Arch Womens Ment Health (2007) 10: 293–297
DOI 10.1007/s00737-007-0204-x
Printed in The Netherlands

Archives of Women's Mental Health

Original contribution

Validation of the Edinburgh Postnatal Depression Scale in an Iranian sample

S. Mazhari, N. Nakhaee

Neuroscience Department, Research Center of Kerman University of Medical Sciences, Kerman, Iran

Received 29 January 2006; Accepted 16 September 2007; Published online 5 December 2007

Summary

Background: Considering the adverse effects of postpartum depression on both mother and infant, a screening instrument for early diagnosis seems to be of importance.

Aims: To assess the psychometric properties of the Persian version of Edinburgh Postnatal Depression Scale (EPDS) on a sample of Iranian postpartum women.

Method: The EPDS was translated and back-translated in the standard method. The questionnaire was completed by 600 postpartum women. Hundred cases with an EPDS score of ≥ 9 and 100 cases with an EPDS of <9 were randomly selected for interview. Sensitivity, specificity, positive likelihood ratio, and receiver operating characteristics were calculated by comparing the EPDS sum score against the DSM-IV diagnoses. The correlation coefficient of the EPDS score with GHQ-12 score was calculated. Principal component analysis and internal consistency were assessed.

Results: The best cutoff scores for major depression were 12/13 with a sensitivity and specificity of 95.3% and 87.9%, respectively. The correlation coefficient of the total score of the Persian version of EPDS with the GHQ-12 total score was 0.76 ($P < 0.001$). A two-factor solution was selected as the most appropriate model based on both values and the score plot. The coefficient alpha for the whole scale was 0.83.

Conclusion: The Persian version of EPDS is a reliable and valid measure for detecting postpartum depression. Furthermore, it seems acceptable to patients and a valid screening instrument for depression in postpartum women.

Keywords: Edinburgh Postnatal Depression Scale; postnatal depression; validation

Introduction

Convergent lines of evidences have shown an increased vulnerability of women for episodes of major depression during postpartum period (Cox et al. 1987; O'Hara and Swain 1996). Epidemiological studies have indicated that 10–15% of mothers suffer from a depressive disorder in the postpartum period (Kumar and Robson 1984; Watson et al. 1984; Cooper et al. 1988; Cox et al. 1993). Behavioural disturbances and cognitive deficits during childhood are common among the children of such depressed mothers (Wrate et al. 1985; Cogill et al. 1986). Other negative aspects of postnatal depression include disturbances in mother–infant interaction, attachment and marital distress, which may have long-term negative effects on the family (Murray 1992; Ballard et al. 1994). Different studies have observed a relationship between weight faltering in the first two years with both elevated maternal EPDS and higher rates of diagnosed depression in mothers (O'Brien et al. 2004). Studies in developing countries have consistently found a correlation between maternal depression and weight faltering, suggestive of maternal depression possibly being a key factor of underprivileged populations (Anoop et al. 2004; Patel et al. 2004).

Recent studies have reported a high prevalence of mental disorders in Iran with higher percentage in women (25.9%) than men (15.9%) (Noorbala et al. 2004). Despite the potentially serious consequences of postpartum depression, it is unfortunate that the high percentages of postpartum depressive cases are not detected due to misdiagnosis or lack of self-awareness. Therefore, a screening instrument for early diagnosis seems to be of importance (Perfetti et al. 2004). The Edinburgh Postnatal Depression Scale (EPDS) is an easy-to-use

Correspondence: Shahrzad Mazhari, M.D., Neuroscience Department, Research Center of Kerman University of Medical Sciences, Jomhoori Islamic Blvd, P.O. Box 67175-113, Kerman, Iran
e-mail: mazhas01@student.uwa.edu.au

scale with 10 self-rating items that is specifically designed to detect depression in postpartum period (Cox et al. 1987). Research has shown that it is valid and reliable as a screening tool for postpartum depression (Cox et al. 1987). Since 1987 numerous validation studies have been carried out with most of them reporting satisfactory psychometric properties (Harris et al. 1989; Murray and Carothers 1990; Boyce et al. 1993; O'Hara et al. 1994; Gubash et al. 1997; Guedeney and Fermanian 1998; Berle et al. 2003).

The aim of the present study was to translate the EPDS into Persian (Iranian language) and to assess the psychometric properties of the Persian version of EPDS on a large sample of Iranian postpartum women.

