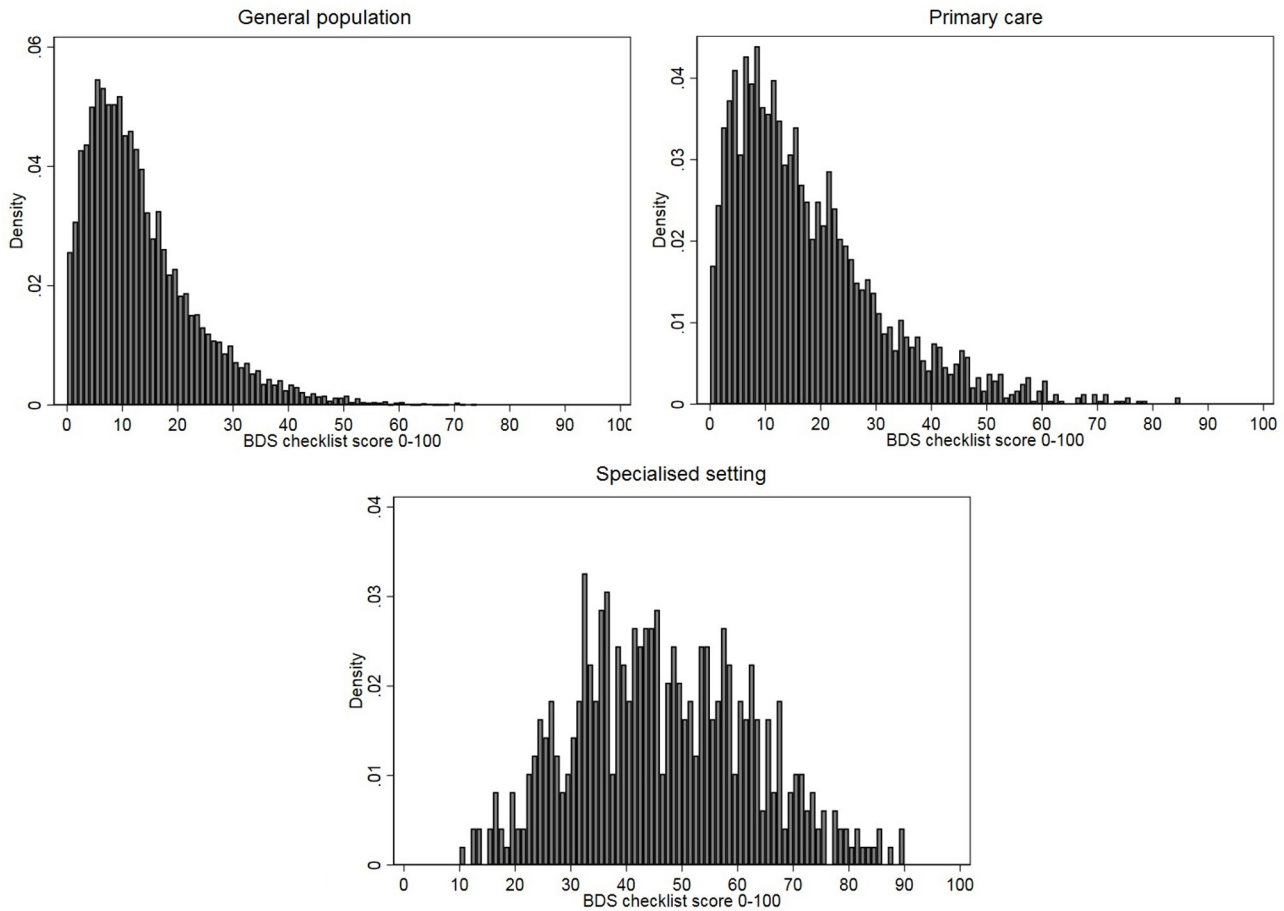
Item	General population (n=9656)			Primary care (n=2480)			Specialised setting (n=492)		
	Missing %	Mean (SD)	Item total correlation*	Missing %	Mean (SD)	Item total correlation*	Missing %	Mean (SD)	Item total correlation*
Palpitations/heart pounding	0.9	0.45 (0.74)	0.427	3.1	0.61 (0.93)	0.544	0.2	1.46 (1.25)	0.48
Precordial discomfort	1.1	0.29 (0.61)	0.41	2.9	0.46 (0.81)	0.517	0.2	1.09 (1.16)	0.391
Breathlessness without exertion	1	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)	0.38
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	0.65 (0.86)	0.403	3.4	0.61 (0.95)	0.391	0.2	1.41 (1.32)	0.357
Abdominal pain	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)	0.43
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 in 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
Pain in the arms or legs	1	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 pain	1.3	0.98 (1.01)	0.572	3.2	1.30 (1.23)	0.584	0.2	2.96 (1.10)	0.485
Pain in the joints	1.6	0.96 (1.07)	0.49	3.9	1.20 (1.27)	0.56	0.2	2.57 (1.32)	0.491
Feeling of paresis or localised weakness	1.4	0.16 (0.55)	0.365	4	0.33 (0.83)	0.46	0.2	1.22 (1.40)	0.481
Backache	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)	0.41	4	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)	0.54	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
<b>Scale</b>									
Total scale missing (%)	0.6			2.7			0.2		
Mean (SD)	13.03 (10.36)			17.33 (13.79)			46.15 (15.91)		
<b>Percentiles</b>									
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		

