


Neuro-11: a new questionnaire for the assessment of somatic symptom disorder in general hospitals

Silin Zeng,^{1,2} Yian Yu,³ Shan Lu ,⁴ Sirui Zhang,¹ Xiaolin Su,¹ Ge Dang,¹ Ying Liu,⁵ Zhili Cai,¹ Siyan Chen,¹ Yitao He,¹ Xin Jiang,⁶ Chanjuan Chen,¹ Lei Yuan,¹ Peng Xie,⁷ Jianqing Shi,^{3,8} Qingshan Geng,⁶ Rafael H Llinas,⁵ Yi Guo^{1,4}

To cite: Zeng S, Yu Y, Lu S, *et al.* Neuro-11: a new questionnaire for the assessment of somatic symptom disorder in general hospitals. *General Psychiatry* 2023;**36**:e101082. doi:10.1136/gpsych-2023-101082

► Additional supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/gpsych-2023-101082>).

JS, QG, RHL and YG contributed equally.,
SZ and YY contributed equally.

SZ and YY are joint first authors.

Received 18 April 2023
Accepted 01 August 2023



© Author(s) (or their employer(s)) 2023. 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

Dr Yi Guo;
xuanyi_guo@163.com

Dr Qingshan Geng;
gengqingshan@szhospital.com

Dr Rafael H Llinas;
rlinas@jhmi.edu

Dr Jianqing Shi;
shijq@sustech.edu.cn

ABSTRACT

Background Somatic symptom disorder (SSD) commonly presents in general hospital settings, posing challenges for healthcare professionals lacking specialised psychiatric training. The Neuro-11 Neurosis Scale (Neuro-11) offers promise in screening and evaluating psychosomatic symptoms, comprising 11 concise items across three dimensions: somatic symptoms, negative emotions and adverse events. Prior research has validated the scale's reliability, validity and theoretical framework in somatoform disorders, indicating its potential as a valuable tool for SSD screening in general hospitals.

Aims This study aimed to establish the reliability, validity and threshold of the Neuro-11 by comparing it with standard questionnaires commonly used in general hospitals for assessing SSD. Through this comparative analysis, we aimed to validate the effectiveness and precision of the Neuro-11, enhancing its utility in clinical settings.

Methods Between November 2020 and December 2021, data were collected from 731 patients receiving outpatient and inpatient care at Shenzhen People's Hospital in China for various physical discomforts. The patients completed multiple questionnaires, including the Neuro-11, Short Form 36 Health Survey, Patient Health Questionnaire 15 items, Hamilton Anxiety Scale and Hamilton Depression Scale. Psychiatry-trained clinicians conducted structured interviews and clinical examinations to establish a gold standard diagnosis of SSD.

Results The Neuro-11 demonstrated strong content reliability and structural consistency, correlating significantly with internationally recognised and widely used questionnaires. Despite its brevity, the Neuro-11 exhibited significant correlations with other questionnaires. A test-retest analysis yielded a correlation coefficient of 1.00, Spearman-Brown coefficient of 0.64 and Cronbach's α coefficient of 0.72, indicating robust content reliability and internal consistency. Confirmatory factor analysis confirmed the validity of the three-dimensional structure ($p < 0.001$, comparative fit index=0.94, Tucker-Lewis index=0.92, root mean square error of approximation=0.06, standardised root mean square residual=0.04). The threshold of the Neuro-11 is set at 10 points based on the maximum Youden's index from the receiver operating characteristic curve analysis. In terms of diagnostic efficacy, the Neuro-11 has an area under the curve of 0.67.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Somatic symptom disorder (SSD) is characterised by a high incidence and low diagnosis rate in general hospital settings, necessitating the development of a rapid screening tool to improve the recognition of SSD in this context.

WHAT THIS STUDY ADDS

⇒ The Neuro-11 Neurosis Scale (Neuro-11) is a screening scale for SSD, employing a three-dimensional structure that incorporates somatic symptoms, negative emotions and adverse events.
⇒ It has been validated to possess good reliability and validity, outperforming the Hamilton Depression Scale and Patient Health Questionnaire 15 items in terms of screening accuracy for SSD.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ A simple and user-friendly scale such as Neuro-11 holds excellent promise for SSD screening, particularly in general hospitals with many outpatient cases in China.
⇒ This approach can potentially enhance the diagnosis and treatment of SSD while reducing the burden on healthcare resources.

Conclusions (1) The Neuro-11 demonstrates robust associations with standard questionnaires, supporting its validity. It is applicable in general hospital settings, assessing somatic symptoms, negative emotions and adverse events. (2) The Neuro-11 exhibits strong content reliability and validity, accurately capturing the intended constructs. The three-dimensional structure demonstrates robust construct validity. (3) The threshold of the Neuro-11 is set at 10 points.

INTRODUCTION

Somatic symptom disorder (SSD) is a mental disorder as defined in the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5),¹ characterised by persistent physical discomfort and excessive symptomatic feelings or behaviours. Since physical

discomfort is the primary symptom of SSD, patients often seek initial care in general hospitals.²⁻³ Research indicates that approximately one-third of patients in general hospitals receive a diagnosis of SSD.⁴⁻⁷ However, due to the intricate and ambiguous diagnostic criteria associated with SSD, the diagnosis rate in general hospitals is relatively low, and such cases are seldom recorded in medical databases.⁸ Consequently, misdiagnosis and underdiagnosis rates among these patients are considerably high, leading to inadequate diagnosis, treatment and significant waste of medical resources.⁹ Currently, there are no established objective biological markers for diagnosing mental illnesses, necessitating psychiatrists to rely on interviews and other subjective methods. Consequently, non-psychiatrists in general hospitals encounter challenges in effectively screening and identifying individuals with SSD.

Psychiatric scales serve as crucial tools for screening and identifying mental illnesses. In general hospitals, widely employed international scales including the Patient Health Questionnaire 9 items (PHQ-9), Generalised Anxiety Disorder 7 items (GAD-7), Hamilton Anxiety/Depression Scale (HAMA/HAMD) for assessing depression and anxiety and the Patient Health Questionnaire 15 items (PHQ-15) for evaluating somatic symptoms. However, due to the considerable overlap of somatic disorders with anxiety and depression,^{3 10 11} many theories currently posit somatic symptoms as manifestations and accompanying features of depression and anxiety disorders.¹²⁻¹⁵ Consequently, anxiety and depression scales are often used to assess patients with somatic symptoms. Despite these scales containing physical symptom items seemingly related to SSD, their application in general hospitals, primarily developed by psychiatrists, presents inevitable disadvantages. First, specific psychiatric scales, such as HAMA/HAMD, can be laborious and time-consuming to implement. Second, numerous psychiatric symptom items, such as those concerning suicidal tendencies, hypochondriasis and hallucinations, may contribute to mental illness stigma. This may result in patients with somatic complaints resisting, feeling perplexed and sensing neglect of their physical symptoms, leading to reduced cooperation when completing these scales. Moreover, an important aspect often overlooked is the influence of adverse events, acknowledged as social and psychological factors, on the occurrence and development of somatic disorders.^{16 17} Unfortunately, existing scales lack relevant items that adequately capture the impact of adverse events on patients' symptoms.

Given the current suboptimal recognition rate of SSD in general hospital settings and the limitations of existing scales, there is a pressing need to swiftly identify patients with this condition using a more practical screening scale. In response to this challenge, our research team has developed a somatic symptom screening scale named the Neuro-11 Neurosis Scale (Neuro-11) (shown in the online supplemental material).¹⁸ The Neuro-11 scale encompasses 11 items organised into three dimensions:

somatic symptoms, negative emotions and adverse events. In previous investigations, we have successfully established the scale's reliability, validity and theoretical framework in the context of somatoform disorders. In the present study, we aimed to further validate and calibrate the reliability, validity and threshold of the Neuro-11 for screening SSD by conducting a comparative analysis with existing standard questionnaires.

