fortune Journals

Review Article

JOURNAL OF PSYCHIATRY AND PSYCHIATRIC DISORDERS

ISSN: 2572-519X

Comparison of Tools for Assessment of Depression and Response to Treatment

Dr. Kirti Mutreja

Abstract

Background: Standardized and validated tools are essential in clinical research and practice to quantify depressive symptoms and evaluate treatment outcomes. Several depression rating scales are available, each with unique strengths and limitations. Many times mental health and psychiatric practitioners in conventional and especially in alternative medical streams are not able to present the clinical results with impact due to lack of knowledge and familiarity of using various standard depression scales.

Objective: To compare commonly used depression rating tools with respect to feasibility, reliability, validity, and applicability in clinical and research contexts, with emphasis on their utility in treatment follow-up.

Methods: Literature from PubMed and indexed journals was reviewed focusing on the Hamilton Rating Scale for Depression (HAMD-17 and HAMD-6), Montgomery–Åsberg Depression Rating Scale (MADRS), Beck Depression Inventory (BDI-II/BDI-40), and the Patient Health Questionnaire-9 (PHQ-9).

Results: HAMD-17 remains widely used but has concerns regarding specificity and inter-rater reliability, leading to the development of shorter versions like HAMD-6. MADRS demonstrates good sensitivity to change but requires trained raters and is time intensive. The BDI provides a comprehensive self-report format but is lengthy and may reduce compliance. PHQ-9, a short, self-administered tool, shows robust psychometric validity, patient compliance, and ease of use, making it highly suitable for both clinical practice and research follow-up.

Conclusion: Among the available scales, PHQ-9 appears to balance validity, reliability, brevity, and patient compliance most effectively, making it a pragmatic choice for outpatient settings and treatment monitoring.

Keywords: Depressive disorder; Depression assessment; Psychiatric status rating scales; Patient health questionnaire; Treatment outcome monitoring.

Introduction

Depression is one of the most prevalent psychiatric disorders worldwide, contributing substantially to disability and disease burden. Accurate measurement of symptom severity and treatment response is fundamental in both clinical care and research. Traditional psychiatric evaluations risk subjectivity; thus, standardized rating scales offer a structured, reproducible, and validated framework for assessment. Several tools exist for depression

Affiliation:
Bhaktivedanta Hospital and Research Institute, Srishti Complex, Mira Road, Thane, Mumbai-401107, Maharashtra, India

***Corresponding author:**
Dr. Kirti Mutreja, Bhaktivedanta Hospital and Research Institute, Srishti Complex, Mira Road, Thane, Mumbai-401107, Maharashtra, India

Citation: Dr. Kirti Mutreja. Comparison of Tools for Assessment of Depression and Response to Treatment. Journal of Psychiatry and Psychiatric Disorders. 10 (2026): 106-109.

Received: May 18, 2026
Accepted: May 25, 2026
Published: June 22, 2026

Dr. Mutreja K., J Psychiatry Psychiatric Disord 2026
DOI:10.26502/jppd.2572-519X0277

measurement, but their applicability depends on ease of administration, comprehensiveness, patient compliance, and sensitivity to change. Observer-rated scales, such as the Hamilton Depression Rating Scale (HAMD-17) and Montgomery–Åsberg Depression Rating Scale (MADRS), are widely used in research but require training. Self-rated tools, such as the Beck Depression Inventory (BDI) and Patient Health Questionnaire-9 (PHQ-9), allow direct patient input and are often more practical in routine practice.

The need of the hour for mental health practitioners, especially in alternative medicine streams, is to present the clinical results with objectivity, clarity and comparitive analysis to create an impact with conventional schools of medicine and patients. But due to lack of knowledge about the available tools, standardised questionnaires available, specific advantages and limitations of each validated scale and required familiarity of using various standard depression scales, many clinicians are not able to contribute to our invaluable science academically. This article compares the relative merits of commonly used scales, aiming to identify the most suitable tool for monitoring treatment response in outpatient and follow-up contexts, and using them to obtain the desired objectivity in treatment.