IQR: 25% percentile and 75% percentile.

\*Item total correlation, corrected for overlap.

exceeds 0.3; (2) if loading on the first-order factors on the second-order factors exceeds 0.5<sup>45</sup>; and (3) if the ratio between the first and second eigenvalues exceeds 3.<sup>46</sup>

All parameters were fulfilled in the general population cohort and in the primary care cohort. For the cohort from specialised setting, the ratio between the first and



**Figure 2** Distribution of the bodily distress syndrome (BDS) total sum score across all three cohorts.

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; however, 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.<sup>6 7</sup> The difference between the results from patients in the specialised settings and from the two other populations may be due to the nature of self-reported measures, where patients in specialised setting still have the opportunity to rate their perceived health as excellent even though they have been referred to specialised medical care due to invalidating physical symptoms. These aspects may result in 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 concerned only one setting at a time.<sup>11 12 14 17</sup> 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.

<b>Table 3</b> Grouping of BDS scores across samples						
Categories of BDS score	General population		Primary care		Specialised setting	
	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

More detailed information about the normative data from each of the three samples and for all three samples pooled together is shown in online supplemental appendix D tables 1–4. BDS, bodily distress syndrome.



The weaknesses of the study include the following: 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 with other measures of physical symptoms, or for the primary care cohort and the cohort from specialised clinical setting with physicians' report. Furthermore, in the linear regression analysis, the assumption of normality of the residuals was not fully met for the primary care cohort and the cohort from specialised clinical care, which is why these results should be interpreted with caution. Finally, as this study had a cross-sectional design, it was not possible to evaluate the responsiveness of the BDS checklist.

### Difference in results compared with 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 to measure the severity of physical symptoms is the PHQ-15.<sup>11 17</sup> 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. PHQ-15 is scored on a 3-point rating scale from 'not bothered at all' (0) to 'bothered a lot' (2), whereas the BDS checklist uses a 5-point rating scale. In one study including a sample from the general Swedish population, factor analyses of the structural properties of PHQ-15 showed a four-factor model, but on the basis of a scree test plot they finally concluded that only one factor should be extracted.<sup>47</sup> Other studies found a bifactor model to have the best fit to PHQ-15.<sup>48 49</sup> Hence, 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 PHQ-15, the SSS-8, the 3-point rating scale is replaced with a 5-point rating option as in the BDS checklist.<sup>12 13 50</sup> However, neither the PHQ-15 nor the SSS-8 is validated for use as a diagnostic categorisation tool for respondents. Other symptom questionnaires resembling the same four-factor structure and the same five-answer categories as the BDS checklist are the 24-item Giessen Subjective Complaints List and its newer shortened version with eight items (G8B-8); however, they have only been established and used in German-speaking countries.<sup>14</sup>

The BDS checklist is, at present, the only symptom checklist providing both diagnostic categorisation 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.<sup>33</sup> 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 specialised 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 BDS total sum score presented here reliably captures FSD illness severity than the distinction in single versus multiorgan BDS (eg, three vs four clusters fulfilled). Nevertheless, a tool which is also able to measure the severity of specific symptom clusters is helpful in specialised 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 physicians' established diagnoses, for example, 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.

