# Using the PHQ-9 self-rating scale to assess depression in primary care

**Ankita Sahni[1], Mark Agius[1,2,*]**

**[1] Clare College Cambridge, Cambridge, UK**

**[2] Department of Psychiatry, University of Cambridge, Cambridge, UK**

***Corresponding author email: ma393@cam.ac.uk**

This paper is an updated version of 'The use of the PHQ-9 self-rating scale to assess depression within primary care' previously published in Psychiatria Danubina, 2017; 29 (Suppl. 3): 615-618

## Abstract

The PHQ-9 is effective in screening patients for depression as well as monitoring progress in a variety of situations. This includes patients with chronic disease, such as diabetes and coronary heart disease, but care must be taken in interpretation of the PHQ-9 in these situations because of changes in specificity and sensitivity; however, proper interpretation is possible. Using the PHQ-9 after a pre-assessment with the PHQ-2 increases its specificity, as well as preventing under-diagnosis. Although it is not suitable as a stand-alone tool for diagnosis, it is a cost-effective, efficient method of screening patients in primary care.

**Keywords:** depression, screening, Patient Health Questionnaire-9, primary care

## Introduction

According to the World Health Organisation (WHO) World Mental Health Survey Consortium 2004 (1), depression is one of the most prevalent mental health disorders. Depression is defined by the Diagnostics and Statistical Manual of Mental Disorders, 4th Edition, Text Revision (DSM-IV-TR) as a loss of interest or pleasure in daily activities which results in impaired functioning. Nine specific symptoms of depression have been described: depressed mood, reduced interest, change in appetite, change in sleep, change in activity, fatigue, feelings of worthlessness, decrease in concentration and risk of committing suicide (2). Depression has many risk factors. They include chronic medical conditions, stress, family history of depression, gender and substance abuse. There are also many social risk factors which contribute to the development of depression. These include low income, job loss, lack of social support and being single, divorced, or widowed (3).

## PHQ-9

The Patient Health Questionnaire-9 (PHQ-9) is a rating scale designed for making the diagnosis of depressive disorders based entirely on the criteria identified by the DSM, 5th Edition (DSM-5) (2). It consists of nine multiple-choice questions. Each question enquires about a single symptom, over a time scale of two weeks. Each question is scored between 0 and 3, 0 denoting "not at all", 1 denoting "some days", 2 denoting "more than half the days", and 3 denoting "nearly every day", respectively. The PHQ-9 is the self-administered version of the mood module of the Primary Care Evaluation of Mental Disorders (PRIME-MD). It has a diagnostic validity which is comparable to the clinician-administered PRIME-MD (4). Studies have also shown and validated its reliability when used as a telephone assessment tool (5). The PHQ-9 has been specifically designed to be used in primary care. It is short, and has been shown to be effective and useful in a variety of patient groups. In the UK, the PHQ-9 has been included in the National Institute of Health and Care Excellence (NICE) guidelines as an assessment tool which can be used for the identification of common mental health disorders. It is currently used as a screening and diagnostic tool for the mental health disorders of depression, anxiety and eating disorders in primary care in the UK. It is used to screen patients who have a history of depression, as well as other at-risk groups, including those patients who have chronic medical conditions. The current guidelines recommend a two-stage process of identification and diagnosis. Patients are first screened for depression using a diagnostic tool, such as the PHQ-2 or the PHQ-9. Those who screen positive should then be interviewed fully before a formal diagnosis can

be made (6).

### Cut-off values

Cut-off points have been proposed for the scores of 5, 10, 15, and 20 for mild, moderate, moderately severe, and severe depression, respectively (7). The optimal cut-off score has, however, been disputed. A 2012 meta-analysis concluded that the PHQ-9 had acceptable diagnostic properties with cut-off scores between 8 and 11 in order to diagnose depression (8). Further meta-analyses have indicated a score of 10 in order to fulfil the minimum criteria for sensitivity and specificity (9, 10) in the diagnosis of Major Depressive Disorder (MDD).

The PHQ-9 is useful as a screening tool in several different settings. This is because it does not have a specified target scenario or audience. The optimal cut-off score may differ, depending on the setting. It is important to ensure that the optimal cut-off scores are being used for the specific settings. One example is in the screening of depression in adolescents. The PHQ-9 has been validated to be used in adolescents and its sensitivity and specificity in this context are similar to those for adults. However, the optimal cut-off point is higher for adolescents (11, 12).

