

Brief Report

The Patient Health Questionnaire depression screener in spinal cord injury

Julia M. P. Poritz¹, Joseph Mignogna ^{2,3,4}, Aimee J. Christie ^{5,6}, Sally A. Holmes^{5,6}, Herb Ames^{5,6}

¹TIRR Memorial Hermann, Houston, Texas, USA, ²VISN 17 Center of Excellence for Research on Returning War Veterans, Waco, Texas, USA, ³Central Texas Veterans Health Care System, Temple, Texas, USA, ⁴Texas A&M College of Medicine, Temple, Texas, USA, ⁵Michael E. DeBakey VA Medical Center, Houston, Texas, USA, ⁶Baylor College of Medicine, Houston, Texas, USA

Context: Although depression is not inevitable following spinal cord injury/dysfunction (SCI/D), it can have a negative impact on rehabilitation. Evidence-based assessment of depression utilizing self-report instruments, such as the Patient Health Questionnaire-9 (PHQ-9), is considered good clinical practice. Although the PHQ-9 has been studied in individuals with SCI/D, little is known about the clinical utility of the Patient Health Questionnaire-2 (PHQ-2). Traditional cutoff scores for the PHQ-2 were examined to explore their operating characteristics as related to PHQ-9 results.

Methods: Archival data were collected for 116 Veterans with SCI/D who completed the PHQ-2 and PHQ-9 as one component of their routine, comprehensive SCI annual evaluation at a Veterans Affairs Medical Center. Logistic regressions were performed to determine the impact of different cutoff scores for the PHQ-2 on the likelihood that participants would endorse clinically significant levels of depressive symptoms on the PHQ-9 (≥ 10).

Results: Using a cutoff score of 3 or greater correctly classified 94.8% of the cases, outperforming the other cutoff scores. A cutoff score of 3 or greater had a sensitivity of 83.3% and specificity of 97.8%, and yielded a positive predictive value of 90.9% and a negative predictive value of 95.7%.

Conclusion: The PHQ-2 shows promise as a clinically useful screener in the community-residing SCI/D population. Findings regarding the presence of suicidal ideation emphasize the importance of routine screening for depressive symptomatology in the SCI/D population. Future research should investigate the role of the PHQ-2 in clinical decision-making and treatment monitoring.

Keywords: Spinal cord injury, Depression, Screening, Patient Health Questionnaire

Introduction

Although variability exists in psychological adjustment following spinal cord injury/dysfunction (SCI/D), depression is the most common form of psychological distress following SCI/D and can have a negative impact on rehabilitation.¹ Specifically, depression in individuals with SCI/D is associated with fewer hours out of bed, fewer days out of the house, less engagement in productive activities, and less planned exercise.^{2,3} Estimates of the prevalence of Major Depressive Disorder in individuals with SCI/D range from about 11% or 12% to 21% 1 year post-injury,⁴⁻⁶ and from about 10% to 18% 5 years post-injury.^{4,6} In

community-residing individuals with SCI/D, about 23% meet the criteria for Major Depressive Disorder.⁷ In comparison, the lifetime prevalence of Major Depressive Disorder in the general U.S. population is 16.6%.⁸ Considering the prevalence and implications of depression in individuals with SCI/D, best practice guidelines recommend evidence-based depression assessment; however, a disconnect remains between best evidence and current practice.⁹

A systematic review of depression measures used in SCI/D research suggested that no one instrument is psychometrically superior. Instead, it was concluded that selection of a depression measure should be made based on other factors, including feasibility, acceptability to patients, ease of administration and scoring, and the ability of the measure to serve additional purposes,

Correspondence to: Herb Ames, Michael E. DeBakey VA Medical Center, 2002 Holcombe Blvd., Houston, Texas 77030 USA. Email: Herbert.Ames2@va.gov

such as monitoring therapy outcomes.¹⁰ Indeed, one of the primary barriers to efficient and effective screening, diagnosis, and treatment of depressive disorders is the fact that many screening measures are considered too long (i.e. 20 plus items) to be practical in routine clinical practice. An additional barrier is that most depression screening measures do not relate to diagnostic criteria for Major Depressive Disorder, requiring clinicians to conduct a separate diagnostic evaluation.⁵

