Assessment and monitoring of depression in patients with coronary artery disease: validation of the Persian version of the PHQ-9

Leila Gholizadeh\textsuperscript{a}, Nazila Shahmansouri\textsuperscript{b}, Mehrdad Heydari\textsuperscript{a}, Patricia M. Davidson\textsuperscript{c}

\textsuperscript{a} Faculty of Health, University of Technology Sydney, Sydney, New South Wales
\textsuperscript{b} Psychiatric Department, Tehran Heart Center, Tehran University of Medical Sciences, Tehran
\textsuperscript{c} School of Nursing, the Johns Hopkins University, Baltimore, MD

* Corresponding Author:
Dr. Leila Gholizadeh
Faculty of Health, University of Technology, Sydney
15 Broadway, Ultimo NSW 2007, Australia
Phone: +612 95144814
Email: Leila.gholizadeh@uts.edu.au

Acknowledgments
This work was supported by University of Technology Sydney under UTS Research Reestabilishment Grants.
Abstract

Background: Depression is common comorbidity in patients with coronary artery disease (CAD), which often remain undetected and untreated. Objective: This study aimed to examine the psychometric properties of the Depression Module of the Patient Health Questionnaire (PHQ-9) on a sample of cardiac patients in Iran. Method: The Persian version of the PHQ-9 was developed and administered to 150 patients with a diagnosis of coronary artery disease (CAD), admitted to a tertiary hospital in Tehran, Iran. The major depression module of the International Neuropsychiatric Interview (MINI) was applied as the gold standard. Results: Cronbach's alpha for the total PHQ-9 was 0.80. The optimal cut-off score of ≥7 showed a sensitivity of 76, specificity of 78 and the area under curve of 0.82. Conclusion: The Persian version of the PHQ-9 has acceptable psychometric properties to screen and detect a current depressive episode in patients with CAD, with recommended cut-off score of ≥7.

Keywords: nursing, coronary artery disease, the PHQ-9, depression, validation study

Impact statement

Cardiovascular disease is a global health concern. The comorbidity of depression adds significantly to the disease burden increasing the risk of future cardiac events and mortality. By cross cultural validation of the PHQ-9, the recommended tool by the American Heart Association, this study facilitates screening and treatment of depression in Persian patients with coronary artery disease.
Introduction

Studies from around the world suggest that between 10% and 50% of patients with cardiac disease experience depression (Lichtman et al., 2008). A considerable number of cardiac diseases in Iran are affected by depression comorbidity, with the reported prevalence rate ranging between 42% and 49% (Bayani et al., 2011; Moradyan, Ebadi, Saeid, & Asiabi, 2013). There is substantial evidence linking depression with increased risk of incident coronary artery disease (CAD) as well as poorer disease prognosis (Lichtman et al., 2008). Despite the high prevalence, adverse patients and disease outcomes and yet cost effective treatability, depression in cardiac patients continues to remain under-recognized and under-treated (Ramamurthy, Trejo, & Faraone, 2013).

In 2008, the American Heart Association’s advisory committee (AHA) recommended that all patients with a diagnosis of cardiovascular disease be routinely screened for depression and treated as necessary (Lichtman et al., 2008). Although, this recommendation has been criticized for being premature due to lack of strong evidence to support the usefulness of depression screening in cardiac patients (Hasnain, Vieweg, Lesnefsky, & Pandurangi, 2011; Ziegelstein, Thombs, Coyne, & de Jonge, 2009), this enterprise has certainly risen awareness towards depression in cardiac patients and has stimulated research on the topic (Beach et al., 2013; Haddad et al., 2013; Holmes, 2011). The results of a recent meta-analysis study suggest that mental health treatments including antidepressants and psychotherapies reduce the risk of recurrent cardiac events and improve depression symptoms in patients with cardiac disease, with no effect on mortality (Rutledge, Redwine, Linke, & Mills, 2013). It is also known that depression screening in combination with a collaborative care intervention, such as referral to mental health services, improve depression symptoms, quality of care and adherence to treatment regimen in patients with cardiac diseases (Colquhoun et al., 2013; Thombs, Ziegelstein, & Whooley, 2008). These findings
provide additional evidence to support screening and appropriate treatment of depression in cardiac patients.