### Benefits of the PHQ-9

The PHQ-9 has been validated for use as a clinical and research tool (7). It has been shown to have overall high specificity. Its sensitivity is, however, suboptimal in certain conditions, such as when it is used in the algorithm scoring method (13). In certain selected subgroups of patients who have a high prevalence of depressive disorder (30 - 40%) the PHQ-9 has a positive predictive value of 85 - 90% (14).

The PHQ-9 is a short and quick assessment, which has easy-to-remember cut-off values. The scores are calculated additively, thus decreasing the room for error. As well as being useful for the identification of depression, the PHQ-9 can be used to grade the severity of depressive symptoms (7, 15-18). Grading severity allows it to be effective for monitoring patients (16-19). The PHQ-9 may be used to monitor treatment responses by observing changes in the scores. End-of-treatment cut-off points can be used to support clinical decision-making (20). The PHQ-9 is potentially effective for identifying non-response to psychiatric treatment (21).

The PHQ-9 has been translated into a variety of languages, including Arabic, Chinese (Cantonese and Mandarin), Thai (22), French, Hindi, and Spanish. It has been found to be effective, and is therefore in use, in several countries worldwide, including China (23), East Africa (24), and Malaysia (25). The PHQ-9 responses in different racial and ethnic groups have also been analysed. The PHQ-9 has been demonstrated to be effective as a screening tool for depression in diverse populations (26).

### Disadvantages of the PHQ-9

An important disadvantage of the PHQ-9 is that, although it has high specificity, it has sub-optimal sensitivity. This shortcoming can be compensated for by the use of the PHQ-2. The PHQ-2 is an ultra-short screening questionnaire which is designed to be used as a first-line measure. If a patient screens positive on the PHQ-2 they are then followed up using the PHQ-9 for further assessment. The PHQ-2 has a very high sensitivity at a cut-off value of 2. This identifies patients who may be depressed. The specificity of the PHQ-9 used subsequently helps to prevent false positives (11, 12).

Some findings have suggested that PHQ-9 assessment might lead to unnecessary diagnosis and recommendation for antidepressant therapy. It is suggested that the PHQ-9 may be more likely to identify a depressive episode than MDD (27). Caution in accurate diagnosis needs to be taken to ensure that the benefits of screening should outweigh the risk of over-treatment with antidepressants.

The DSM-IV-TR specifies several exclusion criteria, such as the absence of manic or hypomanic episodes, which should be screened for in all patients presenting with depressive symptoms. The PHQ-9 does not include these exclusion items. Unless patients are specifically screened for hypomania, this can result in bipolar disorder being misdiagnosed as a unipolar depressive disorder (28). Misdiagnosis can be avoided with proper history taking, which enables clinicians to rule out other causes of depression such as bipolar disorder or bereavement (7).

### Detecting depression in patients with chronic medical conditions

Chronic medical conditions are important risk factors for depression. These conditions can be debilitating and can cause mental distress to patients. For example, people with diabetes have an increased risk of developing depression (29) and it is therefore important that they are assessed for depressive symptoms. The present guidelines are, however, ambiguous, and this leads to low rates of screening (30). The validity of screening tools used in these situations needs also to be assessed. It has been pointed out that several questions listed in the PHQ-9 screen for symptoms which are also categorised as diabetes-related symptoms. This decreases the specificity of the tool, and could result in over-identification of depression in the diabetic population (31, 32). The PHQ-9 performs well in identifying high-risk patients. It is not, however, in its current state, suitable as a stand-alone diagnostic tool for patients with diabetes (33). In these patients, the PHQ-9 tends to identify more patients as moderate to severe than does the Hospital Anxiety and Depression Scale-Depression (HADS-D) (30). It is suggested that the HADS-D may be a more suitable tool for screening patients with diabetes for depression.

Depression is also common in patients with coronary heart disease and has independent associations with increased cardiovascular morbidity and mortality (34). It is suggested that the PHQ-8 may be a more suitable tool for coronary heart disease patients. Item 9 of the PHQ-9 is not an accurate suicide screen and is omitted in the PHQ-8 (35).

