General Hospital Psychiatry 42 (2016) 60–66

Contents lists available at ScienceDirect

General Hospital Psychiatry

journal homepage: http://www.ghpjournal.com

Performance of depression rating scales in patients with chronic kidney disease: an item response theory-based analysis☆,☆☆,★,★★

Marisa Toups, M.D. [a], Thomas Carmody, Ph.D. [a,b], Madhukar H. Trivedi, M.D. [a], A. John Rush, M.D. [c], S. Susan Hedayati, M.D., M.H.Sc. [d,e,*]

[a] Department of Psychiatry, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX, USA, 75390-9119
[b] Division of Biostatistics, Department of Clinical Sciences, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX, USA, 75390-9066
[c] Division of Clinical Sciences, Duke-NUS, 8 College Rd, Singapore, 169857
[d] Division of Nephrology, Department of Medicine, Veterans Affairs North Texas Health Care System, 4500 S. Lancaster Rd MC 111G1, Dallas, TX, USA, 75216
[e] Division of Nephrology, Department of Medicine, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX, USA, 75390-8516

ARTICLE INFO

Article history:
Received 20 May 2016
Revised 18 July 2016
Accepted 20 July 2016

Keywords:
Depression
Chronic kidney disease
Comorbidity
Screening
Assessment

ABSTRACT

Objective: Because there is overlap between somatic symptoms of depression and symptoms of chronic kidney disease (CKD), it is unclear if self-reported depression rating scales can be used accurately in predialysis CKD patients, especially if CKD and other comorbidities are symptomatic. We assessed the performance of two depression scales — the Beck Depression Inventory (BDI) and the Quick Inventory of Depression Symptomatology ($QIDS\text{-}SR_{16}$) — by CKD stage, diagnosis of diabetes and total medical comorbidity burden — using item response theory (IRT) in a sample of 272 predialysis CKD patients.

Methods: We performed IRT by low versus high CKD stage, diabetes versus no diabetes and high (>3 diagnoses) versus low medical comorbidity burden.

Results: IRT models of each rating scale were affected in a limited way by CKD stage, diabetes and medical comorbidity burden. Sleep disturbances on the $QIDS\text{-}SR_{16}$ were more discriminatory for depression in diabetics and those with high comorbidity burden. Pessimism and guilt from the BDI compared to $QIDS\text{-}SR_{16}$ were more discriminatory of depression in the high CKD and high comorbidity groups, respectively.

Conclusions: Overall item differences were modest, and chronic disease severity by CKD stage, diabetes mellitus or other medical comorbidities did not appreciably contribute to differences in scale performance.

1. Introduction

Prevalence of depression is increased in the setting of chronic kidney disease (CKD), such that 20–25% of patients are affected [1–3]. Importantly, depression is associated not only with significant functional impairment but also increased risk of hospitalization, disease progression and death in CKD patients [4,5]. A recent meta-analysis of 22 studies estimated depression increased risk of all-cause mortality in CKD patients by 60% [6]. Most studies analyzing prevalence or sequelae of depression in CKD patients used self-report questionnaires for identification and/or measurement [7,8]. Since somatic symptoms of depression such as decreased concentration, energy, appetite and sleep disturbance overlap with symptoms of CKD, self-reports may not perform as expected and lead to an overestimation of depression diagnoses in CKD patients. In fact, studies suggest that use of self-reports in CKD patients receiving dialysis or at dialysis initiation requires higher cutoff scores for accurate measurement of depression [9–11]. Our previous work demonstrated that in nondialysis CKD patients, cutoff scores for the diagnosis of a major depressive episode (MDE) were equivalent to previously established benchmarks and did not require adjustment [12]. In that analysis, we did not examine clinical differences within the nondialysis

☆ Funding: This work was supported by a Veterans Affairs MERIT grant (CX000217–01) and a grant from the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK, NIH) (R01DK085512) (Susan Hedayati, M.D.), the NIMH (K23MH104768) (Marisa Toups, MD), the Hersh Foundation and the Center for Depression Research and Clinical Care (Madhukar H. Trivedi, M.D). The views expressed here are those of the authors and do not necessarily represent the views of the Department of Veterans Affairs or the NIH.

