Measuring the severity of depression and remission in primary care: validation of the HAMD-7 scale

Roger S. McIntyre, Jakub Z. Konarski, Deborah A. Mancini, Kari A. Fulton, Sagar V. Parikh, Sophie Grigoriadis, Larry A. Grupp, David Bakish, Marie-Josee Filteau, Chris Gorman, Charles B. Nemeroff, Sidney H. Kennedy

An abridged version of this article appeared in the Nov. 22, 2005, issue of *CMAJ*.

Abstract

Background: Symptomatic remission is the optimal outcome in depression. A brief, validated tool for symptom measurement that can indicate when remission has occurred in mental health and primary care settings is unavailable. We evaluated a 7-item abbreviated version (HAMD-7) of the 17-item Hamilton Depression Rating Scale (HAMD-17) in a randomized controlled clinical trial of patients with major depressive disorder being cared for in primary care settings.

Methods: We enrolled 454 patients across 47 primary care settings who met DSM-IV-TR criteria for a major depressive disorder. Of these, 410 patients requiring antidepressant medication were randomized to have their symptoms rated with either HAMD-7 (n = 205) or HAMD-17 (n = 205) as the primary measurement tool. The primary outcome was the proportion of patients who achieved a-priori defined responses to 8 weeks of therapy using each instrument.

Results: Of the 205 participants per group, 67% of those evaluated with HAMD-7 were classified as having responded to therapy (defined as a 50% reduction from the pretreatment score), compared with 74% of those evaluated with HAMD-17 (p = 0.43). The difference between the groups' changes in scores from baseline (pretreatment) to endpoint was significant (p < 0.001), without a main effect of group (p = 0.84) or group-by-time (p = 0.83) interaction. The HAMD-7 test was brief to administer (e.g., 3–4 min for 85% of the primary care physicians evaluated), which facilitated the efficient and structured evaluation of salient depressive symptoms.

Interpretation: The abbreviated HAMD-7 depression scale is equivalent to the HAMD-17 in assessing remission in patients with a major depressive disorder undergoing drug therapy.

Cite this article as *CMAJ* 2005;173(11). DOI:10.1503/cmaj.050786

DOI:10.1503/cmaj.050786

Optimal management of major depressive disorders is enhanced by applying a chronic illness–management model with precise and measurable therapeutic endpoints.[1] In contradistinction to several other chronic medical disorders, biological markers of illness activity in depression do not currently exist. In the interim, therapeutic progress is monitored by evaluating changes in the severity of depressive symptoms and in functional domains. This concatenation of findings is particularly disconcerting in view of the fact that most depressed patients in either primary care or psychiatric settings are not systematically evaluated with objective quantifiable measures — a modifiable deficiency in patient management.[2–6]

The most frequently reported symptomatic outcome measure in clinical trials of antidepressants has been response to treatment, arbitrarily defined as a reduction of 50% or more in total symptom severity from a pretreatment assessment of the patient's depression.[7] A categorical response to therapy that fails to achieve a fully asympomatic remitted state furnishes an unsatisfactory outcome, in that it includes patients with ongoing disease activity that is clinically significant. Patients who show improvement in symptom severity but are not asymptomatic are at risk for developing chronic depression, and continue to be vulnerable to poor outcomes and comorbid medical disorders.[8–10]

Remission is an objective outcome indicated by a quantifiable score with a depressive symptom measurement tool. In antidepressant clinical trials, the 17-item Hamilton Depression Rating Scale (HAMD-17) has been the "gold standard" for use. HAMD-17, however, has not been accepted by clinicians for many reasons,[11,12] notably psychometric deficiencies and the length of time needed to administer it.

Although several brief rating scales for depression that attempt to improve upon the limitations of HAMD-17 have recently been validated and reviewed,[11–18] none that are brief, currently available and use a remission cut-off score that correlates with the most frequently cited definition of remission (a HAMD-17 score ≤ 7)[7] have been validated in both tertiary mental health and primary care settings.

Our broad objective in using HAMD-7 was to improve upon the conceptual and pragmatic deficiencies ascribed to HAMD-17. HAMD-7 was originally derived from analyses of a natural practice database at a tertiary care centre composed of patients diagnosed with a major depressive disorder (n = 248).[14] The HAMD-17 items that were endorsed in a previous study[14] by 70% of depressed patients and were most sensi-

tive to change after 8 weeks of antidepressant efficacy formed the constituent items of HAMD-7 (Appendix 1). A remission cut-off score for HAMD-7 that correlated with HAMD-17 ≤ 7 was also determined (Appendix 2).[1,14] HAMD-7 required minutes to administer and served as an efficient and reliable measure of therapeutic progress and symptomatic remission.

Our main objective in this study was to validate the HAMD-7 scale in a primary care setting by comparing its psychometric properties with those of 2 accepted measurement tools, HAMD-17 and the Montgomery–Asberg Depression Rating Scale (MADRS).

Methods

We identified English- or French-speaking patients 18 years of age or older who met the criteria for a major depressive disorder (a minimum baseline HAMD-17 score 18, whether single-episode or recurrent) as defined in the *Diagnostic and Statistical Manual of Mental Disorders*, fourth edition, text revision (DSM-IV-TR).[19] Patients with current psychiatric or medical comorbidity were eligible only if the comorbid condition was not the primary focus of clinical attention. Eligible patients provided written informed consent.