In practice the PHQ-9 is, however, still used to screen for depression in patients with cardiovascular disease. It has high specificity but poor sensitivity in these patients (36). Lowering the cut-off scores improves the sensitivity of the PHQ-9 while retaining specificity. This way of using the PHQ-9 may be a more useful screen for cardiovascular disease patients (37).

### Why choose the PHQ-9?

There are several other diagnostic tools for depression which have been validated for use. The Beck Depression Inventory, as well as the HADS are commonly used. Unlike these tools, the PHQ-9 is free to use. The PHQ-9 does have equivalent, if not superior, specificity and sensitivity to these other scales (16-18, 38, 39). It is therefore cost effective to use the PHQ-9.

The PHQ-9 is shorter than many other diagnostic tools for depression, such as the HADS and the Composite International Diagnostic Interview (CIDI). A CIDI must be performed by a clinician, whereas the PHQ- 9 and HADS can be administered in a variety of ways, including as a self-assessment (40). This is an advantage because of the time constraints of primary care.

All of these scales demonstrate high internal consistency in their scoring. However, they differ in how they categorise severity (41). Furthermore, the HADS and PHQ-9 do not fully identify the same cases. While they recognise the same prevalence of mild and moderate depression, the PHQ-9 tends to identify more patients as severe. Consequently it is suggested that the two scales are not fully interchangeable (42).

### Limitations of the PHQ-9

The above conclusions are based on references to previous studies. The accuracy of the results of such studies should always be assessed carefully. Depression screening tools need to be accurate in identifying depressed patients that would not otherwise be recognised. The number of previously undiagnosed patients in any particular sample is, by definition, subject to uncertainty. Inclusion of currently diagnosed and treated patients may increase bias by inflating estimates of screening accuracy (43). It is difficult to evaluate the accuracy of studies that do not report sample size calculations, or have small sample sizes. Very few of the studies quoted above meet the criteria needed to determine precise estimates of their accuracy (44).

### Conclusion

Psychiatric illnesses are subject to causal heterogeneity (45). Although the PHQ-9 has efficacy in screening for depression, it is unable to determine predisposing, precipitating, and perpetuating factors. It would consequently be unwise to make a diagnosis using solely the PHQ-9. However, it can be useful in

identifying patients with depression. If used in conjunction with a full biopsychosocial history, the PHQ-9 could allow a more accurate and comprehensive assessment of the patient, allowing better diagnosis and management in primary care.

## References


1. Demyttenaere K, Bruffaerts R, Posada-Villa J, Gasquet I, Kovess V, Lepine J, Angermeyer MC, Bernert S, Morosini P, Polidori G, Kikkawa T. Prevalence, severity, and unmet need for treatment of mental disorders in the World Health Organization World Mental Health Surveys. JAMA. 2004 Jun;291(21):2581-90.

2. The American Psychiatric Association. Diagnostic and statistical manual of mental disorders, text revision (DSM IV- TR). 4th ed. Arlington, VA: American Psychiatric Association; 2000 Jan 1. ISBN: 9780890420249

3. Maurer DM. Screening for depression. Depression. 2012;100:23.

4. Spitzer RL, Kroenke K, Williams JB, Patient Health Questionnaire Primary Care Study Group. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. JAMA. 1999 Nov 10;282(18):1737-44.

5. Pinto-Meza A, Serrano-Blanco A, Penarrubia MT, Blanco E, Haro JM. Assessing depression in primary care with the PHQ-9: can it be carried out over the telephone? Journal of General Internal Medicine. 2005 Aug;20(8):738-42

6. National Collaborating Centre for Mental Health. Depression: The NICE guideline on the treatment and management of depression in adults. London: Royal College of Psychiatrists; 2010 Jul 1. ISBN: 9781904671855

7. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9: validity of a brief depression severity measure. Journal of General Internal Medicine. 2001;16:606–13.

8. Manea L, Gilbody S, McMillan D. Optimal cut-off score for diagnosing depression with the patient health questionnaire (PHQ-9): A meta-analysis. Canadian Medical Association Journal. 2012;184:E191–6.

9. Moriarty AS, Gilbody S, McMillan D, Manea L. Screening and case finding for major depressive disorder using the patient health questionnaire (PHQ-9): A meta-analysis. General Hospital Psychiatry. 2015;37:567–76.