The Patient Health Questionnaire-9 (PHQ-9) and Patient Health Questionnaire-2 (PHQ-2) have been shown to be useful tools for screening depression (Kendrick & Little, 2013; Kroenke, Spitzer, & Williams, 2001, 2003; Lowe, Kroenke, & Grafe, 2005; Rooney et al., 2013). The PHQ-2 consists of the first two items of the PHQ-9. The AHA advocates a two-step screening approach that uses the PHQ-2 and the PHQ-9. The screening procedure starts with the first two questions of the PHQ-9, i.e., the PHQ-2, and continues with the remaining seven PHQ-items if the patient’s response to one or both of the first two questions is positive (Kroenke et al., 2003). Patients screened positive on the PHQ-9 should be referred to an appropriate mental health provider for further assessment and/or treatment (Colquhoun et al., 2013; Lichtman et al., 2008). The validity of the PHQ-9 is well established in primary care (Cameron et al., 2011; Cameron, Crawford, Lawton, & Reid, 2008; Gilbody, Richards, & Barkham, 2007). Although early validation studies support the use of the tool in patients with cardiac disease (Haddad et al., 2013; McManus, Pipkin, & Whooley, 2005; Stafford, Berk, & Jackson, 2007), the usefulness of the PHQ-09 as an screening tool for cardiac patients, particularly patients from culturally and linguistically diverse backgrounds needs further investigation. McManus et al. (2005) established the validity of the PHQ-9 on 1,024 outpatients with CAD in the United States using the Clinical Interview Schedule (CIS) as criterion. Similarly, Stafford et al. (2007) compared the psychometric characteristic of the PHQ-9 with the Hospital Anxiety and Depression Scale (HADS) on a sample of recently hospitalized cardiac patients in Australia and used the MINI as the criterion standard. Both studies found that the PHQ-9 had acceptable properties for screening depression in the target population. In addition, Haddad et al. (2013) found that the PHQ-9 was superior to the
Hospital Anxiety and Depression Scale (the HADS) in detecting depression in cardiac patients in primary care settings (Haddad et al., 2013).

Previous research also confirms the short term and long-term prognostic value of the PHQ-9. Beach et al. (2013) found that each additional point on the PHQ-9 increased the risk of readmission for a cardiac related reason over the subsequent 6 months by 9% (Beach et al., 2013). In another study, patients with higher the PHQ-9 scores at baseline experienced greater adverse events post percutaneous coronary intervention at 1.5 year follow up (Pedersen et al., 2009). Also, patients who were screened positive on the PHQ-9 showed 41% greater risk for adverse cardiovascular outcomes in long term follow ups (Elderon, Smolderen, Na, & Whooley, 2011). Although the available evidence are promising, the sensitivity and specificity of the PHQ-9 to accurately screen, detect and monitor depression, and its prognostic value in cardiac patients from linguistically and culturally diverse backgrounds need further research.

This study aimed to validate the psychometric properties of the PHQ-9 to screen depression in a sample of cardiac patients from Iran. Considering the high prevalence rate of depression in Iranian patients, there is a need for a brief, valid and user-friendly tool to assist clinicians with screening for this comorbidity in the target population. Additionally, the study investigated the diagnostic value of the PHQ-9 using the receiver operating characteristic curve (ROC). In a ROC curve, the true positive rate (sensitivity) is plotted in function of the false positive rate (100-Specificity) for different cut-off points of a parameter (Hajian-Tilaki, 2013).

Method

Participants and Procedures
Using consecutive sampling, 170 patients with a diagnosis of CAD were invited to this prospective study and those interested in the study were screened against the study inclusion and exclusion criteria. Patients with a diagnosis of CAD including acute coronary syndrome (unstable angina and myocardial infarction) and patients who had been admitted for a coronary artery bypass graft surgery (CABG), aged 20 years or above, were recruited from a tertiary hospital in Tehran. The recruitment took place from March 2012 to June 2012. Patients who, according to their medical records, were suffering from any brain injury, dementia, mental illnesses other than depression or had terminal illness were excluded. Out of the 170 recruited participants, 20 patients were inaccessible for the subsequent MINI interview, and they were excluded from the study. The study received ethical approval from a relevant human research ethics committee, the objectives of the study were explained to participants and informed consents obtained. Participants were asked to complete the Persian versions of the PHQ-9 while they were in hospital. Translation and back-translation of the PHQ-9 and the MINI were conducted following an approved guideline (Guillemin, Bombardier, & Beaton, 1993). The Persian versions of translated tools were then reviewed by five psychologists/psychiatrists for clarity. Some participants (31%) were unable to complete the questions themselves because of limited literacy or vision-related issues. In these circumstances, the researcher read out the questions to participants and wrote down their responses.

