ORIGINAL ARTICLE

Reliability and Validity of the Malay Version of Edinburgh Postpartum Depression Scale (EPDS) When Administered to Postpartum Mothers at Two Points in Time

Nursazila Asikin MA[a], Siti Norhani S[b], Nur Arzuar AR[c], Eshaifol Azam O[a]

[a]Integrative Medicine Cluster, Advance Medical Dental Institute, Universiti Sains Malaysia, Kepala Batas, Pulau Pinang, Malaysia

[b] Jalan Damai Health Clinic, Batang Padang District, Tapah, Perak, Malaysia

[c] Regenerative Medicine Cluster, Advance Medical Dental Institute, Universiti Sains Malaysia, Kepala Batas, Pulau Pinang, Malaysia

ABSTRACT

Introduction: Edinburgh Postpartum Depression Scale (EPDS) is a tool used to assess the risk of postpartum depression (PPD). In this study we determined the reliability and validity of the Malay version of EPDS when administered at two different time points in the postpartum period. **Materials and Methods:** This cross-sectional study design was carried out between May and September 2017 at three government primary healthcare clinics located in Batang Padang district, a suburban area of Perak state in Peninsular Malaysia. We recruited a total of 89 women; 41 women were in the early postpartum period (1-30 days) and 48 women were in the late postpartum period (31-120 days). Cronbach's alpha coefficient, inter-item correlation, and corrected item-total correlation were used to assess the internal consistency. The concurrent validity was assessed using Spearman's correlation. The data were analyzed using SPSS version 20 and R 3.4.2. **Results:** The Cronbach's alpha for the first and second group was 0.78 and 0.62, respectively, which indicated satisfactory reliability. At both time periods, removing Item 2 from the scale resulted in a significant increase in Cronbach's alpha (to 0.847 and 0.709, respectively). As expected, the EPDS scores correlated moderately with the BDI-II scores (1–30 days: Spearman's rho = 0.65, $p < 0.01$; 31–120 days: Spearman's rho = 0.73, $p < 0.01$). **Conclusion:** The Malay version of the EPDS is a reliable screening instrument for detecting postpartum depression. It showed reasonability and feasibility and can be used in postpartum clinical settings or for assessing intervention effects in research studies. Furthermore, as our results indicated, removing Item 2 from the Malay version would increase the internal consistency of the EPDS.

KEYWORDS: Malay EPDS, postpartum, postpartum depression syndrome, reliability, concurrent validity

INTRODUCTION

Postpartum depression (PPD) is a common mood disorder observed in the postpartum period. Empirical evidence indicates that the prevalence of PPD at 4–12 weeks after delivery ranges from 13.2% to 37.1%.[1,2] There are additional risk factors associated with the development of PPD amongst women in developing countries, which are related to specific cultural norms[3] and under-resourced healthcare systems.[4]

Many researchers have used the Edinburgh Postnatal Depression Scale (EPDS) to measure PPD. EPDS is a 10-item self-administered screening tool focusing on postpartum mood disorders. The tool has been translated into many languages including Greek,

Corresponding Author
Dr. Eshaifol Azam Omar,
Advance Medical and Dental Institute,
Universiti Sains Malaysia,
Kepala Batas, Pulau Pinang, Malaysia.
Tel: +6013-3645899
Fax: +604-5622349
E-mail: eshaifol@usm.my

French, Chinese and Arabic.[5-8] In 1987, Cox, Holden and Sagovsky established that EPDS is a valid and reliable postpartum depression screening tool.[9] In the last three decades, validity and reliability of this instrument have been extensively tested, confirming its satisfactory psychometric properties.[10,11]

The available Malay language version of the EPDS was developed by Azidah et al in 2004 and was validated on a sample of postpartum Malaysian women in Kelantan, North East of Peninsular Malaysia.[12] The study findings suggested an EPDS cut -off score value of 11.5 for depression with the sensitivity of 72.7% and specificity of 92.6%. The Malay version of the EPDS was also shown to have good internal consistency (Cronbach's alpha = 0.86) and good split-half reliability (Spearman split half coefficient = 0.83). Based on the study conducted by Wan Mahmud and Mohamed, the instrument also showed satisfactory discriminant and concurrent validity.[13]

Although the translated version of the EPDS is consistent with the original scale, variations may occur if it is applied in a different sociocultural environment, as well as when the assessment is carried out at various points in time, for example during the antenatal period, in early or late postpartum period.[14] Therefore, validation of the EPDS in a population should also be conducted at various points in time before the instrument can be reliably used for screening depression. Thus, the aim of the present study was to (i) test the Malay version of the EPDS and establish its reliability in identifying PPD among mothers following childbirth at two different points in time, and (ii) assess the validity of the Malay EPDS.

