**Sensitivity and specificity of Urdu version of the Depression Module of the Patient Health Questionnaire to screen and diagnose major depression in patients with coronary artery disease**

**Leila Gholizadeh, RN, PhDa, Shahzad Ali khanb, PhD, Farnoosh Vahedia, MSc, Patricia M. Davidson, RN, PhDa,c**

a Faculty of Health, University of Technology, Sydney, Australia

b Health Services Academy, Islamabad, Pakistan

c Johns Hopkins University School of Nursing, Baltimore, Maryland, USA

**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](mailto:Leila.gholizadeh@uts.edu.au)

**Acknowledgments**

This work was supported by University of Technology Sydney under UTS Research Re-establishment Grants.

**Abstract**

Background:

The Patient Health Questionnaire (PHQ-9) possesses many characteristics of a good screening tool and has the capacity to be used for screening depression in patients with coronary artery disease (CAD).

Aim:

To examine the psychometric properties and criterion validity of the PHQ-9 to screen and detect depression in patients with CAD in Pakistan

Design:

In this validation study, 150 patients with CAD completed the Urdu version of the PHQ-9. The major depressive episode module of the MINI was used as the gold standard.

Results:

The Urdu version of the PHQ-9 revealed a good internal consistency with Cronbach's alpha of 0.83. Optimal sensitivity (76%) and specificity (76%) were achieved using the cut-off score of PHQ-9 ≥6, with area under the ROC curve of 0.86.

Conclusion:

The Urdu version of the PHQ-9 has acceptable psychometric properties to screen and detect major depression in patients with CAD.

**Key words:** The Patient Health Questionnaire, PHQ-9, depression, cardiac disease, coronary artery disease

# Introduction

Depression is associated with both development and prognosis of coronary artery disease (CAD), and it continues to remain under-recognized and untreated in patients with cardiac disease (Ramamurthy et al., 2013). The American Heart Association’s Advisory Committee (Lichtman et al., 2008) recommends that all patients with a diagnosis of CAD be routinely screened for depression and treated if needed. There is ample of evidence that appropriate management of depression helps reduce cardiac related events and improves depression symptoms, patient compliance with medical regimens, and their quality of life (Colquhoun et al., 2013; Thombs et al., 2008). The evidence is sufficient to support the need for screening and treatment of depression in cardiac patients. Although the available evidence on the psychometric properties of the PHQ-9 is promising, the sensitivity and specificity of the PHQ-9 to accurately screen, detect, and monitor depression in patients with cardiac disease need further investigation, particularly in patients from linguistically and culturally diverse backgrounds. This study aimed to validate the psychometric properties of the Urdu version of the PHQ-9 to screen and detect depression in cardiac patients in Pakistan.

**Background**

To improve screening for depression in busy cardiology practices, a two-step screening approach has been recommended by the American Heart Association’s Advisory Committee ([Lichtman et al., 2008](#_ENREF_1)) and later endorsed by and the Expert Working Group of the National Heart Foundation of Australia (Colquhoun et al., 2013). The screening initiates with the first two questions of the PHQ-9, known as the PHQ-2. Negative answers to the first two questions reliably rule out the possibility of depression, while patients screened positive on the PHQ-2 should be further assessed using the PHQ-9 (Kroenke et al., 2003; McManus et al., 2005). Patients with positive PHQ-9 should be referred to an appropriate mental health service for additional assessment and treatment if needed (Kroenke et al., 2003; Lichtman et al., 2008).

The PHQ-9 has been designed based on the diagnostic criteria for major depressive disorder of the Diagnostic and Statistical Manual- Fourth Edition (DSM-IV). It is a brief, safe, simple, and readily available tool, thus, potentially suitable for screening (Kroenke et al., 2001). The tool has been originally developed on a sample of primary care and obstetrics-gynaecology patients; therefore, the validity of the tool to screen depression in other patient groups needs further investigations (Kroenke et al., 2010). The initial validation study reported a sensitivity of 88% and specificity of 88% (cut off ≥10) (Kroenke et al., 2001). The same cut off provided the optimum sensitivity of 91.7% and specificity of 78.3% in a sample of primary care patients in the UK (Gilbody et al., 2007). In 2009, the US Veterans Health Administration and the US National Institute for Health and Clinical Excellence endorsed the use of PHQ-9 and its shorter version, the PHQ-2, in primary care settings (Department of Veteran Affairs, 2009).

