

The British Journal of Psychiatry

A new depression scale designed to be sensitive to change. S A Montgomery and M Asberg BJP 1979, 134:382-389.

from

Access the most recent version at DOI: 10.1192/bjp.134.4.382

This article cites 0 articles, 0 of which you can access for free at: References

http://bjp.rcpsych.org/content/134/4/382#BIBL

To obtain reprints or permission to reproduce material from this paper, please write Reprints/

permissions to permissions@rcpsych.ac.uk

http://bjp.rcpsych.org/letters/submit/bjprcpsych;134/4/382 You can respond to this article at

http://bjp.rcpsych.org/ on October 18, 2014 Published by The Royal College of Psychiatrists **Downloaded**

A New Depression Scale Designed to be Sensitive to Change

By STUART A. MONTGOMERY and MARIE ASBERG

SUMMARY The construction of a depression rating scale designed to be particularly sensitive to treatment effects is described. Ratings of 54 English and 52 Swedish patients on a 65 item comprehensive psychopathology scale were used to identify the 17 most commonly occurring symptoms in primary depressive illness in the combined sample.

Ratings on these 17 items for 64 patients participating in studies of four different antidepressant drugs were used to create a depression scale consisting of the 10 items which showed the largest changes with treatment and the highest correlation to overall change.

The inter-rater reliability of the new depression scale was high. Scores on the scale correlated significantly with scores on a standard rating scale for depression, the Hamilton Rating Scale (HRS), indicating its validity as a general severity estimate. Its capacity to differentiate between responders and non-responders to antidepressant treatment was better than the HRS, indicating greater sensitivity to change. The practical and ethical implications in terms of smaller sample sizes in clinical trials are discussed.

Introduction

The most common use of psychiatric rating scales is probably for comparison of effects of new drugs to standard treatment. In most trials of antidepressant drugs a difference can be demonstrated between pharmacologically active compounds and placebo (Morris and Beck, 1974) but only rarely are consistent differences found between active drugs even when they are known to have different mechanisms of action.

One possible explanation is that the standard rating scales are not sensitive enough to pick up such differences (Angst, 1972) which is not surprising if the scales were not designed specifically for that purpose. The result has often been that the scales reflect diagnostic features rather than being sensitive to change. We therefore decided to construct a rating scale

for depressive illness where sensitivity and accuracy of change estimates were to be major criteria for the inclusion of items. This work was made possible by the recent introduction of a new, comprehensive psychopathological rating scale or CPRS (Åsberg et al, 1978). The CPRS is composed of 65 scaled items covering a wide range of psychiatric symptoms.

Patients and Ratings

The selection of items for the depression scale was based on ratings on the CPRS of 106 depressed patients, 33 men and 73 women, participating in clinical trials of antidepressant drugs. Thirty-three were out-patients and 73 inpatients. In the majority of cases, two raters participated in the interview. Only patients with a primary depressive illness (Feighner et al,

1972) were included. Inventories (Gurney et al, 1972) were applied to ensure diagnostic and descriptive uniformity. Within this group, a wide variation in patient characteristics was sought and the sample includes endogenous and reactive, psychotic and non-psychotic, bipolar and unipolar out-patients and inpatients from a wide age range (18–69 years). Patients from two countries (England n=54, and Sweden n=52) were included in order to reduce cultural bias in the selection of items.

CPRS scores after four weeks therapy with four antidepressant drugs with different pharmacological profiles were used to study change with treatment. The drugs were mianserin, amitriptyline, maprotiline and clomipramine. Amitriptyline (Tuck and Punell, 1973), maprotiline (Maitre et al, 1971) and clomipramine (Hamberger and Tuck, 1973) block neuronal reuptake of noradrenaline, directly or by means of endogenously formed metabolites. Clomipramine is also a potent serotonin uptake blocker (Asberg et al, 1977). Amitriptyline may also affect serotoninergic neurones (Tuck and Punell, 1973) and has pronounced anti-

cholinergic effects (Mass, 1975); it is also thought to have a stronger sedative effect than the other drugs (Silverstone and Turner, 1974). Mianserin appears to lack uptake inhibiting effects (Ferl et al, 1973) as well as anticholinergic effects (Coppen et al, 1976) and its mechanism of action is not clear.

