Somatic symptom count scores do not identify patients with symptoms unexplained by disease: a prospective cohort study of neurology outpatients

Alan J Carson,1,2 Jon Stone,2 Christian Holm Hansen,3 Rod Duncan,4 Jonathon Cavanagh,5 Keith Matthews,6 G Murray,7 Michael Sharpe8

ABSTRACT

Objective Somatic symptoms unexplained by disease are common in all medical settings. The process of identifying such patients requires a clinical assessment often supported by clinical tests. Such assessments are time-consuming and expensive. Consequently the observation that such patients tend to report a greater number of symptoms has led to the use of self-rated somatic symptom counts as a simpler and cheaper diagnostic aid and proxy measure for epidemiological surveys. However, despite their increasing popularity there is little evidence to support their validity.

Methods We tested the score on a commonly used self-rated symptom questionnaire—the Patient Health Questionnaire (PHQ 15) (plus enhanced iterations including an additional 10 items on specific neurological symptoms and an additional 5 items on mental state) for diagnostic sensitivity and specificity against a medical assessment (with 18 months follow-up) in a prospective cohort study of 3781 newly attending patients at neurology clinics in Scotland, UK.

Results We found 1144/3781 new outpatients had symptoms that were unexplained by disease. The patients with symptoms unexplained by disease reported higher symptoms count scores (PHQ 15: 5.6 (95% CI 5.4 to 5.8) vs 4.2 (4.1 to 4.4) p<0.0001). However, the PHQ15 performed little better than chance in its ability to identify patients with symptoms unexplained by disease. The findings with the enhanced scales were similar.

Conclusions Self-rated symptom count scores should not be used to identify patients with symptoms unexplained by disease.

INTRODUCTION

Patients who present with somatic symptoms unexplained by identifiable organic disease are commonly encountered in all medical settings. Despite the lack of a disease explanation these patients are frequently, disabled and distressed.1 The symptoms are referred to by many names including functional symptoms/disorders, medically unexplained symptoms, somatisation and somatoform symptoms; in this paper we will use the term ‘symptoms unexplained by disease (SUD)’ as it maps most closely to the methodology used in our study.

Patients who have symptoms unexplained by disease often have multiple somatic symptoms. This multiplicity has been recommended as being a useful clue to making the diagnosis.2 Hence it has been proposed that high scores on self-rating scales of somatic symptoms will identify patients likely to have symptoms unexplained by disease. A widely used example of such a scale is the Patient Health Questionnaire (PHQ 15); higher numbers of somatic symptoms on this scale have been consistently found to correlate with poorer outcomes, higher healthcare use and higher rates of depression and anxiety.3

The PHQ15 was initially introduced as diagnostic screening test for symptoms unexplained by disease.4 Later it was suggested that it performed better as “a measure of somatic symptom severity rather than a diagnostic instrument”.5 However, despite this caution from its original authors it is still being widely used as a diagnostic screening test (see table 1).

It is therefore important to know if a high count on a self-rated somatic symptom scale is valid as an indicator of symptoms unexplained by disease. If it is, this would simplify clinical practice and epidemiological research because self-rated scales are quick, cheap and easy to use and avoid the need for a specialist medical assessment. However if it isn’t, the use of such scales may lead to misdiagnosis as well as to misleading prevalence data. A review of the published literature indicates that we have little data to determine the validity of a self-rated somatic symptom scale when compared with the gold standard of a specialist medical examination.6 7 8 27 34

We therefore set out to test the diagnostic accuracy of several self-rated symptom scales using the data set derived from The Scottish Neurological Symptoms Study (SNSS).1 This was a prospective cohort study of 3781 newly attending patients to neurology clinics. All patients were medically assessed and had an 18 month follow-up to ensure diagnostic accuracy.45 A third (1144/3781) of these patients were judged to have symptoms that were either ‘not at all’ or only ‘somewhat’ explained by disease.

The aims of this paper are to: (A) describe the relative prevalence of different somatic symptoms in the sample; (B) test the hypothesis that neurology patients who in the opinion of the assessing physician had symptoms unexplained by disease reported a greater number of symptoms than patients whose symptoms were regarded as ‘explained by disease’; (C) to assess the validity of the total symptom count score on self-rated scales as a method of discriminating between these groups.
METHODS

The SNSS study was a prospective, multicentre study of a representative cohort of consecutive neurology outpatients. Ethical approval for the study was granted by a Multi-centre Research Ethics Committee.

