



Published in final edited form as:

Epilepsy Behav. 2014 August ; 37: 215–220. doi:10.1016/j.yebeh.2014.06.030.

Validation of the Patient Health Questionnaire-9 (PHQ-9) For Depression Screening in Adults with Epilepsy

Jaivir S. Rathore, M.D.^{*,1,2,§,6}, Lara E. Jehi, M.D.^{1,2,4}, Youran Fan, PhD.^{4,5}, Sima I. Patel, M.D.^{1,2,§,7}, Nancy Foldvary-Schaefer, D.O., M.S.^{1,2}, Maya J. Ramirez, PhD.^{2,3,§,8}, Robyn M. Busch, PhD.^{1,2,3}, Nancy A. Obuchowski, PhD.⁵, and George E. Tesar, M.D.^{1,2,3}

Lara E. Jehi: JEHL@ccf.org; Youran Fan: FanY2@ccf.org; Sima I. Patel: sipatel12@gmail.com; Nancy Foldvary-Schaefer: foldvan@ccf.org; Maya J. Ramirez: maya.ramirez@gmail.com; Robyn M. Busch: buschr@ccf.org; Nancy A. Obuchowski: obuchon@ccf.org; George E. Tesar: TESARG@ccf.org

¹Cleveland Clinic Epilepsy Center, Cleveland Clinic Foundation, Cleveland, Ohio, U.S.A

²Neurological Institute, Cleveland Clinic Foundation, Cleveland, Ohio, U.S.A

³Department of Psychiatry and Psychology, Cleveland Clinic Foundation, Cleveland, Ohio, U.S.A

⁴Neurological Institute Center for Outcomes Research and Evaluation, Cleveland Clinic Foundation, Cleveland, Ohio, U.S.A

⁵Department of Quantitative Health Sciences, Cleveland Clinic Foundation, Cleveland, Ohio, U.S.A

⁶Department of Neurology and Rehabilitation, Neuropsychiatric Institute, University of Illinois Medical Center at Chicago (UIC), Illinois, U.S.A

© 2014 Published by Elsevier Inc.

*Corresponding Author: Jaivir S. Rathore, M.D. 912 S. Wood St. 855N (M/C 769) UIC Chicago, IL 60612, U.S.A, Phone: +1-216-925-2499, neurocare9@gmail.com.

§Previous institute.

Authors' contributions

JSR, LEJ and GET are the lead investigators, designed the overall study, selected the variables, and interacted with statisticians for the data analysis. JSR did the literature search and wrote the initial draft of the manuscript and submitted it to GET, who served as the senior author with inputs from other authors. LEJ, the director of epilepsy research at the Cleveland Clinic and dual-board certified neurologist and epileptologist; also served as the mentor to JSR.

YF and NAO (Vice Chair of the Department of Quantitative Health Sciences, Cleveland Clinic) did the statistical data analysis. GET and RMB are the lead consultants for psychiatric/psychological data interpretation and recommendations. NF-S (Director of the Sleep Disorders Center and Staff Epileptologist, Cleveland Clinic), SIP, MJR contributed data from their respective studies. Three clinical fellows, including authors JSR (epilepsy), SIP (epilepsy), and MJR (neuropsychology) administered the M.I.N.I. to study participants. Training and supervision of M.I.N.I. administration were performed by RMB, a board-certified neuropsychologist, and GET, (Chairman, Department of Psychiatry and Psychology, Director of Behavioral Health and Director of General Adult Psychiatry Residency Program, Cleveland Clinic) a senior psychiatrist with dual-board certification in psychiatry and internal medicine. All listed authors were intellectually involved from conceptualization to the completion of this study and contributed in the revision of the manuscript.

Financial Disclosure, Role of the funding source, Conflict of Interest and Ethical Statement.

This study had institutional funding only (Cleveland Clinic Foundation, Cleveland, OH, U.S.A) and there was no external funding source. There is no conflict of interest and no disclosures to report for any one among the authors. The corresponding author (JSR) had full access to all the data in the study and had final responsibility for the decision to submit for publication. We confirm that we have read *Epilepsy & Behavior Journal's* position on issues involved in ethical publication and we solemnly affirm that this report is consistent with those guidelines.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

⁷MINCEP Epilepsy Care, University of Minnesota Medical Center, Minneapolis, Minnesota, U.S.A

⁸Psychology Services, Tampa General Hospital, Tampa, Florida, U.S.A

Abstract

Objective—Assess accuracy and operating characteristics of the Patient Health Questionnaire-9 (PHQ-9) for depression-screening in adults with epilepsy.

Methods—Tertiary epilepsy center patients served as the study population with 237 agreeing to structured interview using the Mini-International Neuropsychiatric Interview (MINI), a “gold standard” instrument developed for rapid diagnosis of neuropsychiatric disorders, including major depressive disorder (MDD); 172 also completed the PHQ-9, and 127 completed both the PHQ-9 and the Neurological Disorders Depression Inventory for Epilepsy (NDDI-E) within two days of the MINI. Sensitivity, specificity, positive & negative predictive values & areas under the ROC curves for each instrument were determined. Cut-points of 10 for the PHQ-9 and 15 for the NDDI-E were used and ratings at or above the cut-points were considered screen-positive. The PHQ-9 was divided into cognitive/affective (PHQ-9/CA) and somatic (PHQ-9/S) subscales to determine comparative depression-screening accuracy.