The Patient Health Questionnaire-9 (PHQ-9) addresses these barriers for several reasons.¹¹ Consisting of only nine items, the PHQ-9 is more acceptable to patients and easier for clinicians to use routinely. Importantly, the nine items parallel the diagnostic criteria for Major Depressive Disorder, making it both a depression severity measure and a diagnostic instrument. Moreover, the PHQ-9 has been validated in an acute SCI/D rehabilitation sample.¹² Despite research demonstrating the PHQ-9's utility, it has not been adopted as the standardized approach to depression assessment during acute SCI/D rehabilitation or at annual evaluations.⁹ This may be due to its length,¹³ or because of the inclusion of an item assessing suicidal ideation. Specifically, it is not uncommon for nurses and primary care physicians to complete their education without having received adequate training in suicide assessment.^{14,15} This lack of training has the potential to make health care professionals feel unprepared to complete a risk assessment if a patient endorses suicidal ideation on a self-report questionnaire.

Research has provided empirical support for use of the PHQ-2, an even shorter screener, in a VA urgent care clinic sample,¹⁶ and in U.S. primary care and obstetrics-gynecology clinic samples.¹⁷ Lowe *et al.* demonstrated that the diagnostic accuracy of the PHQ-2 is comparable with longer measures and that it is sensitive to change and accurately reflects different courses of depression.¹⁸ Since its development, the PHQ-2 has been used in primary care samples across various countries as well as in several other medical populations (Table 1). Although research has been conducted on use of the PHQ-2 in an acute SCI/D rehabilitation population,¹² additional evidence is needed to justify the clinical utility of this depression screener in the community-residing SCI/D population. The present study aims to address this evidence gap using data collected from Veterans with SCI/D who presented for their annual evaluation. It was hypothesized that the results of this study would establish the clinical utility of the PHQ-2 in community-residing individuals with SCI, thereby enabling providers to identify more efficiently which individuals may benefit from additional assessment and treatment.

Methods

Participants and procedure

Archival data were collected from a convenience outpatient sample of Veterans with SCI/D who presented to a Veterans Affairs Medical Center during 2012 for their routine, comprehensive SCI/D annual evaluation. These evaluations were completed by an interdisciplinary team of SCI/D providers. Of 144 Veterans assigned to be assessed by the senior author of this manuscript, 116 (80.6%) completed the PHQ-2 and PHQ-9 among other brief psychological screeners. Participants completed the measures in written or oral format; electronic medical records were reviewed to identify demographic and injury characteristics. This study was approved and monitored for compliance with ethical research practices by an Internal Review Board (H-30172) and the VA Research and Development Committee (12E01.H).

Measures

Patient Health Questionnaire-9 (PHQ-9)¹¹

The PHQ-9 is the nine item depression module from the full Patient Health Questionnaire.⁴⁴ It consists of the nine criteria upon which diagnoses of depressive disorders are based according to the Diagnostic and Statistical Manual of Mental Disorders. The PHQ-9 can assist clinicians in diagnosing depressive disorders as well as determining depressive symptom severity. Participants are instructed to indicate how often over the past 2 weeks they have been bothered by nine symptoms of depression with response options ranging from 'not at all' (0) to 'nearly every day' (3). Responses are summed to obtain a total score that can range from 0 to 27; scores of 5, 10, 15, and 20 serve as cutoff scores that indicate mild, moderate, moderately severe, and severe depression.¹¹ The internal consistency of the data was excellent in both a primary care sample of 3,000 ($\alpha = 0.89$) and in an obstetrics-gynecology sample of 3,000 ($\alpha = 0.86$), and it was demonstrated that the PHQ-9 discriminated well between individuals with and without depression (AUC = 0.95). It was concluded that scoring the PHQ-9 as a continuous measure and using a cutoff score of 10 or higher results in the most accurate predictions of Major Depressive Disorder when compared with independent diagnoses made by a mental health provider.¹¹ In the present study, the data had strong internal consistency, Cronbach's $\alpha = 0.95$.

Patient Health Questionnaire-2 (PHQ-2)¹⁷

The PHQ-2 consists of the first two items from the PHQ-9. In a primary care and obstetrics-gynecology

Table 1 PHQ-2 operating characteristics in medical samples.