MATERIAL AND ETHODS

Study design and setting

This cross-sectional study was carried out between May and September 2017 at three government primary healthcare clinics located in Batang Padang district, a suburban area of Perak state in Peninsular Malaysia. These clinics were selected due to their high postpartum attendance compared to other clinics in the area. Majority of the postpartum mothers who visited the clinics were Malay. While the remaining women were *Orang Asli*, Chinese or Indian, some could read, understand and speak the Malay language fluently.

Sample selection and data collection

Convenience sampling was employed when selecting study participants. The sample size for this study was based on a calculation of the ratio of 7 subjects per item.[15] After considering 25% attrition rate, the sample size of 89 was chosen. Eligible subjects for the study were women who were in the first four months postpartum, showing no signs of depression due to medical illness, and were able to read and understand the Malay language, as well as give consent. Participants were recruited from Jalan Damai, Trolak, and Sungkai Health Clinics. Women were approached at the time of their visit to the clinic for routine postpartum examination or immunization for their infants. A trained female nurse collected the data at two time points: (a) early postpartum (1-30 days postpartum), and (b) late postpartum (31–120 days postpartum). The postpartum women were asked to complete the self -administered Malay version EPDS and the Malay version Beck Depression Inventory-II (BDI-II).

Based from the previous studies, 11.5 cut-off point for the Malay EPDS is valid to detect both minor and major depression.[12,13] Women with EPDS score exceeding 11.5 and a random sample of women with low scoring Malay EPDS, were invited to attend an interview session by the family health specialist at Jalan Damai clinic. The family health specialist who conducted the psychiatric interviews was blind to the EPDS results. The interview was conducted within 3 weeks of the screening, by using the Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I), Clinical Version.

The study received approval from Medical Research and Ethics Committee, Ministry of Health Malaysia (NMRR-16-2457-33280). Permission to collect the data was also obtained from the director of the Health District Office. During the field study, the study rationale was explained to the participants, and informed consent was obtained from all mothers before they answered the self-administered questionnaires.

Study instruments

The tools used in this study included: (a) the Malay EPDS and (b) Malay Beck Depression Inventory-II (BDI -II). The Malay version of the EPDS was first validated in 2004. The cut-off score was 11.5 with

the sensitivity of 72.7% and specificity of 92.6%. When the Malay EPDS was applied to the Malaysian population, it demonstrated a good validity in detecting the postpartum depression.[12] The BDI was developed by Beck et al. in 1961 to measure the presence of depression in normal population and intensity of depressive symptoms in clinical populations.[16] The BDI-II is its revised form and has become the most widely used instrument for detecting depression. BDI-II is short and brief; it consists of 21 items rated on a 4-point scale, ranging from 0 (no symptom) to 3 (severe symptoms), giving the maximum score of 63. BDI-II was translated and validated in the Malay language.[17] However, Item 21 was discarded due to cultural and religious perspectives. Being Muslim and still holding the strong Malay cultural values, participants may present specific barriers, since issues of sexuality will not be revealed, in contrast to Western populations.[18]

Data analysis

Statistical analysis was conducted using SPSS version 20 (SPSS Inc., Chicago, IL) and R 3.4.2. The participants' demographic profile was described using mean and median (for non-normally distributed data), standard deviation (SD), frequency, and percentage. Data were further categorized based on the EPDS score.

