

## The Use of Rating Scales to Identify Post-natal Depression

BRIAN HARRIS, PHILIP HUCKLE, ROGER THOMAS, SANDRA JOHNS and HEDI FUNG

One hundred and forty-seven mothers were screened for major depression at six to eight weeks post-partum. Using predetermined cut-off points, the Edinburgh Postnatal Depression Scale and the Beck scale were compared in their abilities to identify the 15% of subjects who had major depression according to DSM-III criteria. The sensitivity of Edinburgh scale was 95% and its specificity 93%. The performance of the Beck scale was markedly inferior, with a sensitivity of 68% and specificity of 88%.

In the year following delivery, a high incidence of maternal depression has been reported in many studies, the exact incidence measured being partly dependent on the methods used. For example, in a *retrospective* general-practice study, Ryle (1961) found depression in 3% of women post-partum, whereas Dalton (1971) in a *prospective* study carried out in general practice found an incidence of 7%. Pitt (1968), in a prospective study of 305 mothers admitted to the 'lying-in' wards of the London Hospital, found an incidence of depressive illness of approximately 11% in the year following delivery.

Paykel *et al* (1980), conducting a study in a post-natal check-up clinic, found an incidence of 20% mild depression in a population of 120 mothers at six weeks post-partum.

It has been shown that factors associated with this common condition include an increased number of life events in the previous 12 months, previous psychiatric illness (Paykel *et al*, 1980), marital disharmony (Watson *et al*, 1984) and higher levels of neurotic symptoms when well (Cox *et al*, 1982). Other factors include friction with relatives, in-laws, older children or neighbours; physical illness or handicap in family members; financial worries; and poor housing (Kumar & Robson, 1984).

The condition is important in that it may have long-term effects on the children of such mothers, producing, for example, later behavioural disturbances (Wrate *et al*, 1985). Non-accidental injury to the child is also sometimes associated with this problem (Margison & Brockington, 1982).

In view of the high incidence and the nature of the illness, it would be helpful to screen mothers in the community to identify those suffering from the condition, and hence refer them for appropriate help and treatment.

As to methods of screening patients for depression (Williams *et al*, 1980), besides self-report questionnaires there are three types of assessment: firstly, a full clinical interview with a psychiatrist;

secondly, a structured clinical interview such as the Present State Examination (Wing *et al*, 1973); and, thirdly, observer-rated scales. The first two methods are both time-consuming and impractical in screening large numbers of women, the majority of whom are well. Observer-rated scales are more practicable, but they too require some form of training in their application.

Self- and observer-rated scales correlate but do so incompletely as they often measure different variables. However, in the context of puerperal depression, Cox *et al* (1987) appear to have produced a useful self-rated questionnaire. They have devised the ten-item Edinburgh Postnatal Depression Scale (EPDS) as a screening tool. The advantages of such a tool are obvious in that it is a self-report scale and by design can be used by a variety of community health workers. The scale does not replace full psychiatric assessment, but defines a population which needs further evaluation.

In developing their questionnaire, Cox *et al* used two forms of self-report questionnaire on several occasions during pregnancy and the puerperium in a sample of 425 child-bearing women. They compared the Scale for Anxiety and Depression (SAD) of Bedford & Foulds (1978) and their own visual analogue scales. Depression, anxiety and tearfulness on the visual analogue scales were significantly greater one week after delivery than in assessments at other times (during pregnancy, and at 12 weeks and 23 weeks after childbirth). The SAD failed to detect any such change in depression or anxiety, and its validity during pregnancy was also not satisfactory. It was selected by Cox *et al* for their study in preference to other scales because of its separate subscale for anxiety and depression, few somatic features and its relative brevity.