Although the psychometric properties of the PHQ-9 has been studied in different patient groups such as patients with diabetes (Khamseh et al., 2011) and cancer (Thekkumpurath et al., 2011), the recommended cut off points for specific medical conditions are different (Huang et al., 2006; Khamseh et al., 2011; Thekkumpurath et al., 2011). The results of validation studies in patients with cardiac disease suggest the comparability or superiority of the PHQ-9 as a depression screening tool (McMillan et al., 2010; Stafford et al., 2007). Further, the diagnostic validity of the PHQ-9 has been studied in patients with CAD in primary care settings and been found that compared to the Hospital Anxiety and Depression Scale, the PHQ-9 is a superior tool to detect depression in cardiac patients (Haddad et al., 2013). The prognostic value of the PHQ-9 has also been studied and it was found that each additional score on the tool was associated with a 9% increase in the risk of readmission for cardiac events at 6 month follow up (Beach et al., 2013). Similarly, in patients who underwent angioplasty, higher baseline PHQ-9 scores were associated with adverse events at 1.5 year follow up (Pedersen et al., 2009). Also, patients with stable angina who were screened positive on the PHQ-9, had 41% increased risk for adverse cardiac outcomes in long term follow-up (Elderon et al., 2011).

**Methods**

***Participants & Procedures***

Patients with a diagnosis of CAD including unstable angina, myocardial infarction, and those who were admitted for coronary artery bypass graft surgery or percutaneous coronary angioplasty and aged 20 years and above were recruited from a tertiary hospital in Islamabad, Pakistan.. Patients who were suffering from a brain injury, dementia, psychological disease other than depression, or had a terminal illness were excluded from the study. The study obtained approval from relevant human research ethics committees, participants received information about the aims and purpose of the study and consented to participate voluntarily. Translation and back-translation of these tools were conducted following the suggested guidelines (Guillemin et al., 1993). Participants were then asked to complete the Urdu versions of the PHQ-9 and were interviewed using the MINI: Major Depressive Episode Module while they were in hospital. The PHQ-9 was completed by the researcher if the participant was not able to read the questionnaire or requested the researcher to do so.

***Measurements***

***The MINI: Major Depressive Episode Module***

The depression module of the Mini Neuropsychiatric Interview (MINI) was used as the gold standard for concurrent criterion validity assessment. Within 10 days after completing the PHQ-9, participants were contacted via telephone to undertake the MINI interview by a trained clinician who was blind to the participants’ PHQ-9 results. The MINI has been shown to have 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, 1998). It is a brief interview and requires no professional training.

***The Patient Health Questionnaire (PHQ-9)***

The PHQ-9 is a 9-item questionnaire, with responses to each item ranging from 0 to 3 and the total score ranging from 0 to 27 (Kroenke et al., 2001). According to the developers of the tool, PHQ-9 scores of 5, 10, 15, and 20 represent mild, moderate, moderately severe, and severe depression, respectively (Kroenke et al., 2001). A threshold score of 15 or more is recommended for initiating treatment withantidepressants (Thekkumpurath et al., 2011).

***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 by using the MINI: Major Depressive Episode Module as a criterion standard. The sensitivity, specificity, predictive values and likelihood ratios were computed for various cut-off scores and compared against the MINI. The receiver operating characteristic (ROC) curve was generated and the area under the curve (AUC) calculated to further evaluate the specificity and sensitivity of the PHQ-9.

**Results**

Overall, 150 inpatients with a diagnosis of CAD, mainly male (79%), married (99%), and with at least primary school education (80%) completed the study. Mean age of the participants was 54.71±10.86 years. Above half of the participants (53%) were unemployed, and majority (85%) identified their ethnicity as Punjabi. About half of the participants has been hospitalised with diagnosis of either unstable angina or myocardial infarction (49%), other diagnoses included coronary artery bypass graft and percutaneous coronary intervention.

The mean total PHQ-9 score of the sample was 4.68±4.57, with total scores ranging from 0 to 21, median score 3.5, skewness 1.50, and Kurtosis 2.30. The levels and severity of depression of the participants as measured by the PHQ-9 are presented in Table 1.

***Reliability and Item Analysis***

Cronbach's alpha for the Urdu version of PHQ-9 was 0.83. Mean scores for the individual items and correlation of each item with the total PHQ-9 are shown in Table 2. Item 3: trouble falling or staying asleep, or sleeping too much and Item 4: feeling tired or having little energy were the two most highly rated items, while item 9: thoughts that you would be better off dead, or of hurting yourself was the most poorly rated item. All items contributed to the internal consistency of the scale, except for item 9 which did not affect the Cronbach's alpha of total scale if deleted.

***Criterion Validity Analysis***

The criterion validity of the Urdu version of PHQ-9 was investigated using the MINI: Major Depressive Episode Module as the gold standard. Based on the MINI, 33 (14.7%) participants met the diagnostic criteria for major depression. Table 3 demonstrates the sensitivity, specificity, PPV, NPV, PLR, and NLR values for different PHQ-9 cut offs in diagnosing major depression. Using the cut off of ≥6 for diagnosis of major depression, the PHQ-9 showed a sensitivity of 76 and specificity of 76, positive likelihood ratio (PLR) of 3.2, negative likelihood ratio (NLR) of 0.3, positive predictive value (PPV) of 48.0, and negative predictive value (NPV) of 91.84. The area under the curve for the Urdu version of PHQ-9 was 0.86 with 95% CI, 0.78-0.98 (Figure 1).