PARTICIPANTS AND METHODS

Participants

The dataset for this study was collected between April 2020 and December 2021. It consisted of participants who sought outpatient or inpatient care from various departments, including neurology, cardiology, gastroenterology, endocrinology and traditional Chinese medicine, at Shenzhen People's Hospital. The research flowchart is shown in figure 1. The inclusion criteria were as follows: (a) patients who sought outpatient or inpatient care in general hospitals; (b) age range of 18–80 years; (c) participants who voluntarily consented to participate in the clinical trials and provided written informed consent. Exclusion criteria for enrolment included: (a) patients with conditions such as dementia, aphasia or significant consciousness disturbances that would significantly impair their ability to complete scale examinations and participate in diagnostic interviews; (b) participants who are unable to cooperate with the scale assessment and evaluators and (c) individuals with current psychotic disorders, including schizophrenia, major depression, substance abuse or active suicidality. Additional exclusion and discontinuation criteria were (a) participants who exhibit poor compliance and are unable to cooperate with the completion of interviews and (b) individuals who fail to complete all assessments for various reasons.

Methods

Patients who met the predefined inclusion criteria provided informed consent and were instructed to complete the following scales:

- ▶ A custom questionnaire designed to collect demographic characteristics such as sex, age, marital status and education level.
- ▶ Neuro-11, a newly developed scale specifically designed for screening SSD.
- ▶ PHQ-15, a brief self-assessment questionnaire recommended by the DSM-5. It consists of 15 items that assess the severity of physical symptoms experienced in the past month. Previous studies have demonstrated the scale's reliability and validity in assessing the severity of physical discomfort symptoms.¹⁹⁻²¹
- ▶ Medical Outcomes Study Questionnaire Short Form 36 Health Survey (SF-36), a widely used health questionnaire developed by the Boston Health Research Institute. It comprehensively evaluates the quality of life of respondents in the past year across eight domains, including physical functioning (PF), bodily

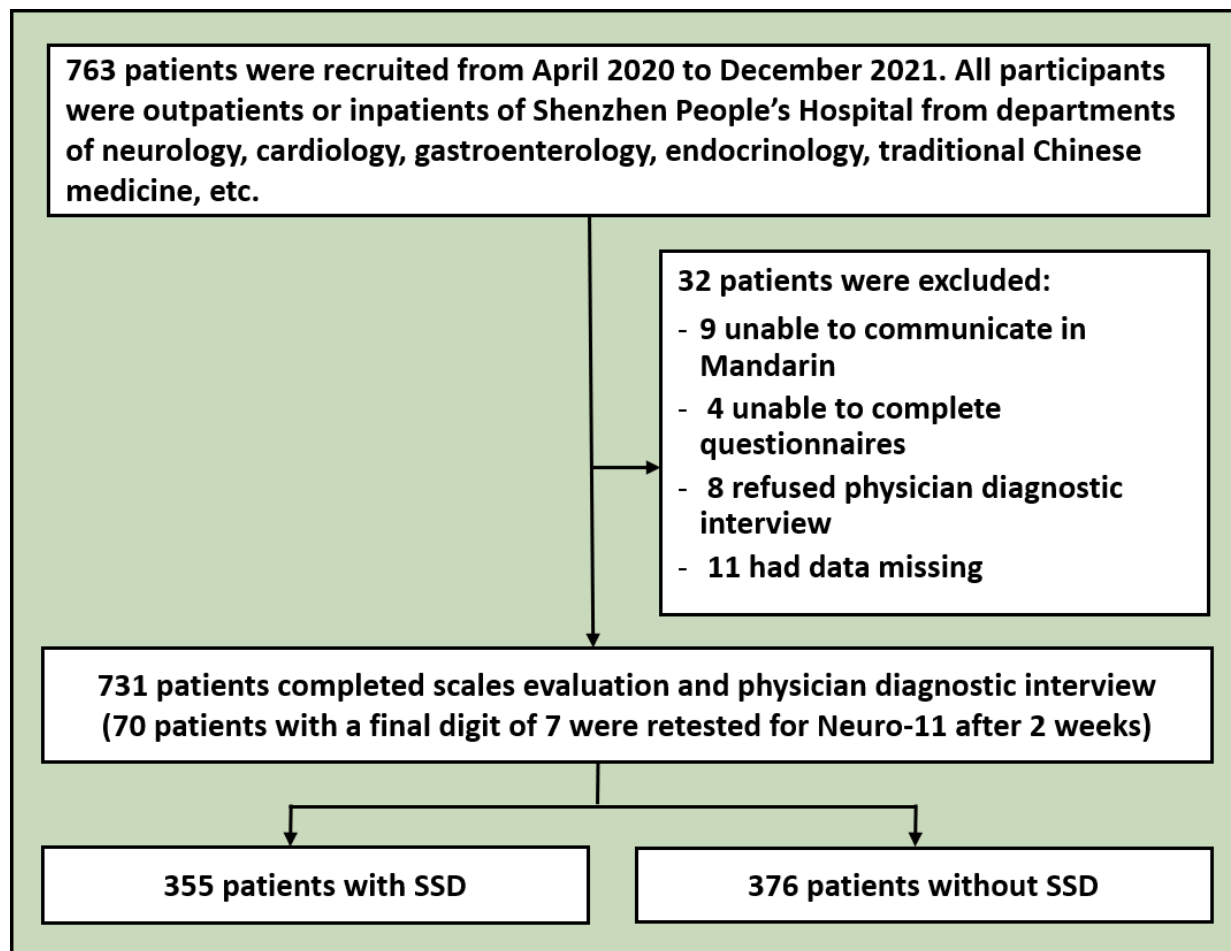


Figure 1 Flowchart of the study. SSD, somatic symptom disorder.

pain (BP), general health (GH), vitality, social functioning (SF), role limitations due to emotional and physical health, emotional well-being and mental health. The SF-36 has shown good reliability and validity.^{22 23}

- ▶ HAMD, a commonly used rating scale for assessing depression developed by Hamilton in 1960. It has demonstrated sensitivity in distinguishing between patients with depression and the general population, along with good reliability and validity.²⁴
- ▶ HAMA, a widely used rating scale for assessing anxiety developed by Hamilton in 1959. It comprises two-factor structures, physical and mental, and is primarily used to evaluate the severity of neurosis and anxiety symptoms experienced in the past week. HAMA has shown good reliability and validity.²⁵

The order in which the Neuro-11, PHQ-15 and SF-36 scales were administered was randomised, while two professional psychological evaluators conducted the HAMD and HAMA scales. The completion of the five scales typically took between 20 and 50 min. After completing the scales, patients underwent interviews conducted by two attending doctors with >5 years of clinical experience. The diagnostic criteria for SSD were based on the DSM-5, and patient assessments were performed using the DSM-5

Structured Clinical Interview. The scale evaluators and doctors were blinded to each other, with the doctors unaware of the patients' scale results. Patients with a final digit of 7 underwent a retest of the Neuro-11 scale after a 2-week interval to assess the test-retest reliability.

Statistical analysis

A total of 763 patients who met the predefined inclusion criteria participated in the survey. The response rate among eligible individuals was 95.8%, resulting in a final sample size of 731 participants. Among them, 355 individuals were diagnosed with SSD. Descriptive data analysis was conducted to compare Neuro-11 scores among different demographic groups. As the Neuro-11 employs discrete values, a non-parametric Mann-Whitney U test was employed for the statistical analysis.

