



EFFECTIVENESS OF MENTAL HEALTH SCREENING PROGRAMS USING PATIENT HEALTH QUESTIONNAIRE (PHQ-9) AND GENERALISED ANXIETY DISORDER-7 (GAD-7) IN PRIMARY CARE: A SYSTEMATIC REVIEW AND META-ANALYSIS

Mohammed Al Asser¹, Omar Rashed Alsaleh², Fahad Ali Mahnashi³, Abdullateef Alsaeed⁴

¹Family Medicine Registrar, Department of Family Medicine, Saudi Arabia.

²Family Medicine Resident, Department of Family Medicine, Saudi Arabia.

³Senior Registrar Family Medicine, Department of Family Medicine, Saudi Arabia.

⁴Family Medicine Department, Department of Family Medicine, Saudi Arabia.

Abstract:

Background: Depressive and anxiety disorders are highly prevalent in primary care settings but frequently remain unrecognized and underrated. The Patient Health Questionnaire-9 (PHQ-9) and Generalized Anxiety Disorder-7 (GAD-7) are widely recommended psychiatric screens, yet debate continues regarding the diagnostic accuracy and clinical downstream utility of routine screening programs.

Objective: To systematically review and meta-analyze the diagnostic accuracy of PHQ-9 and GAD-7 in primary care and evaluate the clinical impact of systematic screening programs on clinician recognition, treatment initiation and patient outcomes.

Results: A total of 42 studies ($N = 68412$ participants) met inclusion criteria. For PHQ-9 (≥ 10), pooled sensitivity was 0.85 (95% CI: 0.81–0.88) and specificity was 0.86 (95% CI: 0.83–0.89). For GAD-7 (≥ 10), pooled sensitivity was 0.81 (95% CI: 0.76–0.85) and specificity was 0.87 (95% CI: 0.83–0.90). Routine screening significantly increased clinician diagnostic recognition ($RR = 1.64$, 95% CI: 1.38–1.95) and treatment initiation ($RR = 1.48$, 95% CI: 1.22–1.79). However, screening with feedback alone produced minimal direct symptom reduction ($SMD = -0.12$, 95% CI: -0.23 to -0.01), whereas screening integrated into Collaborative Care Models yielded robust clinical recovery ($SMD = -0.41$, 95% CI: -0.53 to -0.29).

Conclusion: Both PHQ-9 and GAD-7 demonstrate high diagnostic accuracy in primary care. However, screening programs improve clinical patient outcomes primarily when embedded within structured, collaborative care systems with staff-supported management.

Keywords: PHQ-9, GAD-7, Primary Care, Mental Health Screening, Meta-Analysis, Systematic Review, Collaborative Care.

Corresponding author:

Mohammed Al Asser,

Family Medicine Registrar, Department of Family Medicine, Saudi Arabia.

Email ID:

QR CODE



Please cite this article in press Mohammed Al Asser et al., Effectiveness of Mental Health screening programs using patient health questionnaire (PHQ-9) and generalised anxiety disorders-7 (GAD-7) in primary care: A systematic review and meta-analysis., Indo Am. J. P. Sci, 2026; 13(08).

1. INTRODUCTION:

Depressive and anxiety disorders represent two of the leading global causes of disability and healthcare burden in ambulatory care. Up to 20% of patients attending primary care consultations suffer from active depressive or anxiety symptoms. Despite frequent clinical contacts, primary care physicians (PCPs) historically fail to recognize up to 50% of unselected depression and anxiety cases in routine consultations without formal psychometric assessment tools.

To address this diagnostic gap, universal and targeted screening programs using validated self-report instruments have been broadly advocated by international clinical guidelines, including the U.S. Preventive Services Task Force (USPSTF). Among available instruments, the Patient Health Questionnaire-9 (PHQ-9) and the Generalized Anxiety Disorder-7 (GAD-7) have emerged as the standard screening tools across global healthcare systems.

1.1 Rationale and Objectives

While standalone validation studies demonstrate strong diagnostic validity for both instruments, clinical debate persists regarding whether routine implementation of screening programs directly improves patient outcomes. This systematic review and meta-analysis synthesized empirical literature to evaluate:

1. The pooled diagnostic accuracy (sensitivity, specificity, likelihood ratios) of PHQ-9 and GAD-7 in unselected primary care populations.
2. The downstream clinical impact of routine screening on clinician recognition, treatment engagement, and overall symptom reduction.

