Diagnostic Value in Screening Severe Depression of the Hamilton Depression Rating Scale, Hamilton Anxiety Rating Scale, Beck Depression Inventory Scale, and Zung's Self-Rating Anxiety Scale Among Patients with Recurrent Depression Disorder

Vu Son Tung[1,2], Nguyen-Van Thong[3], Nguyen-Thi Phuong Mai[2], Le-Thi Thao Linh[4], Dang Cong Son[5], Tran-Thi Thu Ha[1,2], Nguyen-Thi Hoa[1,2], Nguyen Thanh Long[1,2], Nguyen-Van Tuan[1,2]

[1]Department of Psychiatry, Hanoi Medical University, Hanoi, Vietnam

[2]National Institute of Mental Health, Bach Mai Hospital, Hanoi, Vietnam

[3]Department of Psychiatry, Can Tho University of Medicine and Pharmacy, Can Tho, Vietnam

[4]School of Preventive Medicine and Public Health, Hanoi Medical University, Hanoi, Vietnam

[5]Department of training and scientific research Institute for Medicine, Pharmacy Science, Technology and Community Health, Hanoi, Vietnam

Corresponding author: Vu Son Tung. Department of Psychiatry, Hanoi Medical University, Hanoi, Vietnam. E-mail: Vusontung269@gmail.com. ORCID: 0000-0002-8971-7010.

ABSTRACT

Background: Depression is a widespread and incapacitating mental health disorder that impacts millions of people worldwide, playing a substantial role in the overall global health challenges. Depression has a big impact on a person's quality of life, cognitive and social functioning, risk of suicide, risk of heart disease and other illnesses, as well as death from all causes. . **Objective:** Objective: It may be challenging to choose the best tools to screen for severe depression in patients with recurrent depression disorder (PRD) considering the diversity of psychological scales in Vietnam. The aim of this study was to evaluate diagnostic value for detect severe depression of four psychological scales including Hamilton Depression Rating Scale (HAM-D), Hamilton Anxiety Rating Scale (HAM-A), Beck Depression Inventory Scale (BECK), and Zung's Self-Rating Anxiety Scale (SAS) by genders and age groups among PRD in Vietnam. **Methods:** This study was conducted at National Institute of Mental Health, Bach Mai Hospital, Vietnam, from 2020 to 2021. There were 109 PRD evaluated with HAM-D, HAM-A, BECK, and SAS by qualified psychiatrists. By analysing Area Under the Curve (AUC) of Receiver Operating Characteristic (ROC) curve, we determined sensitivity, specificity and cut points of four above scales. **Results:** Among four scales, the BECK scale had the best diagnostic effect with the most optimal sensitivity and specificity (61.64% and 75%, respectively). We proposed the new cut-off of HAM-D, HAM-A, BECK, and SAS for detecting severe depression among PRD were 20, 34, 30, and 45, respectively. By genders, the cut points for the HAM-D, HAM-A, BECK, and SAS in males were 20, 27, 34, and 44, respectively, while those figure in females were 14, 34, 30, and 46, respectively. By age groups, adults had cut values for four above scales of 20, 34, 27, and 45, respectively, whereas those for the elderly were 16, 17, 35, and 44, respectively. **Conclusion:** We highly recommended that BECK is the most optimal method to screen severe depression in PRD in Vietnam. It is essential to utilize varied cut values of HAM-D, HAM-A, BECK, and SAS for different genders and age groups.

Keywords: diagnostic value, cut-off, sensitivity, specificity, HAM-D, HAM-A, BECK, SAS, recurrent depression disorder.

1. BACKGROUND

Depression is a widespread and incapacitating mental health disorder that impacts millions of people worldwide, playing a substantial role in the overall global health challenges (1,2). Depression has a big impact on a person's quality of life, cognitive and social functioning, risk of suicide, risk of heart disease and other illnesses, as well as death from all causes (2). Within the diverse spectrum of depressive disorders, recurrent depression stands out as a particularly challenging and recurrent form of this condition (3).