# Discussion

The aim of this study was to assess the validity of the Urdu version of the PHQ-9 to screen and detect depression in patients with CAD. The study demonstrated that the Urdu version of the PHQ-9 has a high internal consistency (Cronbach's alpha of 0.83) and is an appropriate tool to detect depression in cardiac patients. The optimum cut-off threshold for detection of major depression was established with the PHQ-9≥ 6 (sensitivity=76%, specificity= 76%), while the generally recommended cut-off of ≥ 10 reduced the sensitivity of the tool to 45%. A study in the US, which similarly assessed the psychometric properties of the PHQ-9 in a sample of patients with CAD, reported a low sensitivity (sensitivity= 54.3%, specificity= 91.1%) with the cut-off score ≥10. However, the study did not report if the sensitivity of the tool improved when different cut-offs were used (McManus et al., 2005). Similar to our findings, in a sample of outpatients with CAD in Australia, lowering the cut off score of PHQ-9 to ≥ 6 improved the sensitivity of the tool (82.9%) while retained its specificity (78.7%), thus, improving the usefulness of the tool to screen for depression in cardiac patients. In this study, however, the optimal cut-off was established at the PHQ-9≥ 5, with reported sensitivity of 91.4 and specificity of 75.3 (Stafford et al., 2007). A study in the UK, which applied the PHQ-9 to CAD patients at the community, established the optimal cut-off at the PHQ-9 ≥8 (sensitivity=94%; specificity=84%).

In current study, the AUC was 0.86, suggesting that Urdu version of the PHQ-9 discriminates very well between cardiac patients with and without major depression and can; therefore, be used as a diagnostic test to detect depression in these patients. We did not examine the categorical algorithm of the PHQ-9, as the results of previous studies have reported lower sensitivity of the categorical algorithm to detect major depression (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 highly rated items in the scale. Although these somatic symptoms are commonly experienced by patients with cardiac patients, similar findings were also observed by Lotrakul et al. (2008) in a sample of primary care patients. Whereas Item 9: thoughts that you would be better off dead, or of hurting yourself was the most poorly rated item and did not contribute to the internal consistency of the scale. The finding is similar to previous studies (Lotrakul et al., 2008; Razykov et al., 2012) and supports the use of the PHQ-8 in patients with cardiac disease (Razykov et al., 2012).

**Conclusions**

The Urdu version of the PHQ-9 has acceptable psychometric properties for screening and diagnosis of depression in patients with coronary heart disease, with a recommended cut-off score of ≥6 for diagnosis of major depression.

**Acknowledgments**

This study has received financial support from Faculty of Health, University of Technology Sydney.

**References**

Beach, S. R., Januzzi, J. L., Mastromauro, C. A., Healy, B. C., Beale, E. E., Celano, C. M., & Huffman, J. C. (2013). Patient Health Questionnaire-9 score and adverse cardiac outcomes in patients hospitalized for acute cardiac disease. J Psychosom Res*,* 75, 409-413.

Colquhoun, D. M., Bunker, S. J., Clarke, D. M., Glozier, N., Hare, D. L., Hickie, i. B., Tatoulis, J., Thompson, D. R., Tofler, G. H., Wilson, A., & Branagan, M. G. . (2013). Screening, referral and treatment for depression in patients with coronary heart disease. A consensus statement from the National Heart Foundation of Australia. MJA*,* 198, 1-7.

Department of Veteran Affairs, Department of Defense. (2009). VA/DoD clinical practice guideline for management of major depressive disorder (MDD). Washington (DC): Department of Veteran Affairs, Department of Defense.

Elderon, L., Smolderen, K. G., Na, B., & Whooley, M. A. (2011). Accuracy and prognostic value of American Heart Association: recommended depression screening in patients with coronary heart disease: data from the Heart and Soul Study. Circ Cardiovasc Qual Outcomes*,* 4, 533-540.

Gilbody, S., Richards, D., & Barkham, M. (2007). Diagnosing depression in primary care using self-completed instruments: UK validation of PHQ-9 and CORE-OM. Br J Gen Pract*,* 57, 650-652.

Guillemin, Francis, Bombardier, Claire, & Beaton, Dorcas. (1993). Cross-cultural adaptation of health-related quality of life measures: Literature review and proposed guidelines. J Clin Epidemiol*,* 46, 1417-1432.

Haddad, M., Walters, P., Phillips, R., Tsakok, J., Williams, P., Mann, A., & Tylee, A. (2013). Detecting depression in patients with coronary heart disease: a diagnostic evaluation of the PHQ-9 and HADS-D in primary care, findings from the UPBEAT-UK study. PLoS One*,* 8, e78493.