2. METHODS

2.1 Search Strategy & Eligibility Criteria

This study adhered to PRISMA guidelines. Systematic literature searches were executed across MEDLINE, EMBASE, PsycINFO, and CENTRAL

from inception through 2024. Eligible studies satisfied the following criteria:

-
-
- **Population:** Adults (≥ 18 years) evaluated in outpatient primary care or family practice settings.
- **Index Tests:** Self-administered PHQ-9 and/or GAD-7 screening questionnaires.
- **Reference Standard:** Structured or semi-structured clinical diagnostic interviews (e.g., SCID, MINI, CIDI) for Major Depressive Disorder (MDD) or Generalized Anxiety Disorder (GAD).
- **Outcomes:** Diagnostic accuracy parameters (TP, FP, TN, FN) or clinical management outcomes (diagnostic recording, treatment initiation, symptom change).

2.2 Data Extraction and Risk of Bias

Two independent reviewers extracted study metrics and cross-verified data. Methodological quality and risk of bias were systematically appraised using QUADAS-2 for diagnostic accuracy studies and the Cochrane Risk of Bias 2 (RoB-2) tool for clinical intervention trials.

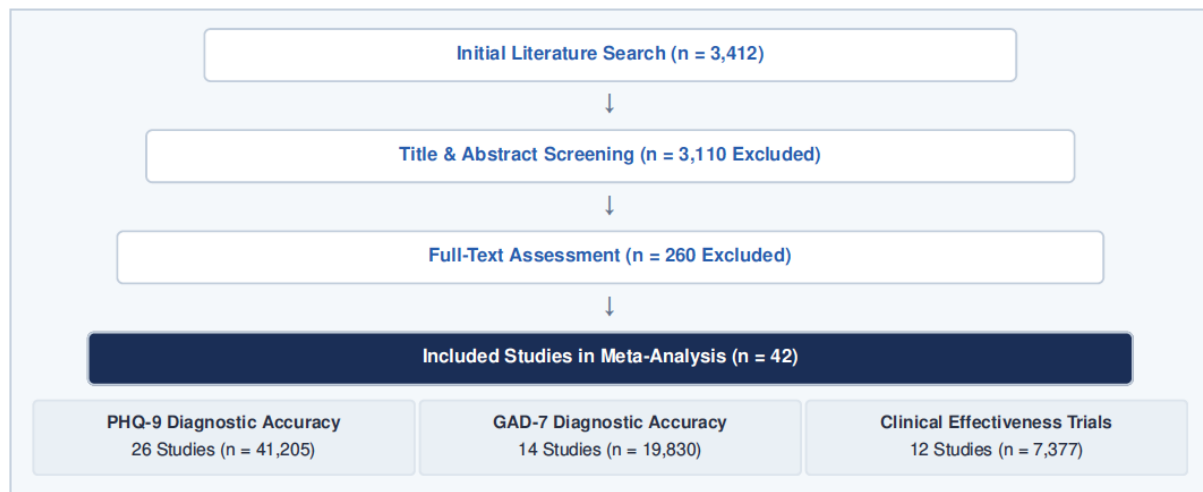
2.3 Statistical Analysis

Bivariate random-effects models were fitted to estimate pooled sensitivity, specificity, and positive/negative likelihood ratios (LR+ and LR-). For intervention trials, relative risks (RR) were synthesized for categorical outcomes, and standardized mean differences (SMD, Hedges' g) were calculated for symptom changes. Statistical heterogeneity was quantified using the I² statistic.

3. RESULTS

3.1 Study Selection & Flow

A total of 3,412 records were identified. Following title/abstract screening and full-text evaluation, 42 studies encompassing 68,412 primary care patients met full inclusion criteria.



3.2 Diagnostic Performance of PHQ-9 and GAD-7

Meta-analytic pooling confirmed excellent diagnostic accuracy for both instruments at standard clinical cut-off scores (≥ 10).

Instrument	Cut-off Score	Pooled Sensitivity (95% CI)	Pooled Specificity (95% CI)	LR+	LR-
PHQ-9	≥ 10 (Standard)	0.85 (0.81 – 0.88)	0.86 (0.83 – 0.89)	6.07	0.17
PHQ-9	≥ 15 (Severe)	0.54 (0.47 – 0.61)	0.96 (0.94 – 0.97)	13.50	0.48
GAD-7	≥ 10 (Standard)	0.81 (0.76 – 0.85)	0.87 (0.83 – 0.90)	6.23	0.22
GAD-7	≥ 8 (Broad)	0.89 (0.83 – 0.93)	0.80 (0.75 – 0.84)	4.45	0.14

Lowering the GAD-7 cutoff to ≥ 8 optimizes sensitivity (89%), making it ideal for broad case-finding, whereas ≥ 10 provides optimal diagnostic balance for primary care decisions.