Two weeks following the completion of all scales, a subgroup of 70 patients was randomly selected to undergo a retest of the Neuro-11. The purpose of this retest was to assess the test-retest reliability and correlation analysis of the Neuro-11 scores. The content reliability of the Neuro-11 data was evaluated using the split-half technique, the Spearman-Brown formula and the Guttman split-half coefficient. Furthermore, the scale was calibrated by examining the correlations between the

Table 1 General demographics and Neuro-11 sum scores of the total subjects (n=731)

	Sample	Neuro-11 sum scores	Z score	P value
Gender				
Male	368 (50.3%)	10.0 (7.0, 15.0)	-5.24	<0.001
Female	363 (49.7%)	13.0 (8.0, 18.0)		
Age (years)				
<45	371 (50.8%)	13.0 (9.0, 17.0)	4.22	<0.001
≥45	360 (49.2%)	10.5 (7.0, 15.8)		
Marital status				
Married	531 (72.6%)	11.0 (7.0, 16.0)	-4.23	<0.001
Others	200 (27.4%)	14.0 (9.0, 18.0)		
Education level				
High school and below	404 (55.3%)	12.0 (7.0, 17.0)	-0.34	0.731
University and above	327 (44.7%)	12.0 (8.0, 16.0)		

Shapiro-Wilk test is used to verify the normal distribution of Neuro-11 scores. The range calculated based on mean (SD) is used to describe the normal scores, and the IQR is used to describe the non-normal scores. Mann-Whitney U test was conducted. P value of 0.05 was used as a significant level.

IQR, interquartile range; Neuro-11, Neuro-11 Neurosis Scale; SD, standard deviation.

total and dimension scores of the Neuro-11 and those obtained from previous surveys. To ensure the suitability of our data for factor analysis, the Kaiser-Meyer-Olkin test value was calculated. Confirmatory factor analysis was then performed to assess the construct validity of the proposed three-dimensional structure of the Neuro-11.

Using physician diagnoses as the reference standard, we determined the threshold for the SSD questionnaire by constructing a receiver operating characteristic (ROC) curve. The SSD questionnaire scores were then categorised based on the established threshold using logistic regression analysis. Statistical analyses were conducted using SPSS V.23.0, AMOS and R software to evaluate the data.

RESULTS

Descriptive analysis

The demographic characteristics for the total 731 subjects are presented in [table 1](#), where the score of Neuro-11 is compared against gender, age, marital status and education level. We found that for the data, the Neuro-11 sum score for females was significantly higher than for males, for younger people was higher than for older people and for married individuals was lower than for others (including unmarried, widowed and divorced). However, there was no significant difference in education level. For the 355 patients with SSD, no significant difference existed between the sum score of Neuro-11 and these four demographic characteristics. The detailed analysis that compared the three dimensions scores of Neuro-11 for these demographics showed some difference and is presented in [table 2](#). In patients with SSD, we found that females had significantly higher scores than males in the somatic symptoms dimension scores, and younger people

or married individuals had lower scores than others in the negative emotions dimension. For education level, subjects in the high school and below group had higher scores in the somatic symptoms dimension than subjects in the university group; these findings are presented in [table 2](#).

Reliability, calibration and structure validation of Neuro-11

Reliability test

We tested the reliability of Neuro-11 with test-retest and split-half methods to ensure the content of the design items was reasonable and valid. Neuro-11 had good reliability, and the sum score of the Neuro-11 in the test-retest was highly correlated ($r=1.00$, $p<0.01$). The results of the split-half reliability test were calculated by dividing the 11 items into two parts according to odd and even items. Neuro-11 had good internal consistency as the non-equal length Spearman-Brown coefficient was 0.64, Guttman Split-half coefficient was 0.69, which was consistent with the results obtained when the scale was first applied to assess somatoform disorders in 2014.¹⁸

Association between Neuro-11 and other questionnaires

Several neuropsychological assessment questionnaires, including the HAMD, HAMA, PHQ-15 and SF-36, are commonly used in general hospitals in China. This study used them as the calibration standard. Spearman's correlation was applied to analyse the association as the score took discrete numbers. Neuro-11 was significantly positively correlated with HAMA ($r=0.772$, $p<0.001$), HAMD ($r=0.748$, $p<0.001$) and PHQ-15 ($r=0.607$, $p<0.001$) and was significantly negatively correlated with SF-36 ($r=-0.429$, $p<0.001$), as shown in [figure 2A](#). The reason for the negative correlation between Neuro-11

Table 2 General demographics and Neuro-11-dimension scores of the patients with SSD (n=355)

	Sample	Body	Z score	Negative emotions	Z score	Adverse events	Z score
Gender							
Male	163 (45.92%)	8.00 (6.00, 11.00)	-2.42	2.00 (1.00, 3.00)	-0.91	3.00 (3.00, 6.00)	1.29
Female	192 (54.08%)	9.00 (6.25, 12.00)		2.00 (1.00, 3.00)		3.00 (3.00, 3.00)	
Age (years)							
<45	215 (60.56%)	6.00 (8.00, 11.00)	-0.51	2.00 (1.00, 3.00)	-2.78**	3.00 (3.00, 6.00)	-0.25
≥45	140 (39.44%)	9.00 (5.25, 12.00)		1.50 (1.00, 3.00)		3.00 (3.00, 6.00)	
Marital status							
Married	237 (66.76%)	8.00 (6.00, 11.00)	-0.55	2.00 (1.00, 3.00)	-3.18***	3.00 (3.00, 6.00)	-0.05
Others	118 (33.24%)	9.00 (6.00, 11.00)		2.00 (1.00, 4.00)		3.00 (3.00, 6.00)	
Education level							
High school and below	176 (49.58%)	9.00 (6.00, 12.00)	2.08	2.00 (1.00, 3.00)	-0.06	3.00 (3.00, 3.00)	-0.55
University and above	179 (50.42%)	8.00 (6.00, 10.00)		2.00 (1.00, 3.00)		3.00 (3.00, 6.00)	

p<0.01 , *p<0.001.
Neuro-11, Neuro-11 Neurosis Scale; SSD, somatic symptom disorder.

and SF-36 is that a high score in SF-36 indicates a healthy condition, but in Neuro-11 it is vice versa.

The correlations between each dimension of Neuro-11 and HAMA, HAMD and SF-36 are presented in figure 2B–D, respectively. The somatic symptoms dimension was highly correlated with the somatic anxiety factor (r=0.643, p<0.001) and the psychogenic anxiety factor (r=0.646, p<0.001). The negative emotion dimension was also highly correlated with the psychogenic anxiety factor (r=0.680, p<0.001), although it was also correlated with the somatic anxiety factor (r=0.445, p<0.001). The adverse event dimension, specially designed for Neuro-11, was also correlated to those two dimensions of HAMA, but the association was relatively weak (r<0.4). Compared with HAMD, the three dimensions of Neuro-11 were all highly correlated (p<0.001) with all the factors of HAMD, except for the weight factor. Compared with SF-36, the three dimensions of Neuro-11 had low correlations (r≈0.0) with the PF and SF factors in SF-36. Factors concerning BP, GH and mental health in SF-36 were highly correlated with Neuro-11. The SF in SF-36 was like the factor of weight in HAMD and may not contribute much information concerning SSD. PF was highly correlated with other physical health factors, for example, role physical (RP), BP, GH, role emotion and healthy transition in SF-36, but it was noteworthy that all three dimensions in Neuro-11 were highly correlated with those physical health factors but not correlated with PF. The correlations within Neuro-11 are presented in figure 2E. We noticed that item 11 had low correlations with all items except items 5 and 7 in the somatic symptoms dimension.