Results—The calculated areas under the ROC curves for the PHQ-9 (n=172) and the PHQ-9/CA and PHQ-9/S sub-scales were 0.914, 0.924, and 0.846, respectively, with the PHQ-9 more accurate than the PHQ-9/S (p=0.002) but no different than the PHQ-9/CA (p=0.378). At cut-points of 10 and 15, respectively, the PHQ-9 had higher sensitivity (0.92 vs 0.87), but lower specificity (0.74 vs 0.89) than the NDDI-E. The areas under the ROC curves of the PHQ-9 and the NDDI-E showed similar accuracy (n=127; 0.930 vs 0.934; p=0.864).

Significance—The PHQ-9 is an efficient & non-proprietary depression screening instrument with excellent accuracy validated for use in adult epilepsy patients as well as multiple other medical populations.

Search Items/Keywords

Epilepsy/Seizures; Depression

Additional Search keywords

PHQ-9; NDDI-E; Screening

INTRODUCTION

Depression is the most common psychiatric disorder in patients with epilepsy ^(1, 2) and is nearly five times more common than in the general population ⁽³⁾. Its presence in patients with epilepsy has repeatedly been shown to compromise quality of life ^(4, 5, 6, 7) to be a significant risk factor for suicide ⁽⁸⁾, and to increase health care utilization ⁽⁹⁾. Yet, co-morbid depression often goes undetected ⁽¹⁰⁾ in patients with epilepsy and even with proper diagnosis, remains under-treated ^(11, 12).

Recognition and accurate diagnosis of depression are fundamental to its successful treatment⁽¹³⁾. The Patient Health Questionnaire (PHQ-9) was developed specifically for making criteria-based diagnoses of depression and other psychiatric disorders commonly seen in primary care. Its diagnostic validity was established in over 11,000 patients: the majority from primary care clinics⁽¹⁴⁾ with the rest from a variety of patient populations, including stroke and traumatic brain injury⁽¹⁵⁾. The PHQ-9 has comparable sensitivity and specificity with other depression screening measures, is more time-efficient than longer measures, and consists of the nine criteria upon which the DSM-IV diagnosis of major depressive disorder (MDD) is based. Additionally, it has established construct validity as a measure of depression severity^(16,17), making it useful for serial monitoring of depressive symptom burden⁽¹⁸⁾ and depression treatment outcome^(19,20).

The primary aim of this study is to measure the diagnostic accuracy of the PHQ-9 in patients with epilepsy using the Mini-International Neuropsychiatric Interview (M.I.N.I.) as a “gold-standard” diagnostic reference⁽²¹⁾. Few other studies evaluated PHQ-9 in patients with epilepsy, including two papers that focused on psychometrics^(22–27). This supports the notion that PHQ-9 can and indeed has been used to screen for depression in epilepsy. The findings in some of these papers further support the idea that somatic and non-somatic symptoms from PHQ-9 may contribute as reported by Mitchell et al⁽²⁵⁾ that four of these symptoms were rated as excellent initial screening questions for depression namely, “Moving or speaking so slowly that other people could have noticed” “Little interest or pleasure in doing things”, “Feeling down depressed or hopeless”, “Trouble concentrating on things such as reading.” The item “Moving or speaking so slowly that other people could have noticed” from the PHQ9 was endorsed in about 90% of depressed patients with epilepsy but only about 6% of non-depressed patients. Having said that, to the best of our knowledge our study is the first to attempt validation of the PHQ-9 in patients with epilepsy.

A secondary aim was to compare the PHQ-9 with the Neurological Disorders Depression Inventory for epilepsy (NDDI-E), an efficient, highly accurate depression screening instrument, developed specifically for use in patients with epilepsy⁽²⁸⁾. The NDDI-E was designed to eliminate the confounding influence of antiepileptic drugs (AED) side-effects such as impaired concentration, reduced energy, and sleep disturbance, similar to the principal somatic criteria for diagnosis of MDD. Unlike the PHQ-9, however, the NDDI-E has been validated for use only in patients with epilepsy. The broader applicability of the PHQ-9 makes it of value not only to clinicians who see patients in settings other than epilepsy centers, but to investigators studying comparative depression epidemiology and management across medical populations. Another potential advantage of the PHQ-9 is its validation for serial use during management of depression^(18, 20).

Another study aim is to examine the differential impact on screen accuracy of the PHQ-9’s somatic (items 3, 4, 5 and 8) and cognitive/affective (items 1, 2, 6, 7, and 9) subscales as previously assessed in cardiac patients^(29, 30). Our hypothesis is that operating characteristics of the cognitive/affective subscale (PHQ-9/CA) would perform better than those of either the somatic subscale (PHQ-9/S) or the PHQ-9 total score (PHQ-9/T), like the NDDI-E; by eliminating the confounding influence of somatic factors such as AED-induced fatigue or impaired concentration that mimics depression.

METHODS

Patient-selection, Standard Protocol Approvals, Registrations, and Patient Consents

Study participants were recruited between 2009 and 2013 from the Cleveland Clinic Epilepsy Center's epilepsy monitoring unit (EMU) and outpatient clinic. The study protocol and informed consent were approved by Cleveland Clinic's Institutional Review Board, and patients provided with written informed consent before study enrollment. Selection criteria included: age of 18 years or older, current diagnosis of electroencephalographically confirmed epilepsy requiring treatment with one or more antiepileptic drugs (AED); and fluency in English with sufficient cognitive capacity to complete the structured diagnostic interview and self-rated depression screening measures. Exclusion criteria included: presence of confusion due to psychosis, delirium, amnesic disorder, intoxication, or postictal state; the development of confusion for any reason during testing procedures; and seizures identified as non-epileptic.

