An examination of the sensitivity of the six-item Hamilton Rating Scale for Depression in a sample of patients suffering from major depressive disorder

Cynthia L. Hooper, MA; David Bakish, MD

Hooper — Institute of Mental Health Research and Psychopharmacology Unit, Royal Ottawa Hospital; Bakish — Institute of Mental Health Research and Psychopharmacology Unit, Royal Ottawa Hospital and Department of Psychiatry, University of Ottawa, Ottawa, Ont.

Objectives: To compare the sensitivity of the 6-item Hamilton Rating Scale for Depression (HRSD6) with the more widely used 17-item Hamilton Rating Scale for Depression (HRSD17) in patients suffering from major depressive disorder, with or without melancholia and/or dysthymic disorder. A secondary objective was to compare the sensitivity of the HRSD6 to the Montgomery–Asberg Depression Rating Scale (MADRS). **Design:** Retrospective analysis of 4 clinical trials that tested antidepressant therapies. **Setting:** Outpatient treatment in a major psychiatric hospital. **Participants:** One hundred and forty-three male and female outpatients meeting the criteria of the DSM-III-R or DSM-IV for major depressive disorder. **Outcome measures:** HRSD17, HRSD6 and MADRS. **Results:** The HRSD6 correlated strongly with the HRSD17, both at baseline and termination of treatment, and for the subgroups of double depression and melancholia. The HRSD6 was also correlated significantly with the MADRS at both measurement times, and for the subgroups. Paired *t*-tests with the HRSD6, HRSD17 and MADRS demonstrated equal sensitivity to change over the course of treatment, both in the full sample and in the dysthymic and melancholic subgroups. **Conclusions:** The HRSD6 appears to be as sensitive to change over treatment as the HRSD17 and the MADRS. A shorter, less time-consuming measure of depression may have utility in clinical practice and research.

Objectifs : Comparer la sensibilité de l'échelle de dépression de Hamilton à six questions (HRSD6 — *Hamilton Rating Scale for Depression*) à l'échelle plus répandue à 17 questions (HRSD17) chez les patients atteints d'un trouble dépressif majeur avec ou sans mélancolie ou trouble dysthymique. L'étude visait aussi à comparer la sensibilité de l'échelle HRSD6 à l'échelle d'évaluation de la dépression de Montgomery–Asberg (MADRS — *Montgomery–Asberg Depression Rating Scale*). **Conception :** Analyse

Correspondence to: Cynthia L. Hooper, MA, Institute of Mental Health Research, Royal Ottawa Hospital, 1145 Carling Ave., Ottawa ON K1Z 7K4; fax 613 729-8126, chooper@rohcg.on.ca

Medical subject headings: antidepressive agents; clinical trials; depressive disorder; dysthymic disorder; psychiatric status rating scales; research

J Psychiatry Neurosci 2000;25(2):178-84.

Submitted Apr. 12, 1999
Revised Aug. 20, 1999
Accepted Aug. 30, 1999

rétrospective de quatre études cliniques au cours de laquelle on a fait l'essai de thérapies contre la dépression. **Contexte :** Traitement en service externe dans un grand hôpital psychiatrique. **Participants :** Cent quarante-trois patients des deux sexes, traités en service externe, qui satisfaisaient aux critères DSM-III-R ou DSM-IV pour ce qui est d'un trouble majeur de la dépression. **Mesures de résultats :** HRSD17, HRSD6 et MADRS. **Résultats :** On a établi un lien solide entre l'échelle HRSD6 et l'échelle HRSD17 à la fois au début et à la fin du traitement et pour les sous-groupes de sujets atteints à la fois de dépression et de mélancolie. On a aussi établi un lien solide entre l'échelle HRSD6 et l'échelle MADRS aux deux points de mesure et pour les sous-groupes. Des tests *t* jumelés réalisés au moyen des échelles HRSD6, HRSD17 et MADRS ont démontré une sensibilité égale au changement pendant le traitement, à la fois dans tout l'échantillon et chez les sous-groupes de sujets atteints de dysthymie et de mélancolie. **Conclusions :** L'échelle HRSD6 semble aussi sensible au changement pendant le traitement que les échelles HRSD17 et MADRS. Une mesure de la dépression plus courte et qui prend moins de temps peut être utile en pratique clinique et en recherche.

