

ORIGINAL ARTICLE

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

Wan Mohd Rushidi Wan Mahmud*, Amir Awang, Iran Herman**, Mahmood Nazar Mohamed**.**

*Department of Psychiatry,
School of Medical Sciences, Universiti Sains Malaysia, Health Campus
16150 Kubang Kerian, Kelantan, Malaysia

**Department of Psychology and Counselling, Faculty of Social and Human Development,
Universiti Utara Malaysia, 06010 Sintok, Kedah**

Increased international collaboration in clinical trials has created a need for cross culturally valid instruments to assess the quality of life and behavioural disorders. Cross cultural studies of depressive symptomatology, in particular, must be preceded by an exhaustive study of the psychometric properties of the instruments to ensure the validity of the comparison. In this article, we examined the validity, reliability and factor structure of the Malay version of the Beck Depression Inventory II (BDI-II) among Malay postpartum women attending selected health centres in Kedah, North West of Peninsular Malaysia. Our findings indicated that the current version of the BDI-II is psychometrically strong and appropriate for use in assessing depressive symptomatology among this group of women.

Key words : BDI-II; validation, psychometric evaluation, postpartum women.

Submitted-2.3.2004, Revised-21.6.2004, Accepted-21.6.2004

Introduction

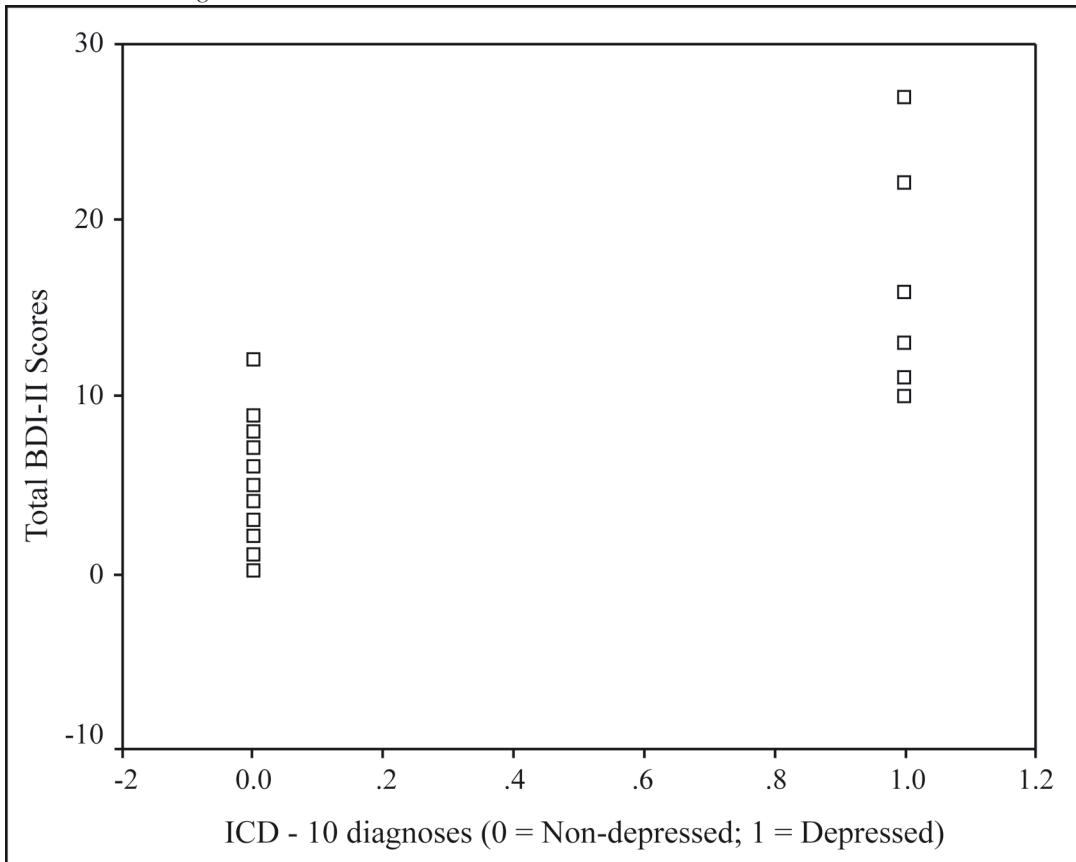
Although postpartum depression affects between 10 to 15% of recently delivered women (1, 2), only a small minority of them is identified by primary health workers (3). In fact, a case of severe postpartum depression can go unrecognized even by those closest to the new mother (4). Better detection can be achieved by administering self-reporting scales during one or more postpartum visits to identify those mothers with substantial depressive symptomatology. Among the depression instruments which have been used and currently being used to evaluate symptoms and screen for postpartum depression include the Edinburgh Postnatal Depression Scale (EPDS) (5), the Beck Depression Inventory (6), the Centre of Epidemiologic Studies Depression Scale (7) and the Postpartum Depression