Reliability

Internal consistency of the instrument was examined using Cronbach's alpha coefficient values, which allowed assessing the relatedness of each item in each and every domain, inter-item correlations and corrected item-total correlations. A Cronbach's alpha of less than 0.6 was considered poor or weak, 0.6–0.8 indicated moderate but satisfactory value, and of 0.8 and above represented a measure of high internal consistency. Corrected item-total correlation values greater than 0.4 were considered acceptable.[19]

Validity

Due to the non-normally distributed data, Spearman's rho, the nonparametric analogue to Pearson's correlation was used to estimate the relationship between two variables. This non-parametric test was carried out between the EPDS and BDI-II scales to establish concurrent validity, which was expected to show good correlation.

RESULTS

Demographic characteristics

This validity and reliability study involved 89 individuals. Most of the participating mothers were Malays 91% (81), and almost two-thirds of the sample (58.43%) were working mothers. The mean age of the participants was 29 years (range 16–39; SD 5.76). About 55.06% (49) of the participants had education up to secondary school level, 24.7% (22) had skill certificate, and 20.2% (18) had a diploma. Table 1 shows the sociodemographic characteristics of the study sample.

Based on their questionnaire responses, 7.3% of mothers in the early postpartum (1–30 days postpartum) and 6.25% in the late postpartum (31-120 days postpartum) period reported to EPDS score more than 11.5. The EPDS score for the whole

Table 1. Socio-demographic characteristic of the respondents

	n (%)
Age (years)	29 (5.76)[#]
Number of children	2 (1.10)*
Race	
Malay	81 (91.0)
Chinese	2 (2.2)
Indian	1 (1.1)
Orang Asli	5 (5.6)
Postpartum day	
1–30	41 (46.1)
31–120	48 (53.9)
Education	
Primary school	4 (4.5)
Secondary school	45 (50.6)
Skill certificate	22 (24.7)
Diploma	18 (20.2)
Occupation	
Professional	23 (25.8)
Non-professional	29 (32.6)
Housewife	37 (41.6)
Household income	
Less than RM1000	18 (20.2)
RM1000-RM2000	28 (31.5)
RM2001-RM3000	16 (18.0)
RM3001-RM4000	14 (15.7)
RM4001-RM5000	8 (9.0)
>RM5000	5 (5.6)
Mode of delivery	
Vaginal	59 (66.3)
Caesarean section	25 (28.1)
Vacuum	5 (5.6)

[#]mean (SD), *median (SD)

sample ranged from 0 to 18, with the mean score of 5.66 (SD 4.05) and 5.04 (SD 3.53) at the first and second time point respectively. The results are shown in Table 2.

Internal consistency

The overall internal consistency of the whole EPDS as measured by the Cronbach's alpha coefficient was 0.70. At the first time point (1–30 days postpartum), the internal consistency was 0.78, declining to 0.62 for the second time point (30–120 days postpartum), which indicated satisfactory reliability. Though the alpha value for the first time point (1–30 days) was high, inter-item correlation suggested that Item 2 (see Figure 1, Item 2 correlations) had a lower correlation with the remaining nine items of the EPDS scale.

As shown in Figure 1, Item 2 is negatively correlated to Item 3 (r=-0.284), Item 5 (r=-0.058), Item 6 (r=-0.079), Item 7 (r=-0.185) and Item 9 (r=-0.225). Moreover, it had extremely lower correlation with Item 8 (r=0.017) and Item 10 (r=0.097) relative to other items. Surprisingly, Item 2 had almost no correlation with Item 4 (r=0.002). These results indicated that Item 2 may not be representative of the same content domain. This finding was further investigated using corrected inter-item correlation and Cronbach's alpha when each item was deleted.

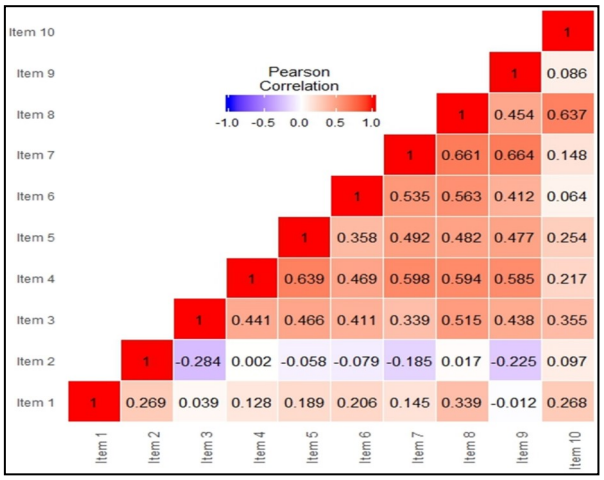

Figure 1. Heat map of the inter-item correlation matrix of 1–30 days EPDS scale.