Diagnostic instrument

The Mini-International Neuropsychiatric Interview (M.I.N.I.) is a validated, structured, rater-administered diagnostic interview intended to render a dichotomous classification of psychiatric diagnoses in patients with pre-existing neurological disorders⁽²¹⁾. Design of the M.I.N.I. was patterned after the Structured Clinical Interview for DSM-III-R (SCID)⁽³¹⁾, but abbreviated for more rapid completion. Disorders included in the M.I.N.I. were selected based on the 12-month prevalence of psychiatric disorders reported in the Epidemiologic Catchment Areas (ECA) Study⁽³²⁾ and the National Comorbidity Survey Replication (NCS-R)⁽³⁾.

A computerized version of M.I.N.I. 6.0.0 was used to gather interview data from Modules A-J (i.e., mood and anxiety disorders), and N (i.e., suicide risk) (dsheehan@health.usf.edu). This study relied exclusively on data from the mood disorders modules, A, B, C and N.

Screening measures

The self-administered nine-item Patient Health Questionnaire-9 (PHQ-9) was designed for use in primary care populations⁽¹⁶⁾. The nine items pertain to the DSM⁽³³⁾ criteria for MDD: (1) Anhedonia, (2) Depressed mood, (3) Trouble sleeping, (4) Feeling tired, (5) Change in appetite, (6) Guilt, self-blame or worthlessness, (7) Trouble concentrating, (8) Feeling slowed down or restless, and (9) Thoughts of being better off dead or hurting oneself. Critical to the diagnosis of MDD is the patient's endorsement of either items (1), (2) or both. Each item is rated on a 4-point scale from 0 to 3 (0 – Never; 1 – Several days; 2 – More than half the time; and 3 – Nearly every day) during the two weeks prior to and including the day of survey completion. The total score ranges from 0 to 27 and a score of 10 or greater represents depressive symptoms of at least moderate severity and is the most commonly-used cut-point when screening for MDD. Other cut-points help stratify the severity of current MDD (minimal, 0–4; mild, 5–9; moderate, 10–14; moderate-to-severe, 15–19; and severe, 20 or more)⁽¹⁶⁾.

The Neurological Disorders Depression Inventory for Epilepsy (NDDI-E)⁽²⁸⁾ is a six-item, patient-rated survey of symptoms lasting at least two weeks prior to and including the day of the survey. The six items (1 – Everything is a struggle; 2 – Nothing I do is right; 3 – Feel guilty; 4 – I'd be better off dead; 5 – Frustrated; and 6 – Difficulty finding pleasure) are rated on a 4-point scale (4 – Always or Often, 3 – Sometimes, 2 – Rarely, and 1 – Never). A score of 15 or more predicts a high likelihood of a current MDD⁽²⁸⁾.

Procedure

Three clinical fellows in accredited fellowships (JSR [Epilepsy], SIP [Epilepsy], and MJR [Neuropsychology]) administered the M.I.N.I. to study participants; each was trained in M.I.N.I.-administration by either GET or RMB. The PHQ-9 and NDDI-E were patient-rated. Whenever possible, the M.I.N.I, PHQ-9 and NDDI-E were administered in random order to eliminate the potential systematic influence of a fixed-testing sequence on patient interview responses and self-ratings.

The prevalence of MDD and other mood disorders was derived from the 237 patients interviewed with the M.I.N.I. One hundred seventy two (172) completed the PHQ-9 within two days of the M.I.N.I. with a complete dataset present in 158 (the group on which sensitivity and specificity analyses were performed). Further analysis was conducted on 127 who completed both the PHQ-9 and the NDDI-E within two days of the M.I.N.I.

Statistical analysis

Receiver operating characteristic (ROC) curves were constructed to characterize the diagnostic accuracy of the PHQ-9 and NDDI-E. Nonparametric methods⁽³⁴⁾ were used to estimate the areas under the ROC curves and construct 95% confidence intervals (CI). A Wald test was used to test the null hypothesis that the ROC curve areas of the PHQ-9 and NDDI-E are equivalent. A significance level of 0.05 (two-tailed) was applied.

The PHQ-9 sensitivity and specificity were assessed at varying cut-points to determine which represented optimum sensitivity and specificity. A score of 15 or more was used to measure the accuracy of the NDDI-E; 95% Wilson CI^(35, 36). McNemar's exact test⁽³⁷⁾ was used to compare the sensitivity or specificity between PHQ-9 at varying cut-points and NDDI-E at cut-point of 15. Positive and negative predictive values (PPV and NPV) were estimated

Based on studies in cardiac patients^(29, 30), the PHQ-9 was deconstructed into its cognitive/affective (items 1, 2, 6, 7, and 9) and somatic (items 3, 4, 5, and 8) domains. Logistic regression models were used to assess the associations between cognitive/affective and somatic symptom scores for detecting depression. Areas under ROC curves for the PHQ-9 total (PHQ-9/T) and somatic (PHQ-9/S) and cognitive/affective (PHQ-9/CA) subscale scores were compared⁽³⁴⁾.

RESULTS

Demographics

Of the 237 participants, 163 (68.8%) were women. Mean age (\pm SD) of the entire sample was 40.6 ± 14.2 years (range, 18–77).

