Psychological Medicine (2011), **41**, 1133–1142. © Cambridge University Press 2010
doi:10.1017/S0033291710001625

ORIGINAL ARTICLE

Empirical foundations for the diagnosis of somatization: implications for DSM-5

J. G. M. Rosmalen*, L. M. Tak and P. de Jonge

Interdisciplinary Center for Psychiatric Epidemiology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands

Background. The aim of this study was to develop empirically validated criteria for the diagnoses of clinically relevant somatization.

Method. This study was performed in a population-representative cohort consisting of 461 males (47.8%) and 503 females (52.2%), with an average age of 55.8 years (S.D. = 11.1). Somatization, anxiety and depression were derived from the Composite International Diagnostic Interview. Mplus was used to perform confirmative factor analyses on the current DSM-IV symptom groups; on alternative symptom clusters previously suggested; and to perform latent class analysis in order to define an empirically derived cut-off for somatization.

Results. The existence of symptom groups as described in DSM-IV was not supported by our data, whereas a differentiation between cardiopulmonary, musculoskeletal, gastrointestinal and general somatic symptoms did fit our data. Latent class analysis revealed two classes characterized by few ($n = 859$) and many ($n = 105$) symptoms. The class of subjects could be approached by a simple cut-off of four functional symptoms (sensitivity 79%, specificity 98%, positive predictive value 82%, negative predictive value 97%) regardless of the number of organ systems involved.

Conclusions. This study in a large population-representative cohort suggests that a simple symptom count can be used as a dimensional diagnosis of somatization. In those instances in which a categorical diagnosis is preferred, a simple cut-off of four out of 43 functional symptoms best fitted our data. We did not find any added value for incorporating the number of symptom clusters into the diagnostic criteria.

Received 3 May 2010; Revised 13 July 2010; Accepted 17 July 2010; First published online 16 September 2010

Key words: Diagnosis, DSM, population cohort, somatization.

Introduction

Functional symptoms are symptoms that cannot be conclusively explained by organic pathology. Clusters of functional symptoms are classified under somatoform disorders in DSM-IV (APA, 1994). A classical example of a disease characterized by various functional somatic symptoms is somatization disorder. The diagnostic criteria for somatization disorder described in the different DSM editions illustrate the difficulties in constructing appropriate diagnostic criteria for clinically relevant somatization. The original criteria involved a lifetime history of 25 unexplained somatic symptoms in addition to attitudinal features. DSM-III criteria required a symptom count of 12 for males and 14 for females, which changed to 13 for both genders in DSM-III-R. The current DSM-IV criteria require eight symptoms originating from four designated symptom groups. For the DSM-V, the diagnosis of complex somatic symptom disorder is proposed to replace the current diagnoses of somatization disorder, undifferentiated somatoform disorder, hypochondriasis and pain disorder (Dimsdale & Creed, 2009). The proposed diagnostic criteria for complex somatic symptom disorder require the presence of somatic symptoms, together with misattributions, excessive concern or preoccupation with symptoms and illness and increased healthcare use.

The diagnostic criteria for somatization highlight a number of issues that require empirical evaluation in the general population. First, whereas the DSM-III criteria required a simple symptom count, a DSM-IV diagnosis of somatization disorder requires a combination of symptoms from several symptom groups: at least four pain symptoms, two gastrointestinal, one sexual and one pseudoneurological. This criterion implies a clustering of symptoms in the different designated symptom groups. Some studies, especially those performed in primary care or community

* Address for correspondence: Dr J. G. M. Rosmalen, Interdisciplinary Center for Psychiatric Epidemiology, University Medical Center Groningen, University of Groningen, PO Box 30.001, 9700 RB Groningen, The Netherlands.
(Email: j.g.m.rosmalen@med.umcg.nl)

settings, have indeed suggested that certain types of symptoms tend to cluster (Swartz *et al*. 1986; Simon *et al*. 1996; Liu *et al*. 1997; Robbins *et al*. 1997; Gara *et al*. 1998; Kroenke *et al*. 1998; Fink *et al*. 2007). However, the latent factors observed in these studies do not completely resemble the DSM-IV symptom groups. Commonly identified factors are gastrointestinal, musculoskeletal, neurological/conversion and cardiopulmonary/autonomic (Swartz *et al*. 1986; Simon *et al*. 1996; Robbins *et al*. 1997; Gara *et al*. 1998; Kroenke *et al*. 1998; Fink *et al*. 2007), although not all studies replicated the existence of all four factors. The individual factors were highly correlated in several of these studies, suggesting that a more general higher order somatization factor explained the majority of the variance in functional symptoms (Liu *et al*. 1997; Deary, 1999; Fink *et al*. 2007). Clinically meaningful symptom clusters could not be identified mainly in studies performed in specialist care settings (Hiller *et al*. 2001; Nimnuan *et al*. 2001*b*; Sullivan *et al*. 2002). It therefore remains unknown whether symptom clusters exist, as is implicitly assumed in DSM-IV, or whether an underlying somatization factor is responsible for symptoms in general, as is assumed in DSM-III and the DSM-5 proposal.