In recent years, the assessment of depression and its recurrent nature has gained significant attention in clinical research (4). Accurate assessment and effective management of recurrent depression are pivotal for providing individuals with the best possible care and support. Various depression assessment scales have been utilized to define cut-off scores that help identify the severity of depressive symptoms in patients (5–8). Among these standardized scales, the Hamilton Anx-

iety Rating Scale (HAM-A) (9), the Hamilton Depression Rating Scale (HAM-D) (10), the Beck Depression Inventory (BECK) (11), and the Zung's Self-Rating Anxiety Scale (SAS) (12) stand out as widely accepted instruments for assessing anxiety and depression in clinical settings. These scales offer structured approaches for quantifying the severity of depressive and anxiety symptoms, aiding clinicians in their diagnostic and therapeutic decision-making processes.

Nonetheless, given the variety of psychological scales available, selecting the most optimal instruments to screen for severe depression may be difficult. In addition, determining cut-off points to reliably distinguish severe depression from milder forms for Vietnamese patients with recurrent depression (PRD) is necessary, because misclassifying severe depression can have serious consequences. Underdiagnosis delays proper treatment, worsening the condition, while overdiagnosis leads to unnecessary treatments and potential harm.

2. OBJECTIVE

Therefore, this study was carried out with the aims to evaluate diagnostic value for detect severe depression of four psychological scales including HAM-D, HAM-A, BECK, and SAS by genders and age groups among patients with recurrent depression disorder in Vietnam.

3. MATERIAL AND METHODS

Study location, time and subjects

From January 2020 to December 2021, we recruited one hundred and nine in patients with recurrent depression disorder at National Institute of Mental Health - Bach Mai Hospital in Vietnam. For inclusion, study participants were diagnosed with recurrent depression disorder (F33) by qualified psychiatrists according to International Classification of Diseases 10th Revision (ICD-10) criteria, and voluntarily participated in the study. We excluded patients with severe medical and surgical conditions such as organic brain diseases (traumatic brain injury, meningoencephalitis, etc.), cancer, emergency conditions, endocrine diseases of the hypothalamus, pituitary gland, and adrenal gland, as well as patients with conditions that impair communication, such as severe dementia, severe mental retardation, congenital deafness, etc. (the limitation of communication ability is not caused by depression).

Questionnaire

We gathered data using four psychological scales:

• Hamilton Depression Rating Scale (HAM-D) (10), which comprises 21 items, is one of the most established and frequently used measures of depression severity in both academic and therapeutic settings. We used a later version (HDRS21), in which, the first 17 of the 21 items are used to calculate the final score (total score ranges from 0 to 52), whereas items 18 to 21 provide extra information not included in the scale (e.g., paranoia and diurnal variation) (13).

• Hamilton Anxiety Rating Scale (HAM-A), which is a clinically validated tool for calculating anxiety symptoms, consists of 14 items with higher scores indicating more severe anxiety symptoms (5). The HAM-A, score of which ranges between 0 and 56 (5), is utilized to measure anxiety symptoms in research on treatments for psychiatric and physical problems, most frequently major depression (14, 15).

• Beck Depression Inventory (BECK), which is a popular 21-item screening instruments questionnaire, may be a valuable technique for detecting depression and assess severity of depression for adolescents and adults, especially in palliative care (11,16). The minimum and maximum of this test are 0 and 63, respectively (11).

• Zung's Self-Rating Anxiety Scale (SAS) is a widely utilized standard comparison tool often employed to screen for anxiety disorders. It consists of 20 items, and ranges in raw score from 20 to 80 on a 4-point scale (17).

Data analysis

For general information, we describe the mean, standard deviation (SD) for quantitative variable, and frequency, percentage for qualitative variable. To evaluate the diagnostic value for severe depression of four scales (HAM-D, HAM-A, BECK, SAS), we used the Receiver Operating Characteristic (ROC) curve to describe Area Under the Curve (AUC), sensitivity, specificity of these scales, then, identified cut-off points of these scales. We also describe sensitivity, specificity and cut points of four scales by genders and by age groups.