Introduction

Several rating scales are available to aid in the measurement of depression. These include clinician- rated scales such as the Hamilton Rating Scale for Depression (HRSD[1]) and the Montgomery-Asberg Depression Rating Scale (MADRS[2]), and patient rated scales, most notably the Beck Depression Inventory.[3] However, the HRSD has emerged as the standard in depression research. There are several versions of the scale, including 17-item,[1] 21-item[1] and 24-item[4] scales. The 17-item HRSD measures the many symptoms of depression, including anxiety, sleep problems, impact on work and activities, and hypochondriasis. The 21-item HRSD includes items on derealization, diurnal variation, obsessive–compulsive symptoms and paranoia, whereas the 24-item HRSD adds information on helplessness, hopelessness and worthlessness. Despite its acceptance by researchers, the HRSD is still not used extensively by nonresearch clinical staff, possibly because of the time required to administer it.[5]

Bech et al[6-8] carried out several item analyses of the HRSD17 and found that it is not a 1-dimensional measure of depression. The scale is strongly influenced by the subtype of depression. Bech et al proposed a 6-item HRSD (the melancholia subscale) as a more consistent measure of "core" depression. These 6 items are: item 1 — depressed mood, item 2 — guilt, item 7 — work and activities, item 8 — psychomotor retardation, item 10 — psychic anxiety and item 13 — somatic symptoms, general. And although item 13 was the least-sensitive item, the HRSD6 seemed to be a more consistent measure of depression, irrespective of subtype.

O'Sullivan et al[5] followed up on this work by comparing the HRSD6 to the other versions of the scale in 164 patients with major depressive disorder, either atypical or typical. They found that not only was the HRSD6 strongly correlated to the other versions of the HRSD both at baseline and after treatment, but it was also as sensitive to change over treatment. The subtype of depression did not influence the results.

This study is a continuation of the work on the HRSD6. Using our database from a number of different clinical trials, we tested whether the HRSD6 was as sensitive to change in depression severity as the HRSD17, in both the total sample and subgroups of double depression (major depressive disorder plus dysthymia) and melancholia. We also examined the correlations between scores on the HRSD6 and on the HRSD17. Finally, we included the MADRS as a comparator to the HRSD6.

Method

One hundred forty-three patients from 4 completed clinical trials of antidepressants in depression were included in the analyses (unpublished date).[9,10] Of these patients, 47 also had a diagnosis of dysthymic disorder, and 46 a diagnosis of melancholia. Data from the records of these patients were combined. They included variables on gender, marital and employment status, severity and episodic nature of the depression, age of onset of the first episode, duration of the current episode, presence of secondary diagnoses, previous psychotropic medication, and family history of psychiatric disorders. The patient total scores on the HRSD17 and MADRS, at baseline and at the termination of treatment were collected, and the HRSD6 total scores were calculated (the total of items 1, 2, 7, 8, 10 and 13; possible range 0 to 22).

Clinical trial description

The 4 clinical trials were similar in terms of inclusion

and exclusion criteria, and study procedures (unpublished data).[9,10] Three trials (ipsapirone CR , BMS 181101, and BIMT 17BS) were double-blind trials; 1 (nefazodone-pindolol) was an open-label trial. All trials required medically healthy or controlled outpatients with a *Diagnostic and Statistical Manual of Mental Disorders*, third edition, revised (DSM-III-R) or *Diagnostic and Statistical Manual of Mental Disorders*, fourth edition (DSM-IV) diagnosis of major depressive disorder, who were not using any other psychotropic medications. Ipsapirone CR disallowed any other psychiatric disorders, BMS 181101 any other Axis 1 diagnoses, BIMT 17BS any concurrent dementia, psychosis or history of mania, and nefazodone-pindolol any diagnosis of schizophrenia or obsessive–compulsive disorder. Trial length ranged from 4 to 9 weeks, and the inclusion age criterion ranged from 17 to 65 years. All clinical trial protocols were approved by a hospital ethics committee, and all participants gave full and informed consent. The majority of study patients were seen by the same physician (D.B.), and all physician-raters had good interrater reliability on both the HRSD17 and MADRS.

Statistical analysis

Pearson product moment correlations were completed to assess the correlations between the 6- and 17-item HRSD and the MADRS for both the complete sample and the subgroups of melancholic and dysthymic patients, and with baseline and termination (last observation carried forward [LOCF]) of treatment scores separated. Two-tailed paired *t*-tests were used to determine change over treatment. All analyses were carried out for both the total sample, and the subgroups of melancholic and dysthymic patients.

Results

Table 1 contains the demographic and clinical information for the total sample of 143 patients (61% female). Of note is that approximately 75% of the sample had recurrent depression, and the mean (5% trimmed) duration of the current episode was more than 1 year.