Second, it is unclear whether an appropriate symptom count threshold for somatization can be set. In contrast to the categorical approach of previous DSM editions, the proposed diagnosis of complex somatic symptom disorders in DSM-5 is based on a dimensional approach (Dimsdale & Creed, 2009). Support for such an approach comes from the linear relationship between the number of somatic symptoms and several indicators of construct validity, including functional impairment, childhood and family risk factors, psychiatric co-morbidity and healthcare use (Kroenke *et al*. 2007). In addition to a dimensional diagnosis, it may be useful to define a categorical approach for clinically relevant somatization, both from a clinical and a research point of view. In that case, it is important to empirically justify a cut-off score.

Third, it is unclear whether the diagnostic criteria should include a threshold for the number of organ systems involved. DSM-IV criteria for somatization disorder included a requirement for symptoms originating from four organ systems, whereas this requirement is not included in the proposed diagnosis of complex somatic symptom disorders in DSM-5 (Dimsdale & Creed, 2009). It has never been formally tested whether or not this threshold is required or which minimum number of organ systems would be appropriate.

The aim of this study was to empirically evaluate criteria for the diagnosis of clinically relevant somatization. We formulated the following questions. First, are there indications for symptom clusters within the functional symptoms, such as defined in the current DSM-IV classification or in some of the alternative proposals? Second, is it possible to define a cut-off for clinically relevant somatization based on population-based empirical data? Third, is there an empirically supported minimum number of organ systems required?

We performed our study in a population cohort using data on somatization derived from the Composite International Diagnostic Interview (CIDI). In accordance with the DSM-5 proposal, we performed our analyses on recent symptoms instead of lifetime symptoms, because detailed inquiry about lifetime symptoms is typically not feasible in busy clinical settings and lifetime recall of functional symptoms is highly inconsistent (Kroenke *et al*. 1997; Simon & Gureje, 1999).

Methods

Population

Our study has been performed in a cohort derived from Prevention of Renal and Vascular End Stage Disease (PREVEND), a population cohort study investigating micro-albuminuria as a risk factor for renal and cardiovascular disease. The recruitment of participants has been extensively described elsewhere (Pinto-Sietsma *et al*. 2000). All inhabitants of the city of Groningen between the ages of 28 and 75 years (85 421 subjects) were asked to send in a morning urine sample and to fill out a short questionnaire on demographics and cardiovascular history. A total of 40 856 subjects (47.8%) responded. After exclusion of subjects with insulin-dependent diabetes mellitus and pregnant women, all subjects with an elevated urinary albumin concentration of $\geqslant 10$ mg/l ($n=7768$), together with a randomly selected control group with a urinary albumin concentration of <10 mg/l ($n=3395$), were invited for further investigations (total $n=11\,163$). Finally, 8592 subjects completed the total screening programme, making up the PREVEND study cohort. Because the PREVEND study population was enriched for albuminuria, this oversampling for albuminuria was counterbalanced in the current substudy. Albuminuria-negative participants and a random sample of albuminuria-positive participants were combined so that a population-representative ratio of albuminuria-positive participants was achieved.

Research assistants handed over invitations in the 2001–2002 wave to 2554 subjects to participate in a substudy, for which additional psychiatric and psychosocial data were collected. Of these 2554

https://doi.org/10.1017/S0033291710001625 Published online by Cambridge University Press

subjects, 1094 (43%) completed the additional measurements. Follow-up measurements in the 2003–2004 wave were completed by a total of 976 participants (89% of the cohort with additional psychiatric and psychosocial data), forming the cohort for the current study. The study was approved by the local medical ethics committee and all subjects gave written informed consent to participate.

Somatization

Somatization was measured by the somatization section of the CIDI. A fully computerized version of the CIDI 2.1 12-month version was applied, suitable for self-administration. Trained interviewers were present for questions and for participants who needed computer help. The probing scheme of the self-administered version is completely identical to the interviewer-administered version; the difference between both versions is that the questions are not read out loud by the interviewer but instead are read on the screen by the participant him/herself. In short, the CIDI somatization section surveys the occurrence of 43 symptoms in the past year. Symptoms are considered present when they meet severity criteria, i.e. provoke a healthcare visit. If these criteria are met, the interview assesses in a hierarchical fashion whether a medical doctor diagnosed a symptom as due to physical illness or injury, or whether a symptom was caused by the use of medication, drugs or alcohol. If these inquiries are negative for these medical explanations, the symptom is scored as a functional symptom. As an additional validation step, we checked all medical diagnoses that participants indicated in the case of medically explained symptoms. In those cases in which the diagnosis involved a functional syndrome (such as irritable bowel syndrome, chronic fatigue syndrome or fibromyalgia), we recoded the symptom as a functional symptom. The CIDI has adequate test–retest reliability and validity (Andrews & Peters, 1998). Complete CIDI data were available for 964 participants (99% of the current study cohort).