We excluded patients if their depression was chronic (2 years of syndromal severity) or resistant to treatment (failure to respond to more than 2 antidepressants of dissimilar classes), or who had a primary clinical diagnosis other than DSM-IV-TR major depressive disorder (e.g., bipolar disorder). We also excluded patients who were pregnant or lactating, had clinically significant and unstable medical diseases (such as cardiovascular or neurological disorders), were judged to be at significant risk for suicide, or met DSM-IV-TR criteria for substance abuse or dependence during the past 3 months.

To act as primary care investigators in the study, we recruited people with office-based practices in any of Canada's 4 most populous provinces (British Columbia, Alberta, Ontario and Quebec). The final selection of primary care investigators was made after consultation with regional psychiatric consultants. A total of 48 primary care investigators (listed in Appendix 3) were identified who practised at 47 sites; each received extensive training on our study protocol and implementation. Details of the training and standards are in Appendix 2.

All primary care investigators were trained in good clinical practice guidelines.[20] Each site was approved by the Central Institutional Review Board (Aurora, Ont.) and the University of Alberta Research Board (Edmonton, Alta.). Upon completion of the study, all primary care investigators were asked to complete confidential questionnaires pertaining to the usefulness and time requirement to administer HAMD-7.

Eligible patients were assigned by means of computer-generated randomization numbers to HAMD-7 or HAMD-17 as the primary symptom-measurement tool before initiating 8 weeks of open-label, flexible-dose antidepressant monotherapy. Randomization was done at visit 2. Patients receiving an antidepressant for the first time or who required a change in antidepressant medication were eligible to participate. Visits after the baseline visit (which was visit 2) were scheduled every 2 weeks. Medications were chosen by the primary care investigators, in consultation with their patients, from the antidepressants available in Canada during the study (2003–2004): venlafaxine XR, citalopram, paroxetine, mirtazapine, fluoxetine, bupropion SR, sertraline and nefazodone. Although concomitant medications were permitted, patients could not be simultaneously enrolled in manual-based psychotherapy (e.g., cognitive behavioural therapy or interpersonal therapy) or receiving electroconvulsive therapy. Symptom severity was evaluated at each visit with either the HAMD-7 or HAMD-17 tool and with the Clinical Global Impression, Improvement or Severity of Illness scales (CGI-I/S). The MADRS test was administered at baseline (visit 2) and endpoint (visit 6).

At the end of the investigations, primary care investigators completed a HAMD-7 Rating Scale Investigator Evaluation Form, which requested an estimate of the average time required to administer HAMD-7 and qualitative comments about their satisfaction with the scale.

The study population (intent-to-treat), comprising all patients with a minimum of one postbaseline assessment, was analyzed with the last observation carried forward (LOCF) statistic. It was estimated that 375 evaluable patients were required to detect a small effect size (0.2) with a 2-tailed, paired t test (total HAMD-7 and HAMD-17 scores at the beginning and end of treatment) with a between-group correlation of 0.25, at a power of 80%. Assuming a 25% attrition of patients, a sample size of 500 patients was targeted initially.

For a comparison of scores evaluated with 2 related, but nevertheless different, measures of depression, a standard score was calculated for all postbaseline visits (i.e., for visits 3–6). For each HAMD-7 and HAMD-17 score acquired after randomization, the standard score (z) was calculated with the baseline group mean and its associated standard deviation (SD), HAMD-7 14.0 (SD 2.93) and HAMD-17 23.1 (SD 5.09), as follows:

$$\text{HAMD-7 z-score} = [(\text{HAMD-7 score}) - (14.0)] \div 2.93$$
$$\text{HAMD-17 z-score} = [(\text{HAMD-17 score}) - (23.1)] \div 5.09$$

We evaluated changes over 8 weeks of antidepressant treatment in standard scores with a repeated measures analysis of variance (ANOVA) statistic with time as the within-subject factor and randomization as the between-subject factor. Differences at each visit were also evaluated with an independent-sample Student's t test. Convergent validity of HAMD-7 was evaluated via correlations between depressive symptom severity, overall change, response (> 50% reduction in pretreatment total MADRS + HAMD-17 scores) and remission of depressive symptoms (i.e., MADRS ≤ 10, HAMD-17 ≤ 7) and the CGI-I/S. A correlation coefficient corresponding to a p value < 0.001 was deemed clinically significant.

Response to therapy was defined as a reduction of 50% from pretreatment in depression symptom severity; remission was defined as a final score on HAMD-17 ≤ 7, HAMD-7 ≤ 3 and MADRS ≤ 10. Categorical data were analyzed with

the χ^2 statistic. Dichotomous variables (i.e., response and remission status) were compared using Fisher's exact test. Pearson's correlation coefficient (r) with a 2-tailed test of significance was employed to quantify the agreement between the MADRS, HAMD-7 and HAMD-17 total scores. Spearman's correlation coefficients (ρ) were employed to compare HAMD-7 and HAMD-17 with the ordinal measures CGI-S and CGI-I. The internal consistency of HAMD-7 was evaluated with the Cronbach's α statistic for all measures of depression severity.

Results

Of 454 patients [164 males (36.1%) and 290 females (63.9%)] enrolled in the study (Fig. 1), a total of 410 were randomized to HAMD-7 ($n = 205$) or HAMD-17 ($n = 205$) as the primary symptom measurement scale (Table 1).