Participating clinics

The majority of neurological services in Scotland are provided by a publically funded National Health Service. Thirty-six of the 38 consultant neurologists working in Scotland participated. Patients were recruited from their general neurology clinics (including their supervised trainee clinics) in the main Scottish neurological centers—Aberdeen, Dundee, Edinburgh and Glasgow and some of their associated peripheral clinics in Airdrie, East Kilbride, Falkirk, Inverness, Perth, Stirling, Vale of Leven and Wishaw—in the period December 2002 to February 2004. The clinics sampled took mainly general practice referrals with patients allocated by medical records staff according to availability of appointment. Tertiary clinics, where patients required a specific diagnosis to attend (such as acute neurovascular and multiple sclerosis clinics) were excluded as were ‘urgent case’ emergency clinics.

Patients

All newly referred patients at the participating neurology outpatient clinics were potentially eligible for inclusion. The exclusion criteria were: age less than 16 years, cognitive or physical impairment of a degree that precluded informed consent, inability to read English, or if the neurologist identified the patient as unsuitable for the study (eg, too distressed, terminally ill). New patients included patients with existing neurological diagnoses who had been referred again from primary care. After complete description of the study to the subjects, written informed consent was obtained.

Procedure

Patients were sent information about the study prior to their appointment with the neurologist. After the consultation they were invited by their neurologist to speak to a research assistant. Consent was obtained from patients willing to participate.

Measures

Physician reported

The neurologists were asked in each case to address the question: ‘To what extent do you think this patient’s clinical symptoms are explained by organic disease?’ Responses were made on a 4-point Likert-type scale: ‘not at all’, ‘somewhat’, ‘largely’ or ‘completely’. Operational criteria were provided to guide the neurologists’ ratings.

Patient reported

A questionnaire was administered immediately after the initial consultation and collected data on demographics and symptoms. Symptoms were measured using the PHQ-15 checklist of the 15 most common physical symptoms that patients present to primary care (excluding upper respiratory tract infections) and with the sexual and menstrual items removed to leave 13 items. We used the original dichotomous version, which asked patients if they had ‘been bothered a lot’ by each of the symptoms over the last month. The scale items referred to stomach pain; back pain; pain in your arms, legs or joints; headaches; chest pain; dizziness; fainting spells; feeling your heart pound or race; shortness of breath; constipation, loose bowels or diarrhoea; nausea, gas or indigestion; feeling tired or having low energy; trouble sleeping. The choice of the original version was pragmatic as we already had pilot data on its performance which predated the release of the updated version in 2002.

In order to see if the inclusion of additional neurological symptoms made a difference we made a second list based on the 10 most common ‘neurological’ symptoms. These are paralysis or weakness; double or blurred vision; difficulty swallowing or lump in throat; difficulty speaking or slurred speech; loss of sensation, numbness or tingling; problems with your memory or concentration; partial or total loss of vision; partial or total loss of hearing; lack of co-ordination or balance; seizure or fit.

Finally, we created a third list of psychological symptoms taken from the Prime MD Questionnaire to see if adding psychological symptoms to the scale improved its accuracy. The scale items were: worrying about a lot of different things; ‘nerves’ or feeling anxious or on edge; feeling down, depressed or hopeless; little interest or pleasure in doing things; an anxiety attack (suddenly feeling fear or panic).

Analysis

In order to maximise the numbers of patients in each category for analysis we dichotomised the neurologists’ ratings of symptoms into ‘unexplained’ (‘not at all’ or only ‘somewhat’ explained by disease) and ‘explained’ (‘largely’ and ‘completely’ explained by disease).

We then compared these groups on: (A) frequency of each individual symptom and (B) three different symptom count measures using (1) the standard PHQ 15 somatic symptoms, (2) the PHQ 15 somatic symptoms plus the additional neurological symptoms, (3) the PHQ 15 somatic symptoms plus the additional neurological plus the psychological symptoms.

Then, in order to determine the ability of symptom counts to discriminate between patients with ‘unexplained symptoms’ and ‘explained symptoms’ we plotted receiver operating characteristic (ROC) curves for each.

Finally, in order to see if the amalgamation of the neurologists’ ratings into two categories had blurred differences between the individual categories, we repeated the analysis using all four neurologist ratings.