Confirmatory factor analyses to test the presence of symptom clusters (question 1)

In order to test previously postulated symptom clusters, we performed confirmatory factor analyses for binary data using Mplus 3.11 (Muthen & Muthen, 2004). We first tested the symptom clusters currently defined in the CIDI DSM-IV scoring algorithm: pain symptoms; gastrointestinal symptoms other than pain; sexual or reproductive symptoms other than pain; pseudoneurological symptoms (see Table 1 for included symptoms). Because of the different definition of the sexual symptom cluster in the CIDI scorings syntax rules, we performed these analyses separately for males and females. Urinary retention and difficulty swallowing or lump in throat were not reported by males and thus not included in the analyses in males. Symptoms that are combined in the CIDI scorings algorithm were included as separate symptoms in these analyses.

In addition, we performed a confirmatory factor analysis using cardiopulmonary, musculoskeletal and gastrointestinal factors resembling those previously reported by Kroenke *et al*. (1998) and Fink *et al*. (2007). We also tested a four-factor model, including the factor general symptoms (see Table 1 for included symptoms). Finally, we performed a confirmatory factor analysis including a second order factor representing a common latent factor underlying the cardiopulmonary, musculoskeletal and gastrointestinal symptom groups. The models were deemed to fit the data well if all of the following goodness-of-fit indices were satisfied: overall χ^2 goodness-of-fit test non-significant; Comparative Fit Index (CFI) >0.95; root mean square error of approximation (RMSEA) $\leqslant 0.05$ (Hu & Bentler, 1999).

Latent class analyses to identify an empirically based cut-off (question 2)

In order to establish an empirically derived threshold for somatization, we applied latent class analysis (LCA) to the CIDI symptoms, using Mplus 3.11 (Muthen & Muthen, 2004). LCA is a statistical model-fitting method identifying different subgroups (classes) of participants within a given dataset, in this study characterized by similar symptom profiles. LCA uses statistical criteria to identify and accurately enumerate the groups that best fit the data. Instead of giving a particular true solution, LCA produces solutions for different numbers of classes with relative fit indices. The Bayesian information criterion (BIC) was used for the goodness-of-fit to determine the optimal number of classes. The BIC is a parsimony index determining how improvements in goodness-of-fit are counterbalanced by increased complexity due to the greater number of parameters. The null model is for one single class, i.e. the whole cohort belonging to the same latent class. This model is rejected when models with two or more parameters result in smaller BIC values. The best model fit is thus indicated by the smallest BIC value. For any given latent class model, each participant has an estimated probability of being a member of each latent class and participants were allocated to the latent classes of which they were most likely to be a member. We performed separate

https://doi.org/10.1017/S0033291710001625 Published online by Cambridge University Press

Table 1. *Included symptoms in the various confirmatory factor analyses and latent class analyses*

	DP	DGI	DSR	DPN	CP	MS	GI	GS	LCA29	LCA23
Abdominal pain	*						*		*	*
Back pain	*					*			*	*
Joint pain	*					*			*	*
Pain in extremities	*					*			*	*
Chest pain	*				*				*	*
Headache	*							*	*	*
Pain during menstruation	*								*	
Pain during urination	*									
Urinary retention				*						
Burning sensation genitals	*								*	
Pain additional sites	*								*	*
Vomiting other than during pregnancy		*								
Vomiting throughout pregnancy			*						*	*
Nausea		*					*		*	*
Diarrhoea		*					*		*	*
Feeling bloated or full of gas		*					*		*	*
Intolerance of several foods		*					*		*	*
Blindness				*						
Blurred vision									*	*
Deafness				*						
Impaired balance				*				*	*	*
Impaired coordination				*				*		
Loss of touch or pain sensation				*		*			*	*
Paralysis				*						
Aphonia				*					*	*
Seizures				*						
Dizziness								*	*	*
Loss of consciousness other than fainting				*						
Dissociative symptoms such as amnesia				*						
Double vision				*					*	*
Shortness of breath					*				*	*
Localized weakness				*		*			*	*
Skin blotches or discoloration									*	*
Bad taste in mouth, excessively coated tongue										
Frequent urination									*	*
Numbness – tingling						*			*	*
Difficulty swallowing or lump in throat				*					*	*
Irregular menses			*							
Excessive menstrual bleeding			*						*	
Sexual indifference			*						*	
Pain during sexual intercourse	*									
Unpleasant sexual intercourse			*						*	
Other sexual problems			*						*	

The first four columns include DSM factors that compose the diagnostic criteria for DSM-IV somatization disorder: DP, DSM Pain; DGI, DSM gastrointestinal other than pain; DSR, DSM sexual/reproductive other than pain; DPN, DSM pseudoneurological. The second four columns include previously suggested symptom clusters: CP, cardiopulmonary factor; MS, musculoskeletal factor; GI, gastrointestinal factor; GS, general symptoms factor. The last two columns summarize which symptoms were included in the latent class analyses with (LCA29) and without (LCA23) the reproductive symptoms.