All the Pearson product moment correlations between the HRSD and MADRS were positive and significantly higher than would be expected by chance ($p < 0.0005$, Table 2). This was true for the total group, the subgroups of melancholic and dysthymic patients, and if the scores were separated into baseline and termination scores. Although the baseline scores had lower correlation coefficients compared with the rest of the values, they were all still significantly stronger than chance. Fig. 1 is a scattergram of the HRSD6 and HRSD17 total scores for baseline and termination visits. It is interesting that the HRSD6 scores ranged from 0 to 17, close to the theoretical range of 0 to 22, whereas the HRSD17 scores ranged from 0 to 37 only (theoretical range 0 to 52).

Paired *t*-tests were completed on the HRSD6, HRSD17 and MADRS scores, comparing baseline to LOCF for both the total sample and the melancholic and dysthymic subgroups. The HRSD6 was able to detect

Table 1: Demographic and clinical information on the sample of 143 patients with major depressive disorder (MDD)

Variable	Finding
Gender, no. (%)	
Male	56 (39.2)
Female	87 (60.8)
Marital status, no. (%)	
Never married	37 (25.9)
Married/ common-law	41 (28.7)
Separated/divorced	31 (21.7)
Not available	34 (23.8)
Employed (n = 123), no. (%)	
Yes	78 (63.4)
No	45 (36.6)
Episode, no. (%)	
First	37 (25.9)
Recurrent	106 (74.1)
Severity of illness (n = 103), no. (%)	
Moderate	95 (92.2)
Severe	8 (7.8)
Mean (and SD) age of onset of first MDD episode, yr	24.7 (11.1)
Mean (and SD) duration of current MDD episode, wk	96.2 (191.8)
5% trimmed, mean (IQR)	64.86 (88)
Previous psychotropic medications (n = 108), no. (%)	
Yes	44 (40.7)
No	64 (59.3)
Secondary diagnoses, past or present (n = 142), no. (%)	
Yes	71 (50)
No	71 (50)
Family history, no. (%)	
Depression (n = 142)	108 (76.1)
Alcohol abuse (n = 53)	28 (52.8)
Substance abuse (n = 54)	19 (35.2)
Anxiety disorder (n = 106)	32 (30.2)
Suicide attempts (n = 122)	16 (13.1)

SD = standard deviation, IQR = interquartile range.

change over treatment in the total sample of patients (Table 3, Fig. 2), and in the subgroups of melancholic and dysthymic patients.

Discussion

Our results indicate that the HRSD6 is strongly correlated to the HRSD17 at baseline and termination of treatment, and for the subgroups of melancholic and dysthymic patients. It is also strongly related to the MADRS at both baseline and termination of treatment. The HRSD6 is able to detect change over treatment as well as do the HRSD17 and MADRS. The weaker correlations found between the baseline measures are most likely a result of the greater homogeneity of the depressed population at the start of treatment. As some

Table 2: Pearson product moment correlations between the total scores for the 6-item and 17-item Hamilton Rating Scale for Depression (HRSD6 and HRSD17) and the Montgomery–Asberg Depression Rating Scale (MADRS)

	HRSD17			MADRS		
HRSD6	Baseline	Termination	Combined	Baseline	Termination	Combined
Total sample						
Baseline	0.50, n = 143			0.40, n = 140		
Termination		0.95, n = 143			0.93, n = 141	
Combined			0.95, n = 286			0.92, n = 281
Melancholic patients						
Baseline	0.61, n = 46					
Termination		0.96, n = 46				
Combined			0.94, n = 92			
Dysthymic patients						
Baseline	0.56, n = 47					
Termination		0.91, n = 47				
Combined			0.94, n = 94			

All correlations are significant at $p < 0.0005$ (2-tailed test).

