Journal of Psychosomatic Research 75 (2013) 60–64

Contents lists available at SciVerse ScienceDirect

Journal of Psychosomatic Research

The factor structure of the PHQ-9 in palliative care

Joseph Chilcot [a,*], Lauren Rayner [b], William Lee [b], Annabel Price [b], Laura Goodwin [b], Barbara Monroe [c], Nigel Sykes [c], Penny Hansford [c], Matthew Hotopf [b]

[a] Health Psychology Section, Psychology Department, Institute of Psychiatry, King's College London, UK
[b] Department of Psychological Medicine, Institute of Psychiatry, King's College London, UK
[c] St Christopher's Hospice, Sydenham, London, UK

ARTICLE INFO

Article history:
Received 26 October 2012
Received in revised form 18 December 2012
Accepted 21 December 2012

Keywords:
PHQ-9
Patient Health Questionnaire
Confirmatory factor analysis
CFA
Symptoms
Depression

ABSTRACT

Objectives: The Primary Care Evaluation of Mental Disorders Patient Health Questionnaire (PRIME-MD PHQ-9) is a common screening tool designed to facilitate detection of depression according to DSM-IV criteria. However, the factor structure of the PHQ-9 within the palliative care population has not been evaluated.
Methods: 300 participants completed the PHQ-9 within one week of referral to a palliative care service. Participants completed the PHQ-9 again four weeks later (n = 213). Confirmatory factor analysis (CFA) and multiple-group CFA were undertaken to test the factor structure of the PHQ-9 and evaluate model invariance over time.
Results: A two-factor model comprising somatic and cognitive–affective latent factors provided the best fit to the data. Multiple-group CFA suggested model invariance over time. Structural equation modelling revealed that follow-up (time 2) cognitive–affective and somatic symptoms were predicted by their baseline (time 1) factors.
Conclusions: The PHQ-9 measures two stable depression factors (cognitive–affective and somatic) within the palliative care population. Studies are now required to examine the trajectories of these symptoms over time in relation to clinical intervention and events.

Introduction

Depression in palliative care is common [1–3] and associated with an array of negative psychosocial and clinical outcomes, including increased disability, poorer prognosis, increased desire for death, and higher mortality [4–6]. Although both antidepressants and psychological therapy have been shown to be effective in treating depression in the context of a life-threatening illness [7,8], depression often goes undetected and untreated in this patient group [9]. The Primary Care Evaluation of Mental Disorders Patient Health Questionnaire (PRIME-MD PHQ-9) is a commonly used screening tool designed to facilitate the detection of depression according to DSM-IV criteria [10,11]. It has been shown to have utility as a measure of change over time [12,13], and validation studies suggest that it is suitable for use in a variety of physical illnesses [14–16]. A PHQ-9 cut-off score of 10 has good sensitivity (0.79 95% CI 0.65 and 0.89) and specificity (0.89 95% CI 0.84 and 0.93) for diagnostic depression in patients with long-term medical conditions [17].

Studies assessing the factor structure of the PHQ-9 in primary care and psychiatric samples have identified one single factor comprising all nine items, which supports the supposition that DSM-IV major depression is a coherent, unidimensional construct [18–21]. However, assessments of depression in physically ill or injured individuals are complicated by the fact that somatic symptoms may be due to their physical condition. Fatigue, sleep disturbance and poor appetite are common symptoms of depression, which are included as diagnostic criteria in DSM-IV and the PHQ-9. Yet these somatic symptoms may also be attributable to pain or disease, and as a result may confound the measurement of the prevalence and course of depression in certain patient groups, particularly those with prevalent physical symptoms and pain.

Recent studies undertaken in people with spinal cord injury (SCI) illustrate this point and suggest that a 2-factor structure yields a better fit [22,23]. Krause and colleagues conducted a confirmatory factor analysis (CFA) of 568 inpatients with traumatic SCI, which showed that a two-factor model (3 somatic items [sleep disturbance, poor energy, appetite change], 6 non-somatic) had a better fit than a single-factor model including all nine PHQ-9 items [23]. In a subsequent study, Krause et al. [24] followed up 227 SCI patients one year after hospital discharge to assess changes in somatic and non-somatic factors. The proportion of patients endorsing the somatic items decreased significantly on follow-up (e.g. appetite change 29.0% to

* Corresponding author at: Health Psychology Section, Institute of Psychiatry, King's College London, 5th floor Bermondsey Wing, Guy's Campus, London Bridge, London SE1 9RT, UK. Tel.: +44 20 718 82597; fax: +44 20 718 80184.
E-mail address: joseph.chilcot@kcl.ac.uk (J. Chilcot).

0022-3999/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jpsychores.2012.12.012

17.3%), whilst endorsement of certain cognitive/ affective symptoms (e.g. anhedonia and feeling down) increased. The authors concluded that ambiguity over the measurement of somatic symptoms in SCI inpatients may obscure the assessment of the natural history of depression in this patient group [24]. However, investigations in both elderly and primary care patients suggest that somatic symptoms predict future depressive episodes and may serve as preclinical indicators of depression [25,26].