Confirmatory factor analysis

The confirmatory factor analysis was conducted to verify the structure of the items with the defined three dimensions based on our dataset. Figure 3 shows the structure

of the confirmatory factor analysis. Table 3 summarises all test results for Neuro-11, indicating the internal consistency and plausibility of this three-dimensional structure. The whole Cronbach’s α coefficient was >0.70, indicating a good internal consistency. The p value of the χ² test was far less than 0.05, meaning the model fitted the data very well. This was also confirmed by other measures where the comparative fit index (CFI) and Tucker-Lewis index (TLI) were >0.90, and the root mean square error of approximation (RMSEA) and the standardised root mean square residual (SRMR) were close to or smaller than 0.05.²⁶ However, our subsequent examination of the subscales’ convergent and discriminant power revealed some minor issues. According to average variance extracted (AVE) and composite reliability (CR) values, the convergent validity of the construct was adequate for dimensions except for the adverse events. Based on the prior defined three-dimensional structure, the somatic symptoms dimension’s AVE value was 0.27, and the CR value was 0.71. The negative emotions dimension’s AVE value was 0.61, and the CR value was 0.76. However, the adverse events dimension’s AVE value was 0.11, and the CR value was 0.08. A detailed examination has found that item 11 contributed little information to the dimension of adverse events loading -0.10 (p=0.175) under the Wald test, as a lower CR (<0.5) indicated that the items did not measure what they were intended to measure, and a low AVE (<0.6) indicated that more errors remained in the items than the variance explained by the intended dimension.²⁷

Determination of the thresholds and classification

Our dataset included 731 subjects; 355 were diagnosed with SSD and 376 were diagnosed with other disorders, for example, somniphathy. In total, 457 patients’ total Neuro-11 scores were ≥10 points, but these included

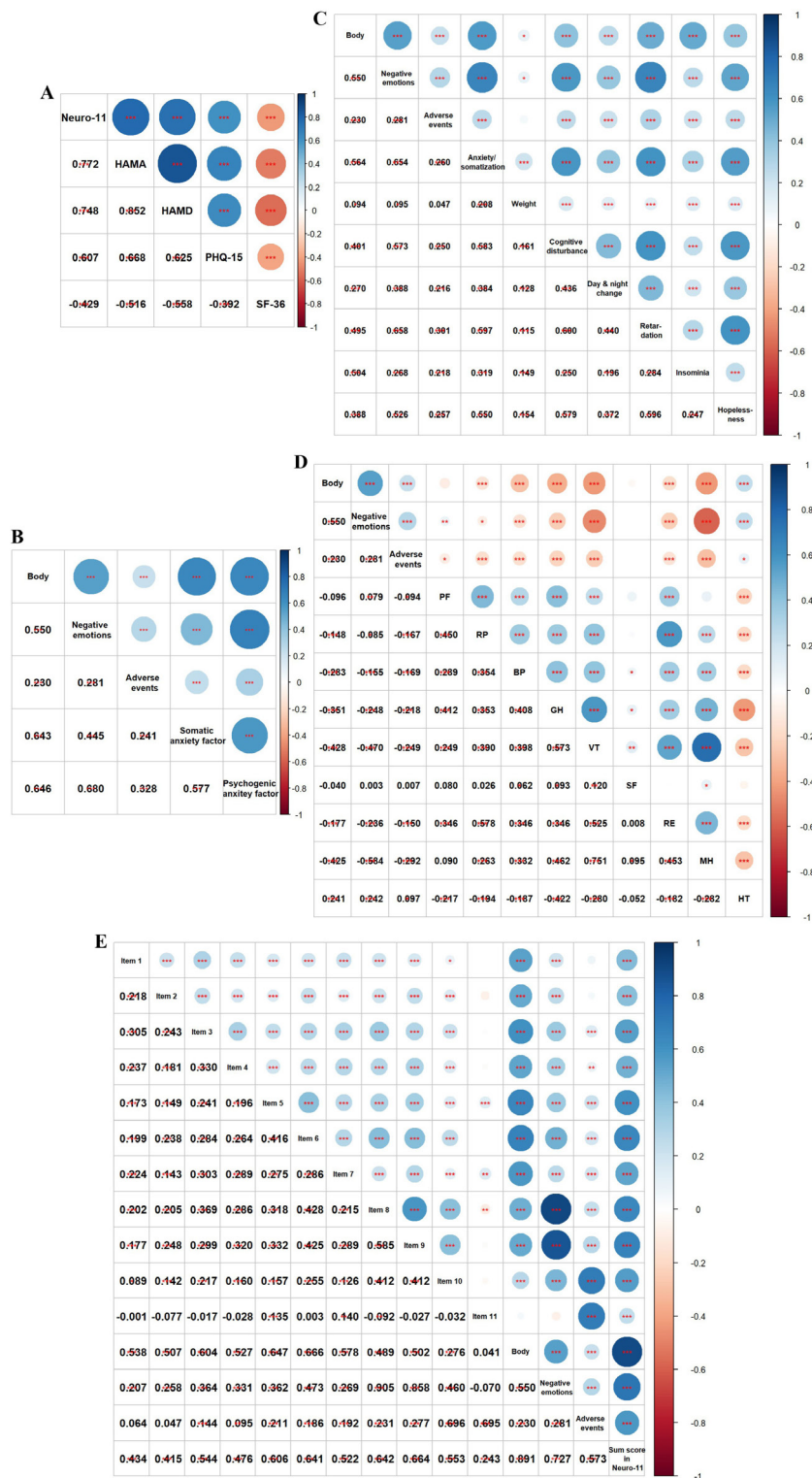


Figure 2 (A) The sum score in Neuro-11 is significantly correlated ($p < 0.001$) with HAMA, HAMD, PHQ-15 and SF-36. The correlation between Neuro-11 with HAMA and HAMD is higher than with the other two questionnaires. (B) The first two dimensions score in Neuro-11 is significantly correlated with the two dimensions of HAMA. The adverse event dimension is also correlated to those two dimensions, but the association is relatively weak ($r < 0.4$). (C) The dimensions of Neuro-11 are all statistically highly correlated with all factors of HAMD, except the weight factor. (D) The dimensions of Neuro-11 have low correlations ($r \approx 0.0$) with PF and SF factors in SF-36. Factors concerning body pain, general health and mental health in SF-36 are highly correlated with Neuro-11 but lowly correlated with PF and SF. (E) The correlation within the items of Neuro-11. Item 11 highly correlates with items 5 and 7 in the body dimension. BP, bodily pain; GH, general health; HAMA, Hamilton Anxiety Scale; HAMD, Hamilton Depression Scale; HT, health transition; MH, Mental Health; Neuro-11, Neuro-11 Neurosis Scale; PF, physical functioning; PHQ-15, Patient Health Questionnaire 15 items; RE, role emotional; RP, role physical; SF, social functioning; SF-36, Short Form 36 Health Survey; VT, Vitality.