Ethical consideration

Our research proposal and protocol were conducted by the Declaration of Helsinki, and approved by the Hanoi Medical University Institutional Ethical Review Board (HMU IRB) (IRB-VN01.001/IRB00003121/FWA 00004148) with the Approval No: 65/GCN-HĐĐNCYSH-ĐHYHN; on 16 April 2020.

4. RESULTS

Table 1 describes some general information of study participants. The majority of study subjects were women (72.5%). Among six age groups, the age group fron 41 to 50 accounted for the highest proportion with 24.8%. The average age of the study population was 48.7 ± 15.1. The mean score of HAM-D, HAM-A, BECK, and SAS were 21.3±8.5, 21.9±11.3, 28.3±14.3, and 43.5±10.2, respectively.

The difference of Receiver Operating Characteristic (ROC) curves and Area Under these Curves (AUC) of four scales is depicted in Figure 1. The AUC distributed between 0.5923 and 0.6834. Among four scales, the BECK scale had

Characteristics		Frequency (n)	Percentage (%)
Gender	Male	30	27.5
	Female	79	72.5
Age group	18-30	16	14.7
	31-40	16	14.7
	41-50	27	24.8
	51-60	25	22.9
	61-70	18	16.5
	71-80	7	6.4
Characteristics		Mean	SD
Age		48.7	15.1
HAM-D (score)		21.3	8.5
HAM-A (score)		21.9	11.3
BECK (score)		28.3	14.3
SAS (score)		43.5	10.2

Table 1. General information of study subjects

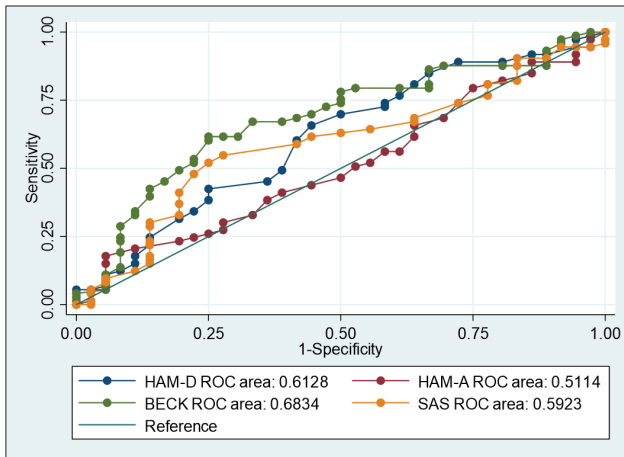**Figure 1. Receiver Operating Characteristic (ROC) analyses, with the HAM-D, HAM-A, BECK, SAS using the diagnosis of ICD-10 as gold standard.**

	AUC	95%CI	
HAM-D	0.6128	0.49829	0.72735
HAM-A	0.5114	0.39686	0.62597
BECK	0.6834	0.57719	0.78963
SAS	0.5923	0.47993	0.70462

Table 2. Area Under the Receiver Operating Characteristic Curve (AUC) and 95% Confidence Interval (CI) of the HAM-D, HAM-A, BECK, SAS

the largest AUC with 0.6834 (95%CI: 0.57719-0.78963), and the HAM-A scale had the smallest area with AUC is 0.5114 (95%CI: 0.39686-0.62597) (Table 2).

Table 3 shows the sensitivity and specificity of HAM-D, HAM-A, BECK and SAS compared to the judgment of the psychiatrist (the gold standard) by genders and age groups. Among four tests, the BECK scale had the best diagnostic effect with its sensitivity was 61.64%, while its specificity was 75%, with a cut-off value of 30. In contrast, the HAM-D≥20 exhibited a sensitivity of 65.75% and a specificity of 55.56%, while the SAS had sensitivity and specificity rates of 52.05% and 75%, respectively, with cut point of 45. Notably, HAM-A with a threshold of ≥34 displayed the lowest sensitivity at 17.81%, but the highest specificity at 94.44%. By genders, the cut-off points for for the HAM-D, HAM-A, BECK, and SAS in males were 20, 27, 34, and 44, respectively, while those figure in females were 14, 34, 30, and 46, respectively. By age groups, adults (aged from 18 to 64 years old) had cut points for the HAM-D, HAM-A, BECK, and SAS of 20, 34, 27, and 45, respectively, whereas those for the elderly (aged ≥65) were 16, 17, 35, and 44, respectively.