LCAs for the entire cohort and for males or females separately, either including the 29 CIDI symptoms eliciting a positive response from $\geqslant$10 participants or including 23 CIDI symptoms after excluding the six reproductive and sexual symptoms (see Table 1 for included symptoms). In line with previous reports (Fink *et al*. 2007), we decided *a priori* to include only those items that were present in at least 10 respondents, since including items with a low prevalence increases the risk of identifying non-replicable

https://doi.org/10.1017/S0033291710001625 Published online by Cambridge University Press

latent classes. To index the amount to which symptoms discriminated the latent classes, we used Cramer's V, which is a correlation coefficient based on the χ^2 statistic. Accordingly, Cramer's V^2 is similar to R^2 in regression models and, in this case, reflects how much of the variability in the dependent variable is explained by latent class membership.

Descriptive analyses and the calculation of sensitivity, specificity and predictive values of various symptom thresholds for latent class membership were analysed using SPSS 16.0 (SPSS Inc., USA). Two-sided p values <0.05 were considered significant.

Results

General characteristics

The current study cohort consists of 461 males (47.8%) and 503 females (52.2%), with an average age of 55.8 years (S.D. = 11.1, minimum 35.9 years, maximum 82.3 years). A total of 583 participants reported at least one functional symptom, while the maximum number of reported functional symptoms was 18. A statistically significant gender difference was found in the total number of symptoms reported [median (IQ range) males 1 (0–1) *v.* females 1 (0–2), $Z = -6.919$, $p < 0.001$]. Since this gender difference might be related to the fact that the CIDI interview includes reproductive and sexual symptoms that are not equally applicable to men and women, we repeated the analysis excluding these symptoms and found that the gender difference remained [median (IQ range) males 1 (0–1) *v.* females 1 (0–2), $Z = -5.559$, $p < 0.001$]. In addition to female reproductive symptoms, significant gender differences existed for a variety of pain symptoms (back, joints, extremities, head, additional sites), dizziness, intolerance of several foods, bad taste in mouth, or excessively coated tongue, difficulty swallowing or lump in throat, sexual indifference and unpleasant sexual intercourse, with women scoring higher for all these symptoms. There was no association between the total number of functional symptoms and age (Spearman's rho = 0.037, $p = 0.256$).

Question 1: Symptom clusters

We first performed a confirmatory factor analysis on the symptom groups that form the core of the current DSM-IV diagnostic criteria for somatization disorder. We used the CIDI scoring rules to define groups of pain symptoms, gastrointestinal symptoms other than pain, sexual or reproductive symptoms other than pain and pseudoneurological symptoms (included symptoms are described in Table 1). Because of the different definition of the sexual or reproductive symptom cluster, we performed these analyses separately for males and females. We did not find a satisfactory fit, for males [χ^2(degrees of freedom (df = 15) = 28.077, $p = 0.0211$; CFI = 0.874; RMSEA = 0.043] nor for females [χ^2(df = 32) = 64.553, $p = 0.0006$; CFI = 0.843; RMSEA = 0.045]. Although, in both cases, RMSEA was <0.05, CFI was not >0.95 and the χ^2 test was significant.

In addition, we performed a confirmatory factor analysis using factors resembling those previously reported (Kroenke *et al*. 1998; Fink *et al*. 2007). Since these factors were defined based on datasets including both males and females, we also tested them on the entire cohort. We defined a cardiopulmonary factor, a musculoskeletal factor and a gastrointestinal factor (included symptoms are described in Table 1). The fit of this model was relatively good [χ^2(df = 31) = 45.067, $p = 0.0492$; CFI = 0.957; RMSEA = 0.022] and this three-factor model fitted our data significantly better than the corresponding one-factor model [χ^2 for difference testing (df = 3) = 28.324, $p < 0.0001$]. We also tested a four-factor model, including a general symptoms factor as suggested previously (Fink *et al*. 2007). This model had a good fit to the data [χ^2(df = 29) = 40.073, $p = 0.0828$; CFI = 0.963; RMSEA = 0.020] and again model fit was significantly better than that of the corresponding one-factor model [χ^2 for difference testing (df = 6) = 29.748, $p < 0.0001$]. Finally, we performed a confirmatory factor analysis including a second order factor representing a common latent factor underlying the cardiopulmonary, gastrointestinal and musculoskeletal symptom groups suggested previously (Fink *et al*. 2007). Also, this model had a relatively good fit to the data [χ^2(df = 31) = 45.067, $p = 0.0492$; CFI = 0.957; RMSEA = 0.022].

Question 2: Empirically-based cut-off

LCA was performed in order to identify different classes of subjects within our dataset and to test whether subjects were classified according to symptom profile or to symptom count. We performed separate LCAs including either 29 or 23 symptoms (in the latter case excluding the reproductive and sexual symptoms) and for the entire cohort or males and females separately. Table 2 shows the BIC values of the LCA solutions of the different models. The best model fit (indicated by the smallest BIC value) in all analyses was achieved with a two-class model.