The mean total scores were 14.0 pretreatment to 5.31 at end point, for patients evaluated with HAMD-7; and 23.1 pretreatment to 8.06 at end point, for those evaluated with HAMD-17. The overall score reduction was highly significant

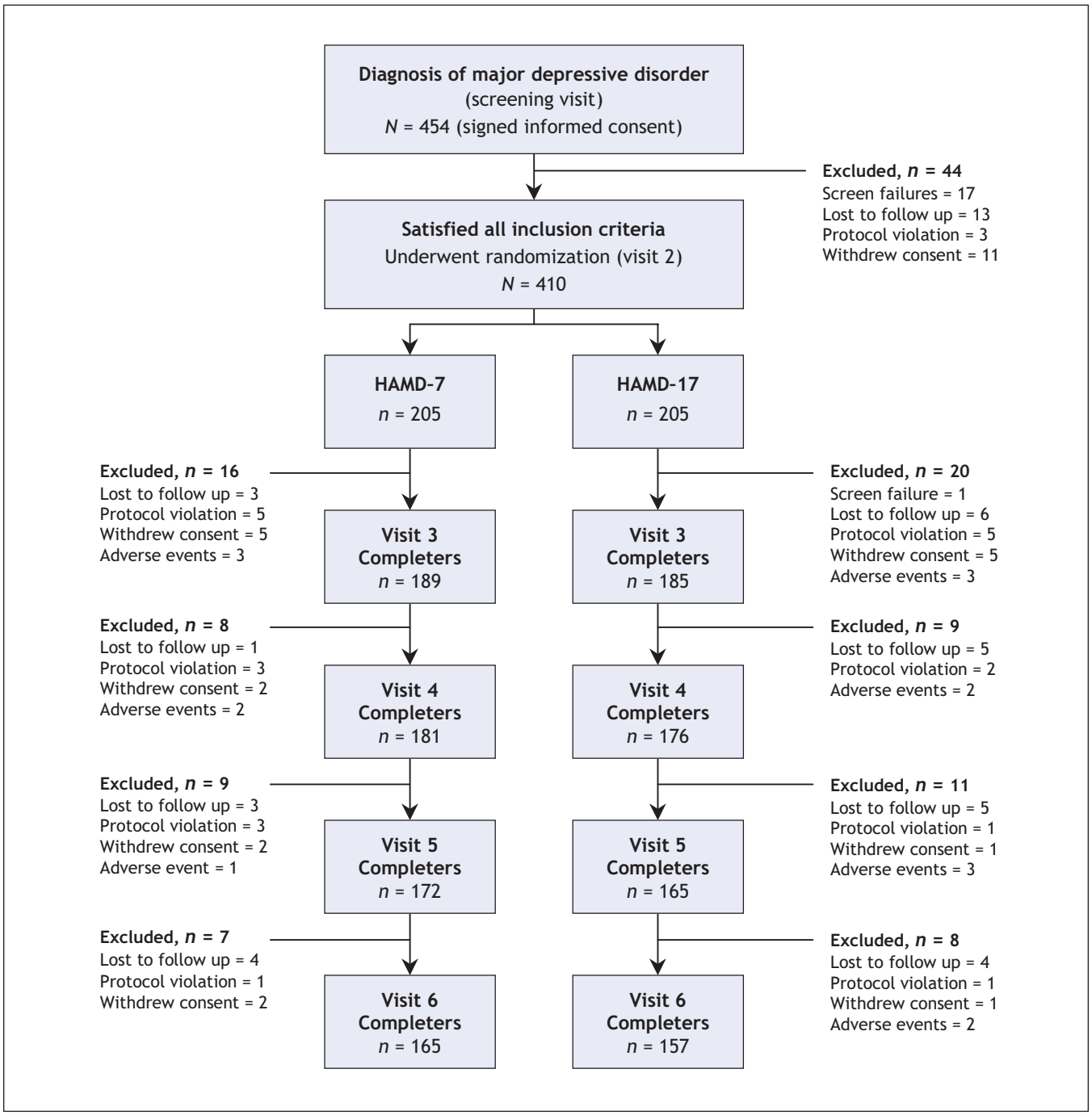

Fig. 1: The evolution of study group numbers, from the original patient cohort at screening through end point (visit 6).

Table 1: Patient characteristics and study outcomes

Characteristic or variable	HAMD-17 *(n = 205)*	HAMD-7 *(n = 205)*
Baseline characteristics		
Age, mean (SD), yr	43.1 (13.0)	42.9 (13.4)
Female, no. (%)	131 (64)	135 (66)
Single episode,* no. (%)	100 (49)	94 (46)
Concomitant medications, no. (%)	116 (57)	112 (55)
CGI-S, mean score (SD)	4.12 (0.77)	4.23 (0.76)
MADRS, mean score (SD)	28.0 (7.6)	29.8 (7.0)
HAMD scores		
Baseline, mean (SD)	23.10 (5.09)	14.00 (2.93)
End point, mean (SD)	8.06 (6.29)	5.31 (4.36)
Study outcomes: patients showing improvement, no. (%)		
Response†: score reduced ≥ 50%	152 (74)	137 (67)
Remission‡: HAMD-17 score ≤ 7 or HAMD-7 score ≤ 3	100 (49)	82 (40)

Note: HAMD = the 17-item or 7-item Hamilton Depression Rating Scale, SD = standard deviation, CGI-S = Clinical Global Impression — Severity of Illness subscale, MADRS = Montgomery–Asberg Depression Rating Scale.
*As opposed to recurrent depressive episodes.
†$p = 0.43$
‡$p = 0.17$

($p < 0.001$), measured with either rating scale (Table 1). Between-group differences in the percentage of patients responding or remitting with therapy in the HAMD-7 group (67% responding and 40% remitting) and the HAMD-17 group (74% and 49%, respectively) were nonsignificant ($p = 0.43$ and 0.17, respectively). There was also a significant pretreatment-to-endpoint change in the standardized HAMD-17 and HAMD-7 ($p < 0.001$), without a main effect of group ($p = 0.84$) or group-by-time interaction ($p = 0.83$), suggesting that sensitivity to change was similar for both scales.