Of the 64 patients for whom scores were available from before and after four weeks of treatment 35 were simultaneously rated on the Hamilton Rating Scale (HRS) (Hamilton, 1967) and on a 7-point scale for global severity of illness.

Construction of the Scale

Parametric statistical methods were used except when dealing with ranked data as recommended by Anderson (1961) and Boneau (1961). The difference between pre-treatment scores has been used as an estimate of change, both for individual and sums of items.

Preliminary item selection

The frequency of scores above zero on all of the 65 items as well as the ranking by inci-

Table I

Frequency of scores above zero and correlation between English and Swedish raters on the 17 most commonly scored items of the CPRS

Item	Frequency of scores above zero in per cent (both samples combined, n = 106)	Interrater reliability between English and Swedish rater (n = 11)
Reported sadness	100	0.96
Apparent sadness	99	0.90
Lassitude	96	0.78
Inner tension	95	0.90
Pessimistic thoughts	94	0.53
Inability to feel	93	0.41
Worrying over trifles	91	0.78
Concentration difficulties	88	0.80
Observed muscular tension	88	0.74
Suicidal thoughts	87	0.71
Fatiguability	86	0.89
Reduced sleep	83	0.76
Indecision	81	0.27
Reported autonomic disturbances	79	0.72
Reported muscular tension	74	0.59
Reduced appetite	73	0.95
Agitation	73	0.62

dence were remarkably similar in the Swedish and the English samples (r = +.88, rho = +.88, P for both <0.001 (Montgomery et al, 1978a). Since the agreement between ratings of English patients by English and Swedish raters was generally high (Table I) it was decided to merge the two samples in the further calculations.

An arbitrary cut-off point of 70 per cent occurrence was used to identify the 17 most common items (Table I) in the total sample of patients. The sum of scores on these 17 items was used as a preliminary estimate of severity of the illness. This estimate was significantly correlated with the HRS scores (r = +.94, P < 0.001) and also with the global scores (r = +.89, P < 0.001) during the fourth treatment week. Before treatment these correlations were slightly lower as would be expected from the more restricted range of scores (HRS, r = +.73, Global, r = +.61, P for both < 0.001).

Table II

Mean change (regardless of direction), and correlation between change on each item and sum of change on the preliminary 17 item scale (n=64). The items are ranked in order of descending sensitivity on both estimates combined

Item	Mean change	Correlation to total change
Apparent sadness	2.12	0.84
Reported sadness	2.19	0.73
Inability to feel	1.98	0.75
Inner tension	1.79	0.73
Suicidal thoughts	2.03	0.64
Lassitude	1.61	0.79
Concentration difficulties	1.79	0.69
Reduced sleep	2.45	0.49
Reduced appetite	1.78	0.63
Pessimistic thoughts	1.69	0.64
Worrying over trifles	1.88	0.46
Fatiguability	1.73	0.58
Muscular tension (reported)	1.41	0.56
Muscular tension (observed)	1.22	0.49
Indecision	1.57	0.44
Autonomic disturbances	1.60	0.40
Agitation	0.99	0.34

Selecting items for sensitivity

Two different estimates of sensitivity were used (Table II). Firstly the mean changes of scores (absolute values) on each of the 17 items after four weeks of treatment were calculated and ranked. The second estimate was the correlation between change on each item and the overall change on the preliminary 17 item scale over the four weeks. These estimates reflect different aspects of sensitivity to change. Ideally, an item should both yield large changes (that can be reliably rated) and be strongly correlated to general amelioration of depression. This is not always the case.

For example, reduced sexual interest yielded large changes but was less well correlated to general outcome. Inclusion of an item like this in a scale might spuriously inflate the change scores.

The summed ranks on both estimates were used to select the 10 most sensitive items for the final depression rating scale, which is shown in the Appendix.

Reliability and Validity of the New Depression Scale

Interrater reliability

Data from the conjoint interviews were used to compute interrater correlations. Comparisons between two English raters, two Swedish raters and one English and one Swedish rater, rating English patients, are given in Table III. The interrater reliability in the simultaneously performed HRS ratings is shown for comparison. The interrater correlations are satisfactory for both scales, for raw scores as well as for difference scores.