We continued with analyses on the total cohort including the 23 symptoms that were not gender specific (results for 29 symptoms are comparable and available upon request). Table 3 shows the proportion (and number) of participants in latent classes 1 and 2 reporting a particular symptom. For all symptoms, the

https://doi.org/10.1017/S0033291710001625 Published online by Cambridge University Press

Table 2. *Goodness-of-fit for the latent class analyses solutions*

	BIC value					
	Symptoms with at least 10 positive responses in the total cohort*			Symptoms with at least 10 positive responses in the total cohort, excluding six reproductive and sexual symptoms		
Class solution	Total cohort	Females	Males	Total cohort	Females	Males
1-class	9851.229	6010.069	3819.341	8150.172	4959.507	3204.632
2-class	9439.473	5822.640	3752.171	7754.167	4757.430	3155.542
3-class	9526.361	5916.439	3859.751	7817.547	4820.862	3253.088
4-class	9683.176	6037.085	3982.936	7908.795	4904.980	3358.085

BIC, Bayesian information criterion (values represent not sample-size adjusted BIC values).
* In males excluding pain during menstruation and excessive menstrual bleeding.

Table 3. *Distribution of symptoms in latent classes based on latent class analysis applied to symptoms with at least 10 positive responses in the total cohort, excluding reproductive and sexual symptoms*

	Proportion (number) reporting symptom			
	Latent class 1 ($n=105$)	Latent class 2 ($n=859$)	Cramer's V	p
Abdominal pain	0.43 (45)	0.04 (34)	0.442	<0.001
Difficulty swallowing or lump in throat	0.29 (30)	0.02 (16)	0.390	<0.001
Feeling bloated or full of gas	0.28 (29)	0.02 (16)	0.380	<0.001
Pain in extremities	0.42 (44)	0.06 (55)	0.364	<0.001
Localized weakness	0.17 (18)	0 (3)	0.356	<0.001
Loss of touch or pain sensation	0.23 (24)	0.01 (12)	0.353	<0.001
Joint pain	0.45 (47)	0.08 (71)	0.347	<0.001
Headache	0.4 (42)	0.07 (60)	0.334	<0.001
Back pain	0.39 (41)	0.07 (62)	0.321	<0.001
Intolerance of several foods	0.16 (17)	0.01 (6)	0.316	<0.001
Dizziness	0.3 (32)	0.04 (38)	0.313	<0.001
Impaired balance	0.18 (19)	0.01 (10)	0.309	<0.001
Chest pain	0.22 (23)	0.03 (25)	0.272	<0.001
Nausea	0.1 (10)	0 (3)	0.248	<0.001
Pain additional sites	0.17 (18)	0.02 (19)	0.242	<0.001
Numbness/tingling	0.14 (15)	0.02 (13)	0.237	<0.001
Shortness of breath	0.13 (14)	0.01 (11)	0.236	<0.001
Blurred vision	0.12 (13)	0.02 (15)	0.197	<0.001
Double vision	0.07 (7)	0.01 (6)	0.161	<0.001
Diarrhoea	0.08 (8)	0.01 (11)	0.142	<0.001
Frequent urination	0.08 (8)	0.02 (19)	0.102	0.002
Aphonia	0.07 (7)	0.02 (17)	0.094	0.004
Skin blotches or discoloration	0.06 (6)	0.02 (14)	0.089	0.006

Symptoms are sorted by Cramer's V; higher values of Cramer's V indicate symptoms that better discriminated the latent classes.

proportion of subjects with a positive response was higher for class 1 members than for class 2 members. There were no specific symptoms characterizing class membership; participants in one of the classes displayed few symptoms ($n=859$) and participants in the other class ($n=105$) presented many symptoms.

We next tested whether we could approach the LCA analysis results with a simple cut-off score solely based on the total number of symptoms. We included all symptoms in this cut-off to test whether our LCA classes could also be approached by a cut-off score that was not restricted to the symptoms included in the

https://doi.org/10.1017/S0033291710001625 Published online by Cambridge University Press

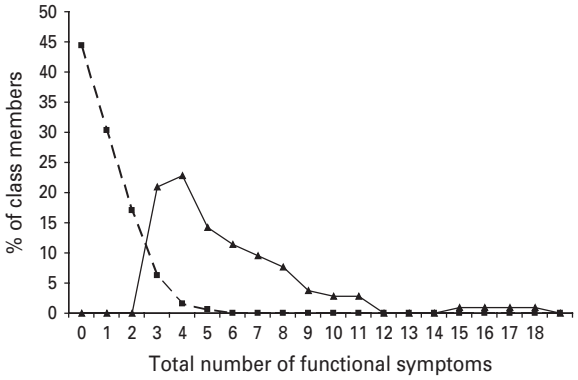

Fig. 1. Proportion of participants in latent class 1 (—) and class 2 (– – –) in relation to functional symptom count. Latent class 1 is characterized by high numbers of functional symptoms, whereas latent class 2 is characterized by low numbers of functional symptoms.