In the study which provided the data reported here, the main objective was to identify major depressive disorder in order to examine its possible association with post-partum autoimmune thyroid

dysfunction and steroid hormone levels (Fung *et al*, 1988). Although the primary aim of the study was not to examine the ability of various questionnaires to identify post-natal depression, we are able to report a comparison of the EPDs with another self-report questionnaire, the Beck Depression Inventory (BDI), in their abilities to identify major depressive disorder. The BDI (Beck *et al*, 1961) consists of 21 questions concerned with depressive symptoms and has been found to correlate significantly with the Hamilton observer-rated scale for depression (Hamilton, 1960). The BDI is subject-rated and has been widely used as a screening instrument for depression. Recent studies have examined its use in medical out-patients (Nielsen & Williams, 1980), in psychiatric out-patients presenting with depression (Harris *et al*, 1984), and in a population of medical students (Zoccolillo *et al*, 1986). It has also been used to identify depression in the puerperium by O'Hara *et al* (1983), but it was found to be unsatisfactory because of its low specificity.

The Montgomery-Asberg Depression Rating Scale (MADRS; Montgomery & Asberg, 1979) was also used in the study reported here. It is composed of ten items concerned mainly with psychic symptoms of depressive illness. Its relative lack of assessment of biological features should make it useful in the post-partum period. Each item is graded 0-6, and this scale is very sensitive to treatment response. Such sensitivity to change should, again on theoretical grounds, make it useful in late pregnancy and the first post-natal year, when changes in mood may be marked.

The Raskin 3 Area Scale for Depression (Raskin *et al*, 1970) was also used. Paykel *et al* (1980) report on its use in a study of 120 mothers at six weeks after delivery. The scale has three components: a verbal report of depressed feelings, worthlessness etc.; observed depression at interview; and secondary symptoms of depression, including insomnia, anorexia and poor concentration.

### Method

Over the course of one year, 147 mothers were assessed at the Caerphilly Miners' Hospital in South Wales. They consisted of 65 antibody-positive women (microsomal and thyroglobulin) and 82 antibody-negative women. These were part of a larger population being studied for post-partum thyroid dysfunction (Fung *et al*, 1988). They had been seen by the physicians at monthly intervals during pregnancy, and also in the post-delivery ward.

The women had originally presented as routine bookings for delivery at the hospital, and had a mean gestation of 16 weeks at presentation. They were unselected in terms of marital, socio-economic and medical problems, apart

from the fact that women with thyroid disorder other than positive antibody status were excluded from the study.

The psychiatric assessment was at a six weeks routine post-natal follow-up clinic, and the psychiatrist was unaware of the antibody status of the women. The mental state of each mother was assessed according to DSM-III criteria for major depression (American Psychiatric Association, 1980) by an experienced psychiatrist between 13.30 h and 15.00 h, and then the Raskin 3 Area Scale for Depression and the Montgomery-Asberg Depression Rating Scale were completed. The whole interview took approximately 40 minutes. Subjects were asked to complete the Edinburgh Postnatal Depression Scale in the clinic, followed by the Beck Depression Inventory. These completed scales were taken home and returned by post. Subjects also completed the Middlesex Hospital questionnaire and Eysenck Personality Inventory (to be reported in a separate paper).

The majority of women were assessed in the clinic, but 49 had afternoon visits at home because of non-attendance. After assessment, each participant was informed that a sum of £2 would be sent by post on receipt of the questionnaires (and saliva samples for a separate study).

Various cut-off points have been suggested for the BDI, but there is no clear agreement from the literature in this regard. We have taken the cut-off point of 11 and over suggested by Nielsen & Williams (1980) as being appropriate when the inventory is used as a screening instrument. This cut-off point defines the lower limits of mild depression, but in an attempt to improve the BDI's performance we have also examined the use of 13 and over (Cavanaugh, 1983) and 21 and over (Cavanaugh *et al*, 1983).

The cut-off point used for the EPDs was 13 and over (Cox *et al*, 1987), 8 and over for the Raskin Scale (Paykel *et al*, 1980), and 12 and over for the MADRS - Snaith *et al* (1986) have shown that mild depression scores between 7 and 18, and a cut-off point of 12 represents the mid-point of the range.