5. DISCUSSION

To the best of our knowledge, this was one of the first studies evaluating the diagnostic value and determining cut-off points by genders and age groups on four common psychological scales in Vietnam (Appendix S1). In addition, this research identified the best scale to detect severe depression

in patients with recurrent depression disorder in Vietnam. By providing clinicians with precise cut-off scores for severe depression conditions by genders and age groups, this research endeavors to enhance the accuracy of recurrent depression diagnosis, facilitate early intervention, and optimize treatment strategies in Vietnam. Furthermore, our findings may contribute to improving the overall management and prognosis of individuals battling the recurrent form of this debilitating condition.

This study suggests that while HAM-D has relatively high sensitivity, it may have limitations in correctly classifying non-cases, as indicated by the lower specificity. Our cut-off score is lower than that of Mark Zimmerman's study: Mark Zimmerman's study also worked on a special group of six hundred twenty-seven outpatients with current major depressive disorder, but the cut-off score was 24 for the comparison of moderate vs. severe depression (18). The cut-off score of the ham-d scale is similar in the general group, men, age group from 18 to 64. Women's cut-off score is significantly lower (14) and the age group > 65 also has a cut-off score of 16, lower than the general group. The sensitivity of the Ham-D scale for women and the over 65 age group is also high at over 85%. Previously, there were authors who analyzed comparisons between subgroups but applied cut-off points to all subgroups without looking for separate cut-off points for each group like our study (19).

Our research indicated that HAM-A had the smallest AUC, and its AUC was smaller than that of a prior study in Arabic (0.873) (20). The difference can be attributed to the different study populations; whereas the previous study concentrated on people with dementia, our investigation included patients

Subgroup	Scale	N	Sensitivity (%)	Specificity (%)	Correctly Classified (%)	Cut-off
All subjects	HAM-D	109	65.75	55.56	21.31	20
	HAM-A	109	17.81	94.44	43.12	34
	BECK	109	61.64	75.00	66.06	30
	SAS	109	52.05	75.00	59.63	45
All males	HAM-D	30	66.67	58.33	63.33	20
	HAM-A	30	50.00	75.00	60.00	27
	BECK	30	55.56	91.67	70.00	34
	SAS	30	61.11	75.00	66.67	44
All females	HAM-D	79	85.45	41.67	72.15	14
	HAM-A	79	14.55	95.83	39.24	34
	BECK	79	60.00	75.00	64.56	30
	SAS	79	47.27	79.17	56.96	46
Aged 18-64	HAM-D	91	65.00	54.84	61.54	20
	HAM-A	91	20.00	93.55	45.05	34
	BECK	91	66.67	67.74	67.03	27
	SAS	91	51.67	70.97	58.24	45
Aged ≥65	HAM-D	18	92.31	60.00	83.33	16
	HAM-A	18	92.31	60.00	83.33	17
	BECK	18	61.54	100.00	72.22	35
	SAS	18	61.54	100.00	72.22	44

Table 3. Receiver Operating Characteristics (ROC) Analyses for HAM-D, HAM-A, BECK and SAS by genders and age groups

with recurrent depression. We discovered that the HAM-A cut point for severe depression was 34, higher than some previous findings (5, 9). In addition, our HAM-A cut-off criteria varied for different age groups and genders. The cut-off for male patients was lower at 27, while the cut-off for females was the same as for the general population. This suggests that males may exhibit heightened sensitivity to anxiety symptoms at lower levels on the HAM-A scale compared to females. By age groups, adults displayed the same cut point as that of the overall population, whereas the elderly displayed an extremely low cut-value of 17.