To test the robustness of the instrument in the hands of untrained raters, ratings were also performed by a trained psychiatrist and a general practitioner or a psychiatric nurse (a detailed account of the procedure is given elsewhere; Montgomery et al, 1978). Also in this setting the interrater correlations were high (Table III).

Validity studies

To test the validity of a scale which estimates the severity of depression a comparison must

TABLE III

Interrater reliability of the new depression scale with different rater pairs. Interrater correlation between simultaneously performed Hamilton scale ratings are shown within brackets for comparison. All correlations are highly significant (P < 0.001)

Rater pair	r	n
Two English raters before treatment during treatment difference	0.89 (HRS: 0.89) 0.95 (HRS: 0.98) 0.90 (HRS: 0.91)	30 (HRS: 25) 20 (HRS: 14) 13 (HRS: 12)
Two Swedish raters before treatment	0.95	22
One English, one Swedish rater various stages of treatment	0.97	11
Psychiatrist and general practitioner various stages of treatment	0.97	17
Psychiatrist and nurse various stages of treatment	0.93	12

be made with an independent measure. An experienced clinician's global judgement as to whether the patient has responded or not is the criterion against which depression scales should be judged. As a preliminary validation of this scale's capacity to identify responders and non-responders to treatment we compared the scale scores with a clinician's global judgement in a sample where there was a clear cut differentiation between responders (18 patients) and non-responders (17 patients).

Point biserial correlations between response category and change scores were calculated for the preliminary 17-item scale, the final 10-item depression scale and the Hamilton Rating Scale. All correlations were highly significant. The correlations were used to determine which of the three scales differentiated better between responders and non-responders to treatment. The best correlation was achieved with the 10-item depression scale (r = +.70), the next best was with the 17-item preliminary scale (r = +.67). Of the three the HRS was the least able to discriminate (r = +.59). Converted into approximate patient numbers needed to achieve equivalent significance at different levels, for a point biserial correlation of this order 28 patients would have been needed to reach a significance at the 0.001 probability

level with the HRS, compared with 19 patients on the 10 item scale.

Discussion

The major requirements of a rating scale for antidepressant treatment effects is that it should be short and easy to apply in a clinical setting, relevant for depressive illness, and provide a sensitive and accurate estimate of change (Hamilton, 1976; Carroll et al, 1973; Asberg et al, 1973).

It would of course be possible to use a very extensive rating scale covering all aspects of depressive illness. A scale of this type might be more likely to pick up unexpected differences in the spectrum of action of drugs. However, the presence of a large number of items that were scored in only a few patients would tend to introduce and increase the random error. More important, the ratings would be cumbersome and time-consuming to undertake. Unskilled raters might have difficulties in covering a large number of items in a single interview. Repeated asking of questions which appear irrelevant to the patient might also be detrimental to clinical rapport and reduce the validity of the information provided.

When reducing the number of items, it is important that those included are relevant to

the illness and indeed occur in the majority of cases. An item may be both of diagnostic importance and likely to change with treatment but because it occurs so infrequently it might diminish the overall sensitivity of the scale. Examples of this type which failed to meet our frequency criterion for consideration are Ideas of Persecution and Compulsive Thoughts. Similar items were included in the initial version of the Hamilton Rating Scale (Hamilton, 1960) but excluded later (Hamilton, 1967).

Sensitivity of the scale in this context refers to its capacity to measure change. Change will be most sensitively recorded on items which are not restricted in range, but when the full width of the scales are used for the ratings. Restriction of the range might occur if a positive score on an item reflects a personality trait which is unlikely to change with short term treatment, rather than a symptom of an illness. It might also occur as a result of central tendency error, in which raters tend to avoid using the extreme ends of the scales (Guilford, 1954).

Change estimates should also be accurate and reflect a change in general severity of depressive illness. Some items may show a high degree of change without this being related to the illness as such. Hospitalization would be expected to have effects, for instance, on sexual interest and general levels of activity. Increased severity of some symptoms (for instance autonomic disturbances) may be side-effects of an antidepressant drug treatment.