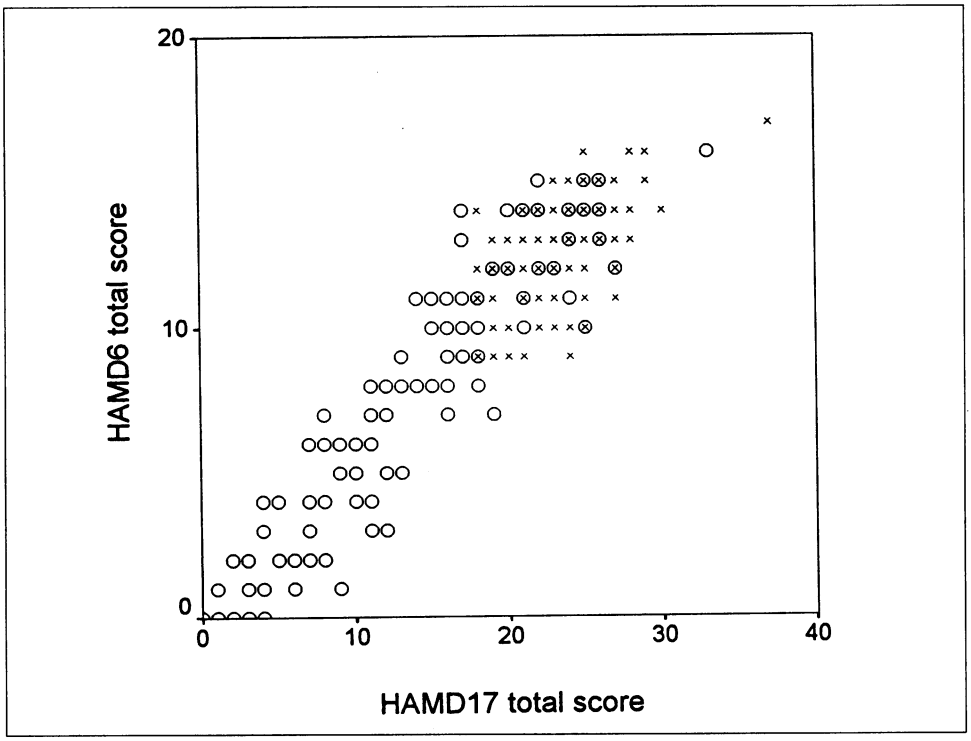

Fig. 1: Scatterplot of Hamilton Rating Scale for Depression (HRSD) total scores, for baseline (crosses) and termination (circles) visits.

patients improved and others did not, there was a greater spread of scores, and stronger correlations resulted.

The ability of the HRSD6 to show comparable results within the total sample and the subgroups of double depression and melancholia, supports the assertion that the HRSD6 measures the "core" depression features. This core depression may be characterized by feelings of sadness, hopelessness and guilt, a decrease in interest or time spent in activities, decreased motor activity and ability to concentrate, a lack of energy, and an increase in tension, irritability and worry. Symptoms such as sleep disturbance, weight change and somatic manifestations of anxiety may be more related to specific sub-

Table 3: Hamilton Rating Scale for Depression (HRSD) and Montgomery–Asberg Depression Rating Scale (MADRS) mean (and standard deviation) total scores at baseline and termination visits for the total sample and for the subgroups of melancholic and dysthymic patients

Scale	Baseline	Termination	Difference	95% confidence interval of the difference	*t* value	*p* value
Total sample						
HRSD17	23.01 (2.82)	10.41 (7.86)	12.60	11.26–13.95	18.509	< 0.0005
HRSD6	12.58 (1.59)	5.85 (4.75)	6.73	5.92–7.54	16.420	< 0.0005
MADRS	27.86 (4.65)	13.81(11.19)	14.05	12.16–15.94	14.691	< 0.0005
Melancholic patients						
HRSD17	22.63 (3.32)	12.48 (8.08)	10.15	7.51–12.80	7.728	< 0.0005
HRSD6	12.46 (1.63)	7.50 (4.94)	4.96	3.44–6.47	6.469	< 0.0005
Dysthymic patients						
HRSD17	23.53 (3.27)	9.98 (6.81)	13.55	11.59–15.52	13.893	< 0.0005
HRSD6	12.72 (1.69)	5.83 (4.42)	6.89	5.66–8.13	11.245	< 0.0005