M.I.N.I.-diagnosed mood disorders (Table I)

Eighty-four patients (35.4%) met criteria for a life-time diagnosis of mood disorder; 72 (30.3%) had current or life-time MDD and 12 (5.1%) had life-time bipolar disorder (BPD). Three of the 72 had a first life-time diagnosis of MDD (single episode: mild, moderate, or severe [296.21–296.23]), leaving 69 with a diagnosis of MDD, recurrent, including 34 in remission [DSM-IV: 296.36] and 35 (14.8% of total) with current MDD or MDD, recurrent: mild, moderate or severe [DSM-IV: 296.31–296.33]. Ten additional major depressive episodes occurred in patients with a history of BPD (none of whom were currently manic or hypomanic) for a total of 45 (19.0%) with current MDD.

Seventeen had a M.I.N.I.-estimated suicide risk of moderate (3.8%) or severe (3.4%) (Table II). None of these were judged to require urgent or emergent clinical suicide management.

Operating characteristics of the PHQ-9 and NDDI-E

The estimated area under the ROC curve of the PHQ-9 was 0.905 (N=172, SE=0.029, 95% Wald CI [0.848, 0.962]) (Figure 1A) and of the NDDI-E was 0.935 (N=127, SE=0.024, 95% Wald CI [0.887, 0.983]) (Figure 1B). The difference did not achieve statistical significance ($p=0.864$). The estimated sensitivity, specificity, PPV, and NPV at each PHQ-9 cut-point from 10 to 15 and corresponding values at the NDDI-E cut-point of 15 are summarized in Table III.

At cut-points of 10 and 15, respectively, the PHQ-9 had a better sensitivity (0.92 vs 0.87, $p = 0.50$ and $N = 30$) and lower specificity (0.74 vs 0.89, $p = 0.03$, $N = 97$) than the NDDI-E. At a cut-point of 13, PHQ-9 specificity equaled that of the NDDI-E, but sensitivity fell to 0.69.

Accuracy of the Somatic and Cognitive/Affective subscales of the PHQ-9

Comparison of the PHQ-9 total score (PHQ-9/T) with its somatic (PHQ-9/S) (items 3, 4, 5, and 8) and cognitive/affective (PHQ-9/CA) (items 1, 2, 6, 7, and 9) subscale scores is depicted in Figure 2. The estimated differences in ROC curve areas between the PHQ-9/S and the PHQ-9/T (0.068, SE=0.016, 95% Wald CI [0.036, 0.10], $p<0.0001$) and the PHQ-9/CA (0.078, SE=0.026, 95% Wald CI [0.028, 0.128], $p=0.002$) were highly significant, whereas the estimated difference between PHQ-9/CA and PHQ-9/T was not (0.011, SE=0.012, 95% Wald CI [-0.013, 0.034], $p=0.378$).

DISCUSSION

The point prevalence of MDD in this sample, 19.0%, is similar to the 17.2% identified by Jones and colleagues⁽¹⁾ and the 17% identified by Gilliam and colleagues⁽²⁸⁾ both of who

used a structured clinical diagnostic interview to identify depression. Notably, 7.2% of the current sample endorsed moderate or severe suicidal ideation although no one was judged to present an imminent risk.

Both PHQ-9 and NDDI-E performed well with no significant differences in their ROC curve areas, 0.930 vs 0.934 ($N = 127$, $p=0.864$), reflecting similar overall test accuracy. At a cut-point of 10, the PHQ-9 sensitivity of 0.92 compared favorably with the 0.87 of the NDDI-E ($n=30$, $p = 0.50$) but the NDDI-E's specificity was higher (0.87 vs 0.74, $n=97$, $p = 0.03$). The PPVs of the PHQ-9 and the NDDI-E were 0.46 and 0.67, respectively, estimated directly from the sample. Table III depicts the relationship between PHQ-9 cut-point and operating characteristics. Raising the PHQ-9 cut-point to 13 achieves greater specificity (0.90) with fewer false-positives (i.e., non-depressed patients incorrectly labeled as depressed), but at the expense of reduced sensitivity (0.69), or more false-negatives (i.e., depressed patients incorrectly labeled as non-depressed). These data reinforce the contention^(28,1) that the NDDI-E has superior specificity when compared to other depression screening measures such as the Beck Depression Inventory (BDI-II)⁽³⁸⁾ or Center for Epidemiologic Studies Depression rating scale (CES-D)⁽³⁹⁾. From this perspective, use of the NDDI-E is likely to enhance clinical efficiency by reducing unnecessary time and attention devoted to patients incorrectly identified as depressed.

The PHQ-9's comparatively low specificity at a cut-point of 10 may warrant increasing it to as high as 15 in patients with epilepsy. In primary care populations, a diagnosis of MDD is much less likely in patients with PHQ-9 scores less than 10 and much greater with scores of 15 or more with scores of 10–14 representing a “gray zone” of increasing specificity and decreasing sensitivity^(16,40). In this sample, increasing the PHQ-9 cut-point to 13 achieves a specificity of 0.90 (comparable to the NDDI-E=0.90). While PHQ-9 sensitivity falls to 0.69, estimated depression symptom-burden of the false-negatives is mild-to-moderate as measured by the mean PHQ-9 of this sub-group ($n=11$; mean \pm SD = 9.27 ± 3.2 , [range, 1–12]). If the PHQ-9 cut-point is increased to 15, the PHQ-9 mean \pm SD changes to 9.58 ± 3.23 ($n=12$; range, 1–13). In our experience patients with PHQ-9 total scores in the 10–14 range often do not view themselves as depressed or in need of antidepressant medication, and rarely require immediate or emergent attention; instead they warrant closer attention at follow-up. Such serial use of the PHQ-9 can be very helpful during longitudinal patient monitoring. Routine, computer-assisted capture of patient-rated health status measures – including the PHQ-9 – can provide valuable information at the point-of-care as well as a growing repository of longitudinal data for future trending⁽⁴¹⁾. Moreover, PHQ-9 has been selected as an indicator of depression quality of care by the National Quality Forum⁽⁴²⁾ and as reported by Griffith, Thompson, Rathore et al., 2014⁽⁴³⁾ that PHQ-9 could also be used in patient reported outcome measures (PROM) enhanced clinical phenotyping (depression) in the electronic health record data for research. The fact that it has been translated into more than 80 languages (www.phqscreeners.com) makes it easy to use and bestows a universal appeal.