Population	Authors/year	PHQ-2 Cutoff score	PHQ-2 Operating characteristics	PHQ-9 Cutoff score	PHQ-9 Operating characteristics
Primary Care	Kroenke <i>et al.</i> (2003) ¹⁷	≥3	Sensitivity = 83%; Specificity = 90%; PPV = 38.4%; AUC = 0.93	Not reported	AUC = 0.95
	Corson <i>et al.</i> (2004) ¹⁹	≥3	Sensitivity = 97%; Specificity = 91%; AUC = 0.94	Not reported	Not reported
	Chen <i>et al.</i> (2009) ²⁰	≥3	Sensitivity = 84%; Specificity = 90%; AUC = 0.92	≥9	Sensitivity = 86%; Specificity = 85%; AUC = 0.92
	Osorio <i>et al.</i> (2009) ²¹	≥3	Sensitivity = 97%; Specificity = 88%; PPV = 81%; NPV = 98%; AUC = 0.97	≥10	Sensitivity = 100%; Specificity = 98%; PPV = 97%; NPV = 100%; AUC = 0.998
	Arroll <i>et al.</i> (2009) ²²	≥2	Sensitivity = 86%; Specificity = 78%	≥10	Sensitivity = 74%; Specificity = 91%;
	Phelan <i>et al.</i> (2010) ²³	≥2	Sensitivity = 75%; Specificity = 67%; AUC = 0.81	≥9	Sensitivity = 88%; Specificity = 80%; AUC = 0.87
	Zuithoff <i>et al.</i> (2010) ²⁴	≥2	Sensitivity = 81%; Specificity = 76%; PPV = 34%; NPV = 96%; AUC = 0.83	≥6	Sensitivity = 82%; Specificity = 82%; PPV = 41%; NPV = 97%; AUC = 0.87
	Liu <i>et al.</i> (2011) ²⁵	≥2	Sensitivity = 88%; Specificity = 82%; AUC = 0.9	≥10	Sensitivity = 86%; Specificity = 93.9%; AUC = 0.96
	Inagaki <i>et al.</i> (2013) ²⁶	≥2	Sensitivity = 77%; Specificity = 95%; PPV = 54%; NPV = 98%; AUC = 0.95	≥4	Sensitivity = 86%; Specificity = 85%; PPV = 32%; NPV = 99%; AUC = 0.93
	Thapar <i>et al.</i> (2014) ²⁷	≥3	Sensitivity = 72.1%; Specificity = 82.1%; PPV = 53.5%; NPV = 91.2%; AUC = 0.87	≥11	Sensitivity = 83.6%; Specificity = 83%; PPV = 58.4%; NPV = 94.7%; AUC = 0.9
	Bhana <i>et al.</i> (2015) ²⁸	≥2	Sensitivity = 60%; Specificity = 84%; AUC = 0.76	≥9	Sensitivity = 49%; Specificity = 94%; AUC = 0.85
	Carey <i>et al.</i> (2015) ²⁹	≥3	Sensitivity = 91%; Specificity = 78%; PPV = 41%; NPV = 98%; AUC = 0.92	≥10	Not reported
	Hanlon <i>et al.</i> (2015) ³⁰	≥1	Sensitivity = 83.3%; Specificity = 60.8%; PPV = 11.7%; NPV = 98.3%; AUC = 0.78	≥5	Sensitivity = 83.3%; Specificity = 74.7%; PPV = 17.1%; NPV = 98.6%; AUC = 0.85
	Suzuki <i>et al.</i> (2015) ³¹	≥3	Sensitivity = 76%; Specificity = 82%; Overall Accuracy = 81%; PPV = 27%; NPV = 98%; AUC = .845	≥11	Sensitivity = 76%; Specificity = 81%; Overall Accuracy = 81%; PPV = 26%; NPV = 97%; AUC = 0.88
Stroke	de Man-van Ginkel, Gooskens, <i>et al.</i> (2012) ³²	≥2	Sensitivity = 100%; Specificity = 77%; PPV = 38%; NPV = 100%	≥10	Sensitivity = 100%; Specificity = 86%; PPV = 50%; NPV = 100%
	de Man-van Ginkel, Hafsteinsdottir, <i>et al.</i> (2012) ³³	≥2	Sensitivity = 75%; Specificity = 76%; AUC = 0.82	≥10	Sensitivity = 80%; Specificity = 78%; AUC = 0.87
Cardiovascular	Turner <i>et al.</i> (2012) ³⁴	≥1	Sensitivity = 77%; Specificity = 63%; AUC = 0.82	≥6	Sensitivity = 85%; Specificity = 63%; AUC = 0.82
	McManus <i>et al.</i> (2005) ³⁵	≥3	Sensitivity = 39%; Specificity = 92%; AUC = 0.84	≥10	Sensitivity = 54%; Specificity = 90%; AUC = 0.86
	Thombs <i>et al.</i> (2008) ³⁶	≥2	Sensitivity = 82%; Specificity = 79%; PPV = 52%; NPV = 94%; AUC = 0.84	≥6	Sensitivity = 83%; Specificity = 76%; PPV = 50%; NPV = 94%; AUC = 0.86
	Wang <i>et al.</i> (2015) ³⁷	≥3	Sensitivity = 85.7%; Specificity = 69.2%; PPV = 57.1%; NPV = 93.6%; AUC = 0.806	≥10	Sensitivity = 87.1%; Specificity = 83.5%; PPV = 58.7%; NPV = 95.6%; AUC = 0.877
HIV/AIDS	Monahan <i>et al.</i> (2008) ³⁸	≥3	Sensitivity = 91%; Specificity = 77%; PPV = 36.9; AUC = 0.91	Not reported	Not reported
Cognitive Impairment	Boyle <i>et al.</i> (2011) ³⁹	≥3	Sensitivity = 78%; Specificity = 71%; AUC = 0.81	≥10	Sensitivity = 89%; Specificity = 71%; AUC = 0.85
Epilepsy	Fiest <i>et al.</i> (2014) ⁴⁰	≥2	Sensitivity = 42.3%; Specificity = 87.3%; PPV = 35.5%; NPV = 90.1%; AUC = 0.75	≥9	Sensitivity = 82.6%; Specificity = 82.2%; PPV = 42.2%; NPC = 96.8%; AUC = 0.88
Migraine	Seo <i>et al.</i> (2015) ⁴¹	≥2	Sensitivity = 66.7%; Specificity = 90.3%; PPV = 74.3%; NPV = 86.6%; AUC = 0.785	≥7	Sensitivity = 79.5%; Specificity = 81.7%; PPV = 64.6%; NPV = 90.5%; AUC = 0.806
Multiple Sclerosis	Amtmann <i>et al.</i> (2015) ⁴²	≥2	Sensitivity = 70.8%; Specificity = 81%	≥10	Sensitivity = 93.8%; Specificity = 61.2%
	Patten <i>et al.</i> (2015) ⁴³	≥3	Sensitivity = 80%; Specificity = 93%; PPV = 64%; NPV = 36%; AUC = 0.943	≥11	Sensitivity = 95%; Specificity = 88.3%; PPV = 55.9%; NPV = 44.1%; AUC = 0.952