### Results

The age range of the population was 17-40 years ( $n = 147$ , mean 24.6 years, s.d. 5.5), with a parity of one to five (49 primigravida, 98 multigravida). At six to eight weeks, 29 mothers were breastfeeding their babies, and 32 were taking the contraceptive pill.

The deliveries of 138 mothers were vaginal deliveries of normal infants; nine mothers had Caesarean sections. No mothers had stillborn or severely handicapped babies.

All the 147 mothers were assessed clinically and scored on the observer-rated scales. These were 21 women who did not return one or both of the self-rated questionnaires, resulting in differing compliance rates.

DSM-III criteria for major depressive disorder were met by 22 women (15%). The majority of these had mild depression, but six were referred to their GPs for further treatment. In addition, there were other depressive states such as 'dysthymic disorder' and 'adjustment disorder with depressed mood', as well as situations where depressive symptoms were present but not of sufficient severity to meet criteria for major depression. No attempt was made to place

TABLE I  
Sensitivity and specificity of four questionnaires in identifying subjects with major depressive disorder, according to DSM-III criteria

Questionnaire	Number completing test	Major depression	Sensitivity <sup>1</sup>	Specificity <sup>2</sup>	'Other depression'	Identification of EPDS 'cases'
EPDS	126	21/22	95%	97/104: 93%	6/13	—
BDI	129	13/19	68%	97/110: 88%	5/13	17/25
Raskin	146	22/22	100%	109/125: 87%	9/15	25/28
MADRS	147	20/22	91%	118/124: 96%	2/15	20/28

1. Percentage of genuine cases identified.

2. Percentage of genuine non-cases identified.

these in diagnostic categories, since the main objective of the study was to examine hormonal parameters in the context of major depression. They are referred to as 'other depression' (Table I).

The individual performance of each scale in identifying depression, in terms of percentage of the population above the predetermined cut-off point, was as follows: EPDS 22%, BDI 19%, MADRS 17%, and Raskin 25%. Misclassification rates for the questionnaires were 6.3%, 14.7%, 5.5%, and 10.8% respectively, i.e. in terms of individuals who did and did not meet criteria for major depression.

Table I illustrates the ability of the four rating scales to identify subjects with major depression.

Out of a total of 22 subjects with major depression, the EPDS identified 21, giving it a sensitivity of 95% (i.e. the percentage of genuine cases identified). There were seven false positives, giving it a specificity of 93% (i.e. the percentage of genuine non-cases identified). Four of the false positives were 'other depression'. Lowering the cut-off point of the EPDS to 10 and over (a possibility suggested by Cox *et al.*, 1987) resulted in a sensitivity of 100% but a lowered specificity of 82% (19 false positives, 11 of which fell into the category of 'other depression').

The BDI identified 13 of 19 subjects with major depression (sensitivity 68%) and had 13 false positives (specificity 88%), indicating a poorer performance than the EPDS. Raising the cut-off point lowered the sensitivity further. At 13 and over (Cavanaugh, 1983) the sensitivity was 63% (12/19) and the specificity 92%, while at 21 and over the sensitivity was 32% (6/19) and the specificity 99%. All of these results with the BDI are clearly unsatisfactory because of the high percentage of false negatives - i.e. missed cases of post-natal depression.

TABLE II  
Major intercorrelations between the depression rating scales (Pearson correlation matrix)

	EPDS	n <sup>1</sup>	Raskin	n <sup>1</sup>	MADRS	n <sup>1</sup>
BDI	0.68*	110	0.62*	129	0.61*	128
EPDS	—	—	0.80*	126	0.79*	125
Raskin	—	—	—	—	0.92*	146

1. Number of mothers assessed.

\* $P < 0.001$ .

The results were also analysed using the Pearson correlation matrix. Results are summarised in Table II. The intercorrelations between the Edinburgh Raskin and Montgomery-Asberg scales were very highly significant, as expected. The association of the Beck scale with the other three questionnaires was also significant, but less so.