Method

The research was carried out in Kerman, which is the centre of the largest province of Iran, located 895 kilometres south of the capital. The participants were 600 women attending their infant's vaccination program in 5 randomly selected urban health centres representing different socioeconomic classes. The research was conducted from May 2004 to May 2005.

Eligibility criteria were as follows: (1) speaking Persian language, (2) no evidence of depression due to medical illness, (3) women's consent.

The questionnaire was administered by a research assistant. She told the mothers that it was a study of their mood after child delivery and that participation was completely voluntary. She also assisted illiterate women in completing the questionnaires. This research was done in two stages. In the first stage, 600 subjects completed the EPDS. In the second stage, a randomized sample of 100 cases with EPDS score of ≥ 9 and 100 cases with EPDS of <9 were referred to the psychiatry clinic. The selected patients were interviewed as soon as they were able to attend the clinic within a two-week period and assured that they would not be required to pay for medical attention. Written informed consent was obtained from all participants.

Clinical interviews were carried out by the research psychiatrist at her clinic. Diagnosis of major and minor depression episodes were made according to DSM-IV criteria, axis 1 (American Psychiatric Association 1996). The research psychiatrist was blind to the EPDS scores and did not know the EPDS results of the participating women. The women referred were also asked to fill out the Persian version of GHQ-12 (Montazeri et al. 2003). For those women reluctant to come for interview in the clinic, the psychiatrist attended the health centre in the women's neighborhood after making necessary arrangements. All efforts were made to keep information private.

The standard forward-backward method was applied to translate the EPDS from English into Persian (Harkness et al. 2002). At first, EPDS was translated into Persian by the authors and then re-translated into English by two independent native English speakers. No significant problematic issues in translation were encountered.

Statistical analysis

Validity

For face validity the acceptability and the meaning of the items were investigated in 20 women with different educational levels.

Criterion validity was judged by comparing the EPDS sum score against the DSM-IV diagnoses (American Psychiatric Association 1996), the sensitivity, specificity, positive likelihood ratio, and receiver operating characteristics (ROC) were calculated for different cutoff scores.

Construct validity was evaluated by the correlation coefficient of the EPDS score with the GHQ-12 score, which was expected to show good correlation (Guedeney and Fermanian 1998; Lee et al. 1998). Meanwhile, the EPDS score was compared among three groups: major depression, minor depression, no depression (according to DSM-IV criteria).

In order to determine whether the items on the scale assessed distinct aspects of depression, we conducted a principal component analysis with varimax rotation on the total sample.

Reliability

Internal consistency was assessed both by Cronbach's alpha coefficient and item-scale correlation (after correction for overlapping). The minimum acceptable level of Cronbach's alpha for self-report questionnaire and item-scale correlation was assumed 0.6 (Nunnally and Bernstein 1994) and 0.4 (Ware and Gandek 1998), respectively. Statistical analyses were done by SPSS 12.0.

Results

Subjects characteristics

The final sample of those participating in the diagnostic interview included 200 postnatal women. The mean age of the mothers was 25.6 ± 4.9. The mean number of children was 1.8 ± 0.9. Six percent were illiterate, 22% had a secondary education or less and the remainder (65%) had a high school diploma or some college education. The majority (84.5%) were not employed. The mean EPDS score for the total subjects ($n = 600$) was 10.1 ± 5.09. Of the 200 interviewed women in the sample, 43 women (21.5%) met the DSM-IV criteria for a major depressive disorder and 26 women (13%) filled the criteria for a minor depressive one. The mean EPDS score of the interviewed sample ($n = 200$) was 10.0 ± 5.4 and the mean of EPDS for the three subgroups (major depression, minor depression, and no depressive disorder) were 17.83 ± 3.75, 11.38 ± 2.97, and 7.14 ± 3.2, respectively, which showed a statistically significant difference ($F\ (2,197) = 171.2$, $P < 0.001$).