RESULTS

Recruitment and follow-up

Patient recruitment took place between December 2002 and February 2004 and is described in figure 1. Of the patients, 3781 participated in the study representing 91% (3781/4161) of those available for recruitment. Neurologists rated 1144 of these patients (30% of the total) as having symptoms ‘not at all’
The frequency of individual symptoms is described in figure 2. Fatigue, headache, ‘pain in arms legs and joints’ and ‘loss of sensation numbness or tingling’ were the most common. Patients with symptoms unexplained by disease had more symptoms of all types except ‘seizures or fits’.

The total symptom count was greater for patients with symptoms unexplained by disease regardless of which combination of symptoms was included in the count (table 2). However, there was substantial overlap in the IQRs (figure 3). This suggests that while the total symptom count was greater in patients with unexplained symptoms, symptom count does not discriminate well between the groups. This poor discrimination was confirmed when we plotted ROC curves. These indicated sensitivities and specificities only just above chance, irrespective of which symptoms were counted or where ‘cut-offs’ were set (figure 4).

The analysis was repeated using the four categories of how explained the symptoms were by disease. Although this approach does not allow ROC curves to be plotted, it can be clearly seen from the box plots in figure 5 that there were not substantive differences between the subcategories of ‘unexplained symptoms’ and ‘explained symptoms’.

**DISCUSSION**

In this large, prospectively recruited sample of new patients attending neurology clinics we found that ‘fatigue’, ‘headache’, ‘pain in arms, legs or joints’ and ‘loss of sensation, numbness and tingling’ were the most common self-reported symptoms. As hypothesised, patients whose symptoms were regarded by the assessing neurologists as being unexplained by disease reported more symptoms on average, than patients whose symptoms were considered to be explained by disease. This relationship was similar for all the different types of symptoms included in the symptom counts and was highly statistically significant.

However, when we examined the diagnostic value of these symptoms counts in identifying patients who had been rated by neurologists as having symptoms unexplained by disease we found that their sensitivities and specificities were poor, in fact little better than chance. That is, the use of symptom counts did not allow us to separate patients with symptoms unexplained by disease from those whose symptoms were considered to be explained by disease. This poor discrimination was confirmed when we plotted ROC curves. These indicated sensitivities and specificities only just above chance, irrespective of which symptoms were counted or where ‘cut-offs’ were set (figure 4).

The analysis was repeated using the four categories of how explained the symptoms were by disease. Although this approach does not allow ROC curves to be plotted, it can be clearly seen from the box plots in figure 5 that there were not substantive differences between the subcategories of ‘unexplained symptoms’ and ‘explained symptoms’.

**DISCUSSION**

In this large, prospectively recruited sample of new patients attending neurology clinics we found that ‘fatigue’, ‘headache’, ‘pain in arms, legs or joints’ and ‘loss of sensation, numbness and tingling’ were the most common self-reported symptoms. As hypothesised, patients whose symptoms were regarded by the assessing neurologists as being unexplained by disease reported more symptoms on average, than patients whose symptoms were considered to be explained by disease. This relationship was similar for all the different types of symptoms included in the symptom counts and was highly statistically significant.

However, when we examined the diagnostic value of these symptoms counts in identifying patients who had been rated by neurologists as having symptoms unexplained by disease we found that their sensitivities and specificities were poor, in fact little better than chance. That is, the use of symptom counts did not allow us to separate patients with symptoms unexplained by disease from those whose symptoms were considered to be explained by disease. This poor discrimination was confirmed when we plotted ROC curves. These indicated sensitivities and specificities only just above chance, irrespective of which symptoms were counted or where ‘cut-offs’ were set (figure 4).

The analysis was repeated using the four categories of how explained the symptoms were by disease. Although this approach does not allow ROC curves to be plotted, it can be clearly seen from the box plots in figure 5 that there were not substantive differences between the subcategories of ‘unexplained symptoms’ and ‘explained symptoms’.

**DISCUSSION**

In this large, prospectively recruited sample of new patients attending neurology clinics we found that ‘fatigue’, ‘headache’, ‘pain in arms, legs or joints’ and ‘loss of sensation, numbness and tingling’ were the most common self-reported symptoms. As hypothesised, patients whose symptoms were regarded by the assessing neurologists as being unexplained by disease reported more symptoms on average, than patients whose symptoms were considered to be explained by disease. This relationship was similar for all the different types of symptoms included in the symptom counts and was highly statistically significant.