### Author affiliations

<sup>1</sup>The Research Clinic for Functional Disorders and Psychosomatics, Aarhus University Hospital, Aarhus, Denmark

<sup>2</sup>Research Unit for General Practice, Aarhus University, Aarhus C, Denmark

<sup>3</sup>Center for Clinical Research and Prevention, Bispebjerg and Frederiksberg Hospital, The Capital Region of Denmark, Denmark

<sup>4</sup>Department of Public Health, Faculty of Health and Medical Science, University of Copenhagen, Copenhagen C, Denmark

<sup>5</sup>Faculty of Medicine, Aalborg University, Aalborg, Denmark

**Acknowledgements** The authors would like to thank senior researcher Grete Moth and data manager Kaare Rud Flarup from the research unit of primary care for providing data from the primary care cohort. The authors are also thankful to the members of the DanFunD steering committee: senior consultants Lene Falgaard Eplow MD PhD, Allan Linneberg MSc PhD, Susanne Brix Pedersen MSc PhD, and Michael Eriksen Benros MD PhD.

**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** The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

**Competing interests** None declared.

**Patient consent for publication** Not required.

**Ethics approval** The current study was carried out in accordance with the relevant guidelines and regulations. Cohort 1 was approved by 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 was 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. The STreSS-5 cohort study was approved by the Danish Data Protection Agency.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** Data are available on reasonable request from the corresponding author.