**Measures**

**Socio-Demographic and Clinical Data**

The following socio-demographic and clinical data were recorded for all patients: age, sex, education, comorbidities, reason for hospital admission, time since first diagnosis of cardiac disease, antidepressant medication and history of psychiatric disease other than depression.
The MINI: Major Depressive Episode Module

The depression module of the Mini Neuropsychiatric Interview (MINI) was used as the gold standard diagnostic tool for depression in this study. The MINI is the most widely used, brief and structured interview to assess the major Axis I psychiatric disorders in DSM-IV and ICD-10. The MINI demonstrates high validity and reliability scores when compared with the Structured Clinical Interview for DSM-III-R (the SCID-P) and the International Classification of Diseases (ICD-10) (Kittirattanapaiboon P & M., 2005; Sheehan et al., 2006). The Major Depressive Episode Module is consisted of maximum 9 questions, assessing depression symptoms over the past two weeks. If the answer to the first two questions is ‘No’, the interview is ended and the participant identified as not having a major depressive episode. If the answer to the either or both questions is ‘Yes’, then the interviewer will proceed with the following 7 questions. A score 5 or more on the 9 questions indicates a current major depressive episode (Sheehan et al., 2006). The MINI has not been validated in an Iranian sample, but been used as the gold standard for validation of another depression assessment tool in this population (Ahmadpanah et al., 2016). The administration of the interview requires no professional training, and the interview can be completed in short time. In the current study, participants were interviewed by a trained clinician, who was blind to the patient’s PHQ-9 results, via telephone within one week of completion of the PHQ-9.

The Patient Health Questionnaire (PHQ-9)

The PHQ-9 is a 9-item instrument which assesses depression symptoms over the past two weeks. Response to each item can range from 0= not at all to 3= nearly every day. Total scores on the PHQ-9 can range from 0- 27 (Kroenke et al., 2001). In the original reference population, a threshold score of 10 or higher, 15 or higher and 20 or higher indicated mild depression, moderate to major depression and severe major depression respectively. In some settings, a threshold score of 15 or more is recommended for initiating treatment with
antidepressants (Arroll et al., 2010). The PHQ-2 consists of the first two questions of PHQ-9, enquiring about the frequency of depressed mood and anhedonia over the past two weeks (Kroenke et al., 2003). Total scores, in the PHQ-2, range from 0-6. In the original reference population, a cut off of 3 or higher was considered positive screening for major depression.

The PHQ-9 is the depression component of the Patient Health Questionnaire which has been developed based on the diagnostic criteria for major depressive disorder in the Diagnostic and Statistical Manual Fourth Edition (DSM-IV). The PHQ-9 has many of the characteristics of a good screening tool; it is brief, simple and readily accessible. This self-report tool can assist a busy clinician to easily screen, diagnose, monitor and measure the severity of depression (Kroenke et al., 2001).

A systematic review of evidence on the psychometric and pragmatic characteristics of the PHQ-9 revealed that the tool possesses good psychometric properties for detecting and monitoring depression in primary care as well as medical specialty populations (Kroenke, Spitzer, Williams, & Lowe, 2010). The initial validation study, conducted on a sample of primary care and obstetrics-gynaecology patients, showed that the tool had good performance in detecting patients with major depression (sensitivity of 88%) and ruling out those without major depression (specificity of 88%), using the PHQ-9 cut-off score of ≥10 (Kroenke et al., 2001). In the United Kingdom primary care setting, the optimum cut-off point was established against the Structured Clinical Interview for DSM (SCID), with a reported sensitivity of 91.7% and specificity of 78.3% (Gilbody et al., 2007). The PHQ-9 has also been successfully validated in a sample of patients from diverse cultural backgrounds including African American, Chinese American, Latino and non-Hispanic white in the United States (Huang, Chung, Kroenke, Delucchi, & Spitzer, 2006).

Data analysis
The Statistical Package for the Social Science 21 (SPSS 21) was used for analysing the data. The internal consistency of the PHQ-9 was measured by Cronbach's alpha coefficient, and the criterion validity of the tool assessed using the MINI: Major Depressive Episode Module. The sensitivity, specificity, predictive values and likelihood ratios were calculated for various cut-off values, ROC curve was constructed and the area under the curve (AUC) calculated to determine the best cut-off score for the Persian version of the PHQ-9.

**Results**

Overall 150 inpatients with a diagnosis of CAD, predominantly male (69%), married (91%) and with at least primary school education (73%) completed the study. Mean age of participants was 61.48±9.01 years. Median time from the first diagnosis of a cardiac condition was 20.5 months, with interquartile range of 46 months. Reasons for hospital admission included medicine therapy, angioplasty, angiography and coronary artery bypass graft (CABG). The demographics and clinical characteristics of participants are presented in Table 1.

[Insert Table 1 here]

The mean PHQ-9 score was 6.16 (SD = 6.28), scores ranged from 0 to 27, with the median score of 4.0 and a skewness of 1.19 (SD = 0.19). The depression status of participants is summarized in Table 2.