Huang, FrederickY, Chung, Henry, Kroenke, Kurt, Delucchi, KevinL, & Spitzer, RobertL. (2006). Using the patient health questionnaire-9 to measure depression among racially and ethnically diverse primary care patients. J Gen Intern Med*,* 21, 547-552.

Khamseh, M. E., Baradaran, H. R., Javanbakht, A., Mirghorbani, M., Yadollahi, Z., & Malek, M. (2011). Comparison of the CES-D and PHQ-9 depression scales in people with type 2 diabetes in Tehran, Iran. BMC Psychiatry*,* 11, 61.

Kittirattanapaiboon P, & M., Khamwongpin. (2005). The Validity of the Mini International Neuropsychiatric Interview (M.I.N.I.)-Thai Version. Journal of Mental Health of Thailand 13, 126-136.

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med*,* 16, 606-613.

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2003). The Patient Health Questionnaire-2: validity of a two-item depression screener. Med care*,* 41, 1284-1292.

Kroenke, K., Spitzer, R. L., Williams, J. B., & Lowe, B. (2010). The Patient Health Questionnaire Somatic, Anxiety, and Depressive Symptom Scales: a systematic review. Gen Hosp Psychiatry*,* 32, 345-359.

Lichtman, Judith H., Bigger, J. Thomas, Blumenthal, James A., Frasure-Smith, Nancy, Kaufmann, Peter G., Lespérance, François, Mark, Daniel B., Sheps, David S., Taylor, C. Barr, & Froelicher, Erika Sivarajan. (2008). Depression and Coronary Heart Disease: Recommendations for Screening, Referral, and Treatment: A Science Advisory From the American Heart Association Prevention Committee of the Council on Cardiovascular Nursing, Council on Clinical Cardiology, Council on Epidemiology and Prevention, and Interdisciplinary Council on Quality of Care and Outcomes Research: Endorsed by the American Psychiatric Association. Circulation*,* 118, 1768-1775.

Lotrakul, M., Sumrithe, S., & Saipanish, R. (2008). Reliability and validity of the Thai version of the PHQ-9. BMC Psychiatry*,* 8, 46.

McManus, David, Pipkin, Sharon S., & Whooley, Mary A. (2005). Screening for Depression in Patients With Coronary Heart Disease (Data from the Heart and Soul Study). Am. J. Cardiol.*,* 96, 1076-1081.

McMillan, D., Gilbody, S., & Richards, D. (2010). Defining successful treatment outcome in depression using the PHQ-9: a comparison of methods. J Affect Disord*,* 127, 122-129.

Pedersen, S. S., Denollet, J., de Jonge, P., Simsek, C., Serruys, P. W., & van Domburg, R. T. (2009). Brief depression screening with the PHQ-2 associated with prognosis following percutaneous coronary intervention with paclitaxel-eluting stenting. J Gen Intern Med*,* 24, 1037-1042.

Ramamurthy, G., Trejo, E., & Faraone, S. V. (2013). Depression treatment in patients with coronary artery disease: a systematic review. Prim Care Companion CNS Disord*,* 15.

Razykov, I., Ziegelstein, R. C., Whooley, M. A., & Thombs, B. D. (2012). The PHQ-9 versus the PHQ-8--is item 9 useful for assessing suicide risk in coronary artery disease patients? Data from the Heart and Soul Study. J Psychosom Res*,* 73, 163-168.

Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E.,Hergueta T, Baker R, Dunbar GC. (1998). The Mini-International Neuropsychiatric Interview (MINI): the development and validation of a structured diagnostic diagnostic psychiatric interview for DSM-IV and ICD-10. J Clin Psychiatry*,* 59, 22-33.

Stafford, Lesley, Berk, Michael, & Jackson, Henry J. (2007). Validity of the Hospital Anxiety and Depression Scale and Patient Health Questionnaire-9 to screen for depression in patients with coronary artery disease. Gen Hosp Psychiatry*,* 29, 417-424.

Thekkumpurath, Parvez, Walker, Jane, Butcher, Isabella, Hodges, Laura, Kleiboer, Annet, O'Connor, Mark, Wall, Lucy, Murray, Gordon, Kroenke, Kurt, & Sharpe, Michael. (2011). Screening for major depression in cancer outpatients. Cancer*,* 117, 218-227.

Thombs, B. D., de Jonge, P., Coyne, J. C., Whooley, M. A., Frasure-Smith, N., Mitchell, A. J., Zuidersma, M., Eze-Nliam, C., Lima, B. B., Smith, C. G., Soderlund, K., et al. (2008). Depression screening and patient outcomes in cardiovascular care: a systematic review. JAMA*,* 300, 2161-2171.