**Supplemental material** This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

#### ORCID iDs

Marie Weinreich Petersen <http://orcid.org/0000-0001-8141-2792>

Andreas Schröder <http://orcid.org/0000-0001-5211-4702>

#### REFERENCES

- Henningsen P, Gündel H, Kop WJ, *et al*. Persistent physical symptoms as perceptual dysregulation: a Neuropsychobehavioral model and its clinical implications. *Psychosom Med* 2018;80:422–31.
- Kroenke K. A practical and evidence-based approach to common symptoms: a narrative review. *Ann Intern Med* 2014;161:579–86.
- Henningsen P, Zipfel S, Sattel H, *et al*. Management of functional somatic syndromes and bodily distress. *Psychother Psychosom* 2018;87:12–31.
- Eliassen 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:e0150664.
- Tomenson B, Essau C, Jacobi F, *et al*. Total somatic symptom score as a predictor of health outcome in somatic symptom disorders. *Br J Psychiatry* 2013;203:373–80.
- Bruusgaard D, Tschudi-Madsen H, Ihlebæk C, *et al*. Symptom load and functional status: results from the Ullensaker population study. *BMC Public Health* 2012;12:1085.
- Creed FH, Davies I, Jackson J, *et al*. The epidemiology of multiple somatic symptoms. *J Psychosom Res* 2012;72:311–7.
- 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. *J Psychosom Res* 2010;68:427–37.
- Eliassen 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.
- Derogatis LR, Lipman RS, Rickels K, *et al*. The Hopkins symptom checklist (HSCL): a self-report symptom inventory. *Behav Sci* 1974;19:1–15.
- Kroenke K, Spitzer RL, Williams JBW. The PHQ-15: validity of a new measure for evaluating the severity of somatic symptoms. *Psychosom Med* 2002;64:258–66.
- Gierk B, Kohlmann S, Kroenke K, *et al*. The somatic symptom scale-8 (SSS-8): a brief measure of somatic symptom burden. *JAMA Intern Med* 2014;174:399–407.
- 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). *J Psychosom Res* 2015;78:352–5.
- 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 (GSB-8). *J Psychosom Res* 2017;95:33–43.
- 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.
- 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;5:129–38.
- Zijlema WL, Stolk RP, Löwe B, *et al*. How to assess common somatic symptoms in large-scale studies: a systematic review of questionnaires. *J Psychosom Res* 2013;74:459–68.
- 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. *Psychosom Med* 2007;69:30–9.
- Budtz-Lilly A, Fink P, Ørnbøl E, *et al*. A new questionnaire to identify bodily distress in primary care: The 'BDS checklist'. *J Psychosom Res* 2015;78:536–45.
- Rosmalen JGM, Tak LM, de Jonge P. Empirical foundations for the diagnosis of somatization: implications for DSM-5. *Psychol Med* 2011;41:1133–42.
- Petersen MW, Schröder A, Jørgensen T, *et al*. The unifying diagnostic construct of bodily distress syndrome (BDS) was confirmed in the general population. *J Psychosom Res* 2020;128:109868.
- 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;18:34.
- 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. *Sci Rep* 2020;10:3273–73.
- 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. *J Psychosom Res* 2020;132:109991.
- Dantoft TM, Ebstrup JF, Linneberg A, *et al*. Cohort description: the Danish study of functional disorders. *Clin Epidemiol* 2017;9:127–39.
- Moth G, Vedsted P, Olesen F, *et al*. *Kontakt- OG sygdomsmønsterundersøgelse KOS 2008: Forskningsenheden for Almen Praksis*, 2010.
- Frolund Pedersen H, Holsting A, Frosthalm L, *et al*. 'Understand your illness and your needs': Assessment-informed patient education for people with multiple functional somatic syndromes. *Patient Educ Couns* 2019;102:1662–71.
- 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;4:378–88.
- Pedersen HF, Agger JL, Frosthalm 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 - CORRIGENDUM. *Psychol Med* 2018;48:2804–10.
- Schröder A, Rehfeld E, Ørnbøl E, *et al*. Cognitive-behavioural group treatment for a range of functional somatic syndromes: randomised trial. *Br J Psychiatry* 2012;200:499–507.
- Fjorback LO, Arendt M, Ørnbøl E, *et al*. Mindfulness therapy for somatization disorder and functional somatic syndromes: randomized trial with one-year follow-up. *J Psychosom Res* 2013;74:31–40.
- Ware JJ, Kosinski M, Gandek B. SF-36 health survey: manual and interpretation guide. *Quality Metric* 1993.
- Rief W, Burton C, Frosthalm 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;79:1008–15.
- 34 Schröder A, Oernboel E, Licht RW, *et al.* Outcome measurement in functional somatic syndromes: SF-36 summary scores and some scales were not valid. *J Clin Epidemiol* 2012;65:30–41.
- 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. *Europ Psychiat Rev* 2010;3:8–12.
- 37 Carstensen TBW, Ørnboel 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.
- 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*. 8th edn. Los Angeles, CA: Muthén & Muthén, 1998.
- 41 Lt H, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equat Model: A Multidiscipl J* 1999;6:1–55.
- 42 Ware J, Kosinski M, Keller SD. A 12-Item short-form health survey: construction of scales and preliminary tests of reliability and validity. *Med Care* 1996;34:220–33.
- 43 Terluin B, van Marwijk HWJ, Adèr 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.
- 44 Beutel ME, Klein EM, Henning M, *et al.* Somatic symptoms in the German general population from 1975 to 2013. *Sci Rep* 2020;10:1595.
- 45 HJT MD. Bifactor analysis in MPlus [Video file]: UK College of Education, 2016. Available: <https://education.uky.edu/edp/apslab/events/2020>
- 46 Cho S-J, Wilmer J, Herzmann G, *et al.* Item response theory analyses of the Cambridge face memory test (CFMT). *Psychol Assess* 2015;27:552–66.
- 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. *Scand J Psychol* 2013;54:112–7.
- 48 Cano-García FJ, Muñoz-Navarro R, Sesé Abad A, *et al.* Latent structure and factor invariance of somatic symptoms in the patient health questionnaire (PHQ-15). *J Affect Disord* 2020;261:21–9.
- 49 Withthöft M, Hiller W, Loch N, *et al.* The latent structure of medically unexplained symptoms and its relation to functional somatic syndromes. *Int J Behav Med* 2013;20:172–83.
- 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.