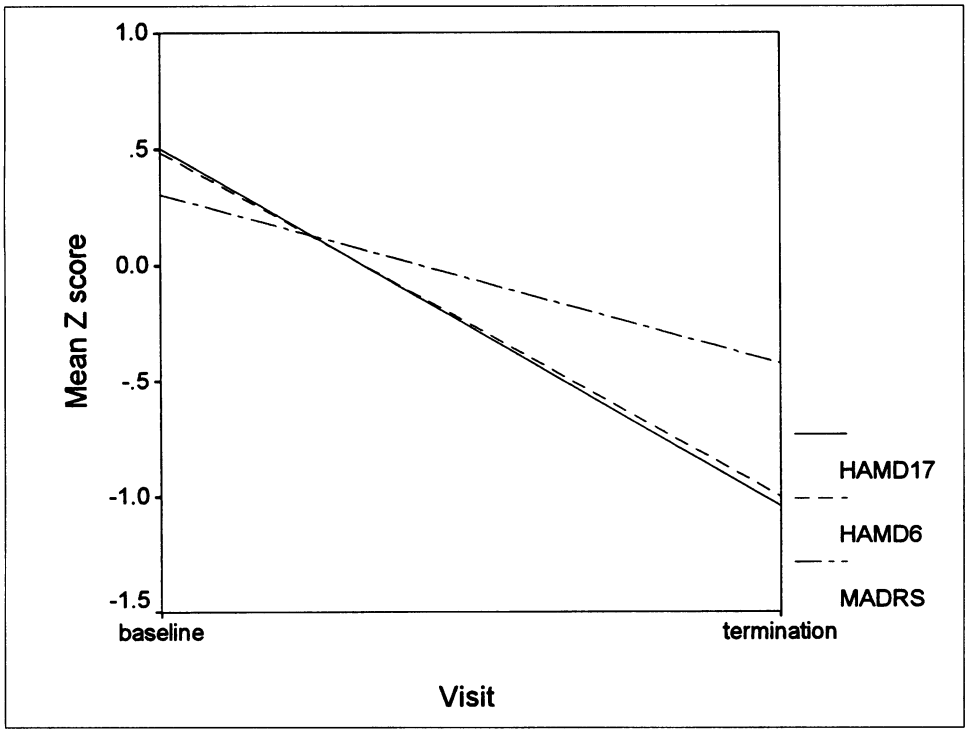

Fig. 2: Mean Z scores for baseline and termination visits for the 6- and 17-item Hamilton Rating Scale for Depression (HRSD6 and HRSD17) and the Montgomery–Asberg Depression Rating Scale (MADRS). Z scores are used for graphic purposes.

types of depression. Other researchers have criticized the HRSD17 as a measure of non-core symptoms. Hammond[11] found the HRSD17 an unreliable measure of depression in the elderly, owing to the high component of items evaluating anxiety and sleep problems. A related problem in using the HRSD17 in elderly patients is the large number of items found to be related to somatic disorders.[12] Linden et al[12] also suggested that several HRSD17 items can be affected by concurrent medication. In fact, Bech et al[6] in their first exploration of the HRSD6 suggested that many of the items of the HRSD17 could be measuring side effects of antidepressants. Symptoms such as insomnia, weight loss, gastrointestinal problems, somatic anxiety and loss of libido can be part of the depressive illness or they can be an adverse event from antidepressants. Finally, a factor analysis by Riskind et al[13] determined that items in the HRSD and HAMA (Hamilton Anxiety Rating Scale[14]) were not measuring only depression and anxiety respectively. In particular, only HRSD items 1 — depressed mood; 2— feelings of guilt; 3 — suicide; 6 — insomnia, late; 7 — work and activities; 8 — retardation; 14 — genital symptoms and 16 — loss of weight were significantly correlated with a diagnosis of major depressive disorder. The rest of the items were either insignificantly correlated or more closely related to a diagnosis of generalized anxiety disorder.

The finding that HRSD6 scores more closely reflected the theoretical range of scores than the HRSD17 scores is interesting. Does it support the assertion that the HRSD17 contains extraneous items that contribute little to the final score? Or does it suggest that the HRSD6 may not be an appropriate measure for severely depressed patients? The current study comprised primarily moderately ill depressed patients. The HRSD6 will have to be tested with severely ill patients to check for ceiling effects before these questions can be answered.

There are several implications of these results. The HRSD6 is a shorter scale for measuring depression than the HRSD17, which makes it more practical for use in clinical settings. This would allow clinicians to objectively rate the progress of their patients with depressive illness. It would also allow them to standardize their treatment approach and measure the efficacy of different treatments in depression. With further research on the reliability and validity of the HRSD6, it may find a place in clinical trial research. With the trend toward greater intensity in clinical trials, involving more frequent visits and measurements, the ability to incorporate a shorter, faster measurement instrument could be useful.

One limitation is that this study only demonstrates equal, not superior, sensitivity of the HRSD6 over the HRSD17 and MADRS. To become fully accepted by scientists, clinicians and regulatory bodies, the HRSD6 may have to demonstrate a measurement advantage over the older, more established scales. It is also important to note that although the HRSD6 may be more useful in measuring change of severity, it cannot replace a thorough diagnostic interview. As well, specific research questions may be better served by using a more detailed measurement instrument. For example, examination of the complete therapeutic effect of a treatment.