Criteria for Selection of Assessment ToolsAim

The tools were evaluated on the following parameters:

1. Ease of administration and brevity
2. Simple language, avoiding technical medical terms
3. Patient compliance for baseline and follow-up assessments
4. Suitability for independent completion (reducing interviewer bias)
5. Comprehensive coverage of depressive symptoms
6. Reliability and validity in both diagnosis and follow-up

Tools for Assessment of Depression

Hamilton Depression Rating Scale (HAMD-17)

The Hamilton Depression Rating Scale, introduced in 1960, remains the most frequently applied clinician-rated tool for evaluating depression severity [1].

Format & Time: 17 items, observer-rated, requiring 20–30 minutes.

- **Advantages:** Comprehensive coverage of depressive and somatic symptoms; historically regarded as the "gold standard."
- **Limitations:** Requires trained clinicians, significant interviewer skill, risk of inter-rater bias, and poor test-retest reliability for some items [2–6]. Only six items focus directly on core depressive symptoms, prompting criticism of its specificity [7–9].

Hamilton Depression Rating Scale – 6 Item Version (HAMD-6)

Developed to improve specificity, HAMD-6 includes six core items (depressed mood, guilt, work/activities, psychomotor retardation, psychic anxiety, and general somatic symptoms) [10, 11].

- **Format & Time:** 6 items, 5–10 minutes.
- **Advantages:** Higher sensitivity to antidepressant effects, improved reliability, and suitability for clinical trials [12,13].
- **Limitations:** Focuses narrowly on core symptoms; may not capture comorbid features like anxiety or somatic complaints.

Montgomery–Åsberg Depression Rating Scale (MADRS)

The MADRS, developed in 1979, was designed to be more sensitive to treatment effects than HAMD [14].

- **Format & Time:** 10 items, clinician-rated, ~20 minutes.
- **Advantages:** Greater sensitivity to symptom change, high internal consistency, and broad international use [15,16].
- **Limitations:** Requires professional training, time-consuming, limited applicability in routine outpatient follow-up [17].

Beck Depression Inventory (BDI-II/BDI-40)

The BDI, a widely used self-report scale, includes 21 or extended 40 items assessing cognitive, affective, and somatic domains of depression.

- **Format & Time:** Self-report; 10–15 minutes.
- **Advantages:** Rich detail on patient's subjective experience; validated in diverse populations.
- **Limitations:** Length may impair compliance; potential influence of literacy and self-awareness; less practical for repeated follow-up in routine care [18].

Patient Health Questionnaire-9 (PHQ-9)

The PHQ-9, derived from the PRIME-MD instrument, operationalizes DSM-IV depression criteria [19].

- **Format & Time:** 9 self-rated items; <5 minutes.
- **Advantages:** Short, simple, free of complex medical terminology, high patient compliance, no rater training required, validated across primary care and psychiatric populations [20–22]. It demonstrates strong reliability, validity, and sensitivity to change.
- **Limitations:** Self-report nature may introduce reporting bias, though large studies support its diagnostic accuracy [21].

Dr. Mutreja K., J Psychiatry Psychiatric Disord 2026
DOI:10.26502/jppd.2572-519X0277

Discussion

The assessment of depression requires tools that are not only psychometrically sound but also feasible for use across varied clinical settings.

HAMD-17 has been the gold standard for decades, especially in clinical trials. However, its reliance on interviewer expertise introduces inter-rater variability. Moreover, the inclusion of anxiety and somatic items dilutes its specificity for depression, limiting utility in purely depressive populations [2–6].

HAMD-6, as a concise derivative, corrects some of these issues by focusing only on core depressive symptoms. It is efficient and sensitive to change, making it suitable for research trials. Yet, its narrow focus risks overlooking comorbidities and atypical presentations [9–11].

MADRS was developed to overcome the insensitivity of HAMD to therapeutic change. It shows excellent internal reliability and responsiveness, making it particularly suitable for evaluating treatment effects in clinical trials [12, 13]. However, it demands trained assessors and remains impractical for high-volume outpatient practice.