As Gliem and Gliem reported, as a rule of thumb, corrected item-total correlation values should be greater than 0.4.[20] In this case, both Item 1 and Item 2 had values below 0.4 (see Table 3). Moreover, the corrected item-total correlation of Item 2 was negative (-0.094). Due to that, Cronbach's alpha was used to determine the importance of retaining these two items in the scale. As presented in Table 3, when each individual item (except for Item 2) was deleted, Cronbach's alpha decreased or stayed at the same level. When Item 2 was deleted, alpha value increased to 0.847, which was 0.067 greater than the value obtained when all ten items were included (0.78).

Using the same approach, the internal consistency of the scale for second time point (30–120 days) was also checked. The same trend was observed, as Item 2 was negatively correlated with Items 4, 5, 6, 7, 9 and 10 (see the light blue areas in Figure 2). In addition, Item 1 also showed negative correlation and remarkably lower correlation with most of the other items. Moreover, Item 3 and Item 4 were negatively correlated with Item 4 and Item 7, respectively. These results indicated that Items 1, 2, 3 and 4 might not be representative of the same content domain.

Table 2. The EPDS score at two points in time, the Cronbach's alpha coefficients indicate the scale reliability and Spearman's rho indicates concurrent validity

	1–30 days postpartum (n = 41) No. (%)	31–120 days postpartum (n = 48) No. (%)
EPDS score		
<11.5	38 (92.7)	45 (93.8)
≥11.5	3 (7.3)	3 (6.25)
Mean (SD)	5.66 (4.05)	5.04 (3.53)
Cronbach's alpha	0.78	0.63
Spearman's rho	0.65*	0.73*

As presented in Table 4, the corrected item-total correlation values for Items 1, 2, 3 and 4 were below 0.40, yet when Item 3 was deleted from the scale, Cronbach's alpha decreased from the previous value of 0.62 to 0.592. On the other hand, when Items 1, 2 and 4 were deleted, an increase in alpha value was noted, though the most significant increase was apparent when Item 2 was removed from the scale. The alpha value increased to 0.709, which was 0.089 higher than the previous value of 0.62.

Concurrent validity

The EPDS scores correlated moderately with the mental health component summary score of

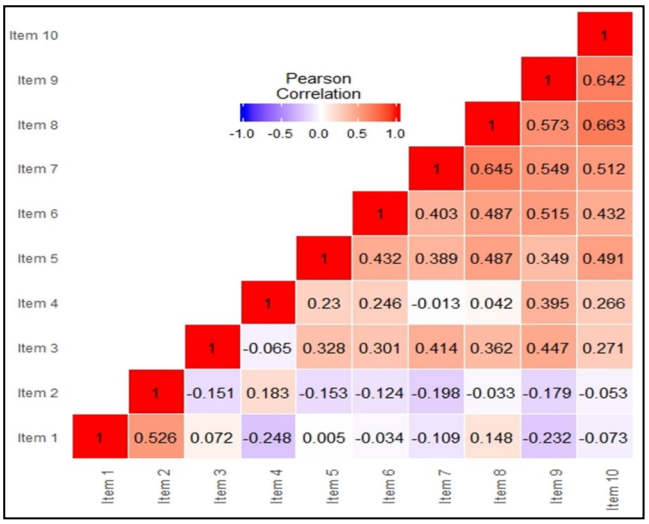

Figure 2. Heat map of the inter-item correlation matrix of 31–120 days EPDS scale

the Malay version of BDI-II. As expected, a significant correlation emerged (1–30 days postpartum: Spearman's rho = 0.65, P < 0.01; 31–120 days postpartum: Spearman's rho = 0.73, P < 0.01).