### Discussion

Many studies have shown that depressive illness occurs at a high rate in the year following delivery, and the figure of 15% for major depression reported here is similar to that reported in other studies (Paykel *et al.*, 1980).

Clinical diagnoses were based on DSM-III criteria for major depression, which requires dysphoric mood or loss of interest or loss of pleasure, together with four other required symptoms. The diagnosis of major depression thus obtained is similar to but less strict than that obtained by the Research Diagnostic Criteria (RDC) for major depressive disorder (Spitzer *et al.*, 1978), which were used by Cox *et al.* (1987) in devising the EPDS. The RDC also allow for 'definite', 'probable' and 'minor' depressive disorder according to severity. This is not the case with DSM-III, and in those instances where insufficient depressive symptoms were present or where other depressive diagnoses such as dysthymic disorder were made, we categorised the states as 'other depression'.

The features of neurotic depression in the puerperium have been described by Pitt (1968) as 'atypical', and include depression varying from day to day, tending to worsen in the evenings, and associated with fatigue, irritability (especially towards the spouse and any other children), disturbance of appetite (usually anorexia), early insomnia and loss of libido.

Some of the symptoms of post-natal depression are probably shared by many 'normal' puerperal women caring for a newborn infant. For example, the demands of regular feeding, 'baby colic', nappy changes, etc. may lead to sleepless nights and irregular

food intake. A problem therefore arises when screening and attempting to identify 'cases' since most depression rating scales give importance to such somatic symptoms. The BDI, for example, gives prominence to lassitude, insomnia and loss of appetite, although these features occur commonly in all illnesses, both physical and mental. Also, as previously stated, these features occur in a large number of normal mothers in the post-natal period, rendering such scales less useful in this context.

Furthermore, as far as depression in general is concerned, Snaith (1987), in discussing the use of rating scales and their essential elements, suggests that the symptom of anhedonia (the loss of the ability to feel or experience pleasure) is the central and most reliable feature of mild depression.

The EPDS was devised from the earlier work of Snaith *et al* (1978) and concentrates on the 'psychic' aspect of depression in the puerperium, containing within it a rating for anhedonia. Its only rating of a biological nature is for 'difficulty sleeping', and clinical experience supports this, since non-depressed mothers are normally able to get to sleep easily even though nights may be disrupted by feeding. The reported sensitivity here of 91% confirms that it is a reliable instrument for screening for post-natal depression. Lowering the cut-off point to 10 and over (a possibility suggested by Cox *et al* (1987)) results in identification of all of our major depression, but lowers the specificity of the EPDS to 82%. However, of the 19 false positives, 11 were 'borderline depressives' and would obviously need further assessment.

Using a cut-off point of 11 and over as suggested by Nielsen & Williams (1980), the BDI showed a sensitivity of 68%, i.e. missing 6 of 19 cases of major depression – a serious deficiency. One-third of its questions concern somatic aspects of depression such as weight loss, features which are obviously less relevant in the context of puerperal depression. Kearns *et al* (1982) have compared the MADRS, the Hamilton, the BDI and the depression subscale of the Wakefield Inventory in patients suffering from depression of sufficient severity to lead to hospital admission. The last two scales had an overall poor performance and the authors recommended discontinuing the use of these particular scales in research. The work reported here supports this finding and the low sensitivity of the BDI makes it unacceptable as a screening instrument for identifying major depression in the puerperium.

In a study seeking primarily to examine the ability of observer-rated scales to identify depressive disorder, as an independent criterion, diagnosis of the disorder would ideally be carried out by an observer

independent of the one rating the scales. Bearing in mind that this was not the case in the study reported here, the results obtained for the two observer-rated scales should be taken with caution.