10. Pettersson A, Bostrom KB, Gustavsson P, Ekselius L. Which instruments to support diagnosis of depression have sufficient accuracy? A systematic review. Nordic Journal of Psychiatry. 2015;69:497–508.

11. Richardson LP, McCauley E, Grossman DC, McCarty CA, Richards J, Russo JE, Rockhill C, Katon W. Evaluation of the patient health questionnaire-9 item for detecting Major Depression among adolescents. Pediatrics. 2010;126:1117–23

12. Richardson LP, Rockhill C, Russo JE, Grossman DC, Richards J, McCarty C, McCauley E, Katon W. Evaluation of the PHQ-2 as a brief screen for detecting major depression among adolescents. Pediatrics. 2010;5(125)

13. Manea L, Gilbody S, McMillan D. A diagnostic metaanalysis of the patient health questionnaire-9 (PHQ-9) algorithm scoring method as a screen for depression. General Hospital Psychiatry. 2015;37:67–75.

14. Wittkampf KA, Naeije L, Schene AH, Huyser J, van Weert HC. Diagnostic accuracy of the mood module of the patient health questionnaire: a systematic review. General Hospital Psychiatry. 2007;29:388–95.

15. Kroenke K, Spitzer RL. The PHQ-9: A new depression diagnostic and severity measure. Psychiatric Annals. 2002;32:509–15.

16. Lowe B. Comparative validity of three screening questionnaires for DSM-IV depressive disorders and physicians' diagnoses. Journal of Affective Disorders. 2004; 78:131-40.

17. Lowe B, Kroenke K, Herzog W, Grafe K. Measuring depression outcome with a brief self-report instrument: Sensitivity to change of the patient health questionnaire (PHQ-9). Journal of Affective Disorders. 2004 ;81:61–6.

18. Lowe B, Unutzer J, Callahan CM, Perkins AJ, Kroenke K. Monitoring depression treatment outcomes with the patient health questionnaire-9. Medical Care. 2004;42:1194–201.

19. Beard C, Hsu KJ, Rifkin LS, Busch AB, Bjorgvinsson T. Validation of the PHQ-9 in a psychiatric sample. Journal of Affective Disorders. 2016;193:267–73.

20. Schueller SM, Kwasny MJ, Dear BF, Titov N, Mohr DC. Cut points on the patient health questionnaire (PHQ-9) that predict response to cognitive–behavioral treatments for depression. General Hospital Psychiatry. 2015;37:470–5.

21. Fowler JC, Patriquin M, Madan A, Allen JG, Frueh BC, Oldham JM. Early identification of treatment nonresponse utilizing the patient health questionnaire (PHQ-9). Journal of Psychiatric Research. 2015; 68:114–9.

22. Lotrakul M, Sumrithe S, Saipanish R. Reliability and validity of the Thai version of the PHQ-9. BMC Psychiatry. 2008;8:46

23. Xiong N, Fritzsche K, Wei J, Hong X, Leonhart R, Zhao X, Zhang L, Zhu L, Tian G, Nolte S, Fischer F. Validation of patient health questionnaire (PHQ) for major depression in Chinese outpatients with multiple somatic symptoms: a multicenter cross-sectional study. Journal of Affective Disorders. 2015;174:636–43.

24. Gelaye B, Williams MA, Lemma S, Deyessa N, Bahretibeb Y, Shibre T, Wondimagegn D, Lemenhe A, Fann JR, Vander Stoep A, Andrew Zhou X-H. Validity of the patient health questionnaire-9 for depression screening and diagnosis in east Africa. Psychiatry Research. 2013;210:653–61

25. Sherina M, Arroll B, Goodyear-Smith F. Criterion validity of the PHQ-9 (Malay version) in a primary care clinic in Malaysia. The Medical Journal of Malaysia. 2012;67:309–15

26. Huang FY, Chung H, Kroenke K, Delucchi KL, Spitzer RL. Using the Patient Health Questionnaire-9 to measure depression among racially and ethnically diverse primary care patients. Journal of General Internal Medicine. 2006;21:547–52.