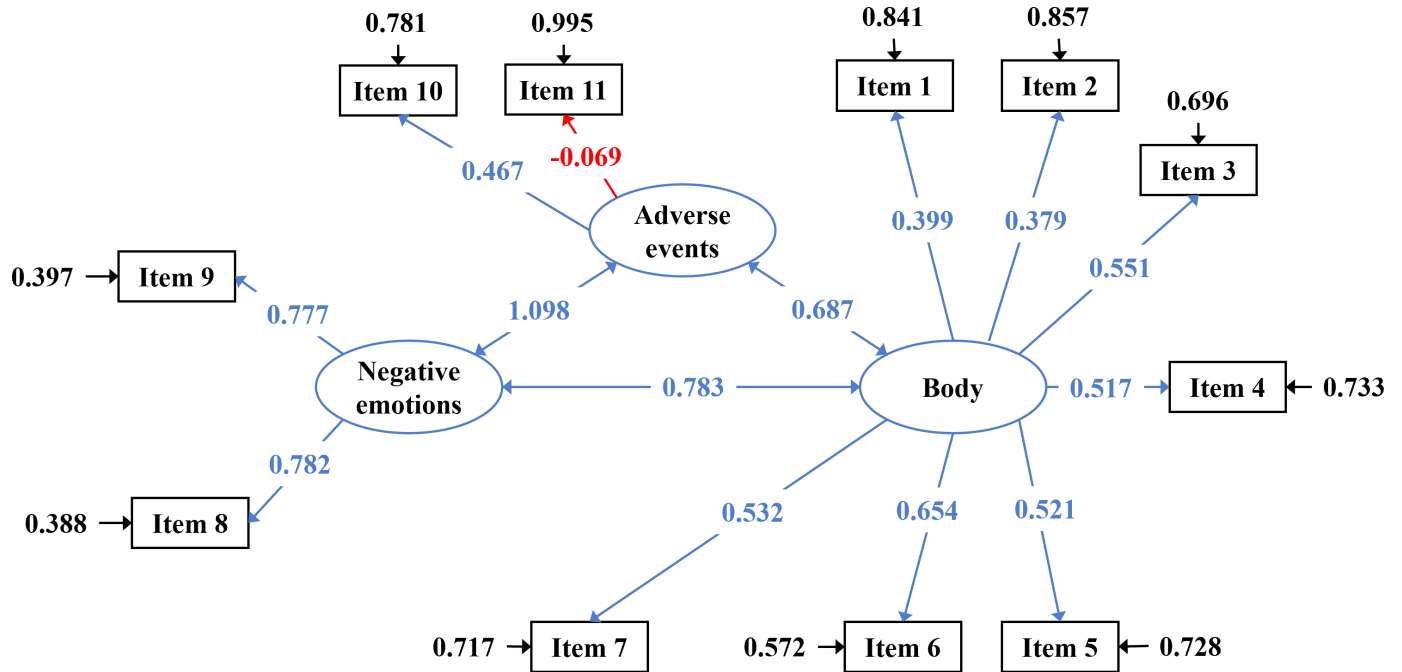


Figure 3 Based on our dataset, the structure diagram of confirmatory factor analysis results of Neuro-11 Neurosis Scale.

scores of 182 patients who had not been diagnosed with SSD. Before we could use the Neuro-11 or other scales to diagnose SSD, we first needed to find a threshold so that patients with scores higher than the threshold could be diagnosed as SSD. We used the gold standard described in the ‘introduction’ section to construct a ROC curve and to calculate Youden’s index.²⁸ The threshold was determined by maximising the index. The threshold of Neuro-11 determined in this way was 10.5 points, 8.5 points for HAMA and 6.5 points for HAMD.

However, in the ‘descriptive analysis’ section, we found that gender, age and marital status had statistically significant effects on neuro-11 sum scores. A subgroup analysis

was performed to examine whether different thresholds of Neuro-11 should be used for the various groups because the features and their interactions may affect the SSD diagnosis. A logistic regression analysis was conducted first to determine the most influential features. We included the demographic characteristics and their interactions as the independent variables and selected the diagnosis of SSD as the response variable. The Akaike information criterion (AIC) was used for feature selection. Our results showed that the most relevant features were gender and age. We then split our data concerning gender and age and conducted the threshold-determining procedure. The results showed that the threshold of Neuro-11

Table 3 Confirmatory factor analysis results of Neuro-11 including the goodness of fit, structure consistency and convergence tests

Dimension (items)	Cronbach’s α coefficient	AVE	CR	Item	Estimate dimension loading	P value (Wald test)
Body (1, 2, ..., 7)	0.71	0.27	0.71	1	0.40	<0.001
Negative emotions (8, 9)	0.76	0.61	0.76	2	0.36	<0.001
Negative events (10, 11)	-0.07	0.11	0.08	3	0.48	<0.001
Whole	0.72			4	0.43	<0.001
Goodness of fit				5	0.59	<0.001
P value (χ^2 test)	<0.001			6	0.66	<0.001
CFI	0.94			7	0.51	<0.001
TLI	0.92			8	0.78	<0.001
RMSEA	0.06			9	0.72	<0.001
SRMR	0.04			10	0.70	0.052
				11	-0.10	0.175

AVE, average variance extracted; CFI, comparative fit index; CR, composite reliability; Neuro-11, Neuro-11 Neurosis Scale; RMSEA, root mean square error of approximation; SRMR, standardised root mean square residual; TLI, Tucker-Lewis index.

Table 4 The results of logistic regression with 10-fold cross-validation

Variables	AUC (95% CI)	Accuracy (95% CI)	Kappa	Sensitivity	Specificity	PPV	NPV
Neuro-11 (dummy by ≥ 10)	0.65 (0.61 to 0.68)	0.64 (0.61 to 0.67)	0.29	0.60	0.71	0.78	0.52
Neuro-11 score	0.67 (0.63 to 0.71)	0.63 (0.59 to 0.66)	0.25	0.64	0.62	0.53	0.72
HAMD score	0.63 (0.59 to 0.67)	0.57 (0.53 to 0.61)	0.13	0.59	0.56	0.39	0.74
HAMA score	0.64 (0.60 to 0.68)	0.60 (0.57 to 0.64)	0.20	0.61	0.60	0.49	0.71
HAMD score+HAMA score	0.64 (0.60 to 0.68)	0.60 (0.57 to 0.64)	0.20	0.62	0.60	0.47	0.73
Neuro-11 (dummy by ≥ 10)+HAMD score+HAMA score	0.66 (0.63 to 0.70)	0.64 (0.61 to 0.68)	0.29	0.60	0.71	0.78	0.52
PHQ-15 score	0.63 (0.59 to 0.67)	0.58 (0.54 to 0.61)	0.15	0.59	0.57	0.44	0.71
HAMD score+HAMA score+PHQ-15 score	0.65 (0.61 to 0.69)	0.61 (0.57 to 0.64)	0.21	0.60	0.62	0.70	0.51
SF-36 score	0.53 (0.49 to 0.57)	0.51 (0.48 to 0.55)	0.01	0.50	0.52	0.24	0.78
Revised Neuro-11 (dummy by ≥ 9)	0.68 (0.64 to 0.71)	0.67 (0.64 to 0.71)	0.35	0.63	0.74	0.79	0.57
Revised Neuro-11 score	0.70 (0.66 to 0.74)	0.64 (0.60 to 0.68)	0.28	0.65	0.63	0.56	0.71

The dependent variable is the whether the patient was diagnosed with SSD or not.

AUC, area under the curve; CI, confidence interval; HAMA, Hamilton Anxiety Scale; HAMD, Hamilton Depression Scale; Neuro-11, Neuro-11 Neurosis Scale; NPV, negative predictive value; PHQ-15, Patient Health Questionnaire 15 items; PPV, positive predictive value; SF-36, Short Form 36 Health Survey; SSD, somatic symptom disorder.

determined in the subgroup of gender was 10.59 and 10.46 for females and males, respectively, and 10.48 and 10.63 for age <45 and ≥ 45 years, respectively. The thresholds in the subgroups showed only slight differences; thus, as no item in Neuro-11 had a decimal point, the threshold of Neuro-11 was set at 10.