Table 1. Trade-off between sensitivity and specificity, with likelihood calculation for detection of major and combined (major and minor) depression using EPDS ($n = 200$)

Cutoff score	Major depression			Combined depression		
	Sensitivity	Specificity	LR+[a]	Sensitivity	Specificity	LR+
7/8	100.0	47.1	1.9	95.2	54.2	2.1
8/9	100.0	63.7	2.8	92.8	72.5	3.4
9/10	97.7	72.0	3.5	89.9	81.7	4.9
10/11	97.7	77.1	4.3	87.0	86.3	6.3
11/12	97.7	84.1	6.1	81.2	91.6	9.7
12/13	95.3	87.9	7.9	73.9	93.1	10.7
13/14	90.7	93.6	14.9	63.8	96.2	16.8
14/15	83.7	95.5	18.6	56.5	96.6	18.2

[a] LR+, Positive likelihood ratio.

Criterion validity

To examine the issue of cutoff (a discriminate level for predicting postnatal depression) sensitivity, specificity, and positive likelihood ratios were analyzed for different levels of EPDS-P score (Table 1). For example, Table 1 shows a score of 9/10 of EPDS identified 98% of women with major depression and 90% with major or minor depression (combined). Whilst EPDS recognized only 28% and 18.3% as ''depressed'' who were certainly disease free (regarding major and combined depression).

Construct validity

The correlation coefficient of the total score of the EPDS-P with the GHQ-12 total score was 0.76 ($P < 0.001$).

Kaiser's measure of sampling adequacy (MSA) was applied to the data prior to factor extraction. It revealed an index of 0.87, indicating data were likely to factor well. A two-factor solution was selected as the most appropriate model based on both values and the score plot, which accounted for 46.7% of the total variance. The factor loadings and communalities are included in Table 2. Loadings under 0.3 were not reported. The factor one consisted of questions one and two, which identified anhedonia. Factor two comprising the remainding eight items included items relating to depressed mood and other nonspecific symptoms of depression.

Table 2. EPDS factors and factor loadings[a]

Item	Communalities	Factor 1	Factor 2
1	0.61	0.76	
2	0.68	0.82	
3	0.52		0.72
4	0.48		0.68
5	0.96		0.67
6	0.34	0.34	0.47
7	0.37		0.56
8	0.49	0.47	0.52
9	0.44	0.36	0.56
10	0.28	0.35	0.39

[a] Loadings less than 0.3 are not mentioned.

Reliability

Internal consistency was acceptable. All items of the two-factor structure demonstrated substantial correlation (all correlation coefficients were higher than 0.4) with their hypothesized scales (satisfactory item-scale correlation). Cronbach's alpha exceeded the recommended minimum of 0.7 (Nunnally and Bernstein 1994), for the whole scale ($\alpha = 0.83$). Breaking the items down into the two factors yielded a coefficient $\alpha = 0.59$ for the ''anhedonia'' factor and 0.82 for the ''other symptoms'' factor.

Discussion

Despite the fact that the GHQ-12 questionnaire (Montazeri et al. 2003) and the Beck Depression Inventory (Hojat et al. 1986) have been translated into Persian for use in Iran, there are some limitations in using them for postpartum depression (Cox et al. 1987). Since the EPDS questionnaire is one of the most prevalent self-report screening instruments for postpartum depression, we investigated its applicability, validity, and reliability (Perfetti et al. 2004). Our study showed the face validity of the questionnaire was substantial due to its acceptability, comprehensibility, absence of annoying questions, and the advantage of it taking a short time to fill in.

We used two different methods to determine validity and reliability suggested by Nunnally and Bernstein (1994). In order to determine criterion validity, beside sensitivity and specificity we also used positive likelihood ratio. Positive likelihood ratio indicates how many

times more likely a test result of a given level is obtained in subjects with the disease than in those without the disease. Positive likelihood ratio surpasses specificity and sensitivity practically, as it is a combination of them. Another advantage of positive likelihood ratio is that when we deal with a numerical range and interval data, a posttest probability can be calculated for each score in different prevalence rates by using a nomogram instead of simply presenting a positive or negative result (Jaeschke et al. 1994).

Our study of the Persian version of EPDS suggested that the best cutoff scores for major depression and combined depression were 12/13 and 11/12, respectively (Table 1). In studies by Cox et al. (1987), Boyce et al. (1993), Lee et al. (1998), and others (Eberhard-Gran et al. 2001), the same cutoff score (i.e., 12/13) has been recommended. It must be mentioned that in this cutoff, 4.7% of the cases remained undiagnosed and 12.1% were falsely labeled as major depressives. Positive likelihood ratio at this cutoff is ''moderate'' (Jaeschke et al. 1994) and if we intend it to reach the ''often conclusive'' category (i.e., higher than 10), the score has to be increased to 13/14, which subsequently results in missing 10% of patients.