clinic sample of 6,000 individuals, the PHQ-2 discriminated well between individuals with and without depression (AUC = 0.93).¹⁷ In the present study, the data had strong internal consistency, Cronbach's alpha = 0.91.

Data analysis

Descriptive statistics were used to detail demographic and injury characteristics. Logistic regressions were performed to determine the impact of different cutoff scores for the PHQ-2 on the likelihood that participants would endorse clinically significant levels of depressive symptoms on the PHQ-9 (≥ 10). Receiver operating characteristic (ROC) curve analysis was also conducted.

Results

Demographic and injury characteristics of the 116 participants are reported in Table 2. The sample was mostly middle-aged, married, White men with incomplete traumatic injuries. These characteristics are comparable to the 2007 fiscal year national data reported on Veterans with SCI/D using VA services.⁴⁵

Using a PHQ-9 cutoff score of 10 or higher as a proxy for diagnosis, the estimated prevalence of Major Depressive Disorder in this sample was 20.7%. In this outpatient, non-psychiatric sample of Veterans with SCI/D, 12% endorsed experiencing suicidal ideation several days or more during the previous 2 weeks.

The average PHQ-2 score was 1.17 ($SD = 1.948$). The average PHQ-9 score was 5.23 ($SD = 7.451$). The

operating characteristics of the PHQ-2 at various cutoff scores are reported in Table 3. Using a cutoff score of 3 or greater outperformed other cutoff scores, correctly classifying 94.8% of the cases. This cutoff score had a sensitivity of 83.3% and a specificity of 97.8%, yielding a positive predictive value of 90.9% and a negative predictive value of 95.7%. Additionally, receiver operating characteristic (ROC) analysis revealed an AUC value of 0.979 for the PHQ-2, suggesting that the PHQ-2 possesses excellent diagnostic accuracy in the present study's sample of community-residing individuals with SCI/D.⁴⁶

Discussion

The present study extended previous research findings by demonstrating that the PHQ-2 shows promise as a clinically useful depression screener in the community-residing SCI/D population. Specifically, a cutoff score of 3 or greater outperformed other scores, but the cutoff score of 2 or greater could be used if the goal is to maximize sensitivity at the expense of specificity. Regardless of the cutoff score selected, the PHQ-2 was shown to possess excellent diagnostic accuracy. Additionally, the estimated prevalence of Major Depressive Disorder in the sample was 20.7%. This is consistent with previous research indicating that in community-residing individuals with SCI/D, about 23% met the criteria for Major Depressive Disorder, also using the PHQ-9 as a proxy for diagnosis.⁷

The results of the present study provide support for the clinical utility of the PHQ-2. Clinical utility is conceptualized as consisting of many factors relevant to clinical practice, including how easily an instrument or intervention can be learned and applied by different practitioners, how it fits in with the idiosyncrasies of a setting, and its compatibility with the values of the culture in which it is applied.⁴⁷ According to Smart's multidimensional model, instruments or interventions that possess clinical utility are appropriate, accessible, practical, and acceptable. In other words, they must be effective (i.e. supported by formal evidence) and relevant (i.e. important for clinical decision-making), easily accessed both in terms of cost and availability, functional in and suitable for the specific environment, and acceptable to clinicians and clients.⁴⁸ Specific to screening for depression, Kroenke asserted that clinically useful measures are brief, easily administered, multipurpose (i.e. can be used for screening, severity assessment, probable diagnosis, and treatment monitoring), free, and easy to score (i.e. single summative score without complex procedures).¹³ Kroenke also emphasized that determining an optimal cutoff score and

Table 2 Demographic and injury characteristics.

Characteristic	Study sample mean (SD) or %
Age (Years)	56.0 (12.4)
Time Since SCI/D (Years)	18.1 (13.3)
Sex	
Male	96.6
Female	3.4
Race/ethnicity	
White	56.0
Black	33.6
Hispanic/Latino	7.8
Native American	1.7
Other	.9
Marital Status	
Married	49.1
Divorced	33.6
Single	14.7
Separated	2.6
Injury Characteristics	
Tetraplegia (AIS A, B, C)	24.2
Paraplegia (AIS A, B, C)	31.7
AIS D	38.8
Traumatic Injury	
Yes	81.0
No	19.0

Table 3 PHQ-2 operating characteristics.

PHQ-2 Cutoff score	Percentage accuracy in classification	Sensitivity	Specificity	Positive predictive value	Negative predictive value
1	84.5%	100%	80.4%	57.1%	100%
2	92.2%	95.8%	91.3%	74.2%	98.8%
3	94.8%	83.3%	97.8%	90.9%	95.7%
4	93.1%	70.8%	98.9%	94.4%	92.9%
5	89.7%	54.2%	98.9%	92.9%	89.2%

showing sensitivity to change are essential aspects of clinical utility. The PHQ-2 meets several of the requirements for depression screening measures recommended by Kroenke (i.e. brief, easily administered, free, easy to score).¹³ Likewise, based on Smart's multidimensional model of clinical utility and specific to the circumstances of the present study, the PHQ-2 was effective (i.e. supported by formal evidence), easily accessed both in terms of cost and availability, and functional in and suitable for the specific environment.⁴⁸

It is important to note both the strengths and the limitations of the present study. Study measures were administered within the context of routine care, suggesting that the findings are more likely to generalize to other real-life clinical settings. Moreover, the measures were administered independently as opposed to taking the first two items from the PHQ-9 to constitute the PHQ-2. Although a convenience sample, participant demographics were similar to the national VA SCI/D sample.⁴⁵ However, the sample was of limited diversity, consisting entirely of Veterans, most of whom were male and many of whom had incomplete traumatic injuries. Thus, further research is needed to determine if the results of the present study may generalize to the broader SCI/D population. Another potential limitation of the present study is that the measures were administered by behavioral health providers, but because of the ease of administration and scoring, the PHQ-2 can be and is routinely administered by other health care providers, such as nurses.^{32,33} This is especially valuable in SCI/D care settings that may not have a full-time behavioral health provider on staff to conduct depression screenings.