DISCUSSION

The aim of the present study was to test the validity and reliability of the Malay version of EPDS as a screening instrument that can be used to identify postpartum depression among mothers who have recently given birth. EPDS validation procedure is essential to establish instrument reliability, especially when a version other than the original one is used. In this study, we found that the Malay version of EPDS was user-friendly. Its reliability, as measured by the internal consistency, was in good agreement with the values reported in the previous validation studies carried out among the postpartum population in Malaysia.[12,13] Concurrent validity findings indicated good correlation with the Malay BDI-II.

In the present study, we also measured the internal consistency of the Malay version of EPDS at different time points. During the postpartum period, there may be some differences in the level of maternal adaptation, be it physical or social adaptation, which may affect mother's psychological wellbeing. However, for internal consistency, the tool should exhibit coherence even when administered on different occasions. In this study, the internal consistency of the scale when administered to the mothers in the first month after childbirth was similar to that obtained when the same questionnaire was completed in the 30–120 day period following childbirth. This is consistent with the findings yielded by a previous study[14] that was conducted with the same purpose. Interestingly, removing Item 2 from the scale in both time periods has resulted in a significant increase in the internal consistency.

Using the cut-off points suggested by the EPDS developers, the findings from this validation study showed a low level of postpartum depression among the mothers as compared to their counterparts in other states of Malaysia. A prospective cohort study using the EPDS cut-off point of 12 conducted on a sample of 1,362 postpartum women in Sabah revealed that the prevalence of postpartum depression was 7.1% at one month and 6.9% at three months.[21] It is assumed that a comparison of the EPDS outcomes with a gold standard, such as psychiatric diagnosis by clinical interview, would

Table 4. Inter-item correlation and Cronbach's alpha if an item was deleted from the EPDS scale (31–120 days)

EPDS Item	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
Item 1	5.49	15.106	.280	.780
Item 2	5.22	16.276	-.094	.847
Item 3	4.83	11.545	.706	.722
Item 4	4.46	13.005	.471	.760
Item 5	4.95	12.248	.618	.738
Item 6	4.98	12.724	.548	.748
Item 7	5.17	13.545	.633	.745
Item 8	5.29	12.362	.780	.721
Item 9	4.93	14.120	.547	.756
Item 10	5.61	15.394	.383	.775

potentially reveal more depression cases in postpartum women. Thus, further assessment using a gold standard test is required.

The limitation of this study is that, the Malay EPDS could not be validated further with the sensitivity and specificity test. This is because, more than a half of the respondents did not attend the 3-weeks follow-up assessments, which finally affect the evaluation to distinguish depressed and non-depressed subjects according to DSM-IV diagnoses.

CONCLUSION

In summary, we can assert that the validity and reliability of the Malay version of the EPDS in identifying postpartum depression is confirmed. Since the Malay version of EPDS has never been examined for test-retest reliability and due to the factorial structure/validity of a sample of postpartum women, these issues should be studied and clarified further. In addition, further research would be needed to decide whether the exclusion of Item 2 from the Malay version of the EPDS is warranted.

Acknowledgements

We would like to acknowledge Dr. Raja Mohd Azim B. Raja Haron, Batang Padang District Health Officer for the permission given to conduct the study. We would also like to thank all staff in Jalan Damai, Sungkai and Trolak Health Clinics who were involved in the study.