Within the group assigned to HAMD-17 as the primary symptom measurement tool, the items encompassed in the HAMD-7 scale were abstracted (HAMD-7A) and noted to highly correlate with HAMD-17 total scores ($p < 0.001$). Pretreatment-to-endpoint change in depressive symptom severity, response rate and remission rate for HAMD-7A and HAMD-17 were all significantly correlated (all $p < 0.001$).

The internal consistency of the HAMD-7, HAMD-7A and HAMD-17 ratings at each postbaseline visit was satisfactory and comparable (Table 2). Comparison with the MADRS depression rating scale demonstrated that HAMD-7, HAMD-7A and HAMD-17 also showed satisfactory convergent validity in depressive symptom severity, overall change, response (50% reduction in pretreatment total MADRS score) and remission of depressive symptoms (MADRS ≤ 10) (Table 3). The estimation of depressive symptom severity and change with treatment was also highly correlated between HAMD-7 and the CGI-I/S (Table 4).

Of 48 physicians, 39 (82%) completed the HAMD-7 Rating Scale Investigator Evaluation Form. Physicians reported a high overall level of satisfaction with HAMD-7, noting that it was brief to administer (3–4 minutes for 85% of respondents), which facilitated the efficient and structured evaluation of salient depressive symptoms.

Table 2: Internal scoring consistency (α)* of the 7- and 17-item versions of the Hamilton Depression Rating Scale (HAMD)

	Visit				
Scale	2†	3	4	5	6
HAMD-7	0.51	0.757	0.792	0.796	0.843
HAMD-7A‡	0.57	0.737	0.808	0.805	0.825
HAMD-17	0.57	0.776	0.819	0.840	0.848

*Expressed as Cronbach's α; all values > 0.70 were considered satisfactory.
†Pretreatment baseline scores were recorded at visit 2.
‡Abstracted 7 items from HAMD-17 that also appear in HAMD-7.

Interpretation

HAMD-7 was as sensitive as HAMD-17 in estimating the severity of depressive symptoms and evaluating the effectiveness of antidepressant treatment in a naturalistic primary care setting. The proportion of patients estimated to have achieved remission with HAMD-7 was statistically similar to the "gold standard" tool, the HAMD-17 rating scale. That the brevity of HAMD-7 did not appear to compromise vital information on patient progress and outcome was indicated by a high correlation with the multidimensional MADRS and CGI-I/S scales, and by acceptable levels of sensitivity and specificity. (Sensitivity, specificity and other psychometric properties are further described in a subsequent companion paper.)

Over the past decade, a fully asymptomatic state of remission has been emphasized as a critical end point in the management of depressed patients. A universally agreed-upon criterion for remission, however, does not currently exist, which belies the clinical utility of the remission concept. Notwithstanding, the proposed definition and operational criteria for remission (HAMD-17 ≤ 7) put forth by the McArthur Foundation group,[21] which is the definition of remission most cited, has served as a useful heuristic.

Several multinational expert guidelines on the management of depressive disorders emphasize remission, an outcome that transcends response, as an achievable and more clinically relevant symptomatic endpoint.[1,22-24] Residual depressive symptoms and incomplete remission are associated with early relapse, shorter duration between depressive episodes, chronicity, poor prognosis of comorbid medical disorders, increased use of medical services, sustained elevation of suicide risk, and psychosocial and functional deficits.[25]

In the absence of a clinically useful and validated biological marker for remission in depression, clinicians are limited to empirically evaluating depressive symptoms and functional domains.[2] Paradoxically, most practitioners do not systematically evaluate patient progress with quantifiable measures. Although it is likely that clinical willingness to carefully track depressive symptoms is affected by multiple

Table 3: Convergent validity (r)* of the 7- and 17-item versions of the Hamilton Depression Rating Scale with the Montgomery–Asberg Depression Rating Scale

	Correlation* with MADRS			
	Symptom severity		Change with treatment	
HAMD version	Pre-treatment	End point	Response rate†	Remission rate‡
HAMD-7	0.576	0.904	0.716	0.662
HAMD-7A§	0.649	0.894	0.747	0.615
HAMD-17	0.672	0.923	0.730	0.710

*Expressed as Pearson's correlation coefficient ($p < 0.001$ for all r values shown).
†Treatment response is an end-point score of 50% or less of pretreatment score.
‡Total remission requires a score of ≤ 3 (HAMD-7, HAMD-7A) or ≤ 7 (HAMD-17).
§Abstracted 7 items from HAMD-17 that also appear in HAMD-7.