Future research should investigate the ways in which the PHQ-2 can address other aspects of clinical utility, such as relevance in clinical decision-making and treatment monitoring. For example, the PHQ-2 could be administered at intervals throughout treatment of depression in a community-residing SCI/D population to determine its sensitivity to change over time and its relevance in clinical decision-making. Clients and clinicians also could be surveyed about their perceptions of the PHQ-2 in order to study an additional component of

Smart's model, which is acceptability to clinicians and clients.⁴⁸ Overall, the results of this study demonstrate the clinical utility of the PHQ-2 in community-residing individuals with SCI/D as a screening measure that reflects the best available evidence while remaining sensitive to the needs of the clinical environment.

Disclaimer statements


Contributor None.


Funding None.

Conflicts of interest The views expressed in this article are those of the authors and do not necessarily reflect the position or policy of the Department of Veterans Affairs, the US government, Baylor College of Medicine, or Texas A&M College of Medicine. None of these bodies played a role in study design; in the collection, analysis and interpretation of data; in the writing of the report; or in the decision to submit the article for publication.

Ethics approval None.

ORCID

Joseph Mignogna  <http://orcid.org/0000-0001-8618-2607>

Aimee J. Christie  <http://orcid.org/0000-0003-2558-966X>

References

- Elliott TR, Frank RG. Depression following spinal cord injury. *Arch Phys Med Rehabil* 1996;77(8):816–23.
- Saunders LL, Krause JS, Focht KL. A longitudinal study of depression in survivors of spinal cord injury. *Spinal Cord* 2012; 50(1):72–7.
- Tate D, Forchheimer M, Maynard F, Dijkers M. Predicting depression and psychological distress in persons with spinal cord injury based on indicators of handicap. *Am J Phys Med Rehabil* 1994;73(3):175–83.
- Arango-Lasprilla JC, Ketchum JM, Starkweather A, Nicholls E, Wilk AR. Factors predicting depression among persons with spinal cord injury 1 to 5 years post injury. *NeuroRehabilitation* 2011;29:9–21.
- Bombardier CH, Richards JS, Krause JS, Tulsy D, Tate DG. Symptoms of major depression in people with spinal cord injury: Implications for screening. *Arch Phys Med Rehabil* 2004;85(11): 1749–56.