The BECK scale has the largest AUC area. Our results are lower than those of Gregory Moullec's study on cardiac outpatients (21). This can be explained by the fact that Gregory Moullec's research subjects were 58 ± 10 years old, 31% women, while our study was mostly women (72.5%) and the average age was also younger with is 48.7 ± 15.1. The cut-off point for the total sample of the BECK scale is 30. This cut-off point remains the same in women but increases slightly in men (34). Moullec's study also found differences between gender subgroups similar to our study: "While the optimal cut-off score was 10 for the total sample (sensitivity 83%, specificity 73%), the analyzes indicated different cut-off scores across covariate subgroups: e.g. sex (women 13; men 10)" (21). In our study, we also looked at the differences between two age groups 18-64 and over 64 years old. The cut-off scores of these two groups are 27 and 35 respectively.

We found that the ZUNG SAS scale is the most stable scale, the cut-off score in the entire study sample and subgroups did not change much (ranging from 44-46). Our study adds evidence that the cut-off score of 40 is too low, similar to the results of some previous studies such as: Debra A Dunstan (2019) show that 50 would appear to provide far greater accuracy; the cut-off score of 40 for the Zung Self-Rating Depression Scale (SDS) should not only be increased but increased beyond the score of 44 suggested by Dunstan (2017) (22, 23).

In this study, AUC is distributed between 0.5923-0.6834. The AUC is a measure of the accuracy of diagnostic tests. The BECK scale has the largest AUC area (0.6834) with a 95% CI of 0.57719-0.78963, indicating strong diagnostic accuracy. In contrast, the HAM-A scale has the smallest AUC (0.5114) with a 95% CI of 0.39686-0.62597, suggesting lower accuracy. These AUC values provide valuable information about the discriminatory power of these scales in identifying the condition of interest. The sensitivity and specificity of various scales, including HAM-D, HAM-A, BECK, and SAS, were calculated in comparison to the psychiatrist's judgment (considered the gold standard). The BECK scale demonstrated the best diagnostic performance with a sensitivity of 61.64% and specificity of 75% using a cutoff of 30. This indicates that the BECK scale is effective in correctly identifying cases and non-cases. SAS and HAM-A had correctly classification rates below 60%, suggesting that these scales may not be as effective in distinguishing between cases and non-cases. However, SAS achieved a sensitivity of 52.05% and specificity of 75% with a cutoff of ≥45, indicating improved performance at a higher cutoff. In contrast, HAM-A reached a sensitivity of 17.81% and specificity of 94.44% with a cutoff of ≥34, indicating high specificity but low sensitivity.

In summary, the study's findings suggest that the BECK scale performs well in diagnosing the condition of interest, especially when using a cutoff of 30. However, it's essential to consider the potential gender bias in the sample and the age distribution's impact on the study's results. Additionally, understanding the trade-off between sensitivity and specificity for different scales is crucial when choosing the most appropriate tool for clinical assessments. Further research may be needed to validate these findings and explore potential adjustments to improve the accuracy of existing scales.

Regarding to the strengths, this research evaluated the diagnostic value and identify cut-off points for severe depression of four common psychological scales (HAM-D, HAM-A, BECK, SAS) in Vietnam. Furthermore, it provides evidence about the best scale in screening severe depression among patients with recurrent depression disorder in Vietnam. This contributes to helping doctors, clinicians and researchers have a basis to classify major depression more accurately in patients with recurrent depression disorder in Vietnam. However, there are a few drawbacks to be aware of. This study has only determined the cut points for diagnosing severe depression. Future research might broaden the subject matter and establish a threshold to precisely categorize different types of depression in Vietnamese patients who experience recurrent depression.

6. CONCLUSION

In conclusion, Beck Depression Inventory (BECK) is a psychological scale with the best sensitivity and specificity to screen for severe depression in patients with recurrent depression in Vietnam. The new cut-off values by genders and age groups of the HAM-D, HAM-A, BECK, and SAS could be used in future studies to identify individuals in Vietnam with recurrent depression disorder who are suffering from severe depression.

REFERENCES

1. World Health Organization. Depressive disorder (depression) [Internet]. [cited 2023 Sep 8]. Available from: https://www.who.int/news-room/fact-sheets/detail/depression
2. Lépine JP, Briley M. The increasing burden of depression. Neuropsychiatr Dis Treat. 2011; 7(Suppl 1): 3–7.
3. Burcusa SL, Iacono WG. Risk for Recurrence in Depression. Clin Psychol Rev. 2007 Dec; 27(8): 959–985.
4. Dunlop BW, Li T, Kornstein SG, Friedman ES, Rothschild AJ, Pedersen R, et al. Correlation between patient and clinician assess-

ments of depression severity in the PREVENT study. Psychiatry Res. 2010 May 15; 177(1): 177–183.