[Insert Table 2 here]

**Reliability and Item Analysis**

Cronbach's alpha for the total PHQ-9 was 0.80. The mean scores for each individual item and correlation of the item with the total PHQ-9 scale are shown in Table 3. Item 2: feeling tired
or having little energy and item 3: trouble falling or staying asleep, or sleeping too much were the two most commonly endorsed items, while item 9: thoughts that you would be better off dead, or of hurting yourself was the least commonly endorsed item. All items contributed to the internal consistency of the scale, except for item 9, which did not affect the Cronbach's alpha for the total PHQ-9, when the item was removed.

Criterion Validity Analysis

The criterion validity of the PHQ-9 was investigated in relation to the Persian version of the MINI: Major Depressive Episode Module. Using the MINI, 33 (22%) participants met the diagnostic criteria for major depression. Using the cut off of ≥7 for diagnosis of major depression, the PHQ-9 had a sensitivity of 76 and specificity of 77, positive predictive value of 48.0, negative predictive value of 91.8, positive likelihood ratio of 3.2 and negative likelihood ratio of 0.3. Table 4 demonstrates the sensitivity, specificity and the predictive values for different PHQ-9 cut offs in diagnosing major depression.

Discussion

The aim of this study was to determine the validity of the Persian version of the PHQ-9 to screen and diagnose depression in cardiac patients in acute care settings. Our study demonstrated a high internal consistency of the Persian version of the PHQ-9 to measure depression in cardiac patients. In the current study, the optimum cut-off score to detect major depression was established with the PHQ-9≥ 7 (sensitivity=75.7, specificity= 78.7), while the generally recommended cut-off of ≥ 10 reduced the sensitivity of the tool to 69.9%. In a
In a study of outpatients with CAD in the US, the cut-off score ≥10 also resulted in a low sensitivity (sensitivity= 54.3%, specificity= 91.1%). In this study, however, authors did not report the operating characteristics of the PHQ-9 when different cut-offs were applied (McManus et al., 2005). In a sample of patients with CAD in Australia, optimum results were obtained using the PHQ-9 cut-off score of ≥ 6, with the reported sensitivity of 82.9% and specificity of 78.7% (Stafford et al., 2007). In our study, however, the PHQ-9 cut-off score of ≥ 6 increased the sensitivity to 82% while decreased the specificity to 73%.

Using the CIS-R as a diagnostic standard, Haddad et al. (2013) reported the optimum operating characteristics with the PHQ-9 cut-point of ≥8 (sensitivity=94%; specificity=84%) in a sample of out-patients with CAD in the United Kingdom. Likewise, in our study, the PHQ-9≥ 8 decreased the sensitivity of the tool to 73%, although the specificity was improved (81%).

In the current study the AUC was 0.80, suggesting that Persian version of the PHQ-9 discriminates well between persons with and without major depression and that the tool can be used to detect depression in cardiac patients, although this value is lower than the AUC of 0.88 reported by previous research (Stafford et al., 2007). We did not examine the categorical algorithm of the PHQ-9, as the results of previous studies suggest lower sensitivity of the tool to detect major depression when the categorical algorithm is applied (Haddad et al., 2013; Stafford et al., 2007). We found that Item 2: feeling tired or having little energy and item 3: trouble falling or staying asleep, or sleeping too much were the two most commonly endorsed items in the scale. These somatic symptoms are expected from patients with cardiac patients and similar findings were observed by Lotrakul et al. (2008) in a sample of out-patients with CAD. Whereas item 9: thoughts that you would be better off dead, or of hurting yourself was the least commonly endorsed item and did not contribute to the internal consistency of the scale. The finding is in line with other previous studies (Lotrakul, Sumrithe, & Saipanish,
2008; Razykov, Ziegelstein, Whooley, & Thombs, 2012), and provides further support that the PHQ-8 is possibly a better choice for cardiac patients.

It is noted that in the current study, the PHQ-9 was administered while participants were in hospital; however, they were mainly interviewed using MINI after discharge from hospital. The transition from hospital setting to home setting might affect participants’ feelings and therefore their responses to the interview questions, although the interviewer explicitly pointed out that the participants should reflect on their feelings over the preceding two weeks.