Table 4: Convergent validity (ρ)* of HAMD scores with CGI subscale ratings

	Correlation of HAMD scores with CGI ratings					
	Visit; HAMD with CGI Severity of Illness					HAMD change‡ with CGI Improvement
Scores	2†	3	4	5	6	
HAMD-7	0.655	0.721	0.781	0.839	0.865	-0.644
HAMD-7A§	0.517	0.723	0.806	0.822	0.868	-0.599
HAMD-17	0.624	0.754	0.818	0.850	0.848	-0.599

Note: HAMD = Hamilton Depression Rating Scale, CGI = the Clinical Global Impression scale.
*Correlations are expressed as Spearman's ρ ($p < 0.001$ for all ρ values shown).
†Pretreatment baseline.
‡Baseline HAMD score (visit 2) minus study end point HAMD score (at visit 6).
§Abstracted 7 items from HAMD-17 that also appear in HAMD-7.

variables, it is likely that time-efficient tools would have greater acceptance in the field.

Limitations of this validation study include the heterogeneity of patients enrolled and treatment assignment. For example, diagnostic criteria for a current major depressive episode was based on clinical judgment, and there was no rigorous control for comorbidity other than the exclusion criteria described above. Although the pretreatment MADRS scores were statistically significantly higher in the group randomized to HAMD-7, the differences between the groups are clinically insignificant. A further limitation is the presumption that the threshold scores of HAMD-17 ≤ 7 or MADRS ≤ 10 are prima facie evidence of depressive episode remission. It has been reported, for example, that depressed patients with HAMD-17 scores ≤ 7 may still manifest clinically significant disease activity.[26] On a further note, we chose HAMD-17 as the primary standard because it has been the most commonly employed and familiar metric both in clinical research on depression and among clinicians. An alternative methodology could have been to compare HAMD-7 to MADRS or to the global psychopathology measure, CGI. Lastly, for various reasons 44 patients withdrew from the study after randomization but before treatment began, and were not included in the analysis. Inclusion of these patients and ascribing them an outcome did not materially change the statistical results.

In a busy primary care setting, self-administered scales[17,27] are an appealing alternative to MADRS, HAMD-17 and other lengthier depression metrics. Several studies, including a meta-analysis, have determined, however, that scales administered by clinicians may be more sensitive to change than self-rated measures, particularly in short-term studies.[28,29] A practical and meaningful marker of remission should simultaneously evaluate both symptomatic and functional outcomes. HAMD-7 is primarily a symptom-measurement tool, inviting the need for additional monitoring of functional outcomes. Moreover, the mean doses of antidepressants in the study were at the lower end of the recommended ranges. However, it should be emphasized that the naturalistic setting, nonstandardization of treatment selection and patient heterogeneity in this study reflect real-world practice.

Conclusion

The HAMD-7 rating scale is the first brief-to-administer depression scale with a remission cut-off score validated in both specialty mental-health and primary care settings. The remission cut-off score (correlating with HAMD-17 ≤ 7) differentiates HAMD-7 from any other brief measure of depression that currently exists. A therapeutic target in the management of depression should be a HAMD-7 score ≤ 3; a vista for future research will be to establish if this objective measure corresponds with an absence of disease activity (e.g., as evinced by neuroimaging and neuroendocrine biomarkers). The routine clinical use of the HAMD-7 scale provides objective quantifiable evidence of depressive symptom severity, antidepressant effectiveness and remission of disease.

This article has been peer reviewed.

From the University Health Network, Toronto, Ont. (McIntyre, Parikh, Grigoriadis, Kennedy); Department of Psychiatry, University of Toronto, Toronto, Ont. (McIntyre, Konarski, Mancini, Fulton, Parikh, Grigoriadis, Grupp, Kennedy); Institute of Medical Science,University of Toronto, Toronto, Ont. (Konarski, Kennedy); Department of Pharmacology, University of Toronto, Toronto, Ont. (Grupp); Ottawa Psychopharmacology Clinic, Ottawa, Ont. (Bakish); Centre de Recherche Laval–Robert Giffard, Beauport, Que. (Filteau); Department of Psychiatry, University of Calgary, Calgary, Alta. (Gorman); and School of Medicine, Emory University, Atlanta, Ga. (Nemeroff).

Competing interests: None declared by Jakub Konarski, Deborah Mancini, Kari Fulton, Sophie Grigoriadis, Larry Grupp, Marie-Josee Filteau or Chris Gorman. Roger McIntyre receives research-grant support and is a consultant and part of the speakers' bureau for Wyeth, AstraZeneca, Janssen, Eli Lilly, Bristol-Myers Squibb, Merck, Pfizer, Organon, Lundbeck, GlaxoSmithKline and Prestwick. Sagar Parikh has received consulting fees and teaching honoraria from Wyeth. Within the past 2 years David Bakish has received grants-in-aid to do research on products made by Eli Lilly, Roche Canada, Wyeth, Lundbeck Canada, Shire Canada, Pfizer Canada, Boehringer Ingelheim Canada, Sanofi Aventis and AstraZeneca, and speaker's fees from Wyeth Canada, Lundbeck Canada, Eli Lilly Canada and Shire Canada. Charles Nemeroff has received grant and research support from Abbott, AstraZeneca, Bristol-Myers Squibb, Eli Lilly, Forest, GlaxoSmithKline, Janssen, National Alliance for Research on Schizophrenia and Depression, National Institute of Mental Health, Organon, Pfizer, Pharmacia-Upjohn, Stanley Foundation/NAMI, and Wyeth-Ayerst; has served as a consultant or speaker for Abbott, Acadia, AstraZeneca, Bristol-Myers Squibb, Cephalon, Corcept, Cypress Biosciences, Eli Lilly, Forest, GlaxoSmithKline, Janssen, Merck, Mindsense, Neurocrine