☆☆ Declaration of interest statement: The authors have no competing interests.

★ Financial Disclosures: A. John Rush, M.D. has received consulting fees from Brain Resource Ltd., H. Eli Lilly, Lundbeck A/S, Medavante, Inc.; National Institute of Drug Abuse, Santium Inc.,Takeda USA,; speaking fees from the University of California at San Diego, Hershey Penn State Medical Center, and the American Society for Clinical Psychopharmacology; royalties from Guilford Publications and the University of Texas Southwestern Medical Center; a travel grant from CINP and research support from Duke-National University of Singapore. Madhukar H. Trivedi, MD is or has been an advisor/consultant and received fees from: Alkermes, AstraZeneca, Cerecor, Eli Lilly & Company, Lundbeck, Naurex, Neuronetics, Otsuka Pharmaceuticals, Pamlab, Pfizer Inc., SHIRE Development and Takeda. Marisa Toups, MD has received fees from Otsuka Pharmaceuticals.

★★ Clinical Trial Registry: N/A

* Corresponding author at: VA North Texas Health Care System, Nephrology Section, 4500 South Lancaster Road MC 111G1, Dallas, TX, 75216-7167. Tel.: +1-214-857-2214; fax: +1-214-857-1514.
E-mail address: susan.hedayati@utsouthwestern.edu (S.S. Hedayati).

http://dx.doi.org/10.1016/j.genhosppsych.2016.07.005
0163-8343/

population, however, so it remains unknown whether illness severity — of CKD itself and due to comorbid illness — may affect the performance of depression self-report measures.

Item response theory (IRT) is a robust method to assess the performance of rating instruments in different samples [13]. IRT models the response characteristics of each item as well as the scale as a whole and, thus, provides a detailed evaluation of performance [14]. To our knowledge, neither the Quick Inventory of Depression Symptomatology Self-Rated (QIDS-SR_{16}) nor the Beck Depression Inventory (BDI) have previously been analyzed using IRT in nondialysis CKD patients, though the BDI has been the subject of IRT analysis in other medical populations. Comparing patients who were above age 65 to younger adults IRT analysis showed that older patients reported more somatic symptoms on the BDI, a trend that was more pronounced as depression severity increased [15]. However, the authors did not attempt to account for the role of physical health in their results. Wardenaar et al. performed a similar IRT analysis of the BDI in subjects postmyocardial infarction and found that mood items were more representative of diagnosed depression, compared to somatic items, which were reported frequently in all subjects [16]. Collectively, these analyses support the use of IRT to better understand the effect of medical comorbidity on the assessment of depression.

Here, we report upon an IRT analysis of the BDI and QIDS-SR_{16} in a large, consecutively recruited cohort with nondialysis Stage 2–5 CKD. The specific objectives were to determine whether CKD stage and presence of other medical comorbidities impact scale performance and, if so, which item(s) account for the differences. Based on previous findings, we would expect that somatic symptoms would drive any difference found in scale performance as disease severity increases.

2. Methods

This study was approved by the Institutional Review Boards of UT Southwestern Medical Center and the North Texas Veterans Affairs Medical Center.

2.1. Participants

This sample was recruited for a prospective observational study designed to identify the prevalence and effects of major depressive disorder (MDD) in the predialysis CKD population [5] and has been reported on previously [12,17]. Participants were adult outpatients with CKD Stages 2, 3, 4 and Predialysis 5, attending the Dallas Veterans Affairs Medical Center Nephrology Clinic, recruited in 2005–2006. Patients with Stage 1 CKD were excluded. Participants were approached consecutively in the clinic waiting area using a schema in which, to avoid selection bias, every sixth patient deemed potentially eligible based on medical record review was chosen for recruitment. All participants completed written, informed consent prior to providing any information to investigators. All subjects receive a formal psychiatric diagnostic interview using the MINI International Neuropsychiatric Interview (MINI) in addition to providing self-assessment of depression symptoms [18].