27. Jerant A, Kravitz RL, Fernandez y Garcia E, Feldman MD, Cipri C, Nishio D, Knoepfler A, Wooddell MK, Baquero V, Franks P. Potential antidepressant overtreatment associated with office use of brief depression symptom measures. The Journal of the American Board of Family Medicine. 2014;27:611–20.

28. Inoue T, Tanaka T, Nakagawa S, Nakato Y, Kameyama R, Boku S, Toda H, Kurita T, Koyama T. Utility and limitations of PHQ-9 in a clinic specializing in psychiatric care. BMC Psychiatry. 2012;12:73.

29. Roy T, Lloyd CE. Epidemiology of depression and diabetes: a systematic review. Journal of Affective Disorders. 2012; 142:S8–S21.

30. Reddy P, Philpot B, Ford D, Dunbar JA. Identification of depression in diabetes: The efficacy of PHQ-9 and HADS-D. British Journal of General Practice. 2010;60:239–45.

31. Twist K, Stahl D, Amiel SA, Thomas S, Winkley K, Ismail K. Comparison of Depressive symptoms in type 2 diabetes using a Two-Stage survey design. Psychosomatic Medicine. 2013;75:791–7.

32. Willborn RJ, Barnacle M, Maack B, Petry N, Werremeyer A, Strand MA. Use of the 9-Item patient health questionnaire for depression assessment in primary care patients with type 2 diabetes. Journal of Psychosocial Nursing and Mental Health Services. 2015;54:56–63.

33. Van der Zwaan GL, van Dijk SEM, Adriaanse MC, van Marwijk HWJ, van Tulder MW, Pols AD, Bosmans JE. Diagnostic accuracy of the patient health questionnaire-9 for assessment of depression in type II diabetes mellitus and/or coronary heart disease in primary care. Journal of Affective Disorders. 2016;190:68–74.

34. Lichtman JH, Bigger JT, Blumenthal JA, Frasure-Smith N, Kaufmann PG, Lesperance F, Mark DB, Sheps DS, Taylor CB, Froelicher ES. Depression and coronary heart disease: Recommendations for screening, referral, and treatment. Focus. 2009;7:406–13.

35. Razykov I, Ziegelstein RC, Whooley MA, Thombs BD. 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. Journal of Psychosomatic Research. 2012;73:163–8.

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. Journal of General Internal Medicine. 2008;23:2014–7.

37. Stafford L, Berk M, Jackson HJ. Validity of the hospital anxiety and depression scale and Patient Health Questionnaire-9 to screen for depression in patients with coronary artery disease. General Hospital Psychiatry. 2007;29:417–24.

38. Thapar A, Hammerton G, Collishaw S, Potter R, Rice F, Harold G, Craddock N, Smith DJ. Detecting recurrent major depressive disorder within primary care rapidly and reliably using short questionnaire measures. British Journal of General Practice. 2013;64:e31–7.

39. Kung S, Alarcon RD, Williams MD, Poppe KA, Jo Moore M, Frye MA. Comparing the Beck Depression Inventory-II (BDI-II) and patient health questionnaire (PHQ-9) depression measures in an integrated mood disorders practice. Journal of Affective Disorders. 2013;145:341–3

40. Spitzer R, Kroenke K, Williams J. 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;282:1737–44.

41. Cameron IM, Crawford JR, Lawton K, Reid IC. Psychometric comparison of PHQ-9 and HADS for measuring depression severity in primary care. British Journal of General Practice. 2008;58:32–6.

42. Hansson M, Chotai J, Nordstom A, Bodlund O. Comparison of two self-rating scales to detect depression: HADS and PHQ-9. British Journal of General Practice. 2009 1;59:283–8.

43. Rice DB, Thombs BD. Risk of bias from inclusion of currently diagnosed or treated patients in studies of depression screening tool accuracy: A cross-sectional analysis of recently published primary studies and meta-analyses. PLOS One. 2016;11:e0150067.

44. Thombs BD, Rice DB. Sample sizes and precision of estimates of sensitivity and specificity from primary studies on the diagnostic accuracy of depression screening tools: A survey of recently published studies. International Journal of Methods in Psychiatric Research. 2016;25:145–52.

45. Maung HH. Diagnosis and causal explanation in psychiatry. Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences. 2016 Dec 1;60:15-24.