The Raskin scale was chosen for use in this study because it had been used in a previous major work on post-natal depression (Paykel *et al*, 1980). It was originally designed as an observer-rated scale for use in monitoring changes in depression occurring with treatment (e.g. response to antidepressant medication). It does not contain an assessment of the subject's "lack of ability to enjoy" (anhedonia) and in this respect differs from the EPDS and MADRS. As far as Snaith is concerned it therefore omits the essential feature of depression (Snaith, 1987). However, it identified all of the subjects with major depression (sensitivity 100%), but its specificity was lower (89%) in that it identified 16 false positives.

The Montgomery-Asberg scale gives some importance to physical symptoms, but is mainly orientated towards psychic symptoms of depression, including apparent and reported sadness, inner tension, impairment of concentration, inability to feel, anhedonia, pessimism and suicidal ideation. Its overall performance is good, possibly because of the relative lack of emphasis on somatic symptoms, since of its ten questions, one deals with appetite and one deals with sleep.

The Edinburgh Postnatal Depression Scale was specifically designed to screen for depression in a community sample, and we have shown that there is a highly significant correlation between the EPDS and two observer-rated scales for depression, both of which have been used extensively in screening for depression in community samples. The value of the EPDS can be increased still further by lowering the cut-off point to 10 and over. In conclusion it can be stated that the EPDS is a valuable screening test which performs as well as two observer-rated scales, and better than the patient-rated Beck Depression Inventory. It is confirmed that the latter is unsuitable as a screening instrument for post-natal depression.

## References

- AMERICAN PSYCHIATRIC ASSOCIATION (1980) *Diagnostic and Statistical Manual of Mental Disorders* (3rd edn). Washington, DC: American Psychiatric Association.
- BECK, A. T., WARD, C. H., MENDELSON, M., MOCK, J. E. & ERBAUGH, J. (1961) An inventory for measuring depression. *Archives of General Psychiatry*, 4, 561–571.
- BEDFORD, A. & FOULDS, G. (1978) *Delusions – Symptoms – States, Inventory, State of Anxiety and Depression (Manual)*. Windsor: National Foundation for Education Research.
- CAVANAUGH, S. VON A. (1983) The prevalence of emotional and cognitive dysfunction in a general medical population, using the MMSE, GHQ and BDI. *General Hospital Psychiatry*, 5, 15–24.



