**Review of validation studies of the Edinburgh postnatal depression scale**

*Eberhard-Gran M, Eskild A, Tambs K, Opjordsmoen S, Samuelsen S O*

**Authors' objectives**
To review validation studies of the Edinburgh postnatal depression scale (EPDS).

**Searching**
MEDLINE and Science Citation Index Expanded were searched from 1987 (when the EPDS was launched) to October 2000. The search terms were 'Edinburgh postnatal depression scale', 'EPDS' and 'validation'. Reference lists in identified publications were screened for additional studies. No language restrictions were reported.

**Study selection**
Study designs of evaluations included in the review
Validation studies were eligible for inclusion.

Specific interventions included in the review
Studies of the EPDS were eligible for inclusion. In the included studies, the cut-off value of the EPDS ranged from 8.5 to 12.5. Most of the studies applied the EPDS as part of a clinical interview; it was unclear how it was applied in those studies in which it was not part of a clinical interview.

Reference standard test against which the new test was compared
No inclusion criteria relating to the reference standard were specified. The reference standards used in the included studies were the American Psychiatric Association's DSM-III, DSM-III-R and DSM-IV criteria for depression, the Research Diagnostic Criteria, the Comprehensive Psychopathological Rating Scale (CPRS-Depression), the Schedule of Affective Disorders and Schizophrenia, the Present State Examination (PSE-ID-Catego) and the OCD-10. Some studies used both major and minor depressive disorder in their definition of postpartum depression; others used a more strict definition that included only major depression, or major and moderate depression.

Participants included in the review
No inclusion criteria relating to the participants were specified. The time since delivery ranged from 4 days to 6 months.

Outcomes assessed in the review
No inclusion criteria relating to the outcomes were specified. The outcomes reported in the review were the sensitivity, specificity and predictive values.

How were decisions on the relevance of primary studies made?
The authors did not state how the papers were selected for the review, or how many reviewers performed the selection.

**Assessment of study quality**
No formal assessment of quality was undertaken.

**Data extraction**
The authors did not state how the data were extracted from the primary studies. The 95% confidence interval around the sensitivity and specificity were calculated according to the Clopper-Pearson exact method (see Other Publications of Related Interest).

**Methods of synthesis**
How were the studies combined?
A narrative synthesis of the study results was presented. The results of the individual studies were presented as ranges in sensitivity and specificity, and positive predictive values were calculated based on a disease prevalence of 13%. The authors stated that positive and negative predictive values of the EPDS were calculated assuming a prevalence of postpartum depression of 5, 10, 15, 20 and 25%.

How were differences between studies investigated?
The authors did not report a method for assessing any differences between the studies.

Results of the review
Eighteen studies were included. The total sample size was unclear (the individual study size ranged from 53 to 147).

Estimates of sensitivity ranged from 65% (specificity 96%) to 100% (specificity 80 to 96%). The specificity ranged from 49% (sensitivity 96%) to 100% (sensitivity 96%). Studies that used a high EPDS cut-off tended to show lower sensitivity and higher specificity. The inclusion of minor depressive disorders in the definition of postpartum depression tended to lower estimates of sensitivity.

Authors' conclusions
Most studies showed a high sensitivity of the EPDS. Uncertainty about the comparability between sensitivity and specificity estimates of different EPDS versions remains, because of the differences in study design and large confidence intervals.

CRD commentary
This review presented a clearly stated objective, although very few details of the inclusion criteria were reported. A limited literature search, which was restricted to two databases and did not include attempts to locate unpublished studies, was undertaken. The review may, therefore, be subject to publication bias and may not have identified all available published literature. In addition, using the term 'validation' as a search term is likely to have restricted the search and might have led to important studies being missed: not all studies evaluating the accuracy of the EPDS would be likely to include the term validation in a searchable field. No details of the review process were reported, so it is not possible to determine whether appropriate steps were taken to minimise bias. In addition, since no formal quality assessment was undertaken, it is not possible for the reader to assess the validity of the primary studies.

No attempts were made to formally synthesise the results of the primary studies, or to investigate heterogeneity between the studies. Although it may not have been appropriate to pool the study results given the heterogeneity between studies, the presentation of individual study results on a receiver operating characteristic curve or Forest plot, or further attempts to synthesise the results in a narrative, would have helped the interpretation of the results of this review. Some study details were tabulated; however, there were insufficient details to get a clear picture of the study design of the individual studies or the women included in the studies. The authors' conclusions appear to follow from the results presented, but should be interpreted with extreme caution due to the limitations of this review, as discussed.

Implications of the review for practice and research
Practice: The authors stated that when the EPDS is applied to the general population, a substantial proportion of the cases identified as depressed with the EPDS are false positive cases; this might also result in a falsely high prevalence estimate.

Research: The authors did not report any implications for further research.

Funding
Norwegian Research Council.
Bibliographic details

PubMedID
11722298

Other publications of related interest

Indexing Status
Subject indexing assigned by NLM

MeSH
Confidence Intervals; Depression, Postpartum /diagnosis /psychology; Female; Humans; Personality Inventory /statistics & numerical data; Psychometrics; Reproducibility of Results; Sweden

AccessionNumber
12001002416

Date bibliographic record published
31/01/2005

Date abstract record published
31/01/2005

Record Status
This is a critical abstract of a systematic review that meets the criteria for inclusion on DARE. Each critical abstract contains a brief summary of the review methods, results and conclusions followed by a detailed critical assessment on the reliability of the review and the conclusions drawn.