The importance of screening psychiatric disorders in primary health care is clear (Halverson and Chan 2004). In Iran's health system, health centres are major places where women routinely receive prenatal care, vaccinations, and advice on family planning. Maternal and child health care are considered a central aspect of primary health care provided by urban and rural health centres and the target population has free access to services. These services are provided by health workers but not physicians. Health workers refer all those cases suspected of suffering from a disease to a physician. This questionnaire is used as a screening tool for early detection of depression, and reflects the national health policy's emphasis on a preference for preventive measures (Asadi-Lari et al. 2004). Therefore, applying a lower cutoff may be preferred if it is designed to be used by nonphysician health care providers to decrease the number of missed cases. We suggest the cutoff of 11/12 in the more peripheral levels of the health care system where unskilled health personnel are the first to provide health services and see to the referral of patients to advanced health facilities. Obtaining posttest probability of postpartum depression in the clinical setting requires calculating positive likelihood ratio, information about pretest probability (i.e., prevalence of the disease in the community) with use of a nomogram (Deeks and Altman 2004). The area under the curve was 98% for major depression and 93% for combined depression, which is acceptable and comparable with other similar studies (Garcia-Esteve et al. 2003; Berle et al. 2003). The mean score of EPDS in major depression was in the range of 17–18 and comparable to other studies (Garcia-Esteve et al. 2003; Lee et al. 1998). Desirable correlation between EPDS score and GHQ-12 and also statistically significant difference of EPDS score for the three subgroups (major depression, minor depression, no depression) showed substantial construct validity.

Two-factor models were found in our factor analysis as well as two other studies (Guedeney and Fermanian 1998; Pop et al. 1992). Since both of these factors are depressive in nature (Cox et al. 1987), this finding is not applicable and unidimensional factor seems to be preferable (Palmer et al. 2002). According to Nunnally and Bernstein (1994), the test-retest method was not used for calculating reliability, but the other two methods yielded very good results that were strongly in favor of internal consistency.

A limitation of this study was response bias. It could be an issue because illiterate respondents had to be asked verbally, but it was inevitable to ensure the generalizability of the results to the population.

In conclusion, the data obtained from this study show that the Persian version of EPDS is a reliable and valid means of detecting postpartum depression. In addition, it seems acceptable to patients and can be used for quick screening in puerperal women.

Acknowledgement

This research was supported by the Neuroscience Research Centre, Kerman University of Medical Sciences. We thank Professor JL Cox for his permission to use EPDS and the women taking part.

References

American Psychiatric Association (1996) Diagnostic and statistical manual of mental disorders (DSM-IV), 4th edn. American Psychiatric Association, Washington, DC

Anoop S, Saravanan B, Joseph A, Cherian A, Jacob KS (2004) Maternal depression and low maternal intelligence as risk factors for malnutrition in children: a community based case-control study from South India. Arch Dis Child 89: 325–329

Asadi-Lari M, Sayyari AA, Akbari ME, Gray D (2004) Public health improvement in Iran: lessons from the last 20 years. Public Health 118: 395–420

Ballard C, Davis R, Cullen PC, Mohan RN, Dean C (1994) Prevalence of postnatal psychiatric morbidity in mothers and fathers. Br J Psychiatry 164: 782–788

Berle JO, Aarre TF, Mykletun A, Dahl AA, Holsten F (2003) Screening for postnatal depression. Validation of the Norwegian version of the

Edinburgh Postnatal Depression Scale, and assessment of risk factors for postnatal depression. J Affect Disord 76(1–3): 151–156

Boyce P, Stubbs J, Todd A (1993) The Edinburgh Postnatal Depression Scale: validation for an Australian sample. Aust N Z J Psychiatry 27: 472–476

Cogill SR, Coplan HL, Alexandra H, Robson KM, Kumar R (1986) Impact of maternal postnatal depression on cognitive development of young children. BMJ 292: 1165–1167

Cooper PJ, Murray L (1995) The course and recurrence of postnatal depression. Br J Psychiatry 166: 191–195

Cox JL, Holden JM, Sagovsky R (1987) Detection of postnatal depression. Development of the 10-item Edinburgh Postnatal Depression Scale. Br J Psychiatry 150: 782–786