To our knowledge, the factor structure of the PHQ-9 in palliative care patients has not been studied, but there is on-going debate regarding the predictive validity of somatic symptoms of depression in this population. In a recent cross-sectional survey of depression in 300 patients newly referred to specialist palliative care, Rayner et al. [1] assessed the ability of somatic symptoms to predict cases of major depressive disorder (MDD) defined according to the PHQ-9. The positive predictive values of sleep disturbance, poor appetite and fatigue were low (<24%) indicating that somatic symptoms are relatively poor predictors of depression in this patient group [1]. As expected, the prevalence of depression in this patient group was high at 19%. However, when followed up four weeks later 69% of the patients depressed at baseline no longer met criteria for MDD. The high remission rate echoes previous studies in palliative care, which have also indicated a higher rate of remission of depression than that observed in other settings [27,28]. One explanation for this high remission rate is that improved symptom control upon referral to specialist palliative care ameliorated patients' somatic symptoms (sleep disturbance, poor appetite and fatigue), resulting in fewer individuals meeting the PHQ-9 criteria for MDD [29].

The present study sought to identify the factor structure of the PHQ-9 using CFA, and to examine change in somatic and non-somatic depressive symptoms four weeks after referral to specialist palliative care. Given that past analyses in physically ill populations suggest that two latent factors underlie the PHQ-9 [23,30], it was hypothesised here that a two-factor model comprising cognitive–affective and somatic factors would provide the best fit to the data, and that the factor structure would be invariant over time.

Methods

Design and procedure

We conducted a cross sectional study with a four week follow-up (see Rayner et al. [1]). Patients completed face-to-face interviews (time 1) including completion of the PHQ-9 within one-week of referral to a palliative care service, and then again by telephone 4-weeks later (time 2).

Participants

The sample comprises 300 patients newly referred to a palliative care service in south London. We excluded individuals who were too frail to participate and those with cognitive impairment (as determined by the Abbreviated Mental Test [31]), and recruited a consecutive sample.

Two hundred and thirteen patients completed the PHQ-9 at both time 1 and time 2. Between time 1 and time 2, 43 deaths occurred. Other reasons for non-follow-up were; patients declined ($n = 13$), too ill ($n = 15$), family declined ($n = 4$), unable to contact ($n = 7$), discharged ($n = 3$) and other ($n = 2$). The median interval between the first and second assessments was 28 days (interquartile range = 27–32).

Assessment of depression symptoms: PHQ-9

Depression symptoms were assessed using the PRIME-MD PHQ-9 [10,11]. The PHQ-9 assesses the nine symptoms which make up the diagnosis of DSM-IV major depressive disorder (MDD), for which each item is rated on whether the symptom has bothered the respondent "not at all", "several days", "more than half of the days" or "nearly every day" in the last two weeks. Criteria for MDD are met if the patient reports having experienced low mood or loss of interest within a total of five or more symptoms. Symptoms are scored if they are experienced for more than half of the days in the previous two weeks with the exception of the item about suicidality, which counts towards the criteria if experienced at all.

The PHQ-9 has high sensitivity and specificity for depression in a variety of physical illnesses [14–17] and has good test–retest characteristics [32]. Furthermore, administering the PHQ-9 via telephone has good concordance with face-to-face interview [33].

Statistical methods

CFA of the PHQ-9 (time 1 data, $n = 300$) was evaluated in Mplus 5.21 [34] using weighted least-squares with mean and variance adjustment estimation (WLSMV). WLSMV was employed due to the ordinal response format of the PHQ-9 and the skewed distribution of several items. The WLSMV estimator weights the fit function by the variances/covariances and kurtosis of the data to adjust for violations of multivariate normality, producing unbiased consistent and efficient parameter estimates and standard errors in this situation [35]. Sartorra–Bentler corrected chi-square and fit indices were evaluated to determine adequacy of the model fit. A non-significant chi-square test would suggest that the hypothesised and observed models do not differ significantly, thus supporting the fit of the model to the data. However because the chi-square statistic is sensitive to sample size, several fit indices were also evaluated. The comparative fit index (CFI) is an incremental fit index that compares the proposed model with the null model, and uses an approach based on the non-central chi-square distribution [36]. A CFI exceeding 0.90 indicates good fit, although CFIs ≥ 0.95 are preferred [37]. Root mean square error of approximation (RMSEA) was also evaluated where a value of <0.08 is considered to demonstrate approximate fit [38], although values < 0.05 are preferred. As suggested from a review of fit indices for estimators such as WLSMV, weighted root mean square residual (WRMR; <1.0) was also evaluated [39].

In addition to a one-factor model with all 9 items loading onto one latent factor (termed "depression"), a two factor model (somatic vs. cognitive–affective) as suggested by Krause et al. [23] was also evaluated. The best fitting model as determined by examination of both the chi-square statistic and fit indices was selected for model modification to see whether the fit could further be improved. Modification involved checking factor loadings to ensure their significance and examining Mplus derived modification statistics.

Model invariance between time 1 and time 2 was evaluated using a multiple-group CFA approach [40,41]. A baseline model was run with factor loadings and item thresholds freely estimated at both time points, and scale factors fixed to 1 and means fixed to zero (non-invariant model). A second model was then evaluated with factor loadings and item thresholds constrained to be equal (invariant model). Model fit between the variant and invariant models was compared using the DIFFTEST procedure in Mplus. A non-significant DIFFTEST chi-square suggests that the factor structure is invariant over time.

Finally, a structural equation model (SEM) was evaluated to establish whether cognitive–affective and somatic symptoms at time 1 predicted symptoms at time 2. The SEM utilised latent factors with a full measurement model for the PHQ-9 as derived from the CFA. The latent factors (cognitive–affective and somatic) at time 2 were regressed onto their time 1 factors. Like the CFA, the SEM used WLSMV estimation and evaluated the same fit indices as described above.