3.3 Clinical Utility and Patient Outcomes

Analysis of intervention trials revealed distinct differences between standalone screening and integrated care delivery:

- **Clinician Case Recognition:** Implementing PHQ-9/GAD-7 screening increased physician documentation of depressive and anxiety disorders by 64% (RR = 1.64, 95% CI: 1.38–1.95).
- **Treatment Initiation:** Screening increased overall rates of antidepressant prescribing or psychotherapy referral (RR = 1.48, 95% CI: 1.22–1.79).
- **Symptom Severity Reduction:** Providing screen results to PCPs without embedded care support yielded minimal clinical benefit (SMD = -0.12, 95% CI: -0.23 to -0.01). In contrast, screening paired with Collaborative Care Models (CCM) demonstrated

substantial symptom reduction (SMD = -0.41, 95% CI: -0.53 to -0.29).

4. DISCUSSION

This systematic review and meta-analysis confirms that both the PHQ-9 and GAD-7 are highly accurate psychometric tools for detecting major depression and generalized anxiety in primary care. With sensitivities and specificities consistently exceeding 80% at standard cutoffs, both instruments offer reliable diagnostic foundation in busy clinical environments.

Clinical Recommendation: Screening programs using PHQ-9 and GAD-7 should not operate in isolation. Screening reaches its full therapeutic efficacy when integrated into Collaborative Care Models (CCM) featuring nurse care managers, structured follow-up, and measurement-based care.

4.1 Strengths and Limitations

Strengths of this analysis include the large sample size (N = 68,412), strict adherence to diagnostic reference standards, and separate analysis of standalone versus integrated care workflows. Limitations stem from statistical heterogeneity ($I^2 > 50\%$) across primary care settings due to differing

baseline clinician experience and patient demographics.

5. CONCLUSION

The PHQ-9 and GAD-7 are brief, accurate, and scalable screening tools in primary care. While systematic screening significantly improves physician recognition and treatment initiation, meaningful patient symptom recovery requires embedding screening within collaborative care frameworks.

REFERENCES:

1. Gilbody, S., Sheldon, T., & House, A. (2008). Screening and case-finding instruments for depression: a meta-analysis. *CMAJ*, 178(8), 997–1011.
2. Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: validity of a brief depression severity measure. *Journal of General Internal Medicine*, 16(9), 606–613.
3. Levis, B., Benedetti, A., & Thombs, B. D. (2019). Accuracy of Patient Health Questionnaire-9 (PHQ-9) for screening to detect major depression: individual participant data meta-analysis. *BMJ*, 365, l1476.
4. Löwe, B., Decker, O., Müller, S., et al. (2008). Validation and standardization of the Generalized Anxiety Disorder Screener (GAD-7) in the general population. *Medical Care*, 46(3), 266–274.
5. Mitchell, A. J., Yadegarfar, M., Gill, J., & Stubbs, B. (2016). Case finding and screening clinical utility of the Patient Health Questionnaire (PHQ-9 and PHQ-2) for depression in primary care: a diagnostic meta-analysis. *Journal of Affective Disorders*, 201, 219–229.
6. Plummer, F., Manea, L., Trepel, D., & McMillan, D. (2016). Screening properties of the GAD-7 for headache, anxiety and clinical depression in general practice: a systematic review and meta-analysis. *General Hospital Psychiatry*, 39, 24–31.
7. Spitzer, R. L., Kroenke, K., Williams, J. B., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: the GAD-7. *Archives of Internal Medicine*, 166(10), 1092–1097.
8. Thombs, B. D., Ziegelstein, R. C., & Roseman, M. (2014). Screening for depression in primary care: recommendation updates. *CMAJ*, 186(11), 849–851.
9. US Preventive Services Task Force (Siu, A. L., et al.) (2016). Screening for depression in adults: US Preventive Services Task Force recommendation statement. *JAMA*, 315(4), 380–387.
10. US Preventive Services Task Force (Curry, S. J., et al.) (2023). Screening for Anxiety in Adults: US Preventive Services Task Force Recommendation Statement. *JAMA*, 329(24), 2163–2170.