BDI empowers patients to report their subjective experience and has been validated extensively across cultures. It provides rich detail but, in practice, its length and literacy dependence may reduce compliance, especially in follow-up studies (18).

PHQ-9 stands out for its brevity, reliability, and ease of administration. It is widely used in primary care, psychiatric clinics, and community surveys. Its self-report nature removes interviewer bias and enhances patient empowerment. Critically, PHQ-9 is not only diagnostic but also sensitive enough to track changes over time, aligning it with the requirements of follow-up studies [16–20]. The PHQ-9 offers an optimal balance between brevity and validity. Numerous studies have validated its utility in screening, diagnosis, and follow-up [18–21]. Its ability to reduce treatment time and minimize costs further enhances its clinical suitability. Compared to HAMD-17, the PHQ-9 provides a practical and reliable tool for outpatient practice and epidemiological studies.

Depression rating scales are indispensable for quantifying illness severity and monitoring treatment outcomes. Observer-rated scales such as HAMD-17 and MADRS remain standards in clinical research but require training and are time-intensive [6, 7, 15]. HAMD-6 improves specificity but sacrifices comprehensiveness [12, 13]. The BDI, despite its comprehensive scope, is lengthy and less practical for clinical settings.

Taken together, **observer-rated scales remain valuable in controlled trials and inpatient settings,** while **self-rated scales, particularly PHQ-9, are more practical in outpatient, primary care, and epidemiological research.**

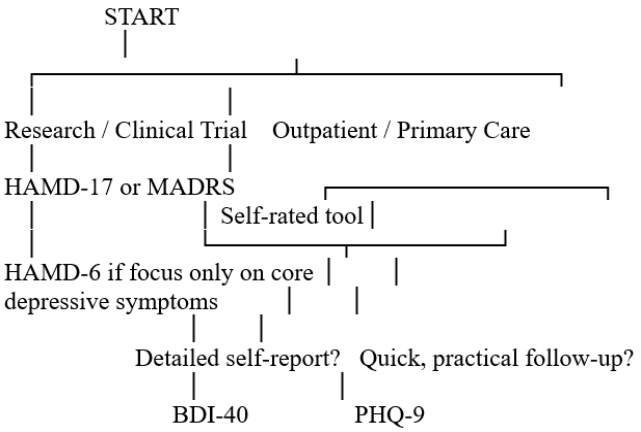

Figure 1: Choosing a Depression Assessment Tool

Table 1: Comparative Summary Table

Tool	Type	Items	Time	Strengths	Limitations	Best suited for
HAMD-17	Clinician-rated	17	20–30 min	Comprehensive; historical gold standard	Requires training; poor inter-rater reliability; includes non-depressive symptoms	Clinical trials, inpatient psychiatry
HAMD-6	Clinician-rated	6	5–10 min	Focus on core depression symptoms; sensitive to change	Narrow scope; misses comorbid/ atypical symptoms	Research, monitoring core symptoms
MADRS	Clinician-rated	10	~20 min	High sensitivity to change; reliable	Training required; time-consuming	Antidepressant trials, academic studies
BDI (21/40)	Self-rated	21–40	10–15 min	Patient-centered; validated widely	Length; literacy dependent; compliance issues	Clinical research, detailed patient assessment
PHQ-9	Self-rated	9	<5 min	Brief; reliable; widely validated; cost-free; patient-friendly	Self-report bias; less detail than longer scales	Outpatient care, follow-up, epidemiology

Citation: Dr. Kirti Mutreja. Comparison of Tools for Assessment of Depression and Response to Treatment. Journal of Psychiatry and Psychiatric Disorders.10 (2026): 106-109.

Dr. Mutreja K., J Psychiatry Psychiatric Disord 2026
DOI:10.26502/jppd.2572-519X0277

Conclusion

The choice of depression rating scale must consider the clinical setting, purpose, and feasibility. While HAMD-17 and MADRS retain value in controlled research environments, their limitations constrain use in routine clinical follow-up. The BDI provides depth but at the cost of time and compliance. PHQ-9 emerges as the most pragmatic option, offering brevity, reliability, and patient compliance, making it highly suitable for clinical follow-up, primary care, and large-scale studies.

