Journal of Affective Disorders 219 (2017) 1–8

Contents lists available at ScienceDirect

Journal of Affective Disorders

journal homepage: www.elsevier.com/locate/jad

Research paper

Factor structure and longitudinal measurement invariance of PHQ-9 for specialist mental health care patients with persistent major depressive disorder: Exploratory Structural Equation Modelling

Boliang Guo[a], Catherine Kaylor-Hughes[a], Anne Garland[b], Neil Nixon[b], Tim Sweeney[b], Sandra Simpson[c], Tim Dalgleish[e], Rajini Ramana[d], Min Yang[f], Richard Morriss[a,*]

[a] CLAHRC-EM, School of Medicine, University of Nottingham, Nottingham, United Kingdom
[b] Adult Mental Health Directorate, Nottinghamshire Healthcare Trust, Nottingham, United Kingdom
[c] Nottinghamshire Healthcare Foundation Trust
[d] Cambridge and Peterborough Partnership NHS Foundation Trust
[e] Medical Research Council Cognition and Brain Sciences Unit, Cambridge
[f] West China School of Public Health, Sichuan University, PR China

ARTICLE INFO

Keywords:
PHQ-9
Factor structure
Measurement equivalence/invariance
Exploratory Structural Equational Modelling
Major depressive disorder
Chronic depression

ABSTRACT

Background: The Patient Health Questionnaire-9 (PHQ-9) is a widely used instrument for measuring levels of depression in patients in clinical practice and academic research; its factor structure has been investigated in various samples, with limited evidence of measurement equivalence/invariance (ME/I) but not in patients with more severe depression of long duration. This study aims to explore the factor structure of the PHQ-9 and the ME/I between treatment groups over time for these patients.

Methods: 187 secondary care patients with persistent major depressive disorder (PMDD) were recruited to a randomised controlled trial (RCT) with allocation to either a specialist depression team arm or a general mental health arm; their PHQ-9 score was measured at baseline, 3, 6, 9 and 12 months. Exploratory Structural Equational Modelling (ESEM) was performed to examine the factor structure for this specific patient group. ME/I between treatment arm at and across follow-up time were further explored by means of multiple-group ESEM approach using the best-fitted factor structure.

Results: A two-factor structure was evidenced (somatic and affective factor). This two-factor structure had strong factorial invariance between the treatment groups at and across follow up times.

Limitations: Participants were largely white British in a RCT with 40% attrition potentially limiting the study's generalisability. Not all two-factor modelling criteria were met at every time-point.

Conclusion: PHQ-9 has a two-factor structure for PMDD patients, with strong measurement invariance between treatment groups at and across follow-up time, demonstrating its validity for RCTs and prospective longitudinal studies in chronic moderate to severe depression.

1. Introduction

The Patient Health Questionnaire-9 (PHQ-9) is a 9-item self-reported scale measuring the symptoms of major depression derived from the Diagnostic and Statistical Manual, (fourth edition (DSM-IV)) (Kroenke et al., 2001; Spitzer et al., 1999). It can help clinicians quickly evaluate the severity of a person's mood and has been applied in various patient populations such as coronary heart disease (de Jonge et al., 2007), spinal cord injury (Krause et al., 2010), diabetes (Zhang et al., 2013), and primary care (Baas et al., 2011; Petersen et al., 2015); the scale has also been used to measure depression in the general population (Yu et al., 2012).

Recently the PHQ-9 was used as a depression measure for secondary care patients with persistent major depressive disorder (PMDD) in a pragmatic clinical trial conducted in the UK (Morriss et al., 2010, 2016). As a well validated and frequently used instrument, the PHQ-9's underlying factor structure has been explored for various patient populations already. However no study has yet investigated the factor

Abbreviations: CFA, confirmatory factor analysis; CFI, comparative fit index; EFA, exploratory factor analysis; ESEM, exploratory structural equation modelling; ME/I, measurement equivalence/invariance; NNFI, non-normal fit index; PMDD, persistent major depressive disorder; RMSEA, root mean square error of approximation; SDS, specialist depression service

* Correspondence to: Institute of Mental Health, University of Nottingham, Triumph Road, Nottingham NG7 2TU, United Kingdom.
E-mail address: richard.morriss@nottingham.ac.uk (R. Morriss).

http://dx.doi.org/10.1016/j.jad.2017.05.020
Received 29 September 2016; Received in revised form 5 April 2017; Accepted 6 May 2017
Available online 08 May 2017
0165-0327/

structure for specialist mental health patients with persistent or chronic moderate to severe unipolar depressive disorder. Understanding the factor structure of PHQ-9 for secondary care patients with PMDD could help to understand precisely what is being measured by this instrument to aid the interpretation of studies such as randomised controlled trials of interventions or large scale mechanistic or epidemiological investigations in this population of patients. Additionally it could help understand the underlying dimensions and mechanism of long term unipolar depressive disorder (Elhai et al., 2012).

Studies that have explored the factor structure of PHQ-9 have shown heterogeneous findings (Petersen et al., 2015), with the number of underlying factors varying between one and two (Baas et al., 2011; Krause et al., 2010; Richardson and Richards, 2008). These differences might be due to the different patient populations, physical and mental co-morbidities, research design and analyses, e.g. using exploratory factor analysis (EFA) compared to confirmatory factor analysis (CFA). Two-factor structure items have generally been grouped into two types: somatic (e.g. sleep difficulties, appetite changes and fatigue) and non-somatic/affective items (e.g. depressed mood, feeling of worthlessness and suicidal thoughts). However, even with the two-factor structure, there are still some inconsistent item-factor mapping patterns across studies (Elhai et al., 2012; Petersen et al., 2015). Patients with PMDD are more likely to have other axis 1 psychiatric disorders; in particular: generalised anxiety disorder, social phobia, post-traumatic stress disorder and hypochondriasis as well as more atypical depression and treatment resistance (Rush et al., 2012), which is in itself associated with melancholia and a number of personality traits (Bennabi et al., 2015). Clinically, melancholia is associated with more complete loss of pleasure, low energy, walking and talking more slowly and less reactivity of mood among features measured by the PHQ-9 (Parker et al., 2013).