Cox JL, Murray D, Chapman G (1993) A controlled study of the onset, duration and prevalence of postnatal depression. Br J Psychiatry 163: 27–31

Eberhard-Gran M, Eskild A, Tambs K, Opjordsmoen S, Samuelsen SO (2001) Review of validation studies of the Edinburgh Postnatal Depression Scale. Acta Psychiatry Scand 104: 243–249

Garcia-Esteve L, Ascaso C, Ojuel J, Navarro P (2003) Validation of the Edinburgh Postnatal Depression Scale (EPDS) in Spanish mothers. J Affect Disord 75: 71–76

Gubash R, Abou-Saleh MT, Daradkeh TK (1997) The validity of the Arabic Edinburgh Postnatal Depression Scale. Soc Psychiatry Psychiatr Epidemiol 32: 474–476

Guedeney N, Fermanian J (1998) Validation study of the French version of the Edinburgh Postnatal Depression Scale (EPDS): new results about use and psychometric properties. Eur Psychiatry 13: 83–89

Halverson J, Chan C (2004) Screening for psychiatric disorders in primary care. WMJ 103(6): 46–50

Harkness JA, Van De Vijver FJR, Mohler PP (2002) Cross-cultural survey methods. Wiley, Hoboken, NJ

Hojat M, Shapurian R, Mehryar AH (1986) Psychometric properties of a Persian version of the short form of the Beck Depression Inventory for Iranian college students. Psychol Rep 59: 331–338

Jaeschke R, Guyatt GH, Sackett DL (1994) For the Evidence-Based Medicine Working Group. Users' guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for patients? JAMA 271: 703–770

Kumar R, Robson K (1984) A prospective study of emotional disorders in child-bearing women. Br J Psychiatry 144: 35–74

Lee DT, Yip SK, Chiu HF, Leung TY, Chan KP, Chau ID, Leung HC, Chung TK (1998) Detecting postnatal depression in Chinese women. Validation of the Chinese version of the Edinburgh Postnatal Depression Scale. Br J Psychiatry 172: 433–437

Montazeri A, Harirchi AM, Shariati M, Garmaroudi G, Ebadi M, Fateh A (2003) The 12-item General Health Questionnaire (GHQ-12): translation and validation study of the Iranian version. Health Qual Life Outcomes 13: 66–69

Murray L (1992) The impact of postnatal depression on infant development. J Child Psychol Psychiatry 33: 543–547

Murray L, Carothers AD (1990) The validation of the Edinburgh Postnatal Depression Scale on community sample. Br J Psychiatry 157: 288–290

Noorbala AA, Bagheri Yazdi SA, Yasamy MT, Mohammad K (2004) Mental health survey of the adult population in Iran. Br J Psychiatry 184: 70–73

Nunnally JC, Bernstein IH (1994) Psychometric theory, 3rd edn. McGraw-Hill, New York

O'Brien L, Heycock E, Hanna M, Jones PW, Cox JL (2004) Postnatal depression and faltering growth: a community study. Pediatrics 113: 1242–1247

O'Hara MW, Swain AM (1996) Rates and risk of postpartum depression: a meta analysis. Rev Psychiatry 8: 37–54

Palmer RF, Graham JW, Taylor B, Tatterson J (2002) Construct validity in health behavior research: interpreting latent variable models involving self-report and objective measures. J Behav Med 25: 525–550

Patel V, Rahman A, Jacob KS, Hughes M (2004) Effect of maternal mental health on infant growth in low income countries: new evidence from South Asia. BMJ 328: 820–823

Perfetti J, Clark R, Fillmore CM (2004) Postpartum depression: identification, screening, and treatment. WMJ 103(6): 56–63

Pop VJ, Komproe IH, van Son MJ (1992) Characteristics of the Edinburgh Post Natal Depression Scale in The Netherlands. J Affect Disord 26: 105–110

Ware JE Jr, Gandek B (1998) Methods for testing data quality, scaling assumptions, and reliability: the IQOLA Project approach. International Quality of Life Assessment. J Clin Epidemiol 51: 945–952

Watson JP, Elliott SA, Rugg AJ, Brough DI (1984) Psychiatric disorder in pregnancy and the first postnatal year. Br J Psychiatry 144: 453–462

Wrate RM, Rooney AC, Thomas PF, Cox JL (1985) Postnatal depression and child development: a three-year follow-up study. Br J Psychiatry 146: 622–627