References

1. Hamilton M. A rating scale for depression. J Neurol Neurosurg Psychiatry 23 (1960): 56-62.
2. Bech P. Fifty years with the Hamilton scales for anxiety and depression: a tribute to Max Hamilton. Psychother Psychosom 78 (2009): 202-211.
3. Carrozzino D, Patierno C, Fava GA, et al. The Hamilton Rating Scales for Depression: a critical review of clinimetric properties of different versions. Psychother Psychosom 89 (2020): 133-150.
4. Bagby RM, Ryder AG, Schuller DR, et al. The Hamilton Depression Rating Scale: has the gold standard become a lead weight? Am J Psychiatry 161 (2004): 2163-2177.
5. O'Hara MW, Rehm LP. Hamilton rating scale for depression: reliability and validity of judgments of novice raters. J Consult Clin Psychol 51 (1983): 318-319.
6. Müller MJ, Dragicevic A. Standardized rater training for the Hamilton Depression Rating Scale (HAMD-17) in psychiatric novices. J Affect Disord 77 (2003): 65-69.
7. Rehm LP, O'Hara MW. Item characteristics of the Hamilton Rating Scale for Depression. J Psychiatr Res 19 (1985): 31-41.
8. Craig TJ, Richardson MA, Pass R, et al. Measurement of mood and affect in schizophrenic inpatients. Am J Psychiatry 142 (1985): 1272-1277.
9. Lecrubier Y, Bech P. The HAM-D6 is more homogenous and as sensitive as the HAM-D17. Eur Psychiatry 22 (2007): 252-255.
10. Bech P, Gram LF, Dein E, et al. Quantitative rating of depressive states. Acta Psychiatr Scand 51 (1975): 161-70.
11. Timmerby N, Andersen JH, Søndergaard S, et al. A systematic review of the clinimetric properties of the 6-item version of the Hamilton Depression Rating Scale (HAM-D6). Psychother Psychosom 86 (2017): 141-149.
12. Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. Br J Psychiatry 134 (1979): 382-389.
13. Williams JBW, Kobak KA. Development and reliability of a structured interview guide for the Montgomery Asberg Depression Rating Scale (SIGMA). Br J Psychiatry 192 (2008): 52-58.
14. Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. JAMA 282 (1999): 1737-1744.
15. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 16 (2001): 606-613.
16. Ma S, Yang J, Yang B, et al. The Patient Health Questionnaire-9 vs. the Hamilton Rating Scale for Depression in assessing major depressive disorder. Front Psychiatry 12 (2021): 747139.
17. Wu Y, Levis B, Riehm KE, et al. Equivalency of the diagnostic accuracy of the PHQ-8 and PHQ-9: a systematic review and individual participant data meta-analysis. Psychol Med 50 (2020): 2816.
18. Costantini L, Pasquarella C, Odone A, et al. Screening for depression in primary care with Patient Health Questionnaire-9 (PHQ-9): a systematic review. J Affect Disord 279 (2021): 473-483.
19. Manea L, Gilbody S, McMillan D. Optimal cut-off score for diagnosing depression with the Patient Health Questionnaire (PHQ-9): a meta-analysis. CMAJ 184 (2012): 191-196.
20. He C, Levis B, Riehm KE, et al. The accuracy of the Patient Health Questionnaire-9 algorithm for screening to detect major depression: an individual participant data meta-analysis. Psychother Psychosom 89 (2020): 25-37.
21. Trivedi MH. Evaluating and monitoring treatment response in depression using measurement-based assessment and rating scales. J Clin Psychiatry 74 (2013): e14.
22. Trivedi MH. Tools and strategies for ongoing assessment of depression: a measurement-based approach to remission. J Clin Psychiatry 6 (2009): 26-31.

Citation: Dr. Kirti Mutreja. Comparison of Tools for Assessment of Depression and Response to Treatment. Journal of Psychiatry and Psychiatric Disorders.10 (2026): 106-109.