Biosciences, Novartis, Organon, Otsuka, Pfizer, Pharmacia-Upjohn, Sanofi, Somerset, Vela, and Wyeth-Ayerst; and is a stockholder in Corcept. Sidney Kennedy has received research support from Pfizer, AstraZeneca, Organon and Boehringer Ingelheim; is on the speakers' bureaus of Lundbeck, Organon, Wyeth-Ayerst and GlaxoSmithKline; and serves on advisory boards for Pfizer, the Lundbeck Foundation, Eli Lilly, GlaxoSmithKline and Servier.

Contributors: Roger McIntyre and Sidney Kennedy were involved in developing the concept and design of the study, and manuscript preparation and critical revision. Deborah Mancini was active in the critical revision of the manuscript and in data acquisition and management. Jakub Konarski participated in manuscript preparation and critical revision, and data acquisition and management. Charles Nemeroff contributed to the preparation and critical revision of the manuscript. Chris Gorman, David Bakish and Larry Grupp participation in study design and critical manuscript review. Charles Nemeroff contributed to the preparation and critical revision of the manuscript. Marie-Josee Filteau and Kari Fulton did critical manuscript revision and data acquisition. Sagar Parikh and Sophie Grigoriadis also did critical revision of the manuscript.

Acknowledgements: We thank R. Michael Bagby for his valuable comments about the development of the HAMD-7 scale; Ed Vidgen, Robyn Beck, Isaac Chen and Barry McFarlane for their technical assistance; and all of the primary care investigators for their feedback.

This investigation was supported by Wyeth Canada Inc.

REFERENCES

1. Thase ME, Sloan DM, Kornstein SG. Remission as the critical outcome of depression treatment. *Psychopharmacol Bull* 2002;36(4 Suppl 3):12-25.
2. Keller MB. Past, present, and future directions for defining optimal treatment outcome in depression: remission and beyond. *JAMA* 2003;289(23):3152-60.
3. Von Korff M, Katon W, Unutzer J, Wells K, Wagner EH. Improving depression care: barriers, solutions, and research needs. *J Fam Pract* 2001;50(6):E1.
4. Glassman AH, Shapiro PA. Depression and the course of coronary artery disease. *Am J Psychiatry* 1998;155(1):4-11.
5. Wells KB, Stewart A, Hays RD, Burnam MA, Rogers W, Daniels M, Berry S, Greenfield S, Ware J. The functioning and well-being of depressed patients: results from the Medical Outcomes Study. *JAMA* 1989;262(7):914-9.
6. Kessler RC, Berglund P, Demler O, Jin R, Koretz D, Merikangas KR, Rush AJ, Walters EE, Wang PS. The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). *JAMA* 2003;289(23):3095-105.
7. Frank E, Prien RF, Jarrett RB, Keller MB, Kupfer DJ, Lavori PW, et al. Conceptualization and rationale for consensus definitions of terms in major depressive disorder: remission, recovery, relapse, and recurrence. *Arch Gen Psychiatry* 1991;48(9):851-5.
8. Judd LL, Akiskal HS, Maser JD, Zeller PJ, Endicott J, Coryell W, et al. A prospective 12-year study of subsyndromal and syndromal depressive symptoms in unipolar major depressive disorders. *Arch Gen Psychiatry* 1998;55(8):694-700.
9. Brown ES, Varghese FP, McEwen BS. Association of depression with medical illness: Does cortisol play a role? *Biol Psychiatry* 2004;55(1):1-9.
10. Paykel ES. Remission and residual symptomatology in major depression. *Psychopathology* 1998;31(1):5-14.
11. Faries D, Herrera J, Rayamajhi J, DeBrota D, Demitrack M, Potter WZ. The responsiveness of the Hamilton Depression Rating Scale. *J Psychiatr Res* 2000;34(1):3-10.
12. Bagby RM, Ryder AG, Schuller DR, Marshall MB. The Hamilton Depression Rating Scale: Has the gold standard become a lead weight? *Am J Psychiatry* 2004;161(12):2163-77.
13. Bech P, Gram LF, Dein E, Jacobsen O, Vitger J, Bolwig TG. Quantitative rating of depressive states. *Acta Psychiatr Scand* 1975;51(3):161-70.
14. McIntyre R, Kennedy S, Bagby RM, Bakish D. Assessing full remission. *J Psychiatry Neurosci* 2002;27(4):235-9.
15. Maier W, Philipp M, Heuser I, Schlegel S, Buller R, Wetzel H. Improving depression severity assessment — I. Reliability, internal validity and sensitivity to change of three observer depression scales. *J Psychiatr Res* 1988;22(1):3-12.
16. Gibbons RD, Clark DC, Kupfer DJ. Exactly what does the Hamilton Depression Rating Scale measure? *J Psychiatr Res* 1993;27(3):259-73.
17. Spitzer RL Kroenke K, Williams JB. 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(18):1737-44.
18. Zimmerman M, Posternak MA, Chelminski I. Using a self-report depression scale to identify remission in depressed outpatients. *Am J Psychiatry 2004*;161:1911-3.
19. *Diagnostic and statistical manual of mental disorders*. 4th ed, text revision. Washington: American Psychiatric Association; 2000.
20. Papakostas GI, Petersen T, Hughes ME, Nierenberg AA, Alpert JE, Fava M. Anxiety and somatic symptoms as predictors of treatment-related adverse events in major depressive disorder. *Psychiatry Res* 2004;126(3):287-90.
21. Frank E, Kupfer DJ, Perel JM, Cornes C, Mallinger AG, Thase ME, et al. Comparison of full-dose versus half-dose pharmacotherapy in the maintenance treatment of recurrent depression. *J Affect Disord* 1993;27(3):139-45.
22. American Psychiatric Association. Practice guideline for the treatment of patients with major depressive disorder (revision). *Am J Psychiatry* 2000;157(4 Suppl):1-45.
23. Anderson IM, Nutt DJ, Deakin JF; British Association for Psychopharmacology. Evidence-based guidelines for treating depressive disorders with antidepressants: a revision of the 1993 British Association for Psychopharmacology guidelines [review]. *J Psychopharmacol* 2000;14(1):3-20.
24. Kennedy SH, Lam RW, Cohen NL, Ravindran AV. Clinical guidelines for the treatment of depressive disorders. IV. Medications and other biological treatments. *Can J Psychiatry* 2001;46(Suppl 1):38S-58S.
25. Keller MB. Remission versus response: the new gold standard of antidepressant care. *J Clin Psychiatry* 2004;65(Suppl 4):53-9.
26. Nierenberg AA, Keefe BR, Leslie VC, Alpert JE, Pava JA, Worthington JJ 3rd, et al. Residual symptoms in depressed patients who respond acutely to fluoxetine. *J Clin Psychiatry* 1999;60(4):221-5.
27. Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J. An inventory for measuring depression. *Arch Gen Psychiatry* 1961;4:561-71.
28. Lambert MJ, Masters KS, Astle D. An effect-size comparison of the Beck, Zung, and Hamilton Rating Scales for Depression: a three-week and twelve-week analysis. *Psychol Rep* 1988;63(2):467-70.
29. Sayer NA, Sackeim HA, Moeller JR, Prudic J, Devanand DP, Coleman EA, Kiersky JE. The relations between observer-rating and self-report of depressive symptomatology. *Psychol Assess* 1993;5:350-60.