In summary, the HRSD6 scale demonstrated equal sensitivity to the more frequently used HRSD17, as well as strong correlations with the HRSD17 and the MADRS. The HRSD6 has several advantages over the HRSD17, including faster administration and a reduced influence of antidepressant adverse effects. It may have a role in the standardization of clinical practice and in antidepressant clinical trials.

Acknowledgements

Three of the clinical trials included in this report were funded independently by Bristol-Myers Squibb Pharmaceutical Research Institute, Miles Canada Inc., and Boehringer Ingelheim (Canada). We also thank the entire psychopharmacology team for their excellent work.

References

1. Hamilton M. A rating scale for depression. *J Neurol Neurosurg Psychiatry* 1960;23:56-62.

2. Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. *Br J Psychiatry* 1979;134:382-9.

3. Beck AT, Ward CH, Mendelson M, Mock EJ. An inventory for measuring depression. *Arch Gen Psychiatry* 1961;4:53-63.

4. Guy W. *ECDEU assessment manual for psychopharmacology*. Washington: US Department of Health, Education, and Welfare; 1976.

5. O'Sullivan RL, Fava M, Agustin C, Baer L, Rosenbaum JF. Sensitivity of the six-item Hamilton Depression Rating Scale. *Acta Psychiatr Scand* 1997;95:379-84.

6. Bech P, Gram LF, Dein E, Jacobsen O, Vitger J, Bolwig TG.

Quantitative rating of depressive states. *Acta Psychiatr Scand* 1975;51:161-70.

7. Bech P, Bolwig TG, Kramp P, Rafaelsen OJ. The Bech–Rafaelsen Mania Scale and the Hamilton Depression Scale. Evaluation of homogeneity and interobserver reliability. *Acta Psychiatr Scand* 1979;59:420-30.

8. Bech P, Allerup P, Gram LF, Reisby N, Rosenberg R, Jacobsen O, et al. The Hamilton Depression Scale: evaluation of objectivity using logistic models. *Acta Psychiatr Scand* 1981;63:290-9.

9, Bakish D, Hooper CL, Thornton MD, Wiens A, Miller CA, Thibaudeau CA. Fast onset: an open study of the treatment of major depressive disorder with nefazodone and pindolol combination therapy. *Int Clin Psychopharmacol* 1997;12:91-7.

10. Lapierre YD, Silverstone P, Reesal RT, Saxena B, Turner P, Bakish D, et al. A Canadian multicenter study of three fixed doses of controlled-release ipsapirone in outpatients with moderate to severe major depression. *J Clin Psychopharmacol* 1998;18:268-73.

11. Hammond MF. Rating depression severity in the elderly physically ill patient: reliability and factor structure of the Hamilton and the Montgomery–Asberg depression rating scales. *Int J Geriatr Psychiatry* 1998;13(4):257-61.

12. Linden M, Borchelt M, Barnow S, Geiselmann B. The impact of somatic morbidity on the Hamilton Depression Rating Scale in the very old. *Acta Psychiatr Scand* 1995;92(2):150-4.

13. Riskind JH, Beck AT, Brown G, Steer RA. Taking the measure of anxiety and depression. Validity of the reconstructed Hamilton scales. *J Nerv Ment Dis* 1987;175(8):474-9.

14. Hamilton M. The assessment of anxiety states by rating. *Br J Med Psychol* 1959;32:50-5.

THE WG DEWHURST LABORATORIES

In October 1999, the laboratories of the Neurochemical Research Unit, Department of Psychiatry, University of Alberta, Canada, were named in honour of Dr. William G. Dewhurst.

A founding Vice-President and then President of the Canadian College of Neuropsychopharmacology (CCNP), Dr. Dewhurst was a pioneer in the study of the involvement of trace amines such as tryptamine and 2-phenylethylamine in affective disorders. During his distinguished career, he has exerted great influence on the development of psychiatry, mental health services and research in biological psychiatry and neuropsychopharmacology, both within Canada and internationally. The Neurochemical Research Unit, in particular, has become a centre of excellence for research in the etiology and pharmacotherapy of psychiatric disorders, analytical neurochemistry and drug metabolism. In recognition of his outstanding contribution to neuropsychopharmacology in Canada, Dr. Dewhurst was awarded the CCNP Medal in 1993. Last year's dedication of the laboratories in his name is indeed a fitting tribute to one of Canada's eminent psychiatrists and neurochemists.