Of the 272 participants, 5 did not have complete diagnostic data and were excluded, leaving 267. A partially overlapping group was missing items on either the QIDS-SR_{16}, BDI or both and was excluded from the respective analyses. Therefore, 263 participants were eligible for the QIDS-SR_{16} analysis and 252 for the BDI analysis.

2.1.1. Group definitions

We divided the participants into low (2 or 3) and high stage (4 or 5) CKD. Stage 2–3 CKD is mild-to-moderate and asymptomatic, while Stages 4–5 is considered severe and symptomatic [19]. CKD stage was defined at baseline using the National Kidney Foundation guidelines [19]: Stage 2, estimated glomerular filtration rate (eGFR) 60–89 mL/min/1.73 m^2 and other evidence of kidney disease manifest by either pathologic abnormality of kidney on biopsy or markers of kidney damage; Stage 3, eGFR 30–59; Stage 4, eGFR 15–29; and Stage 5, eGFR<15. From the 267 participants who had data on CKD stage and at least one depression scale, 118 (44.2%) had low stage and 149 (55.8%) had high stage CKD.

We chose diabetes for individual analysis because it is the most common cause of CKD in the United States [20] and was the most common comorbidity (55.6%) in our sample. Diabetes was defined by medical record diagnosis of either Type I or II diabetes or use of insulin or other medications indicated for normalizing blood sugar. By this definition, diabetes was present in 148 participants (55.6%) from the 266 who had data on diabetes diagnosis and at least one depression scale.

We used an a priori definition of *high burden* of medical comorbidity as >3 and *low burden* as ≤3 diagnoses present in the medical record for each individual subject. The threshold was chosen pragmatically; there is no consensus on the definition of medical burden in the context of depression diagnosis and treatment [21]. The median number of comorbidities was 3; therefore, a split at 3 produced two groups of comparable size. Tracked medical comorbidities included hypertension, diabetes mellitus, congestive heart failure, coronary artery disease, cerebrovascular and peripheral vascular disease, lung disease, liver disease, nonskin malignancy and infection with human immunodeficiency virus. Each diagnostic category was used only once; for example, if a patient had two types of malignancy, this was counted as one comorbidity. Of the 267 participants with data from at least one scale and complete medical record data, 111 (41.6%) had more than three medical comorbidities and were classed as high burden.

2.2. Depression assessments

The BDI [22] has 21 items rated using a 0–3-point Likert scale and summed for the final score, with a maximum of 63. Higher scores correspond to more severe symptoms. The QIDS-SR_{16}[23] contains sixteen items covering the nine DSM-IV symptom domains for MDD. Three domains have multiple items: sleep has four items rating initial, middle and terminal insomnia and hypersomnia; appetite has four items rating increased appetite, decreased appetite and increased and decreased weight; and psychomotor has items assessing increased and decreased psychomotor activity. Only the highest rated item in each of these domains is used in the final score, for a range of 0–27. For this analysis, only the scored item was used in the IRT model; therefore, the magnitude of the dysfunction in those domains is modeled rather than the specific symptom profile.

2.3. Statistical analysis

Baseline demographic and clinical characteristics were compared between groups using *t* tests for continuous outcomes and chi-square tests for categorical outcomes.

Before applying IRT to a sample of assessment scores, it first must be shown that the scale(s) of interest are unidimensional in the analyzed sample. To do this, we used exploratory factor analysis based on polychoric correlations and parallel analysis (PA), as modified by Glorfeld, to determine the number of factors [24]. Simulation studies support the accuracy of this method over other methods, such as the scree test, to determine the best number of factors [24,25]. PA is conducted by generating simulated random data sets with the same number of items and observations as the real data set. The eigenvalues derived from these datasets determine the size of eigenvalues which may occur by chance. We created 5000 simulated data sets and took the 95th percentile of the eigenvalue distribution as the threshold for the eigenvalues derived from the real data. If only one factor derived from the real data set has an eigenvalue larger than this threshold, the scale is considered to be unidimensional.

Next, IRT was applied to each scale and for each condition using MULTILOG Version 7 software [26] to run a four-parameter model