- 6 Hoffman JM, Bombardier CH, Graves DE, Kalpakjian CZ, Krause JS. A longitudinal study of depression from 1 to 5 years after spinal cord injury. *Arch Phys Med Rehabil* 2011;92(3):411–8.
- 7 Fann JR, Bombardier CH, Richards JS, Tate DG, Wilson CS, Tenkin N. Depression after spinal cord injury: Comorbidities, mental health service use, and adequacy of treatment. *Arch Phys Med Rehabil* 2011;92(3):352–60.
- 8 Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the national comorbidity survey replication. *Arch Gen Psychiatry* 2005;62(6):593–602.
- 9 Elliott TR. Studying depression following spinal cord injury: Evidence, policy, and practice. *J Spinal Cord Med* 2015;38(5):584–6.
- 10 Kalpakjian CZ, Bombardier CH, Schomer K, Brown PA, Johnson KL. Measuring depression in persons with spinal cord injury: A systematic review. *J Spinal Cord Med* 2009;32(1):6–24.
- 11 Kroenke K, Spitzer RL, Williams JBW. The PHQ-9: Validity of a brief depression severity measure. *J Gen Intern Med* 2001;16(9):606–13.
- 12 Bombardier CH, Kalpakjian CZ, Graves DE, Dyer JR, Tate DG, Fann JR. Validity of the Patient Health Questionnaire-9 in assessing Major Depressive Disorder during inpatient spinal cord injury rehabilitation. *Arch Phys Med Rehabil* 2012;93(10):1838–45.
- 13 Kroenke K. Enhancing the clinical utility of depression screening. *Can Med Assoc J* 2012;184(3):281–2.
- 14 Bolster C, Holliday C, Oneal G, Shaw M. Suicide assessment and nurses: What does the evidence show? *Online J Issues Nurs* 2015;20(1):2.
- 15 Sudak D, Roy A, Sudak H, Lipschitz A, Maltzberger J, Hendin H. Deficiencies in suicide training in primary care specialties: A survey of training directors. *Acad Psychiatry* 2007;31(5):345–9.
- 16 Whooley MA, Avins AL, Miranda J, Browner WS. Case-finding instruments for depression: Two questions are as good as many. *J Gen Intern Med* 1997;12(7):439–45.
- 17 Kroenke K, Spitzer RL, Williams JBW. The Patient Health Questionnaire-2: Validity of a two-item depression screener. *Med Care* 2003;41(11):1284–92.
- 18 Lowe B, Kroenke K, Grafe K. Detecting and monitoring depression with a two-item questionnaire (PHQ-2). *J Psychosom Res* 2005;58(2):163–71.
- 19 Corson K, Gerrity MS, Dobscha SK. Screening for depression and suicidality in a VA primary care setting: 2 items are better than 1 item. *Am J Manag Care* 2004;10(11):839–45.
- 20 Chen S, Chiu H, Xu B, Ma Y, Jin T, Wu M, et al. Reliability and validity of the PHQ-9 for screening late-life depression in Chinese primary care. *Int J Geriatr Psychiatry* 2009;25(11):1127–33.
- 21 Osorio FD, Mendes AV, Crippa JA, Loureiro SR. Study of the discriminative validity of the PHQ-9 and PHQ-2 in a sample of Brazilian women in the context of primary health care. *Perspect Psychiatr Care* 2009;45(3):216–27.
- 22 Arroll B, Goodyear-Smith F, Crengle S, Gunn J, Kerse N, Fishman T, et al. Validation of PHQ-2 and PHQ-9 to screen for major depression in the primary care population. *Ann Fam Med* 2010;8(4):348–53.
- 23 Phelan E, Williams B, Meeker K, Bonn K, Frederick J, LoGerfo J, et al. A study of the diagnostic accuracy of the PHQ-9 in primary care elderly. *BMC Fam Pract* 2010;11:63–71.
- 24 Zuithoff NPA, Vergouwe Y, King M, Nazareth I, van Wezep MJ, Moons KGM, et al. The Patient Health Questionnaire-9 for detection of major depressive disorder in primary care: Consequences of current thresholds in a cross-sectional study. *BMC Fam Pract* 2010;11:98–105.
- 25 Liu S, Yeh Z, Huang H, Sun F, Tjung J, Hwang L, et al. Validation of patient health questionnaire for depression screening among primary care patients in Taiwan. *Compr Psychiatry* 2011;52(1):96–101.
- 26 Inagaki M, Ohtsuki T, Yonemoto N, Kawashima Y, Saitoh A, Oikawa Y, et al. Validity of the Patient Health Questionnaire (PHQ)-9 and PHQ-2 in general internal medicine primary care at a Japanese rural hospital: A cross-sectional study. *Gen Hosp Psychiatry* 2013;35(6):592–7.