input dataset on which the definition of the classes was based. Fig. 1 depicts the proportion of participants in the different latent classes in relation to functional symptom count. Based on this figure, we tested sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) of cut-off three, four and five symptoms to determine class membership, including all 43 symptoms to determine whether participants scored above or below the cut-off. A cut-off of three has a high sensitivity and specificity (1.00 and 0.92, respectively), but the PPV is low (0.59; NPV = 1.00). A cut-off of five, on the other hand, has a good PPV and NPV (0.92 and 0.95, respectively), but is not very sensitive (sensitivity 0.56, specificity 0.99). A threshold of four symptoms is the optimal cut-off in terms of sensitivity and predictive value. A simple cut-off score of four out of 43 symptoms correctly identified 79% of the LCA class 1 participants (and 98% of the LCA class 2 participants) while correctly classifying 97% of participants scoring below and 82% of participants scoring about the cut-off as LCA classes 2 and 1, respectively. When we based our cut-off score on the 23 prevalent symptoms, not on all symptoms, the optimal cut-off in terms of sensitivity (1.00), specificity (0.97), PPV (0.82) and NPV (1.00) was three symptoms.

Question 3: The multiple organ system requirement

We tested the number of symptom clusters involved in participants scoring above and below the cut-off of four symptoms, using the cardiopulmonary, gastrointestinal, musculoskeletal and general symptom clusters that were found to fit our empirical data. Of participants scoring above the cut-off, in 10.9% only one symptom cluster was involved, whereas in 47.5% two symptom clusters were involved, in 32.7% three and in 8.9% all four clusters were involved.

Table 4 summarizes the co-morbidity with DSM-IV common mental disorders. High co-morbidity is found, since participants scoring above the cut-off of four symptoms have a more than four times higher risk of having any anxiety and depression disorder than participants scoring below the cut-off. All specific diagnoses more often occur in participants scoring above the cut-off; with one remarkable exception, the simple phobia of the blood- or injection-injury type is absent in somatizers but not in controls (0.0 *v.* 0.5%), whereas all other specific phobias are more prevalent in the somatizers than in the controls (animal type 2.0 *v.* 0.1, natural environment type 2.0 *v.* 0.8, situational type 2.0 *v.* 1.0). When comparing the number of involved symptom clusters, the increased co-morbidity with depression and anxiety is evident from one symptom cluster onwards, reaching its maximum at two symptom clusters.

Discussion

This study in a large population-representative cohort suggests that a simple symptom count can be used as a dimensional diagnosis of somatization, as suggested in the DSM-5 proposal. In those instances in which a categorical diagnosis is preferred, a simple cut-off of four out of 43 functional symptoms best fitted our data. We did not find any added value for incorporating the number of symptom clusters into the diagnostic criteria.

A major strength of our study is the use of a population sample including approximately equal numbers of both genders covering a wide age range. Moreover, we performed structured psychiatric interviews in all participants; thus, not selecting subjects on the basis of a screening questionnaire. In addition, we focused on recent symptoms, whereas past epidemiological studies of somatization relied on lifetime symptoms. There are also a few limitations to discuss. First, we used self-reported presence of functional symptoms, possibly underestimating the real prevalence if people tend to seek a physical reason for their complaints. There is no consensus about whether symptoms ascribed to functional syndromes should be classified as somatization symptoms. The classification of these symptoms as either somatization symptoms or medically explained symptoms did not essentially influence the current results. Second, the utility of both factor analyses and LCA is critically dependent on the input dataset. The fact that pseudoneurological symptoms do not cluster with one of the large clusters such as gastrointestinal or cardiovascular symptoms may be explained by their low prevalence. However,

https://doi.org/10.1017/S0033291710001625 Published online by Cambridge University Press

Table 4. *DSM-IV common mental disorders in both groups* (%)

	Total ($n=964$)	<4 functional symptoms ($n=863$)	$\geqslant 4$ functional symptoms ($n=101$)	1 cluster ($n=11$)	2 clusters ($n=48$)	3 clusters ($n=33$)	4 clusters ($n=9$)
Any depression or anxiety disorder	11.8	8.8	37.6	27.3	37.5	39.4	44.4
Major depression	7.0	5.1	22.8	27.3	22.9	21.2	22.2
Dysthymia	0.4	0	4.0	0	2.1	6.1	11.1
Generalized anxiety disorder	2.7	1.9	9.9	18.2	8.3	9.1	11.1
Panic disorder	0.9	0.6	4.0	0	0	9.1	11.1
Agoraphobia without history of panic disorder	0.5	0.3	2.0	0	2.1	0	11.1
Agoraphobia with or without history of panic disorder	0.8	0.5	4.0	0	2.1	3.0	22.2
Social phobia	1.5	0.9	5.9	0	6.2	9.1	0
Simple phobia	2.3	2.0	5.0	0	4.2	6.1	11.1

despite the fact that we used the CIDI interview, we were able to confirm previously suggested symptom clusters that were defined using the SCAN interview (Fink *et al*. 2007) and using a 15-symptom checklist derived from the PRIME-MD (Kroenke *et al*. 1998). Third, only 43% of approached people agreed to participate. Previous analyses indicated no differences in gender, age or neuroticism between those who participated and those who refused (Tak *et al*. 2010), making it unlikely that selection bias essentially influenced our results.