However, when we examined the diagnostic value of these symptoms counts in identifying patients who had been rated by neurologists as having symptoms unexplained by disease we found that their sensitivities and specificities were poor, in fact little better than chance. That is, the use of symptom counts did not allow us to separate patients with symptoms unexplained by disease from those whose symptoms were considered to be explained by disease. This poor discrimination was confirmed when we plotted ROC curves. These indicated sensitivities and specificities only just above chance, irrespective of which symptoms were counted or where ‘cut-offs’ were set (figure 4).
disease from those whose symptoms were explained by disease. Notably these findings were robust to reanalysis using different combinations of symptoms and using the more fine-grained rating of how explained the symptoms were by disease.

Our study has the strength of being of a large and representative sample of neurology patients who had received a full specialist diagnostic assessment as well as symptom assessment using the study measures. Importantly, we have been able to follow this cohort over time and have confidence in the neurologists’ rating of ‘unexplained by disease’ as this diagnosis has a very low rate of subsequent diagnostic revision.45 However, our study also has limitations. Although almost all Scottish neurologists participated in the study, not all their clinics were sampled and specialised clinics such as neurovascular and memory clinics were not included; consequently patients with these disorders may be under-represented. Similarly we cannot be certain that Scottish neurological practice is similar to neurological practice round the world, although the prevalence rates of the common neurological disorders in patients attending the clinics sampled were similar to those reported in North America and Europe.49

Our study was in secondary care, which could have altered the performance of the measure (developed for primary care). It may also be that patients attending neurology clinics pose particular difficulties diagnostically, potentially limiting the generalisation of our results to other clinical populations. Ideally our results should be replicated in other non-neurological populations.

A small number of previous studies have measured the utility of the PHQ 15 against a medical examination diagnosis of symptoms unexplained by disease; often operationalised as DSM-IV somatoform disorders. In a study of 10 507 patients in primary care,34 of whom 19% were clinically ‘somatisers’ the

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Comparison of mean symptom counts (using different measures) between ‘unexplained’ cases and ‘explained’ controls</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Maximum symptom count score</td>
</tr>
<tr>
<td></td>
<td>Unexplained*</td>
</tr>
<tr>
<td>PHQ-15‡</td>
<td>13</td>
</tr>
<tr>
<td>PHQ-15+10 neurological symptoms</td>
<td>23</td>
</tr>
<tr>
<td>PHQ-15+10 neurological symptoms+5 PRIME MD symptoms</td>
<td>28</td>
</tr>
</tbody>
</table>

*Unexplained=Not at all explained/Somewhat explained by disease.†Explained=Largely explained/Completely explained by disease.‡PHQ-15=the maximum score is 13 as the 2 items on sexual problems and menstrual difficulties were removed. PHQ, Patient Health Questionnaire.
PHQ 15 supplemented by data from the Whiteley-7 hypochondriasis score had sensitivity of only 33% and a specificity of 82%; the accuracy of primary care diagnosis, which was the comparator standard, is unknown. We previously examined accuracy of primary care diagnoses in relation to neurological symptoms unexplained by disease and found that they correlated very poorly with specialist opinion.46 However, one cannot simply assume that specialist opinion is more likely to be accurate. Perhaps more importantly there is a lot of anecdotal evidence suggesting a so-called ‘neurophobia’ among many doctors—that is, they feel a lack of confidence in neurological diagnosis. It may well be that data from neurology paint an unduly negative picture and primary care practitioners perform better in other fields such as orthopaedics, cardiology or respiratory medicine.

A study of 906 Dutch primary care patients27 reported a sensitivity of 78% for the PHQ (at a cut-off of three or more severe symptoms) against a SCID-1 diagnosis of a somatoform disorder. The observed positive predictive value of 20% is not clinically useful, although some would argue that a negative predictive value of 3% does have utility as a screening test. However, in this study the SCID-1 only diagnosed a somatoform disorder in 32% of patients who had a clinical diagnosis of ‘unexplained somatic complaints’ made by their family physician. The SCID user’s guide, with reference to somatoform diagnoses, cautions against making diagnostic judgements using the SCID. The advice given is “Needless to say this judgement can be difficult (or impossible) to make during an interview and usually requires contact with the subject’s general medical provider or a review of the subject’s medical records”.