Correspondence to: *Dr. Roger S. McIntyre, Mood Disorders Psychopharmacology Unit, University Health Network, 399 Bathurst St., Toronto ON M5T 2S8; fax 416 603 5368; roger.mcintyre@uhn.on.ca*

Editor's take

- It is important to determine when patients being treated with major depressive disorder recover or enter remission.
- In this RCT, patients being treated with pharmacologic agents for depression were randomized to receive ongoing assessments with a standard 17-item research questionnaire, the HAMD-17, or a shorter clinical version of the Hamilton Depression Rating Scale, the HAMD-7. The shorter version was as effective as the longer version in detecting remissions.

Implications for practice: The 7-item HAMD-7 measure of depression can be used to determine when patients with major depressive disorders are in remission.

Appendix 1: The 7-Item Hamilton Depression Rating Scale (HAMD-7)

1. Depressed mood (sadness, the blues, weepiness) • Have you been feeling down or depressed this past week? • How often have you felt this way, and for how long?	[ ] Absent [ ] Indicated only on questioning [ ] Spontaneously reported verbally [ ] Communicates nonverbally (facial expression, posture, voice, tendency to weep) [ ] Patient reports *virtually only* these feeling states in spontaneous verbal and nonverbal communication
2. Feelings of guilt (self-criticism, self-reproach) • In the past week, have you felt guilty about something you've done, or that you've let others down? • Do you feel you're being punished by being sick?	[ ] Absent [ ] Self-reproach (letting people down) [ ] Ideas of guilt or rumination over past errors or sinful deeds [ ] Present illness seen as punishment; delusions of guilt [ ] Hears accusatory or denunciatory voices or experiences threatening visual hallucinations
3. Interest, pleasure, level of activities (work and activities of daily living) • Are you as productive at work and at home as usual? • Have you felt interested in doing things that usually interest you?	[ ] No difficulty [ ] Fatigue, weakness or thoughts of incapacity (related to activities, work or hobbies) [ ] Loss of interest in activities (directly reported or indirectly through listlessness, indecision and vacillation) [ ] Decrease in actual time spent in activities or in productivity [ ] Stopped working because of current illness
4. Tension, nervousness (psychological anxiety) • Have you been feeling more tense or nervous than usual this week? • Have you been worrying a lot?	[ ] No difficulty [ ] Subjective tension and irritability [ ] Worrying about minor matters [ ] Apprehensive attitude apparent in face or speech [ ] Fears expressed without questioning
5. Physical symptoms of anxiety (somatic anxiety) • How much have these things been bothering you in this past week? DON'T RATE IF SYMPTOMS ARE CLEARLY DUE TO MEDICATION: • In the past week, have you had any of these symptoms? — Gastrointestinal: dry mouth, gas, indigestion, diarrhea, cramps, belching — Cardiovascular: heart palpitations, headaches — Respiratory: hyperventilation, sighing — Having to urinate frequently — Sweating	[ ] Absent [ ] Mild [ ] Moderate [ ] Severe [ ] Incapacitating
6. Energy level (somatic symptoms) • How has your energy been this past week? • Have you felt tired? • Have you had any aches or pains or felt any heaviness in your limbs, back or head?	[ ] None [ ] Heaviness in limbs, back or head (backache, headache, muscle aches; loss of energy and fatigability) [ ] Any clear-cut symptom rates 2 points
7. Suicide (ideation, thoughts, plans, attempts) • Have you any thoughts life is not worth living or you'd be better off dead? • Have you thoughts of hurting or killing yourself? • Have you done anything to hurt yourself?	[ ] Absent [ ] Feels life is not worth living [ ] Wishes to be dead (or any thoughts of possible death to self) [ ] Suicidal ideas or gestures [ ] Attempts at suicide (any serious attempt rates 4)
HAMD-7 score ≥ 3 indicates full remission. HAMD-7 score ≤ 4 indicates non/partial response.	**Total score:**