- 27 Thapar A, Hammerton G, Collishaw S, Potter R, Rice F, Harold G, et al. Detecting recurrent major depressive disorder within primary care rapidly and reliably using short questionnaire measures. *Br J Gen Pract* 2014;64(618):e31–7.
- 28 Bhana A, Rathod SD, Selohilwe O, Kathree T, Petersen I. The validity of the Patient Health Questionnaire for screening depression in chronic care patients in primary health care in South Africa. *BMC Psychiatry* 2015;15:118–26.
- 29 Carey M, Boyes A, Noble N, Waller A, Inder K. Validation of the PHQ-2 against the PHQ-9 for detecting depression in a large sample of Australian general practice patients. *Aust J Prim Health* 2015. Advance online publication. doi:10.1071/PY14149
- 30 Hanlon C, Medhin G, Selamu M, Breuer E, Worku B, Hailemariam M, et al. Validity of brief screening questionnaires to detect depression in primary care in Ethiopia. *J Affect Disord* 2015;186:32–9.
- 31 Suzuki K, Kumei S, Ohhira M, Nozu T, Okumura T. Screening for Major Depressive Disorder with the Patient Health Questionnaire (PHQ-9 and PHQ-2) in an outpatient clinic staffed by primary care physicians in Japan: A case control study. *PLoS One* 2015;10(3):119–47.
- 32 de Man-van Ginkel JMD, Gooskens F, Schepers VPM, Schuurmans MJ, Lindeman E, Hafsteinsdottir TB. Screening for poststroke depression using the patient health questionnaire. *Nurs Res* 2012;61(5):333–41.
- 33 de Man-van Ginkel JMD, Hafsteinsdottir T, Lindeman E, Burger H, Grobbee D, Schuurmans M. An efficient way to detect poststroke depression by subsequent administration of a 9-item and a 2-item patient health questionnaire. *Stroke* 2012;43(3):854–6.
- 34 Turner A, Hambridge J, White J, Carter G, Clover K, Nelson L, et al. Depression screening in stroke: A comparison of alternative measures with the Structured Diagnostic Interview for the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (Major Depressive Episode) as criterion standard. *Stroke* 2012;43(4):1000–5.
- 35 McManus D, Pipkin SS, Whooley MA. Screening for depression in patients with coronary heart disease: Data from the heart and soul study. *Am J Cardiol* 2005;96(8):1076–81.
- 36 Thombs BD, Ziegelstein RC, Whooley MA. Optimizing detection of major depression among patients with coronary artery disease using the patient health questionnaire: Data from the heart and soul study. *J Gen Intern Med* 2008;23(12):2014–7.
- 37 Wang L, Lu K, Li J, Sheng L, Ding R, Hu D. Value of patient health questionnaires (PHQ)-9 and PHQ-2 for screening depression disorders in cardiovascular outpatients. *Zhonghua Xin Xue Guan Bing Za Zhi* 2015;43(5):428–31.
- 38 Monahan PO, Shacham E, Reece M, Kroenke K, Ong'or WO, Omollo O, et al. Validity/reliability of PHQ-9 and PHQ-2 depression scales among adults living with HIV/AIDS in western Kenya. *J Gen Intern Med* 2008;24(2):189–97.
- 39 Boyle LL, Richardson TM, He H, Xia Y, Tu X, Boustani M, et al. How do the PHQ-2, the PHQ-9 perform in aging services clients with cognitive impairment? *Int J Geriatr Psychiatry* 2011;26(9):952–60.
- 40 Fiest KM, Patten SB, Wiebe S, Bulloch AGM, Maxwell CJ, Jette N. Validating screening tools for depression in epilepsy. *Epilepsia* 2014;55(10):1642–50.
- 41 Seo J, Park S. Validation of the Patient Health Questionnaire (PHQ-9) and PHQ-2 in patients with migraine. *J Headache Pain* 2015;16:65.
- 42 Amtmann D, Bamer AM, Johnson KL, Ehde DM, Beier ML, Elzea JL, et al. A comparison of multiple patient reported outcome measures in identifying major depressive disorder in people with multiple sclerosis. *J Psychosom Res* 2015;79(6):550–7.
- 43 Patten SB, Burton JM, Fiest KM, Wiebe S, Bulloch AGM, Koch M, et al. Validity of four screening scales for major depression in MS. *Mult Scler* 2015;21(8):1064–71.
- 44 Spitzer RL, Kroenke K, Williams JBW. Validation and utility of a self-report version of PRIME-MD: The PHQ primary care study. *JAMA* 1999;282(18):1737–44.
- 45 Curtin CM, Suarez PA, Di Ponio LA, Frayne SM. Who are the women and men in Veterans Health Administration's current spinal cord injury population? *J Rehabil Res Dev* 2012;49(3):351–60.