We used Neuro-11 with a threshold of 10 to classify whether or not each of the 731 subjects had SSD. For comparison, we also used a logistic regression model with the score of Neuro-11 as a covariate to conduct classification. The performance of the statistical models was investigated by a 10-fold cross-validation. By comparing this simple method with a logistic regression using the Neuro-11 score (denoted by the Neuro-11 score in [table 4](#)) as a covariate, we found the former was even better than the latter in terms of accuracy, kappa, specificity and positive predictive value, although the area under the curve (AUC) was slightly lower. This indicates that using Neuro-11 with a threshold of 10 is a reliable assessment method, and the performance is consistent with the logistic regression model, but we should bear in mind that Neuro-11 is easy to use in practice. According to the confirmatory factor analysis results, item 11 should be treated carefully when Neuro-11 is used to diagnose SSD. We then removed item 11 and used the remaining 10 items (denoted by revised Neuro-11 in [table 4](#)) to conduct the SSD classification. With the revised dataset, the performance was even slightly better. We also compared our Neuro-11 with other questionnaires, including HAMA, HAMD, PHQ-15 and SF-36. The results shown in [table 4](#) were based on a logistic regression model but with each of those scores as a covariate. We found the performance of Neuro-11 was consistently better than the others when they were used to classify SSD. The ROC curve of different approaches is presented in [figure 4](#).

DISCUSSION

Main findings

This study introduces the Neuro-11, a concise self-rating scale comprising 11 items designed to capture the clinical features of somatic symptoms and related disorders across three dimensions: somatic symptoms, negative emotions and adverse events. The items in the Neuro-11 are organised into three sections. The first section focuses on commonly experienced somatic symptoms such as muscle pain, vertigo, headache, palpitations, shortness of breath, loss of appetite, sleep disturbances, fatigue and difficulty concentrating. These symptoms represent manifestations across various bodily systems. Patients in general hospital settings often emphasise the psychological aspects of their somatic symptoms, so including somatic symptoms in the first dimension aims to reflect the multi-system clinical characteristics of SSDs and enhance patient cooperation. The second section of the scale assesses core symptoms of depression and anxiety, including loss of interest, propensity for crying and worry or fear. These two negative emotion items were chosen to be more acceptable to non-psychiatric patients and physicians than scales containing obscure and stigmatising psychiatric symptoms, such as suicidal tendencies found in the HAMD scale. The third section addresses adverse life events and chronic illnesses. Per diagnostic criteria for somatic symptoms and related disorders, these mental disorders are often associated with life-stress events,^{29 30} and the magnitude of adverse events holds significant importance. Previous studies have consistently demonstrated a strong association between neurosis and the occurrence of adverse events. Patients with a history of adverse events are at a significantly higher

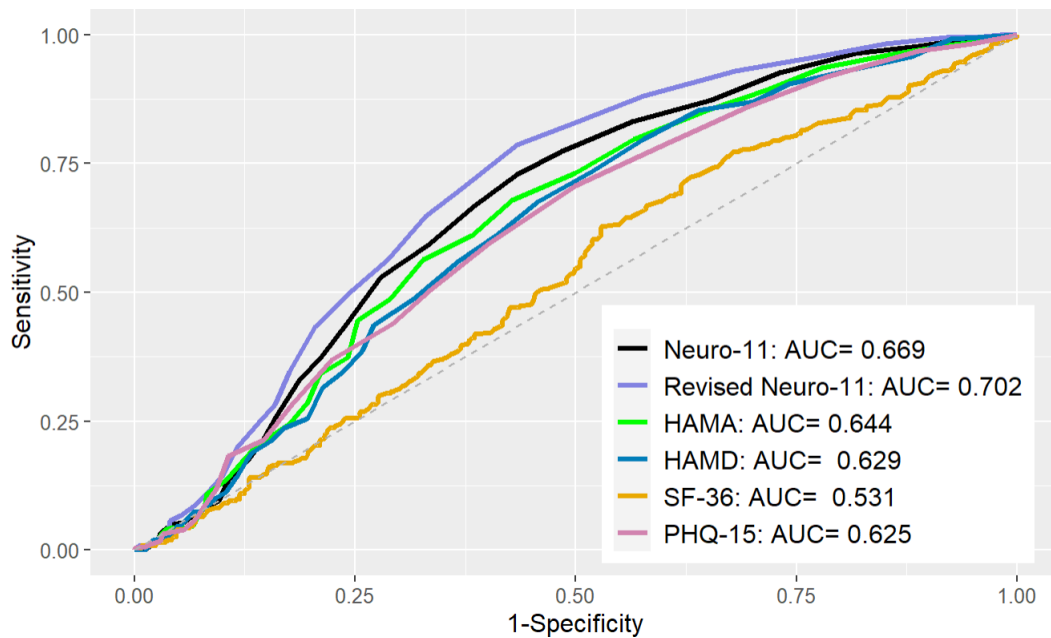


Figure 4 The plot of receiver operating characteristic curves. The AUC for Neuro-11 Neurosis Scale (Neuro-11)—with or without item 11—is the highest compared with other questionnaires. AUC, area under the curve; HAMA, Hamilton Anxiety Scale; HAMD, Hamilton Depression Scale; PHQ-15, Patient Health Questionnaire 15 items; SF-36, Short Form 36 Health Survey.

risk of developing neurosis than normal controls.^{31–35} Furthermore, numerous studies have confirmed the close relationship between somatic symptoms and chronic diseases, including cardiovascular disease,^{36,37} cancer³⁸ and common medical conditions treated in primary care settings, such as migraines and asthma.³⁹ Item 11 in this dimension, related to chronic disease, provides clinicians with important information. The three dimensions of the Neuro-11 exhibit inter-relatedness, as confirmed by confirmatory factor analysis, supporting its theoretical construct. While there may be limited available data regarding the reliability and validity of the adverse event dimension, based on this dataset, it appears that the 11th item concerning chronic diseases did not have statistical effectiveness in diagnosing SSD, potentially due to sample influences on test results. However, considering the extensive literature highlighting the strong association between chronic illness and psychiatric symptoms, we have decided to retain this item for future research using more comprehensive datasets or revised entries.

In the calibration correlation analysis, significant positive correlations were observed between the Neuro-11 and the total scores of HAMA, HAMD and PHQ-15. Likewise, the three dimensions of the Neuro-11 exhibited significant correlations with the corresponding dimensions of the other scales. These findings indicate that the evaluation performance of the Neuro-11 is comparable to that of widely used and effective scales in the field. Notably, the adverse event dimension of the Neuro-11 showed only weak correlations with the two factors of HAMA, suggesting that the adverse event dimension provides unique

and valuable information not captured by HAMA. Regarding the correlation analysis with HAMD, it was found that the correlation between weight change and the Neuro-11 was weaker, implying that weight change may be a manifestation of depressive symptoms with higher specificity compared with SSD.^{40–42} Additionally, the Neuro-11 demonstrated a significant negative correlation with the total score of the SF-36, indicating its ability to evaluate the severity of patients' conditions. Specifically, the energy (Vitality, VT) and mental health (Mental Health, MH) domains of the SF-36 showed strong correlations with the Neuro-11 (>0.40), suggesting that the Neuro-11 scale primarily assesses impairments in energy and mental health domains among patients with SSD. Based on physician diagnosis as the gold standard, a cut-off score of 10 points on the Neuro-11 indicated the presence of SSD. Although gender and age differences were observed in the total score of the Neuro-11, subgroup analysis revealed that these differences had minimal impact on the cut-off value. Therefore, a cut-off score of 10 points was deemed effective for both genders and different age groups. However, it is important to note that the AUC of the Neuro-11 scale was 0.67, which may be considered relatively low compared with our previous study on somatoform disorders, where Neuro-11 achieved an AUC of 0.89. This discrepancy may be attributed to the diagnostic criteria for SSD, which place greater emphasis on psycho-behavioural changes rather than solely the presence of somatic symptoms, leading to more stringent criteria compared with somatoform disorders.⁴³