The 10 items included in the new depression scale are all core symptoms of depressive illness. A few characteristic symptoms are, however, not included. Motor retardation (called Slowness of Movement in the full CPRS) is perhaps the most conspicuous omission. It was excluded from the primary selection, since it occurred in relatively few patients (63 per cent of the English patients and 69 per cent of the Swedish patients). Clinical experience shows that motor retardation only occurs in a proportion of patients and this is indeed the background of one of the classifications of depressive illness into retarded and non-retarded forms. In any case the mean change in Slowness of Movement was smaller than for any item included in the scale

and the correlation of change to overall amelioration was comparatively low (r = +.57).

The large number of rating scales available to clinical investigators is a problem in psychiatric research (Pichot, 1972) and the comparability between scales is rarely known. It is therefore important that a new rating scale should be shown to have clear advantages over existing instruments before it is accepted for research purposes. Our scale appears to have certain advantages, even when compared to the most widely used depression rating scale, the Hamilton Rating Scale.

Our depression scale has fewer items to score than the HRS (10 vs 17) and a reduction in reliability might be expected. However, we have demonstrated equally high reliabilities with the HRS and the 10 item scale. The scale can be used by trained nurses and psychologists as well as by psychiatrists (Montgomery et al, 1978b). It appears to be a more precise measure of change than the HRS. This means that significant differences between treatments may be revealed with smaller numbers of patients. In clinical trials this is important for ethical reasons, since fewer patients need to be exposed to possibly inferior treatments.

There is a definite need for simple, clinically oriented rating scales to measure treatment effects in other psychiatric syndromes such as mania, schizophrenia and anxiety states. To our knowledge, the strategy we employed for arriving at an empirically founded scale has not previously been used and should prove valuable in constructing further scales.

Acknowledgement

This study was supported by the Swedish Medical Research Council, 21P-4676; 27X-05015, the Bank of Sweden Tercentenary Fund, 74/110, the Gurli Wehtje Fund for Depression Research, S.E. Thames Regional Health Authority, Grant No. LOR/76/24, and Guy's Hospital Medical School.

References

Anderson, N. H. (1961) Scales and statistics: parametric and non-parametric. *Psychological Bulletin*, 58, 305-16. Angst, J. (1972) Concluding discussion. In *Depressive Illness* (ed. P. Kielholz). Bern: Huber.

Asberg, M., Kragh-Sorensen, P., Mindham, R. H. S. & Tuck, J. R. (1973) International reliability and communicability of a rating scale for depression. *Psychological Medicine*, 3, 458-65.