Our data fail to provide empirical support for the designated symptom clusters in the DSM-IV. It is important to note that we performed our analysis on symptoms experienced in the previous 12 months, whereas the DSM-IV diagnostic criteria refer to lifetime symptoms. We present the first study, which formally tested the DSM-IV symptom groups using confirmatory factor analyses. Our results are in agreement with exploratory analyses that also did not replicate the DSM-IV clusters in a general population cohort (Liu *et al*. 1997). When using previously suggested symptom clusters that did fit our data, almost 90% of participants scoring above our cut-off have symptoms derived from more than one symptom cluster, suggesting that there is little additional value for a minimum required number of organ systems. Similarly, data on psychiatric co-morbidity also do not give indications for such a multiple organ system threshold. Clinically significant somatization is known to be accompanied by psychiatric co-morbidity (de Waal *et al*. 2004; Haug *et al*. 2004). An abrupt increase in psychiatric co-morbidity beyond a certain threshold number of involved symptom clusters could thus be regarded as indicative of a dichotomy between states of health (innocent symptoms) and disease (clinically relevant somatization). Our data do not support such a symptom cluster threshold. Despite the fact that previously suggested clusters fit our data, LCA revealed that participants are clustered based on symptom count instead of symptom profile. This is in agreement with several earlier studies that found highly correlated symptom factors, suggesting that a general higher order somatization factor explained the majority of the variance in functional symptoms (Liu *et al*. 1997; Deary, 1999; Fink *et al*. 2007).

Our data underline the validity of a dimensional approach of diagnosing. Nonetheless, if a categorical approach is preferred, our data indicate that a simple cut-off of four symptoms best distinguishes somatizers from non-somatizers. Although the mere counting of physical complaints as a basis for the classification has been criticized in the past (Fink, 1996), it has been shown that the number of bodily symptoms is still an important feature for the prediction of course and outcome (Jackson *et al*. 2006; Kroenke *et al*. 2007). Our results seem in agreement with results on multisomatoform disorder (MSD), which is defined as three or more currently bothersome unexplained physical complaints (from a 15-symptom checklist), plus a history of chronic somatization (i.e. unexplained symptoms that were usually present for at least 2 years) (Kroenke *et al*. 1997). In fact, when we based our cut-off score on the 23 prevalent symptoms, not on all symptoms, which might be more practical in clinical situations, the appropriate cut-off would be three symptoms. Despite differences in the number of symptoms and the time-frame, there are remarkable similarities between our cut-off and MSD. It is interesting that MSD was present in 8% (Spitzer *et al*. 1994; Jackson & Kroenke, 2008) to 19% (Dickinson *et al*. 2003) of primary care patients, compared with a

comparable proportion of 11% scoring above our cut-off in our population cohort. In addition, psychiatric co-morbidity is strikingly similar. Major depression was present in 21% of patients with MSD and 23% of patients above our cut-off, generalized anxiety disorder was present in 11% with MSD and 10% of our somatizers and panic disorder was present in 2% with MSD and 4% of patients above our cut-off (Jackson & Kroenke, 2008). It appears that our LCA-based cut-off might identify the patients who fulfil the diagnostic criteria for MSD. Unfortunately, not all MSD symptoms are surveyed in the CIDI and it is thus not possible to calculate the agreement.

Further research should validate these results in other populations and using other interviews. Given the fact that, in some specialties, functional symptoms outnumber the medically explained symptoms (Nimnuan *et al*. 2001*a*), the importance of these validations is not restricted to psychiatric settings. The observation that the majority of participants included in the subgroup with high levels of functional symptoms do not show a depression or anxiety disorder underlines the broader importance of these findings.

Acknowledgements

This study was financially supported by the Netherlands Organization for Scientific Research (Pionier 900–00–002).

Declaration of Interest

None.

References

Andrews G, Peters L (1998). The psychometric properties of the composite international diagnostic interview. *Social Psychiatry and Psychiatric Epidemiology* **33**, 80–88.

APA (1994). *Diagnostic and Statistical Manual of Mental Disorders* (4th edn). Washington, DC: American Psychiatric Association.

Deary IJ (1999). A taxonomy of medically unexplained symptoms. *Journal of Psychosomatic Research* **47**, 51–59.

de Waal MWM, Arnold IA, Eekhof JAH, van Hemert AM (2004). Somatoform disorders in general practice – prevalence, functional impairment and comorbidity with anxiety and depressive disorders. *British Journal of Psychiatry* **184**, 470–476.