More recently Korber and colleagues commented enthusiastically on a sensitivity of 80% and specificity of 58% using the PHQ 15 in a German primary care cohort of 308 patients.8 Finally in a recent systematic review of symptoms counts the PHQ15 was recommended for a variety of purposes including measuring a ‘comparable construct’ to somatisation.50 In summary, despite enthusiastic endorsements, studies showing comparison against the gold standard of specialist medical assessment with follow-up were notably lacking.

Our own data, in which patients had the benefit of a prospectively conducted specialist assessment, suggested a much poorer performance of somatic symptom counts. While somatic symptom counts can identify patients who are high users of

![Figure 3](image_url) Box plots of the total number of symptoms presented separately for patients with symptoms unexplained and explained by disease. The box plots are repeated for each of the three different versions of the symptoms questionnaire. (Bars show IQRs).

![Figure 4](image_url) ROC curves showing the diagnostic qualities of the total symptom count for identifying symptoms unexplained by organic disease. The ROC curves are shown for each of the three versions of the symptoms questionnaire.
medical services,10 28 function as a measure of severity in patients with a diagnosis of somatisation disorder27 and potentially allow for targeted symptom management programmes.40 51 52 we have found that a self-rated symptom count scores, such as the PHQ 15, did not have diagnostic utility in identifying patients with symptoms unexplained by disease.

The recently published DSM-5 has changed the somatoform category. A new diagnosis of somatic symptom disorder (SSD) has been introduced, which is designed to make the concept of symptoms unexplained by disease less central to the diagnosis. It does this by defining the disorder on the basis of an excessive reaction to symptoms, whether explained by disease or not. DSM-5 has not eradicated the importance of identifying symptoms unexplained by disease however. First, the diagnosis of SSD still requires a judgement about whether the patient’s concern about their symptoms is ‘disproportionate to their seriousness’. Second, in such cases the degree to which the symptoms are explained by disease is important in guiding medical treatment; for example, while patients with inflammatory bowel disease and irritable bowel disease may merit a diagnosis of SSD, one may benefit from immunosuppressant therapy and possibly surgical intervention whereas the other would not. It remains the case that the distinction of whether or not there is pathophysiological disease is the core purpose of medical assessment and we would suggest the one of central importance to patients. Third, and especially important to neurological symptoms, are the modifications in DSM-5 to conversion (functional) disorder; which emphasise that this is a positive diagnosis that should be made on the basis of physical signs on examination that are inconsistent or incongruent with neurological disease, and not simply because tests are normal. SNSS data as a whole could be viewed as supporting this change in definition as our original report on diagnostic accuracy43 showed a very low level of revision when diagnoses were made on such lines whereas this report suggests attempts to define a psychopathological measure of ‘somatisation’ using symptom counts are not successful.

In conclusion, our findings confirm the common clinical impression that patients with symptoms unexplained by disease have more symptoms. However, on an individual level, self-rated symptom counts perform poorly as a diagnostic test, and should not be used to identify patients with symptoms unexplained by disease (or ‘medically unexplained symptoms’). Accurate diagnosis based on a thorough clinical assessment remains the cornerstone of good medical practice. Clinical judgments and epidemiological surveys that use symptom counts as an indicator of ‘medically unexplained symptoms’ should be interpreted with caution.

Author affiliations
1Department of Psychiatry, University of Edinburgh, Edinburgh, UK
2Department of Clinical Neurosciences, University of Edinburgh, Edinburgh, UK
3Department of Medical Statistics, University of Edinburgh, Edinburgh, UK
4Department of Neurology, University of Otago, Christchurch, New Zealand
5Department of IHW, University of Glasgow, Glasgow, UK
6Division of Neuroscience, University of Dundee, Dundee, UK
7Department of Public Health Sciences, University of Edinburgh Medical School, Edinburgh, UK
8Department of Psychiatry, University of Oxford, Oxford, UK

Contributors All the authors contributed to study design, execution and writing of the manuscript.

Competing interests None.

Ethics approval MREC Scotland.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

Figure 5 Box plots of the total number of symptoms presented separately for each of the four organicity groups and repeated for each of the three different versions of the symptoms questionnaire. PHQ, Patient Health Questionnaire.