- 46 Murphy JM, Berwick DM, Weinstein MC, Borus JF, Budman SH, Klerman GL. Performance of screening and diagnostic tests: Application of receiver operating characteristic analysis. *Arch Gen Psychiatry* 1987;44(6):550–5.
- 47 Beutler LE, Howard KI. Clinical utility research: An introduction. *J Clin Psychol* 1998;54(3):297–301.
- 48 Smart A. A multi-dimensional model of clinical utility. *Int J Qual Health Care* 2006;18(5):377–82.

Appendix

The Patient Health Questionnaire-9¹¹

Response Options:

Not at all (0); Several days (1); More than half the days (2); Nearly every day (3)

Over the last 2 weeks, how often have you been bothered by any of the following problems?

1. Little interest or pleasure in doing things
2. Feeling down, depressed, or hopeless
3. Trouble falling or staying asleep, or sleeping too much
4. Feeling tired or having little energy
5. Poor appetite or overeating
6. Feeling bad about yourself – or that you are a failure or have let yourself or your family down
7. Trouble concentrating on things, such as reading the newspaper or watching television
8. Moving or speaking so slowly that other people could have noticed? Or the opposite – being so fidgety or restless that you have been moving around a lot more than usual
9. Thoughts that you would be better off dead or of hurting yourself in some way