5. Matza LS, Morlock R, Sexton C, Malley K, Feltner D. Identifying HAM-A cutoffs for mild, moderate, and severe generalized anxiety disorder. Int J Methods Psychiatr Res. 2010 Dec; 19(4): 223–232.

6. Zimmerman M, Martinez JH, Young D, Chelminski I, Dalrymple K. Severity classification on the Hamilton depression rating scale. J Affect Disord. 2013 Sep 5; 150(2): 384–388.

7. McDowell I. Measuring Health: A Guide to Rating Scales and Questionnaires. Oxford University Press; 2006. 765 p.

8. Dunstan DA, Scott N, Todd AK. Screening for anxiety and depression: reassessing the utility of the Zung scales. BMC Psychiatry. 2017 Sep 8; 17: 329.

9. Hamilton MAX. The assessment of anxiety states by rating. Br J Med Psychol. 1959;

10. Hamilton M. A Rating Scale for Depresion. J Neurol Neurosurg Psychiatry. 1960 Feb; 23(1): 56–62.

11. Beck AT. Manual for the Beck Depression Inventory-II. No Title [Internet]. 1996 [cited 2023 Sep 5]; Available from: https://cir.nii.ac.jp/crid/1370564063990947215

12. Zung WW. A rating instrument for anxiety disorders. Psychosomatics. 1971; 12(6): 371–379.

13. Hedlund JL, Vieweg BW. The Hamilton rating scale for depression: a comprehensive review. J Oper Psychiatry. 1979; 10(2): 149–165.

14. DeBattista C, Solomon A, Arnow B, Kendrick E, Tilston J, Schatzberg AF. The Efficacy of Divalproex Sodium in the Treatment of Agitation Associated With Major Depression. J Clin Psychopharmacol. 2005 Oct; 25(5): 476.

15. Shelton RC, Keller MB, Gelenberg A, Dunner DL, Hirschfeld R,

Thase ME, et al. Effectiveness of St John's wort in major depression: a randomized controlled trial. JAMA. 2001 Apr 18; 285(15): 1978–1986.

16. Richter P, Werner J, Heerlein A, Kraus A, Sauer H. On the validity of the Beck Depression Inventory. A review. Psychopathology. 1998; 31(3): 160–168.

17. Dunstan DA, Scott N. Norms for Zung's Self-rating Anxiety Scale. BMC Psychiatry. 2020 Feb 28; 20(1): 90.

18. Zimmerman M, Martinez JH, Young D, Chelminski I, Dalrymple K. Severity classification on the Hamilton Depression Rating Scale. J Affect Disord. 2013 Sep 5; 150(2): 384–388.

19. Mottram P, Wilson K, Copeland J. Validation of the Hamilton Depression Rating Scale and Montgommery and Asberg Rating Scales in terms of AGECAT depression cases. Int J Geriatr Psychiatry. 2000 Dec; 15(12): 1113–1119.

20. Feghali Y, Koubaissy H, Fares Y, Abbas L. Cross-Cultural Adaptation and Validation of the Arabic Version of the Rating Anxiety in Dementia Scale. Clin Gerontol. 2019 Oct 29; 43.

21. Moullec G, Plourde A, Lavoie KL, Suarthana E, Bacon SL. Beck Depression Inventory II: determination and comparison of its diagnostic accuracy in cardiac outpatients. Eur J Prev Cardiol. 2015 May; 22(5): 665–672.

22. Dunstan DA, Scott N. Clarification of the cut-off score for Zung's self-rating depression scale. BMC Psychiatry. 2019 Dec; 19(1): 177.

23. Dunstan DA, Scott N, Todd AK. Screening for anxiety and depression: reassessing the utility of the Zung scales. BMC Psychiatry. 2017 Sep 8; 17(1): 329.