The purpose for subdividing the PHQ-9 into somatic and cognitive affective components is meant to address the problem that inspired development of the NDDI-E, namely, eliminating the potential confound of epilepsy-specific factors, in particular, sedating effects of anti-

epileptic drugs (AED) simulating somatic features of depression. Our hypothesis that the PHQ-9 cognitive/affective subscale score (PHQ-9/CA) would be more accurate than the PHQ-9 total score (PHQ-9/T) or the somatic subscale score (PHQ-9/S) proved correct in part. As predicted the PHQ-9/S was less accurate than the other two measures (Figure 2). In contrast, the predicted superiority of the PHQ-9/CA proved incorrect. Nevertheless, the PHQ-9 includes all 9 DSM symptoms for depression, allowing it to be potentially used as a diagnostic measure for likely MDD. Although antiepileptic medications could potentially confound some of the somatic symptoms included in the PHQ-9, it is very much likely that such symptoms may also be accounted to a certain extent by depression. Indeed, this confounding by somatic symptoms has been considered for other medical illnesses, and some empiric data support that while the exclusive-etiological approach identifies the most severe and persistent depressions, the inclusive approach is the most sensitive and reliable approach and is an intermediate predictor of persistent depression^(44,45). Thus, along with the cognitive symptoms, one must still inquire about the somatic symptoms that constitute core criteria for MDD. There is a wealth of literature on the cut-points for the ordinal categories of PHQ-9, and the fact that a 3–5 point change on the PHQ-9 is clinically significant, one should consider these issues when deciding whether to use the full PHQ-9, or a partial subset of symptoms (**Link to Supplementary data: PHQ-2 as Ultra-brief Depression Screener: A Comparison Among PHQ-2, PHQ-9 and NDDI-E**). Insufficient power to demonstrate a difference may account for this finding, since it seems logical to expect the somatic items to detract from the accuracy of the PHQ-9 total score. Further study of larger samples using similar methodology or principles of Item Response Theory (IRT)⁽⁴⁶⁾ may be useful.

The accuracy of a screening measure is clearly important, but, however accurate it may be must not be mistaken as a diagnostic instrument; clinical confirmation is essential. Also, its effect on the management and outcome of a disorder such as MDD is different than other medical disorders. First, the “gold standard” diagnostic reference itself relies on subjective report. Moreover, delivering a diagnosis of depression does not necessarily ensure patient acceptance. This is often true for patients who prioritize management of their epilepsy, have depressive symptoms that are mild and not necessarily disabling, or are troubled by the stigma of a psychiatric diagnosis and its treatment. Patients in the PHQ-9 mild or “gray-zone” can be informed of possible depression not necessarily requiring immediate treatment, but worthy of closer observation.

Potential limitations of this study deserve consideration. The patient sample was derived from a tertiary epilepsy center and may therefore not be typical of a community sample of patients with less complex epilepsy syndromes. Also, patients were examined by three different individuals at different intervals without assuring inter-rater reliability, and recruitment strategies were not necessarily uniform across the three studies, one of which recruited females only. These potential limitations are balanced by the study’s strengths, including a comparatively large sample derived from a diverse population of local, regional, national and international patients at a level-4 epilepsy center.

CONCLUSIONS

Depression is a common psychiatric co-morbidity that adversely affects epilepsy outcomes. Recognition of this problem and the growing attention of professional, public and regulatory bodies to healthcare outcomes demand the formation of integrated, interdisciplinary healthcare teams. Standardization of care and procedures across healthcare systems is an ideal that creates efficiencies, uniformity and the opportunity to amass large repositories of reliable, clinically relevant data. Depression screening measures such as the PHQ-9, applicable in a wide variety of neurological and non-neurological disorders; are an essential feature of disease management strategies focused on standardization and continuous improvement of healthcare delivery across both tertiary and primary care settings. This study, which confirms the accuracy in epilepsy patients of the already widely-used PHQ-9, is a small but very important step towards achieving this goal.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