**Conclusion:**

The Persian version of the PHQ-9 has acceptable psychometric properties for screening and diagnosing depression in patients with CAD, with a recommended cut-off score of \( \geq 7 \) for diagnosis of major depression.
References:


doi:10.1161/circoutcomes.110.960302


doi:http://dx.doi.org/10.1016/j.jpsychores.2010.10.009


doi:[http://dx.doi.org/10.1016/j.amjcard.2005.06.037](http://dx.doi.org/10.1016/j.amjcard.2005.06.037)


Rutledge, T., Redwine, L. S., Linke, S. E., & Mills, P. J. (2013). A Meta-Analysis of Mental Health Treatments and Cardiac Rehabilitation for Improving Clinical Outcomes and Depression Among Patients With Coronary Heart Disease. Psychosomatic Medicine, 75(4), 335-349. doi:10.1097/PSY.0b013e318291d798


Table 2: Depression status of study participants according to the PHQ-9

<table>
<thead>
<tr>
<th>PHQ-9 scores</th>
<th>n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–4 (no depression)</td>
<td>79 (53)</td>
</tr>
<tr>
<td>5–9 (mild depression)</td>
<td>32 (21)</td>
</tr>
<tr>
<td>10–14 (moderate depression)</td>
<td>19 (13)</td>
</tr>
<tr>
<td>15–19 (moderately severe depression)</td>
<td>13 (9)</td>
</tr>
<tr>
<td>20–27 (severe depression)</td>
<td>7 (5)</td>
</tr>
</tbody>
</table>
Table 3: PHQ-9 item level values and item-total correlations (n=150)

<table>
<thead>
<tr>
<th>The PHQ-9 item</th>
<th>Mean</th>
<th>SD</th>
<th>Corrected item-total correlation</th>
<th>Alpha if item deleted</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Little interest or pleasure in doing things</td>
<td>0.82</td>
<td>1.19</td>
<td>0.49</td>
<td>0.80</td>
</tr>
<tr>
<td>2. Feeling down, depressed or hopeless</td>
<td>0.84</td>
<td>1.15</td>
<td>0.68</td>
<td>0.78</td>
</tr>
<tr>
<td>3. Trouble falling or staying asleep, or sleeping too much</td>
<td>0.98</td>
<td>1.27</td>
<td>0.47</td>
<td>0.81</td>
</tr>
<tr>
<td>4. Feeling tired or having little energy</td>
<td>1.07</td>
<td>1.24</td>
<td>0.66</td>
<td>0.78</td>
</tr>
<tr>
<td>5. Poor appetite or overeating</td>
<td>0.60</td>
<td>1.10</td>
<td>0.48</td>
<td>0.81</td>
</tr>
<tr>
<td>6. Feeling bad about yourself or that you are a failure</td>
<td>0.44</td>
<td>0.93</td>
<td>0.56</td>
<td>0.80</td>
</tr>
<tr>
<td>7. Trouble concentrating on things</td>
<td>0.62</td>
<td>1.04</td>
<td>0.52</td>
<td>0.80</td>
</tr>
<tr>
<td>8. Moving or speaking so slowly that other people could have noticed</td>
<td>0.60</td>
<td>1.06</td>
<td>0.47</td>
<td>0.81</td>
</tr>
<tr>
<td>9. Thoughts that you would be better off dead, or of hurting yourself</td>
<td>0.16</td>
<td>0.60</td>
<td>0.38</td>
<td>0.821</td>
</tr>
</tbody>
</table>
Table 4: The performance of various PHQ-9 cut-off scores in detecting major depression

<table>
<thead>
<tr>
<th>Cut-off score</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Positive Predictive Value</th>
<th>Negative Predictive Value</th>
<th>Positive Likelihood Ratio</th>
<th>Negative Likelihood Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>≥ 6</td>
<td>81.8</td>
<td>73.5</td>
<td>46.5</td>
<td>93.4</td>
<td>3.0</td>
<td>0.2</td>
</tr>
<tr>
<td>≥ 7</td>
<td>75.7</td>
<td>76.9</td>
<td>48.0</td>
<td>91.8</td>
<td>3.2</td>
<td>0.3</td>
</tr>
<tr>
<td>≥ 8</td>
<td>72.7</td>
<td>81.2</td>
<td>52.1</td>
<td>91.3</td>
<td>3.8</td>
<td>0.3</td>
</tr>
<tr>
<td>≥ 9</td>
<td>69.7</td>
<td>82.9</td>
<td>53.4</td>
<td>90.6</td>
<td>4.0</td>
<td>0.3</td>
</tr>
<tr>
<td>≥ 10</td>
<td>69.7</td>
<td>86.0</td>
<td>58.9</td>
<td>90.9</td>
<td>5.1</td>
<td>0.3</td>
</tr>
</tbody>
</table>
Figure 1: The Receiver Operating Characteristics (ROC) curve of the PHQ-9 versus the MINI for diagnosis of major depression.