- RINGBERGER, V. A., SJÖQVIST, F., THORN, P., TRÄSKMAN, L. & TUCK, J. R. (1977) Monoamine metabolites in the cerebrospinal fluid and serotonin uptake inhibition during treatment with clomipramine. Clinical Pharmacology and Therapeutics, 21, 201-7.
- Montgomery, S., Perris, C., Schalling, D. & Sedvall, G. (1978) A comprehensive Psychopathological Rating Scale. Acta Psychiatrica Scandinavica, Supplement 272.
- BONEAU, C. A. (1961) A note on measurement scales and statistical tests. American Psychologist, 16, 160-261.
- CARROLL, B. J., FIELDING, J. M. & BLASHKI, T. G. (1973) Depression rating scales. A critical review. Archives of General Psychiatry, 28, 361-6.
- COPPEN, A., GUPTA, R., MONTGOMERY, S., GHOSE, K., BAILEY, J., BURNS, B. & DE RIDDER, J. J. (1976) Mianserin hydrochloride: a novel antidepressant. British Journal of Psychiatry, 129, 342-5.
- Feighner, J. P., Robins, E., Guze, S. B., Woodruff, R. A., Winokur, G. & Munoz, R. (1972) Diagnostic criteria for use in psychiatric research. *Archives of General Psychiatry*, 26, 57-63.
- FERL, P. J., QUANTOCK, D. C. & VAN DER BURG, W. J. (1973) The human pharmacology of GB94—a new psychotropic agent. European Journal of Clinical Pharmacology, 5, 166-73.
- Guilford, J. P. (1954) Psychometric methods. 2nd Ed., p 278. New York: McGraw-Hill.
- Gurney, C., Roth, M., Garside, R. F., Kerr, T. A. & Schapira, K. (1972) Studies in the classification of affective disorders. The relationship between anxiety states and depressive illnesses—II. British Journal of Psychiatry, 121, 162-6.
- Hamberger, B. & Tuck, J. R. (1973) Effect of tricyclic antidepressants on the uptake of noradrenaline and 5-hydroxytryptamine by rat brain slices incubated in buffer or human plasma. European Journal of Clinical Pharmacology, 5, 229-35.
- Hamilton, M. (1960) A rating scale for depression. Journal of Neurology, Neurosurgery and Psychiatry, 23, 56-62.
- —— (1967) Development of a rating scale for primary depressive illness. British Journal of Social and Clinical Psychology, 6, 278-96.
- —— (1976) Comparative value of rating scales. British Journal of Clinical Pharmacology, 3, 58-60.
- Mass, J. W. (1975) Biogenic amines and depression. Biochemical and pharmacological separation of two types of depression. Archives of General Psychiatry, 32, 1357-61.
- MAITRE, L., STAEHELIN, M. & BEIN, H. J. (1971) Blockade of noradrenaline uptake by 34276 B a, A new anti-depressant drug. *Biochemical Pharmacology*, 20, 2169-86.
- Montgomery, S., Asberg, M., Traskman, L. & Montgomery, D. (1978a) Cross-cultural studies on the use of the CPRS in English and Swedish depressed patients. *Acta Psychiatrica Scandinavica*, Supplement 272, 33-9.

- JÖRNESTEDT, L., THOREN, P., TRÄSKMAN, L., MCAULEY, R., MONTGOMERY, D. & SHAW, P. (1978b) Reliability of the CPRS between the disciplines of psychiatry, general practice, nursing and psychology. *Acta Psychiatrica Scandinavica*. Supplement 272, 29-32.
- MORRIS, J. B. & BECK, A. T. (1974) The efficacy of antidepressant drugs—a review of research (1958 to 1972) Archives of General Psychiatry, 30, 667-74.
- Pichot, P. (1972) The problem of quantifying the symptomatology of depression. In *Depressive Illness* (ed. P. Kielholz). Bern: Huber.
- SILVERSTONE, T. & TURNER, P. (1974) Drug Treatment in Psychiatry, p 117. London: Routledge & Kegan Paul.
- Tuck, J. R. & Punell, G. (1973) Uptake of ³H-5-hydroxytryptamine and ³H-noradrenaline by slices of rat brain incubated in plasma from patients treated with clomipramine, imipramine or amitriptyline. *Journal of Pharmacy and Pharmacology*, 25, 573-4.

APPENDIX

Montgomery and Åsberg (MADRS) Depression Rating Scale.

The rating should be based on a clinical interview moving from broadly phrased questions about symptoms to more detailed ones which allow a precise rating of severity. The rater must decide whether the rating lies on the defined scale steps (0, 2, 4, 6) or between them (1, 3, 5).

It is important to remember that it is only on rare occasions that a depressed patient is encountered who cannot be rated on the items in the scale. If definite answers cannot be elicited from the patient all relevant clues as well as information from other sources should be used as a basis for the rating in line with customary clinical practice.

The scale may be used for any time interval between ratings, be it weekly or otherwise but this must be recorded.

Item List

- 1. Apparent sadness
- 2. Reported sadness
- 3. Inner tension
- 4. Reduced sleep
- 5. Reduced appetite
- 6. Concentration difficulties
- 7. Lassitude
- 8. Inability to feel
- 9. Pessimistic thoughts
- 0. Suicidal thoughts

1. Apparent Sadness

Representing despondency, gloom and despair, (more than just ordinary transient low spirits) reflected in speech, facial expression, and posture. Rate by depth and inability to brighten up.

0 No sadness.

1

 Looks dispirited but does brighten up without difficulty.