1. Jones JE, Bell B, Fine J, et al. A controlled prospective investigation of psychiatric comorbidity in temporal lobe epilepsy. *Epilepsia*. 2007; 48:2357–60. [PubMed: 17651410]
2. Tellez-Zenteno JF, Patten SB, Jetté N, et al. Psychiatric comorbidity in epilepsy: a population-based analysis. *Epilepsia*. 2007; 48:2336–2344. [PubMed: 17662062]
3. Kessler RC, McGonagle KA, Zhao S, et al. Lifetime and 12-month prevalence of DSM-II-R psychiatric disorders in the United States: results from the National Comorbidity Survey. *Arch Gen Psychiatry*. 1994; 51:8–19. [PubMed: 8279933]
4. Boylan LS, Flint LA, Labovitz DL, et al. Depression but not seizure frequency predicts quality of life in treatment-resistant epilepsy. *Neurology*. 2004; 62:258–61. [PubMed: 14745064]
5. Gilliam F, Hecimovic H, Sheline Y. Psychiatric comorbidity, health, and function in epilepsy. *Epilepsy Behav*. 2003; 4(suppl 4):S26–30. [PubMed: 14654425]
6. Jehi L, Tesar G, Obuchowski N, et al. Quality of life in 1931 adult patients with epilepsy: seizures do not tell the whole story. *Epilepsy Behav*. 2011 Dec; 22(4):723–7. [PubMed: 22019018]
7. Vickrey BG, Berg AT, Sperling MR, et al. Relationships between seizure severity and health-related quality of life in refractory localization-related epilepsy. *Epilepsia*. 2000; 41:760–4. [PubMed: 10840410]
8. Christensen J, Vestergaard M, Mortensen PB, et al. Epilepsy and risk of suicide: a population-based case-control study. *Lancet Neurol*. 2007; 6:693–698. [PubMed: 17611160]
9. Cramer JA, Blum D, Fanning K, et al. Impact Project Group: the impact of comorbid depression on health resource utilization in a community sample of people with epilepsy. *Epilepsy & Behav*. 2004; 5:337–342.
10. Gilliam FG, Santos J, Vahle V, et al. Depression in epilepsy: ignoring clinical expression of neuronal network dysfunction? *Epilepsia*. 2004; 45:28–33. [PubMed: 15186342]
11. Franco, E.; Tesar, GE. Epileptologist Adherence to Depression Treatment Guidelines. 64th Annual Meeting of the American Epilepsy Society; December 3–7, 2010; San Antonio, TX. (Poster presentation)
12. Weigartz P, Seidenberg M, Woodard A, et al. Co-morbid psychiatric disorder in chronic epilepsy: recognition and etiology of depression. *Neurology*. 1999; 53(5 suppl2):S3–S88.
13. Pignone MP, Gaynes BN, Rushton JL, et al. Screening for depression in adults. A summary of the evidence for the U.S. Preventive Services Task Force. *Ann Intern Med*. 2002; 136:765–776. [PubMed: 12020146]

14. Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study: Primary Care Evaluation of Mental Disorders. Patient Health Questionnaire. *JAMA*. 1999; 282(18):1737–44. [PubMed: 10568646]
15. Gilbody S, Richards D, Brealey S, Hewitt C. Screening for depression in medical settings with patient Health Questionnaire (PHQ): a diagnostic meta-analysis. *J Gen Intern Med*. 2007; 22:1596–1602. [PubMed: 17874169]
16. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9: validity of a brief depression severity measure. *J Gen Internal Med*. 2001; 16:606–613. [PubMed: 11556941]
17. Kroenke K, Spitzer RL, Williams JBW, Löwe B. The Patient Health Questionnaire somatic, anxiety, and depressive symptom scales: a systematic review. *Gen Hosp Psychiatry*. 2010; 32:345–359. [PubMed: 20633738]
18. Lowe B, Kroenke K, Herzog, et al. Measuring depression outcomes with a brief self-report instrument: sensitivity to change of the PHQ-9. *J Affective Dis*. 2004; 81:61–66.
19. Katzelnick DJ, Duffy FF, Chung H, et al. Depression outcomes in psychiatric clinical practice: using a self-rated measure of depression severity. *Psychiatric Services*. 2011; 62:929–935. [PubMed: 21807833]
20. Lowe B, Schenkel I, Comey-Doebbeling C, et al. Responsiveness of the PHQ-9 to psychopharmacological depression treatment. *Psychosomatics*. 2006; 47:62–67. [PubMed: 16384809]
21. Sheehan DV, Lecrubier Y, Sheehan KH, et al. The Mini-International Neuropsychiatric Interview (MINI). The development and validation of a structured diagnostic interview for DSM-IV and ICD-10. *J Clin Psychiatry*. 1998; 59(suppl 20):22–33. [PubMed: 9881538]
22. Seminario NA, Farias ST, Jorgensen J, et al. Determination of prevalence of depression in an epilepsy clinic using a brief DSM-IV-based self-report questionnaire. *Epilepsy Behav*. 2009; 15:362–366. [PubMed: 19525150]
23. Haut SR, Katz M, Masur J, et al. Seizures in the elderly: Impact on mental status, mood, and sleep. *Epilepsy & Behavior*. 2009; 14(3):540–544. [PubMed: 19189862]
24. Margrove K, Mensah S, Thapar A, et al. Depression screening for patients with epilepsy in a primary care setting using the Patient Health Questionnaire-2 and the Neurological Disorders Depression Inventory for Epilepsy. *Epilepsy & Behavior*. 2011; 21(4):387–390. [PubMed: 21742562]
25. Mitchell AJ, Ioannou N, Rampling JM, et al. Which symptoms are indicative of depression in epilepsy settings? An analysis of the diagnostic significance of somatic and non-somatic symptoms. *Journal of Affective Disorders*. 2013; 150(3):861–867. [PubMed: 23668901]
26. Thompson AW, Miller JW, Katon W, et al. Sociodemographic and clinical factors associated with depression in epilepsy. *Epilepsy & Behavior*. 2009; 14(4):655–660. [PubMed: 19233316]
27. Walker ER, Engelhard G, Thompson NJ. Using Rasch measurement theory to assess three depression scales among adults with epilepsy. *Seizure*. 2012; 21:437–443. [PubMed: 22608975]
28. Gilliam FG, Barry JJ, Hermann BP, et al. Rapid detection of major depression in epilepsy: a multicenter study. *Lancet Neurol*. 2006; 5:399–405. [PubMed: 16632310]
29. de Jonge P, Mangano D, Whooley MA. Differential association of cognitive and somatic depressive symptoms with heart rate variability in patients with stable coronary heart disease: Findings from the Heart and Soul Study. *Psychosom Med*. 2007; 69:735–739. [PubMed: 17942844]
30. Smolderen KG, Spertus JA, Reid KJ, et al. The association of cognitive and somatic depressive symptoms with depression recognition and outcomes after myocardial infarction. *Circ Cardiovasc Qual Outcomes*. 2009; 2:328–337. [PubMed: 20031858]
31. Spitzer RL, Williams JB, Gibbon M, et al. The Structured Clinical Interview for DSM-III-R (SCID). I: History, rationale, and description. *Arch Gen Psychiatry*. 1992; 49:624–629. [PubMed: 1637252]
32. Regier DA, Myers JK, Kramer M, et al. The NIMH Epidemiologic Catchment Area program. Historical context, major objectives, and study population characteristics. *Arch Gen Psychiatry*. 1984; 41:934–41. [PubMed: 6089692]

33. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. 3. Washington DC: American Psychiatric Association; 1987. Revised
34. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic (ROC) curves: A Nonparametric Approach. *Biometrics*. 1988; 44:837–845. [PubMed: 3203132]
35. Brown LD, Cai TT, DasGupta A. Interval estimation for a binomial proportion. *Statistical Science*. 2001; 16:101–133.
36. Wilson EB. Probable inference, the law of succession, and statistical inference. *J Am Statistical Assoc*. 1927; 22:209–212.
37. Zhou, X.; Obuchowski, NA.; McClish, DK. *Statistical Methods in Diagnostic Medicine*. 2. John Wiley & Sons, Inc; 2011.
38. Beck, AT.; Steer, RA.; Brown, GK. *Manual for the Beck Depression Inventory-II*. San Antonio, TX: Psychological Corporation; 1996.
39. Radloff LS. The CES-D Scale: a self-report depression scale for research in the general population. *Appl Psychol Measurement*. 1977; 1:385–401.
40. Manea, Laura, MSc; Gilbody, Simon, PhD; McMillan, Dean, PhD. Optimal cut-off score for diagnosing depression with the Patient Health Questionnaire (PHQ-9): a meta-analysis. *CMAJ*. Feb 21; 2012 184(3):E191–E196. [PubMed: 22184363]
41. Katzan, I.; Speck, M.; Dopler, C., et al. The Knowledge Program: an innovative, comprehensive electronic data capture system and warehouse. *Ann Symp Proc/AMIA Symposium*; 2011; p. 683-92.
42. Basch, Ethan, MD, MSc; Torda, Phyllis, MA; Adams, Karen, PhD. Standards for Patient-Reported Outcome–Based Performance Measures. *JAMA*. 2013; 310(2):139–140. [PubMed: 23839744]
43. Griffith, SD.; Thompson, NR.; Rathore, JS., et al. Incorporating patient-reported outcome measures into the electronic health record for research: Application using the Patient Health Questionnaire-9 (PHQ-9). *Quality of Life Research*; Jun. 2014 In press
44. Koenig HG, George LK, Peterson BL, et al. Depression in medically ill hospitalized older adults: prevalence, characteristics, and course of symptoms according to six diagnostic schemes. *Am J Psychiatry*. 1997 Oct; 154(10):1376–83. [PubMed: 9326819]
45. Simon GE, Von Korff M. Medical co-morbidity and validity of DSM-IV depression criteria. *Psychol Med*. 2006 Jan; 36(1):27–36. [PubMed: 16202189]
46. Kendel F, Wirtz M, Dunkel A, et al. Screening for depression: Rasch analysis of the dimensional structure of the PHQ-9 and the HADS-D. *J Affective Dis*. 2010; 122:241–246.

Highlights

- Validation of PHQ-9 for depression screening in adults with epilepsy.
- Diagnostic accuracy of PHQ-9, PHQ-2 & NDDI-E tested using “gold standard” MINI test.
- High powered prospective study at level 4 epilepsy center with diverse cohort.
- Offers free non-proprietary alternative to screen depression in epileptic adults.
- Highlights relatively under reported & treated psychiatric comorbidity in epilepsy

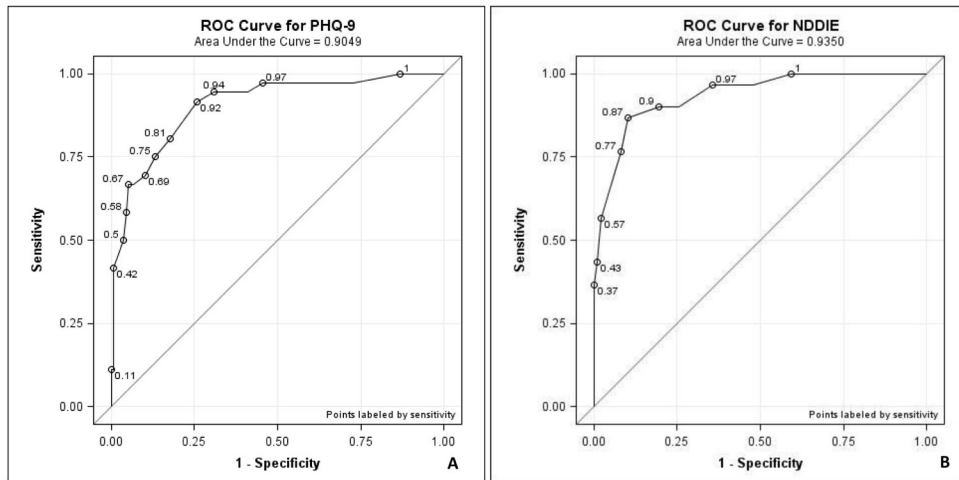


Figure 1.