The PHQ-15 is widely recognised as an effective measure for assessing somatic symptoms.⁴⁴ In our study, Neuro-11 demonstrated the highest AUC when compared with HAMA, HAMD and PHQ-15 concurrently in diagnosing SSD. Furthermore, categorical logistic regression models using Neuro-11 scores as variables indicated that Neuro-11 was more effective in diagnosing SSD compared with using HAMD, HAMA, PHQ-15 alone, or a combination of HAMD, HAMA and PHQ-15. These findings suggest that the Neuro-11 can replace multiple scales in screening patients with SSD in general hospital settings, resulting in significant time, labour and cost savings. It is worth noting that the sensitivity and specificity of the HAMA and HAMD scales were also analysed to determine their optimal cut-off values, with approximately 8 points identified for the HAMA and HAMD scales in this population (results not shown). This implies that the SSD population in our study had only mild levels of anxiety and depression. While it is recommended to include anxiety and depression scales in the assessment of patients with somatic complaints to avoid false negatives,⁴⁵ there remains a need to develop new, applicable scales in line with the evolving diagnostic culture and disease understanding. Nonetheless, the HAMA and HAMD scales are still widely used in China.⁴⁶ In terms of administration, the Neuro-11 offers the advantages of simplicity and ease of implementation as it is self-assessed, requiring a brief evaluation time of approximately 2–5 min. These characteristics make it highly suitable for use in general hospital settings.

Limitations

This study has several limitations that should be acknowledged. First, the sample used in this research was drawn from a single centre, which may limit the generalisability of the findings. Further investigation with larger and more diverse samples across multiple centres is needed to validate the reliability and validity of the Neuro-11. Second, the design concept of the Neuro-11 primarily focuses on the characteristics of Chinese patients in general hospital settings, who often exhibit prominent physical symptoms and relatively weaker psychological symptoms. However, previous studies have indicated that there may be psychometric differences among populations in different countries.^{20 47} Therefore, it is important to conduct revalidation studies to ensure the generalisability of the Neuro-11 to diverse populations in various countries. Third, as the Neuro-11 scale is a self-rating scale that relies on the recall of symptom duration and frequency for evaluation, the potential for recall bias should be considered. Lastly, when evaluating changes in illness severity, it is important to consider the impact of objective records of adverse events in the third dimension on the total score, warranting further investigation.

Implications

The Neuro-11, as a concise and efficient neuropsychological scale, offers the advantage of capturing the psychological state of patients while minimising the time required for face-to-face consultations. This feature enhances the comprehensive diagnosis and treatment of patients and is well-received by both physicians and patients in general hospital settings. Consequently, using the Neuro-11 in diagnosing and treating individuals with somatic symptoms would provide valuable support. The Neuro-11 has the potential to become the preferred questionnaire for physicians in general hospitals, serving as a diagnostic tool for SSD or for routine screening purposes.

Author affiliations

¹Department of Neurology, Shenzhen People's Hospital, The second Affiliated Hospitals of Jinan University, The first Affiliated Hospitals of Southern University of Science and Technology, Shenzhen, Guangdong, China

²Jinan University, Guangzhou, Guangdong, China

³Department of Statistics and Data Science, Southern University of Science and Technology, Shenzhen, Guangdong, China

⁴Institute of Neurological Diseases, Shenzhen Bay Laboratory, Shenzhen, Guangdong, China

⁵Department of neurology, The Johns Hopkins University School of Medicine, Baltimore, Maryland, USA

⁶Department of Geriatrics, Shenzhen People's Hospital, The second Affiliated Hospitals of Jinan University, The first Affiliated Hospitals of Southern University of Science and Technology, Shenzhen, China

⁷NHC Key Laboratory of Diagnosis and Treatment on First Affiliated Hospital of Chongqing Medical University, Chongqing, Chongqing, China

⁸National Center for Applied Mathematics, Shenzhen, Guangdong, China

Contributors SZ: formulation or evolution of overarching research goals and aims; development or design of methodology; reproducibility of results; writing the initial draft. YY: application of statistical; data presentation. SL: revision; editing. SZ: experiments and other research outputs. XS: data collection. GD: data collection. YL: data collection; editing. ZC: data collection. SC: data collection. YH: data collection. XJ: data collection. CC: data collection. LY: data collection. PX: editing. JS: editing. QG: editing. RHL: editing. YG: guarantor; ideas; formulation or evolution of overarching research goals and aims; acquisition of the financial support for the project leading to this publication.

Funding This research was supported by the following funds: Shenzhen Science and Technology Innovation Commission (KCFZ20201221173400001; KCXFZ20201221173411032; SGDX20210823103805042); Natural Science Fund of Guangdong Province (2021A1515010983); Shenzhen Key Medical Discipline Construction Fund (no. SZXK005).

Competing interests None declared.

Patient consent for publication Not applicable.

Ethics approval This study was approved by Shenzhen People's Hospital Research Ethics Committee (no. KY-LL-2021564-01). Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available on reasonable request.

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 iD

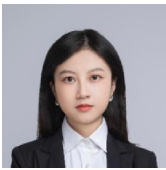
Shan Lu <http://orcid.org/0000-0001-7297-5649>