3

4 Appears sad and unhappy most of the time.

5

6 Looks miserable all the time. Extremely despondent.

2. Reported sadness

Representing reports of depressed mood, regardless of whether it is reflected in appearance or not. Includes low spirits, despondency or the feeling of being beyond help and without hope.

Rate according to intensity, duration and the extent to which the mood is reported to be influenced by events.

Occasional sadness in keeping with the circumstances.

1

2 Sad or low but brightens up without difficulty.

3

4 Pervasive feelings of sadness or gloominess. The mood is still influenced by external circumstances.

5

6 Continuous or unvarying sadness, misery or despondency.

3. Inner tension

Representing feelings of ill-defined discomfort, edginess, inner turmoil, mental tension mounting to either panic, dread or anguish.

Rate according to intensity, frequency, duration and the extent of reassurance called for.

0 Placid. Only fleeting inner tension.

1

 Occasional feelings of edginess and illdefined discomfort.

3

4 Continuous feelings of inner tension or intermittent panic which the patient can only master with some difficulty.

5

6 Unrelenting dread or anguish. Overwhelming panic.

4. Reduced sleep

Representing the experience of reduced duration or depth of sleep compared to the subject's own normal pattern when well.

0 Sleeps as usual.

1

2 Slight difficulty dropping off to sleep or slightly reduced, light or fitful sleep.

3

4 Sleep reduced or broken by at least two hours.

5

6 Less than two or three hours sleep.

5. Reduced appetite

Representing the feeling of a loss of appetite compared with when well. Rate by loss of desire for food or the need to force oneself to eat.

0 Normal or increased appetite.

1

2 Slightly reduced appetite.

3

4 No appetite. Food is tasteless.

5

6 Needs persuasion to eat at all.

6. Concentration difficulties

Representing difficulties in collecting one's thoughts mounting to incapacitating lack of concentration.

Rate according to intensity, frequency, and degree of incapacity produced.

0 No difficulties in concentrating.

1

2 Occasional difficulties in collecting one's thoughts.

3

4 Difficulties in concentrating and sustaining thought which reduces ability to read or hold a conversation.

5

6 Unable to read or converse without great difficulty.

7. Lassitude

Representing a difficulty getting started or slowness initiating and performing everyday activities.

Hardly any difficulty in getting started. No sluggishness.

1

2 Difficulties in starting activities.

3

4 Difficulties in starting simple routine activities which are carried out with effort.

5

6 Complete lassitude. Unable to do anything without help.

8. Inability to feel

Representing the subjective experience of reduced interest in the surroundings, or activities that normally give pleasure. The ability to react with adequate emotion to circumstances or people is reduced.

0 Normal interest in the surroundings and in other people.

2

2 Reduced ability to enjoy usual interests.

3

Loss of interest in the surroundings. Loss of feelings for friends and acquaintances.

5

6 The experience of being emotionally paralysed, inability to feel anger, grief or pleasure and a complete or even painful failure to feel for close relatives and friends.

9. Pessimistic thoughts

Representing thoughts of guilt, inferiority, self-reproach, sinfulness, remorse and ruin.

0 No pessimistic thoughts.

2 Fluctuating ideas of failure, self-reproach or self depreciation.

3

4 Persistent self-accusations, or definite but still rational ideas of guilt or sin. Increasingly pessimistic about the future.

5

6 Delusions of ruin, remorse or unredeemable sin. Self-accusations which are absurd and unshakable.

10. Suicidal thoughts

Representing the feeling that life is not worth living, that a natural death would be welcome, suicidal thoughts, and preparations for suicide.

Suicidal attempts should not in themselves influence the rating.

- 0 Enjoys life or takes it as it comes.
- Weary of life. Only fleeting suicidal thoughts.
- 4 Probably better off dead. Suicidal thoughts are common, and suicide is considered as a possible solution, but without specific plans or intention.
- Explicit plans for suicide when there is an opportunity. Active preparations for suicide.

* Stuart A. Montgomery, B.Sc., M.D., M.R.C.Psych., Senior Lecturer, Academic Department of Psychiatry, Guy's Hospital Medical School, London, S.E. 1,

Marie Asberg, M.D., Karolinska Institute, Stockholm, Sweden

* Correspondence.

(Received 24 April; revised 30 August 1978)