A. the PHQ-9 receiver operating characteristic (ROC) curve. Points on the curve from left to right represent sensitivity and specificity at PHQ-9 scores of 23, 18, 17, 16, 15, 13, 12, 11, 10, 9, 6, and 1, respectively.

B. ROC curve for the NDDI-E. Points on the curve from left to right represent sensitivity and specificity at NDDI-E cut-points of 19, 18, 17, 16, 15, 14, 12, and 9, respectively.

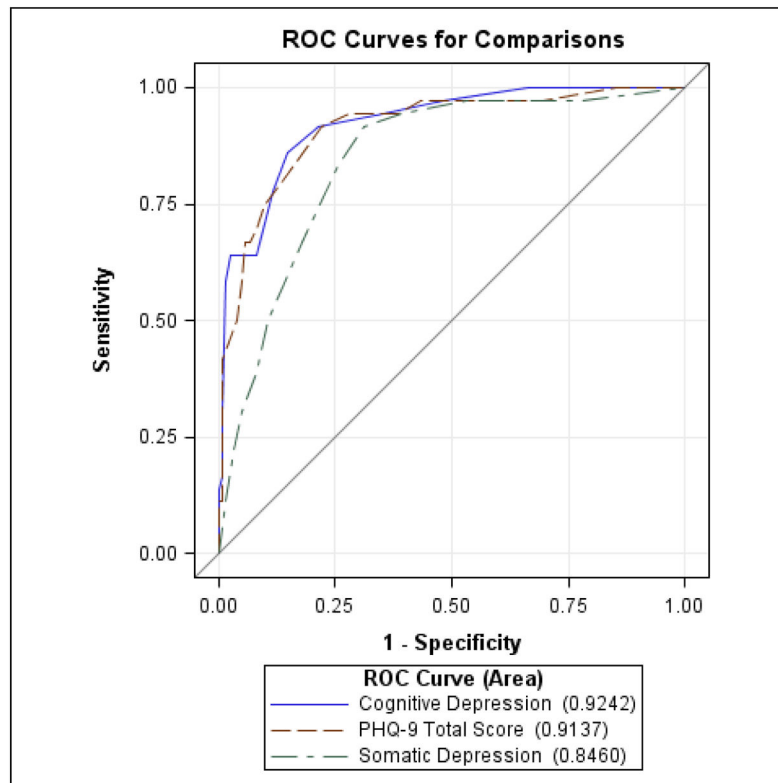


Figure 2.

Comparison of areas under the ROC curves of PHQ-9 total and somatic and cognitive/ affective subscale scores. The estimated difference in ROC area between PHQ-9 total score and cognitive/affective is 0.011 with standard error as 0.012 and 95% Wald confidence interval as (-0.013, 0.034) and the difference is not significant ($p = 0.378$). The estimated difference in ROC area between PHQ-9 total score and the somatic components only is 0.068 with standard error as 0.016 and 95% Wald confidence interval as (0.036, 0.10) and the difference is statistically significant ($p < 0.0001$). The estimated difference in ROC area between cognitive/affective and somatic is 0.078 with standard error as 0.026 and 95% Wald confidence interval as (0.028, 0.128) and the difference is statistically significant ($p = 0.002$).

Table I

M.I.N.I diagnoses of MDD and BPD among 237 study subjects.

Diagnosis		n	%
Major depressive disorder (MDD)	Current	35	14.8
	Life-time	72	30.3
Bipolar disorder (BPD)	Current	10	4.2
	Life-time	12	5.1

M.I.N.I = Mini-International Neuropsychiatric Interview. ²¹

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table II

M.I.N.I.-estimated suicide risk and concurrent PHQ-9 scores (Mean \pm SD)

	M.I.N.I.-estimated suicide risk							
	None		Low		Moderate		Severe	
	N	%	n	%	n	%	n	%
	179	76.2	39	16.1	9	3.8	8	3.4
PHQ-9 (Mean \pm SD)	6.87 \pm 6.12		12.33 \pm 6.81		15.11 \pm 5.86		19.0 \pm 4.75	

Table III

PHQ-9 and NDDI-E sensitivity, specificity, PPV and NPV at various cut-points.

Cut-point	Sensitivity		Specificity		PPV	NPV
	Point estimate	Wilson 95% CI	Point estimate	Wilson 95% CI		
PHQ-9						
10	0.92	[0.782, 0.971]	0.74	[0.663, 0.809]	0.46	0.97
11	0.81	[0.650, 0.902]	0.82	[0.751, 0.878]	0.52	0.95
12	0.75	[0.589, 0.863]	0.87	[0.801, 0.915]	0.57	0.94
13	0.69	[0.531, 0.820]	0.90	[0.835, 0.938]	0.61	0.93
14	0.67	[0.503, 0.798]	0.93	[0.879, 0.965]	0.70	0.92
15	0.67	[0.503, 0.798]	0.95	[0.898, 0.975]	0.75	0.92
NDDI-E						
15	0.87	[0.703, 0.947]	0.89	[0.822, 0.945]	0.67	0.97

PPV = positive predictive value, NPV= negative predictive value