- , CLARK, D. C. & GIBBONS, R. D. (1983) Diagnosing depression in the hospitalised medically ill. *Psychosomatics*, **24**, 809–815.
- COX, J. L., CONNOR, Y. & KENDELL, R. E. (1982) Prospective study of the psychiatric disorders of childbirth. *British Journal of Psychiatry*, **140**, 111–117.
- , CONNOR, Y. M., HENDERSON, I. P., *et al* (1983) Prospective study of the psychiatric disorders by self-report questionnaire. *Journal of Affective Disorders*, **5**, 1–7.
- , HOLDEN, J. M. & SAGOVSKY, R. (1987) Detection of postnatal depression – development of the 10-item Edinburgh Postnatal Depression Scale. *British Journal of Psychiatry*, **150**, 782–786.
- DALTON, K. (1971) Prospective study into puerperal depression. *British Journal of Psychiatry*, **118**, 689–692.
- FUNG, H. Y. M., KOLOGLU, M., COLLISON, K., *et al* (1988) Postpartum mood: thyroid dysfunction in Mid Glamorgan. *British Medical Journal*, **296**, 241–244.
- HAMILTON, M. (1960) A rating scale for depression. *Journal of Neurology, Neurosurgery and Psychiatry*, **23**, 56–62.
- HARRIS, B., YOUNG, J. P. R. & HUGHES, W. (1984) Appetite and weight change in patients presenting with depressive illness. *Journal of Affective Disorders*, **6**, 331–339.
- KEARNS, N. P., CRUICKSHANK, K. J., MCGUIGAN, K. J., *et al* (1982) A comparison of depression rating scales. *British Journal of Psychiatry*, **141**, 45–49.
- KUMAR, R. & ROBSON, K. M. (1984) A prospective study of emotional disorders in childbearing women. *British Journal of Psychiatry*, **144**, 35–37.
- MARGISON, F. & BROCKINGTON, I. F. (1982) In *Motherhood and Mental Illness* (eds I. F. Brockington & R. Kumar). London: Academic Press/Grune and Stratton.
- MONTGOMERY, S. A. & ASBERG, M. (1979) A new depression scale designed to be sensitive to change. *British Journal of Psychiatry*, **134**, 382–389.
- NIELSEN, A. C. & WILLIAMS, T. A. (1980) Depression in ambulatory medical patients. *Archives of General Psychiatry*, **37**, 999–1004.
- O'HARA, M. W., REHM, L. P. & CAMPBELL, S. B. (1983) Postpartum depression: a role for social network and life stress variables. *Journal of Nervous and Mental Disease*, **171**, 336–341.
- PAYKEL, E. S., EMMS, E. M., FLETCHER, J. & RASSABY, E. S. (1980) Life events and social support in puerperal depression. *British Journal of Psychiatry*, **136**, 339–346.
- PITT, B. (1968) Atypical depression following childbirth. *British Journal of Psychiatry*, **114**, 1325–1335.
- RASKIN, A., SCHULTERBRANDT, J., REATIG, N. & MCKISON, J. (1970) Differential response to chlorpromazine, imipramine and placebo – a study of sub-groups of hospitalised depressed patients. *Archives of General Psychiatry*, **23**, 164–173.
- RYLE, A. (1961) The psychological disturbances associated with 345 pregnancies in 137 women. *Journal of Mental Science*, **107**, 279–282.
- SNAITH, R. P. (1987) Concepts of mild depression. *British Journal of Psychiatry*, **150**, 387–393.
- , CONSTANTOPOULOS, A. A., JARDINE, M. Y. & MCGUFFIN, P. (1978) A clinical scale for the self assessment of irritability. *British Journal of Psychiatry*, **132**, 164–171.
- , HARROP, F. M., NEWBY, D. A. & TEALE, C. (1986) Grade scores of the Montgomery-Asberg Depression and the Clinical Anxiety scales. *British Journal of Psychiatry*, **148**, 599–601.
- SPITZER, R. L., ENDICOTT, J. & ROBINS, S. (1978) Research Diagnostic Criteria. *Archives of General Psychiatry*, **35**, 773–782.
- WATSON, J. P., ELLIOT, S. A., RUGG, A. J. & BROUGH, D. I. (1984) Psychiatric disorder in pregnancy and the first postnatal year. *British Journal of Psychiatry*, **144**, 453–462.
- WILLIAMS, P., TARNPOLSKY, A. & HAND, D. (1980) Case definition and case identification in psychiatry epidemiology – review and assessment. *Psychological Medicine*, **10**, 101–114.
- WING, T. K., COOPER, J. E. & SARTORIUS, N. (1973) *Present State Examination*. Cambridge: Cambridge University Press.
- WRATE, R. M., ROONEY, A. C., THOMAS, P. F. & COX, J. L. (1985) Postnatal depression and child development – a three-year follow up study. *British Journal of Psychiatry*, **146**, 622–627.
- ZOCCOLILLO, M., MURPHY, G. E. & WETZEL, R. D. (1986) Depression among medical students. *Journal of Affective Disorders*, **11**, 91–96.

\*Brian Harris, BSc, MBBS, MRCPsych, DPM, *Consultant Psychiatrist, Department of Psychological Medicine, University Hospital of Wales, Cardiff*; Philip Huckle, MB, ChB, *Registrar in Psychiatry, Roger Thomas, MB, ChB, MRCPsych, Consultant Psychiatrist, Whitchurch Hospital, Cardiff*; Sandra Johns, MB, ChB, MRCPsych, *Consultant Psychiatrist, Cefn Coed Hospital, Swansea*; Hedi Fung, MB, ChB, MD, *Registrar, Obstetrics, King's College Hospital Group, London*

\*Correspondence