Given that PHQ-9 has been used as a depression outcome measure in various studies, establishing the measurement invariance/equivalence (ME/I) across groups is a logical prerequisite to conducting substantive cross-group and/or follow-up time comparisons (Vandenberg and Lance, 2000). Measurement invariance of PHQ-9 was made between male and female (Petersen et al., 2015) and across ethnic groups (Baas et al., 2011); a further study by Richardson and Richards (2008) also reported that the PHQ-9 factor structure was relatively stable across follow-up times for patients with spinal cord injury. However, this was performed using only exploratory factor analysis (EFA) on PHQ-9 measures collected at each follow-up time and comparing the factor loading by eye to draw their conclusion. No formal statistical tests were applied to justify the cross-time measurement invariance (Vandenberg and Lance, 2000). Hence, conclusions on measurement invariance across follow-up time requires further examination.

To make group comparisons of PHQ-9 development across follow-up time points, longitudinal between-group measurement invariance should be established before making any valid inference based on comparing PHQ-9 scores between treatment and control groups across measurement time. Nevertheless, no study has yet investigated the between group measurement invariance across follow-up time.

EFA and CFA have previously been used to investigate PHQ-9 factor structure (Petersen et al., 2015; Yu et al., 2012). However, both EFA and CFA have methodological limits (Asparouhov and Muthén, 2009; Marsh et al., 2014). Using EFA, it is impossible to incorporate latent EFA factors into subsequent analyses and it is not easy to test measure invariance across groups and/or times (Marsh et al., 2014). With CFA, each item is strictly allowed to load on one factor and all non-target loadings are constrained to zero. In applied research, it is generally justifiable by theory and/or item contents that item(s) could cross load on different latent factors. Restrictive zero loading typically results in inflated CFA factor correlation and leads to biased estimates in CFA modelling when other variables are included in the model (Marsh et al., 2014). The latest methodology development integrates the best features of both EFA and CFA together as Exploratory Structural Equational Modelling (ESEM), applying EFA rigorously to specify more appropriately the underlying factor structure together with applying the advanced statistical methods typically associated with CFAs (Marsh et al., 2014). Hence, ESEM will be performed to test the factor structure of PHQ-9 for secondary care patients with PMDD. Measurement invariance tests of PHQ-9 factor structure, i.e. between treatment group invariance at and across follow-up time, will also be conducted using ESEM.

In summary, the factor structure and measurement invariance between treatment groups across follow-up time of PHQ-9 for secondary care patients with PMDD will be explored. This study will apply methodologically rigorous ESEM modelling to explore the factor structure of the PHQ-9 scale and measurement invariance between treatment groups at and across follow-up time points.

2. Method

2.1. Patients and instruments

Participants ($N = 187$) were drawn from a multicentre pragmatic randomised controlled trial (RCT) evaluating outcomes of a specialist mood disorders team for treatment seeking adults in secondary mental health care services (Special depression service, SDS) compared to treatment as usual (TAU). At the time of recruitment participants were receiving treatment in secondary mental health services from community mental health teams, out-patient and in-patient units in three mental health trusts across Nottinghamshire, Derbyshire and Cambridgeshire in the UK.

Participants were eligible for the study if they were: thought by the referrer to have primary unipolar depression; aged 18 years or over; able and willing to give oral and written informed consent to participate in the study; had been offered or received direct and continuous care from one or more health professionals in the preceding 6 months and currently be under the care of a secondary care mental health team; had a diagnosis of major depressive disorder with a current major depressive episode according to the structured clinical interview for DSM-IV (SCID) (First et al., 1997); met five of nine NICE criteria for symptoms of moderate depression; had a score of ≥ 16 on the 17-item Hamilton Depression Rating Scale ($HDRS_{17}$) (Williams et al., 2008); and had a Global Assessment of Functioning (GAF) (American Psychiatric Association, 1994) score ≤ 60. Participants were not included if they were: in receipt of emergency care for suicide risk, at risk of severe neglect, or a homicide risk, but were not excluded because of such risk provided the risk was adequately contained in their current care setting and the primary medical responsibility for care was with the referral team; did not speak fluent English; were pregnant; had unipolar depression secondary to a primary psychiatric or medical disorder, except when bipolar disorder was identified by the research team after referral with unipolar depression because an SDS would be expected to manage bipolar depression in clinical practice ($n = 8$, 4.3%).

Of the total 187 patients, 93(49.7%) patients were allocated to the treatment arm and 94 (50.3%) to treatment as usual (TAU) arm. (See Fig. 1 for CONSORT diagram of participant flow through the study). The primary outcome measures were $HDRS_{17}$ and GAF which were measured at baseline, 6 and 12 month follow up time points (Morriss et al., 2010); the secondary outcome measures included Beck Depression Inventory version I (BDI-I) (Beck et al., 1961), Patient Health Questionnaire (PHQ-9) (Kroenke et al., 2001), Quick Inventory of Depressive Symptomatology Self-Report (QIDS-SR) (Rush et al., 2006), the modified Social Adjustment Scale (SAS-M) (Cooper et al., 1982), Patient Doctor Relationship Questionnaire (PDRQ) (Van der Feltz-Cornelis et al., 2004) and the EQ-5D-3 L (Euroqol Group, 1990). The study design and data collection procedures have been described in the published protocol (Morriss et al., 2010); more details about the