Journal of Psychosomatic Research 74 (2013) 459–468

Contents lists available at SciVerse ScienceDirect

Journal of Psychosomatic Research

Review

How to assess common somatic symptoms in large-scale studies: A systematic review of questionnaires ☆

Wilma L. Zijlema [a,b], Ronald P. Stolk [a], Bernd Löwe [d], Winfried Rief [e], BioSHaRE, Peter D. White [c], Judith G.M. Rosmalen [b,]*

[a] Department of Epidemiology, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands
[b] Interdisciplinary Center Psychopathology and Emotion regulation (ICPE), University of Groningen, University Medical Center Groningen, Groningen, the Netherlands
[c] Wolfson Institute of Preventive Medicine, Barts and the London Medical School, Queen Mary University of London, London, United Kingdom
[d] Department of Psychosomatic Medicine and Psychotherapy, University Medical Center Hamburg-Eppendorf and Schön Klinik Hamburg-Eilbek, Hamburg, Germany
[e] Department of Clinical Psychology and Psychotherapy, University of Marburg, Marburg, Germany

ARTICLE INFO

Article history:
Received 21 January 2013
Received in revised form 22 March 2013
Accepted 23 March 2013

Keywords:
Common somatic symptoms
Functional somatic symptoms
Large-scale studies
Questionnaires
Reliability
Validity

ABSTRACT

Objective: Many questionnaires for assessment of common somatic symptoms or functional somatic symptoms are available and their use differs greatly among studies. The prevalence and incidence of symptoms are partially determined by the methods used to assess them. As a result, comparison across studies is difficult. This article describes a systematic review of self-report questionnaires for somatic symptoms for use in large-scale studies and recommends two questionnaires for use in such studies.
Methods: A literature search was performed in the databases Medline, PsycINFO and EMBASE. Articles that reported the development, evaluation, or review of a self-report somatic symptom measure were included. Instrument evaluation was based on validity and reliability, and their fitness for purpose in large scale studies, according to the PhenX criteria.
Results: The literature search identified 40 questionnaires. The number of items within the questionnaires ranged from 5 to 78 items. In 70% of the questionnaires, headaches were included, followed by nausea/upset stomach (65%), shortness of breath/breathing trouble (58%), dizziness (55%), and (low) back pain/backaches (55%). Data on validity and reliability were reported and used for evaluation.
Conclusion: Questionnaires varied regarding usability and burden to participants, and relevance to a variety of populations and regions. Based on our criteria, the Patient Health Questionnaire-15 and the Symptom Checklist-90 somatization scale seem the most fit for purpose for use in large-scale studies. These two questionnaires have well-established psychometric properties, contain relevant symptoms, are relatively short, and are available in multiple languages.

Introduction

A symptom is a self-reported bodily sensation or mental experience that is perceived by a person as a change from normal health [1]. A British study showed that the mean number of symptoms reported in the general population is between three and four in the last two weeks [2] and a recent Norwegian study found a strong association between the number of symptoms and functional status [3]. Symptoms may be mediated by a change in bodily function or may be associated with disease. It is commonly the case that a symptom cannot be conclusively explained by organic pathology. These symptoms are referred to as functional somatic symptoms (FSS). The more symptoms reported, the more likely symptoms are functional in nature [4,5]. Patients who frequently complain of physical symptoms that either lack a demonstrable organic basis, or that are judged to be in excess of what would be expected based on medical findings, are thought to be suffering from the process of somatization. This view is qualified by the knowledge that future research may find medical explanations for some of these FSS. Somatization refers to a tendency to experience and communicate somatic distress in response to psychosocial stress and seek medical help for it [6]. Functional somatic symptoms are common [7–9], disabling [10–12], and costly [13], and patients often feel misunderstood, guilty and even ashamed [14].

The occurrence of symptoms is often assessed by using a self-report symptom questionnaire. Many different questionnaires are available and the use of these questionnaires differs greatly among current

---

☆ Institution where the work was performed: Interdisciplinary Center Psychopathology and Emotion regulation (ICPE), University of Groningen, University Medical Center Groningen, P.O. Box 30.001, 9700 RB, Groningen, the Netherlands.

* Corresponding author at: Interdisciplinary Center Psychopathology and Emotion regulation (ICPE), University of Groningen, University Medical Center Groningen, P.O. Box 30.001, 9700 RB, Groningen, the Netherlands. Tel.: +31 50 3611242; fax: +31 50 3619722.
E-mail address: j.g.m.rosmalen@umcg.nl (J.G.M. Rosmalen).

0022-3999/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jpsychores.2013.03.093

studies, complicating the comparison of studies. Firstly, the questionnaires differ greatly in the number of symptoms questioned. Secondly, there is a large variety in the type of symptoms included in the questionnaires. Previous studies have suggested that certain types of symptoms cluster together [15–17]. Although not found in all studies, the following four clusters are commonly reported: cardiopulmonary (including autonomic symptoms), gastrointestinal, musculoskeletal, and general symptoms. Thirdly, some of these questionnaires assess symptoms in general, while others focus on medically unexplained symptoms. Being certain that a symptom is medically unexplained can be difficult, has low inter-rater reliability [18,19], and would be impossible to assess in large-scale studies. Fourthly, the time frame of assessment varies largely. Some questionnaires are based on life-time symptoms; but several researchers suggest that recall of lifetime symptoms is unreliable and inconsistent [20–22]. Others address time frames of between a week and a month. Fifthly, some questionnaires inquire only about the categorical presence of symptoms, while others inquire about symptom severity; both symptom diversity and severity may be important [9].