REFERENCES

1. Giardinelli L, Innocenti A, Benni L, Stefanini MC, Lino G, Lunardi C, Svelto V, Afshar S, Bovani R, Castellini G, Faravelli C. Depression and anxiety in perinatal period: prevalence and risk factors in an Italian sample. Archives of women's mental health. 2012 Feb 1;15(1):21-30.
2. Veisani Y, Delpisheh A, Sayehmiri K, Rezaeian S. Trends of postpartum depression in Iran: a systematic review and meta-analysis. Depression research and treatment. 2013;2013.
3. Abdollahi F, Lye MS, Zain AM, Ghazali SS, Zarghami M. Postnatal depression and its associated factors in women from different cultures. Iranian journal of psychiatry and behavioral sciences. 2011;5(2):5.
4. Hung KJ, Tomlinson M, le Roux IM, Dewing S, Chopra M, Tsai AC. Community-based prenatal screening for postpartum depression in a South African township. International Journal of Gynecology & Obstetrics. 2014 Jul 1;126(1):74-7.
5. Vivilaki VG, Dafermos V, Kogevinas M, Bitsios P, Lionis C. The Edinburgh Postnatal Depression Scale: translation and validation for a Greek sample. BMC Public Health. 2009 Dec;9(1):329.
6. Guedeney N, Fermanian J. Validation study of the French version of the Edinburgh Postnatal Depression Scale (EPDS): new results about use and psychometric properties. European psychiatry. 1998 Jan 1;13(2):83-9.
7. Lee DT, Yip SK, Chiu HF, Leung TY, Chan KP, Chau IO, Leung HC, Chung TK. Detecting postnatal depression in Chinese women. Validation of the Chinese version of the Edinburgh Postnatal Depression Scale. The British Journal of Psychiatry. 1998 May 1;172 (5):433-7.
8. Ghubash R, Abou-Saleh MT, Daradkeh TK. The validity of the Arabic Edinburgh postnatal depression scale. Social psychiatry and psychiatric epidemiology. 1997 Nov 1;32(8):474 -6.
9. Cox JL, Holden JM, Sagovsky R. Detection of postnatal depression: development of the 10-item Edinburgh Postnatal Depression Scale. The British journal of psychiatry. 1987 Jun;150 (6):782-6.
10. Aydin N, Inandi T, Yigit A, Hodoglugil NN. Validation of the Turkish version of the Edinburgh Postnatal Depression Scale among women within their first postpartum year. Social Psychiatry and Psychiatric Epidemiology. 2004 Jun 1;39(6):483-6.
11. Wickberg B, Hwang CP. The Edinburgh postnatal depression scale: validation on a Swedish community sample. Acta Psychiatrica Scandinavica. 1996 Sep 1;94(3):181-4.
12. Kadir AA, Nordin R, Ismail SB, Yaacob MJ, Rushidi WMWM. Validation of the Malay Version of Edinburgh Postnatal Depression Scale for Postnatal Women in Kelantan, Malaysia. Asia Pac Fam Med. 2004;3(September 2015):9–18.
13. Mahmud WM, Awang A, Mohamed MN. Revalidation of the Malay version of the Edinburgh postnatal depression scale (EPDS) among Malay postpartum women attending the Bakar Bata health Center in Alor Setar, Kedah,

north west of peninsular Malaysia. The Malaysian journal of medical sciences: MJMS. 2003 Jul;10(2):71.
14. Montazeri A, Torkan B, Omidvari S. The Edinburgh Postnatal Depression Scale (EPDS): translation and validation study of the Iranian version. BMC psychiatry. 2007 Dec;7(1):11.
15. Barlett JE, Kotrlik JW, Higgins CC. Organizational research: Determining appropriate sample size in survey research. Information technology, learning, and performance journal. 2001 Apr 1;19(1):43.
16. Beck AT, Steer RA, Carbin MG. Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. Clinical psychology review. 1988 Jan 1;8(1):77-100.
17. Mahmud WM, Awang A, Herman I, Mohamed MN. Analysis of the psychometric properties of the Malay version of Beck Depression Inventory II (BDI-II) among postpartum women in Kedah, North West of Peninsular Malaysia. The Malaysian journal of medical sciences: MJMS. 2004 Jul;11(2):19.
18. Muhktar F, Oei TP. Exploratory and confirmatory factor validation and psychometric properties of the Beck Depression Inventory for Malays (BDI-Malay) in Malaysia. Malaysian Journal of Psychiatry. 2008;17(1).
19. Downing SM. Reliability: on the reproducibility of assessment data. Medical education. 2004 Sep 1;38(9):1006-12.
20. Gliem JA, Gliem RR. Calculating, interpreting, and reporting Cronbach's alpha reliability coefficient for Likert-type scales. Midwest Research-to-Practice Conference in Adult, Continuing, and Community Education. 2003; (1992):82–8.
21. Yusuff AS, Tang L, Binns CW, Lee AH. Prevalence and risk factors for postnatal depression in Sabah, Malaysia: a cohort study. Women and Birth. 2015 Mar 1;28(1):25-9.