Appendix 2: Details of the training and standards used

All primary care investigators demonstrated high interrater reliability (κ_w) on the primary measures of depressive symptoms: the 7-item and 17-item Hamilton Depression Rating Scales (HAMD-7 0.83, HAMD-17 0.98), the Montgomery-Asberg Depression Rating Scale (MADRS 0.89), and the Clinical Global Impression, Severity of Illness subscale (CGI-S 0.80).

For all patients, all rating items were scored by 2 of the investigators (R.S.M. and S.H.K.) before the primary care investigators were trained. A single expert standard was derived by agreement, and for all items, a single related rating was made available:

- Item ratings within the range of the expert standard ± 1 were assumed acceptable and received a credit of 1
- Ratings that fell outside this range received a credit of 0.

Weighted kappa values were calculated with this formula:

$$\kappa_w = \frac{(p_o - p_c)}{(1 - p_c)}$$

For each item, the observed agreement (p_o) was calculated as the proportion of participants with a credit of 1. For example, for item 1, n = xx (with ratings within the expert standard ± 1) and total n = yy:

$$p_o = \frac{xx}{yy}$$

For each item, the chance agreement (p_c) was calculated by the actual chance for agreement, with the expert standard ± 1. In case of the HAMD-17 measurement tool, 5-point and 3-point scales were applied; depending on the expert standard, p_c can receive the following values.

Accepted range of chance agreement (p_c)

- For 5-point scale items (items 1–3, 7–11 and 15):
  - 0, acceptable ratings 0 or 1, $p_c = 2/5 = 0.4$
  - 1, acceptable ratings 0 or 1 or 2, $p_c = 3/5 = 0.6$
  - 2, acceptable ratings 1 or 2 or 3, $p_c = 3/5 = 0.6$
  - 3, acceptable ratings 2 or 3 or 4, $p_c = 3/5 = 0.6$
  - 4, acceptable ratings 3 or 4, $p_c = 2/5 = 0.4$
- For 3-point scale items (items 4–6, 12–14, 16 and 17):
  - 0, acceptable ratings 0 or 1, $p_c = 2/3 = 0.67$
  - 1, acceptable ratings 0 or 1 or 2, $p_c = 3/3 = 1.0$
  - 2, acceptable ratings 1 or 2, $p_c = 2/3 = 0.67$

The higher κ_w values noted with the HAMD-17 versus those with the HAMD-7 scale may be inconsistent with previous literature (Bagby et al,[12] *Am J Psychiatry* 2004;161:2163-77). This issue is being explored in a separate manuscript.

Appendix 3: The 48 primary care investigators who participated in studying the 7-item version of the Hamilton Depression Rating Scale (HAMD-7), by province

British Columbia: Drs. Brian Carlson, Coquitlam; Michael Golbey, Kelowna; John Kelly, Victoria; Douglas Leitner, Penticton; Paul Murray, Prince George; Anthony Nielson, Victoria; Anthony Ocana, Vancouver; and Marianne Russell, Coquitlam

Alberta: Drs. George Barr, Calgary; John Bromley, Red Deer; and Edward Papp, Edmonton

Ontario: Drs. Norman Abramson, Mississauga; Murray Awde, London; John Axler, Toronto; Yee Ling Chang, Toronto; Arif Chaudhri, Etobicoke; Chin Chung, Willowdale; Ronald Cox, Brampton; John Dawson, Richmond Hill; Larry Deutch, Ottawa; Giuseppe D'Ignazio, Hawkesbury; Alan Greenspoon, Hamilton; Steven Grossman, Richmond Hill; Margaret Grunebaum, North York; Tommy Hong, Mississauga; Alan Kaplan, Richmond Hill; Dennis Kavalsky, Hamilton; James Kim, Brampton; Christiane Kuntz, Ottawa; Douglas Mah, Mississauga; Krisanne Mendelssohn, Scarborough; Paul Perlon, Richmond Hill; Maryam Rostami, Milton; Andre Roch, Sudbury; Irving Siegel, Markham; Eric Silver, Toronto; Laurie Wells, Dundas; Linda Yolles, Toronto; Lauren Zeilig, Toronto; and Paul Ziter, Windsor

Quebec: Drs. Denis Beaulieu, Val-Belair; Guy Chouinard, Charlesbourg; Jean Paul Czitrom, Montreal; Claude Laroche, Montreal; Suzanne Pelchat, Loretteville; Jean Ross, Sillery; and Julie Ross, Charlesbourg