Given the heterogeneity of scales with respect to content, scaling, and dimensionality, severity scores from different FSS or somatic symptom scales are incomparable. Using cut-off-scores, each scale identifies a unique subgroup that differs in various aspects from subgroups identified by other scales. It would be very useful to have a gold standard measure for the assessment of both FSS and somatic symptoms in general in large-scale studies. Agreement regarding the use of such an instrument can facilitate systematic comparisons or meta-analytical studies and thereby contribute to the understanding of the etiology of functional somatic symptoms.

To the best of our knowledge, an overview of the currently available symptom questionnaires has not been reported. This article describes a systematic review of self-report questionnaires for common somatic symptoms for use in large-scale studies. It will conclude with a recommendation for which symptom questionnaires are the best to use. This recommendation will be based on validity and reliability on the one hand, and applicability in large-scale studies on the other hand. The first aspects will be evaluated in terms of number and types of symptoms included, the response scale, time frame covered, and data on validity and reliability. The second aspects will be evaluated using the PhenX (Phenotypes and eXposures) criteria, according to which a measure should be well-established, easy in its use, of low burden to participants, relevant for future use, and applicable to a variety of populations and regions [23].

Methods

Search strategy

A literature search was performed in the databases Medline, EMBASE, and PsycINFO on the 15th of October 2012. A search term was formulated for searching the databases, which contained a combination of somatoform disorder or synonyms and questionnaire or synonyms and symptoms. For Medline, the following search term was used: (“somatoform disorders/classification” [MeSH Major Topic] OR “somatoform disorders/diagnosis” [MeSH Major Topic] OR “somatoform disorders/epidemiology” [MeSH Major Topic] OR “functional somatic symptoms” [Title/Abstract]) AND (questionnaire [Title/Abstract] OR screen* [Title/Abstract] OR “self report” [Title/Abstract] OR “index” [Title/Abstract]) AND symptoms. For EMBASE and PsycINFO, comparable search terms were used. The search was conducted without language restrictions. Personal files of the authors were also reviewed for relevant articles. Additional searches were performed using the search engine Google and the authors of the questionnaires were contacted to obtain supplemental information of the symptom questionnaires, if needed.

Screening and selection procedure

The titles and abstracts of the retrieved articles were screened by two independent researchers. The articles were chosen for the development, evaluation, or review of somatic symptom or somatization questionnaires. Additionally, the measure had to be a self-report symptom checklist. Interviews were excluded, as these are not suitable for use in large studies. Furthermore, the questionnaires chosen had to include symptoms from more than one symptom cluster; not just symptoms of the gastrointestinal tract or cardiopulmonary system. When the symptom questionnaire was a sub-scale derived from a larger questionnaire, the symptom subscale had to have been separately validated and used. There were no criteria for the target population of the questionnaire. Discrepancies between the two researchers were resolved by consensus. Full articles were then obtained for all included studies. Based on the full text, articles that still fulfilled the inclusion criteria were included in the review.

Data extraction

Authors, year of publication, name of questionnaire, purpose of questionnaire, questionnaire instructions, list of symptoms, answering scale, and language of questionnaire, were extracted for every questionnaire. Data extraction from papers describing the validation of a questionnaire also included validity data, characteristics about the population used (clinical or general, gender and age distribution, nationality or race), and number of participants. Data extraction from articles written in languages other than English was done by native speakers.

Instrument evaluation

The questionnaires were evaluated according to the following criteria; the first set of criteria concerned the validity and reliability of the instrument. Firstly, we examined the type of symptoms included. We assumed that the proportion of questionnaires including a specific symptom reflected expert knowledge on the importance of that specific symptom for the underlying construct. To ensure that the questionnaire was not too restrictive with regard to the type of symptom, we evaluated whether the questionnaires included at least one symptom from each of the following symptom clusters identified in previous studies [15–17]: cardiopulmonary (including autonomic symptoms), gastrointestinal, musculoskeletal, and general symptoms. Since being certain that a symptom is medically unexplained can be difficult, has low inter-rater reliability [18,19], and would be impossible to assess in large-scale studies, we proposed that self-report symptom questionnaires should preferably question symptoms in general, as opposed to medically unexplained symptoms. Secondly, we noted the time frame covered by the questionnaire. Studies have shown that the recall of lifetime symptoms is unreliable [20–22]. We therefore do not recommend the use of lifetime as a recall period. Thirdly, we assessed the response scale. We looked at the format of the questionnaires' responses: severity, frequency, and the number of response categories. Fourthly, we recorded the psychometric characteristics. The internal consistency was assessed, which reflects whether items in a questionnaire are correlated, thus measuring the same concept. The factor structure of a questionnaire indicates which symptom clusters are present within the questionnaire. Correlations of the symptom questionnaire with other related constructs, for example health anxiety and illness behavior, also inform the validity of the questionnaire. In addition, the test–retest reliability of a questionnaire is an important indicator for the stability of the questionnaire. It is likely that symptom reporting fluctuates over time. Therefore, short time intervals for test–retest reliability would be most appropriate, and an interval of 3–4 weeks is commonly used [24]. We therefore chose to report the test–retest reliability for time intervals no longer than 1 month.