Dickinson WP, Dickinson LM, de Gruy FV, Candib LM, Main DS, Libby AM, Rost K (2003). The somatization in primary care study: a tale of three diagnoses. *General Hospital Psychiatry* **25**, 1–7.

Dimsdale J, Creed F (2009). The proposed diagnosis of somatic symptom disorders in DSM-V to replace somatoform disorders in DSM-IV – a preliminary report. *Journal of Psychosomatic Research* **66**, 473–476.

Fink P (1996). Somatization – beyond symptom count. *Journal of Psychosomatic Research* **40**, 7–10.

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

Gara MA, Silver RC, Escobar JI, Holman A, Waitzkin H (1998). A hierarchical class analysis (HICLAS) of primary care patients with medically unexplained somatic symptoms. *Psychiatry Research* **81**, 77–86.

Haug TT, Mykletun A, Dahl AA (2004). The association between anxiety, depression, and somatic symptoms in a large population: The HUNT-II Study. *Psychosomatic Medicine* **66**, 845–851.

Hiller W, Cuntz U, Rief W, Fichter MM (2001). Searching for a gastrointestinal subgroup within the somatoform disorders. *Psychosomatics* **42**, 14–20.

Hu L, Bentler PM (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling* **6**, 1–55.

Jackson J, Fiddler M, Kapur N, Wells A, Tomenson B, Creed F (2006). Number of bodily symptoms predicts outcome more accurately than health anxiety in patients attending neurology, cardiology, and gastroenterology clinics. *Journal of Psychosomatic Research* **60**, 357–363.

Jackson JL, Kroenke K (2008). Prevalence, impact, and prognosis of multisomatoform disorder in primary care: a 5-year follow-up study. *Psychosomatic Medicine* **70**, 430–434.

Kroenke K, Sharpe M, Sykes R (2007). Revising the classification of somatoform disorders: key questions and preliminary recommendations. *Psychosomatics* **48**, 277–285.

Kroenke K, Spitzer RL, de Gruy FV, Hahn SR, Linzer M, Williams JBW, Brody D, Davies M (1997). Multisomatoform disorder – an alternative to undifferentiated somatoform disorder for the somatizing patient in primary care. *Archives of General Psychiatry* **54**, 352–358.

Kroenke K, Spitzer RL, de Gruy III FV, Swindle R (1998). A symptom checklist to screen for somatoform disorders in primary care. *Psychosomatics* **39**, 263–272.

Liu G, Clark MR, Eaton WW (1997). Structural factor analyses for medically unexplained somatic symptoms of somatization disorder in the Epidemiologic Catchment Area study. *Psychological Medicine* **27**, 617–626.

Muthen LK, Muthen BO (2004). *Mplus User's Guide*. Muthen & Muthen: Los Angeles, CA.

Nimnuan C, Hotopf M, Wessely S (2001*a*). Medically unexplained symptoms – an epidemiological study in seven specialties. *Journal of Psychosomatic Research* **51**, 361–367.

Nimnuan C, Rabe-Hesketh S, Wessely S, Hotopf M (2001*b*). How many functional somatic syndromes? *Journal of Psychosomatic Research* **51**, 549–557.

Pinto-Sietsma SJ, Janssen WMT, Hillege HL, Navis G, de Zeeuw D, de Jong PE (2000). Urinary albumin excretion is associated with renal functional abnormalities in a

nondiabetic population. *Journal of the American Society of Nephrology* **11**, 1882–1888.

Robbins JM, Kirmayer LJ, Hemami S (1997). Latent variable models of functional somatic distress. *The Journal of Nervous and Mental Disease* **185**, 606–615.

Simon G, Gater R, Kisely S, Piccinelli M (1996). Somatic symptoms of distress: an international primary care study. *Psychosomatic Medicine* **58**, 481–488.

Simon GE, Gureje O (1999). Stability of somatization disorder and somatization symptoms among primary care patients. *Archives of General Psychiatry* **56**, 90–95.

Spitzer RL, Williams JBW, Kroenke K, Linzer M, de Gruy FV, Hahn SR, Brody D, Johnson JG (1994). Utility of new procedure for diagnosing mental disorders in primary-care – the Prime-Md-1000 Study. *Journal of the American Medical Association* **272**, 1749–1756.

Sullivan PF, Smith W, Buchwald D (2002). Latent class analysis of symptoms associated with chronic fatigue syndrome and fibromyalgia. *Psychological Medicine* **32**, 881–888.

Swartz M, Blazer D, Woodbury M, George L, Landerman R (1986). Somatization disorder in a United States southern community – use of a new procedure for analysis of medical classification. *Psychological Medicine* **16**, 595–609.

Tak LM, Janssens KAM, Dietrich A, Slaets JPJ, Rosmalen JGM (2010). Age-specific associations between cardiac vagal activity and functional somatic symptoms: a population-based study. *Psychotherapy and Psychosomatics* **79**, 179–187.