REFERENCES

- American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*. 5th edn. American Psychiatric Association, 2013.
- Scamvougeras A, Howard A. Somatic symptom disorder, medically unexplained symptoms, somatoform disorders, functional neurological disorder: how DSM-5 got it wrong. *Can J Psychiatry* 2020;65:301–5.
- de Waal MWM, Arnold IA, Eekhof JAH, *et al*. Somatoform disorders in general practice: prevalence, functional impairment and comorbidity with anxiety and depressive disorders. *Br J Psychiatry* 2004;184:470–6.
- Cao J, Wei J, Fritzsche K, *et al*. Prevalence of DSM-5 somatic symptom disorder in Chinese outpatients from general hospital care. *Gen Hosp Psychiatry* 2020;62:63–71.
- Schaefer R, Höner C, Salm F, *et al*. Psychological and behavioral variables associated with the somatic symptom severity of general hospital outpatients in China. *Gen Hosp Psychiatry* 2013;35:297–303.
- Xiong N, Zhang Y, Wei J, *et al*. Operationalization of diagnostic criteria of DSM-5 somatic symptom disorders. *BMC Psychiatry* 2017;17:361.
- Cao J, Wei J, Fritzsche K, *et al*. Detecting DSM-somatic symptom disorder in general hospitals in China: B-criteria instrument has better accuracy—A secondary analysis. *Front Psychiatry* 2022;13:935597.
- Ou H. The historical changes and research progress of somatic symptom and related disorders. *Chin J Psychiatry* 2021;54:66–70.
- Dunphy L, Penna M, El-Kafsi J. Somatic symptom disorder: a diagnostic dilemma. *BMJ Case Rep* 2019;12:e231550.
- Huang W-L, Chang S-S, Wu S-C, *et al*. Population-based prevalence of somatic symptom disorder and comorbid depression and anxiety in Taiwan. *Asian J Psychiatr* 2023;79:103382.
- Löwe B, Levenson J, Depping M, *et al*. Somatic symptom disorder: a scoping review on the empirical evidence of a new diagnosis. *Psychol Med* 2022;52:632–48.
- Vaccarino AL, Sills TL, Evans KR, *et al*. Prevalence and association of somatic symptoms in patients with major depressive disorder. *J Affect Disord* 2008;110:270–6.
- Li X, Zhang H, Han X, *et al*. Predictive potential of somatic symptoms for the identification of subthreshold depression and major depressive disorder in primary care settings. *Front Psychiatry* 2023;14:999047.
- Wang J, Guo W-J, Mo L-L, *et al*. Prevalence and strong association of high somatic symptom severity with depression and anxiety in a Chinese inpatient population. *Asia Pac Psychiatry* 2017;9.
- Kroenke K. Patients presenting with somatic complaints: epidemiology, psychiatric co-morbidity and management. *Int J Methods Psychiatr Res* 2003;12:34–43.
- Garnefski N, van Rood Y, de Roos C, *et al*. Relationships between traumatic life events, cognitive emotion regulation strategies, and somatic complaints. *J Clin Psychol Med Settings* 2017;24:144–51.
- Reddy B, Chaturvedi SK, Desai G, *et al*. Health-related life events in patients with somatic symptom disorders: a case control study. *Int J Soc Psychiatry* 2019;65:265–70.
- Zeng S, Huang P, Cai Z, *et al*. A novel self-report screening scale to predict somatoform disorders. *Chinese Journal of Stroke* 2020;15:468–79.
- 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.
- Zhang L, Fritzsche K, Liu Y, *et al*. Validation of the Chinese version of the PHQ-15 in a tertiary hospital. *BMC Psychiatry* 2016;16:89.
- Toussaint A, Hüsing P, Kohlmann S, *et al*. Detecting DSM-5 somatic symptom disorder: criterion validity of the patient health Questionnaire-15 (PHQ-15) and the somatic symptom Scale-8 (SSS-8) in combination with the somatic symptom disorder-B criteria scale (SSD-12). *Psychol Med* 2020;50:324–33.
- Lin Y, Yu Y, Zeng J, *et al*. Comparing the reliability and validity of the SF-36 and SF-12 in measuring quality of life among adolescents in China: a large sample cross-sectional study. *Health Qual Life Outcomes* 2020;18:360.
- Scott KM, Tobias MI, Sarfati D, *et al*. SF-36 health survey reliability, validity and norms for New Zealand. *Aust N Z J Public Health* 1999;23:401–6.
- Fava GA, Kellner R, Munari F, *et al*. The hamilton depression rating scale in normals and depressives. *Acta Psychiatr Scand* 1982;66:26–32.
- Maier W, Buller R, Philipp M, *et al*. The hamilton anxiety scale: reliability, validity and sensitivity to change in anxiety and depressive disorders. *J Affect Disord* 1988;14:61–8.
- Markus KA. Principles and practice of structural equation modeling. In: *Kline RB*. Taylor & Francis, 2012.
- Fornell C, Larcker DF. Evaluating structural equation models with unobservable variables and measurement error. *J Mark Res* 1981;18:39–50.
- Fluss R, Faraggi D, Reiser B. Estimation of the Youden index and its associated cutoff point. *Biom J* 2005;47:458–72.
- Saunders NR, Gandhi S, Chen S, *et al*. Health care use and costs of children, adolescents, and young adults with somatic symptom and related disorders. *JAMA Netw Open* 2020;3:e2011295.
- Barsky AJ, Orav EJ, Bates DW. Somatization increases medical utilization and costs independent of psychiatric and medical comorbidity. *Arch Gen Psychiatry* 2005;62:903–10.
- von dem Knesebeck O, Lehmann M, Löwe B, *et al*. Causal attributions for somatic symptom disorder. *J Psychosom Res* 2020;129:109910.
- Fairbrass KM, Costantino SJ, Gracie DJ, *et al*. Prevalence of irritable bowel syndrome-type symptoms in patients with inflammatory bowel disease in remission: a systematic review and meta-analysis. *Lancet Gastroenterol Hepatol* 2020;5:1053–62.
- Hickie I, Davenport T, Wakefield D, *et al*. Post-infective and chronic fatigue syndromes precipitated by viral and non-viral pathogens: prospective cohort study. *BMJ* 2006;333:575.
- Momsen AH, Nielsen CV, Nielsen MBD, *et al*. Work participation and health-related characteristics of sickness absence beneficiaries with multiple somatic symptoms. *Public Health* 2016;133:75–82.
- Bonvanie IJ, Janssens KAM, Rosmalen JGM, *et al*. Life events and functional somatic symptoms: a population study in older adolescents. *Br J Psychol* 2017;108:318–33.
- Kohlmann S, Gierk B, Hümmelgen M, *et al*. Somatic symptoms in patients with coronary heart disease: prevalence, risk factors, and quality of life. *JAMA Intern Med* 2013;173:1469–71.
- Kop WJ, Synowski SJ, Gottlieb SS. Depression in heart failure: biobehavioral mechanisms. *Heart Fail Clin* 2011;7:23–38.
- Teunissen S, Wesker W, Kruitwagen C, *et al*. Symptom prevalence in patients with incurable cancer: a systematic review. *J Pain Symptom Manage* 2007;34:94–104.
- Kop WJ, Toussaint A, Mols F, *et al*. Somatic symptom disorder in the general population: associations with medical status and health care utilization using the SSD-12. *Gen Hosp Psychiatry* 2019;56:36–41.
- Jung SJ, Woo H-T, Cho S, *et al*. Association between body size, weight change and depression: systematic review and meta-analysis. *Br J Psychiatry* 2017;211:14–21.
- Weissenburger J, Rush AJ, Giles DE, *et al*. Weight change in depression. *Psychiatry Res* 1986;17:275–83.
- Carter FA, Bulik CM, Joyce PR. Direction of weight change in depression. *J Affect Disord* 1994;30:57–60.
- Claassen-van Dessel N, van der Wouden JC, Dekker J, *et al*. Clinical value of DSM IV and DSM 5 criteria for diagnosing the most prevalent somatoform disorders in patients with medically unexplained physical symptoms (MUPS). *J Psychosom Res* 2016;82:4–10.
- 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.
- Bagby RM, Ryder AG, Schuller DR, *et al*. The hamilton depression rating scale: has the gold standard become a lead weight?. *Am J Psychiatry* 2004;161:2163–77.
- Fu C-W, Tan AW, Sheng F, *et al*. The prevalence of anxiety symptoms and depressive symptoms in patients with somatic disorders in urban China: a multi-center cross-sectional study. *Int J Psychiatry Med* 2007;37:185–99.
- Horton M, Pery AE. Screening for depression in primary care: a rasch analysis of the PHQ-9. *BJPsych Bull* 2016;40:237–43.



Silin Zeng, Master of Medicine, graduated from the Department of Neurology, Jinan University School of Clinical Medicine in China in 2015 and was a visiting scholar at the Institute of Psychiatry and Behavior, Stanford University, in 2018. Since 2017, she has been attending a doctoral program in the Department of Neurology at Shenzhen People's Hospital. Her main research interests include diagnosing and treating neuropsychiatric diseases, developing physical symptom scales, researching brain network mechanisms of neuropsychiatric disorders, and exploring non-invasive neuromodulation. She has published several related papers, participated in the writing of the book "Application of TMS in the Diagnosis and Treatment of Nervous System Diseases" in 2021, obtained a patent for a psychological analysis software system in 2022, and participated in several clinical studies on the diagnosis and treatment of neuropsychiatric disorders and cerebrovascular disease using high-density electroencephalogram and transcranial magnetic stimulation.



Yian Yu received a bachelor's degree in statistics from the Department of Statistics and Data Science of Southern University of Science and Technology, China in 2020. She was recommended to continue to study for a doctorate. She is now in the third year of a PhD in statistics. Her main research interests include statistical modelling for tackling real data problems, especially in the framework of functional data. She has published one paper in Statistics in Medicine as the first author.