Screening Scale (8).

Beck Depression Inventory (BDI) is one of the most widely used measures of depressive symptoms both in adolescents and in adults (9). The original version of this 21-item self-rating scale was introduced in 1961 (6) and later revised and copyrighted by the Centre for Cognitive Therapy in 1978 and published in 1979 (10). Although the instrument was initially developed to measure intensity or depth of the depressive symptomatology in patients with psychiatric disorders (6), it has now been widely used as a screening instrument to detect depression both in clinical practice and in research projects (11).

In postpartum research, previous evaluations of the instrument have been very limited and yielded disappointing results. Harris et al (12) for example showed that BDI was only able to identify 13 out of

Figure 1: Scatter plot - the total scores on the Malay version of the BDI-II versus ICD-10 diagnoses



19 cases of major depression in a series of 147 postpartum mothers using a pre-determined cut-off score of 10/11. In another study involving 120 women, Whiffen (13) reported that BDI was not sensitive to minor depression, leading to substantial false negatives. It was also suggested that the somatic items of BDI might render the scale unsuitable for postpartum women who may have a few somatic symptoms even when they are not depressed (12, 13).

More recently however, Lee and colleagues (14) found better psychometric performance of BDI in detecting depression among recently delivered Chinese women. Using 10/11 as the optimal cut-off score, the sensitivity, specificity, positive predictive value and negative predictive value were 82%, 89%, 50% and 97% respectively. The area under the Receiver Operating Characteristics (ROC) curve was 0.95; and their BDI and EPDS scores were moderately correlated (Spearman correlation = 0.78, $p < 0.001$). More carefully chosen cut-off score using the ROC curve and the fact that Asian women are more likely to present their depression with somatic symptoms were cited as the main reasons for the good performance of the instrument in their study (14).

BDI-II is the most recent version of the Beck Depression Inventory (15). It was first introduced in 1996 and is relatively new to the psychological assessment literature. The instrument was upgraded in order to make its symptoms content more reflective of the diagnostic criteria that are described by the American Psychiatric Association (16) for major depressive disorders. Although BDI-II is still composed of 21 items, it now contains four new symptoms namely 'agitation', 'worthlessness', 'concentration difficulty' and 'loss of energy' in place of 'weight loss', 'body image change', 'work difficulty', and 'somatic preoccupation'. The upgraded instrument continues to be scored by summing the highest rating for each of the 21 symptoms and a 4-point scale ranging from 0 to 3 is still employed for each item but the time frame for the rating has been changed from one to two weeks (15).

To the best of our knowledge, the psychometric performance of this new instrument has never been evaluated in the postpartum period. Considering the usefulness of its predecessor in detecting depression after delivery among Asian Chinese population (14), we attempted to determine validity, reliability and factor structure of BDI-II

among our postpartum Malay women attending selected health centres in Kedah, North West of Peninsular Malaysia.

Methodology

Stage I

The English version of Beck Depression Inventory II was translated into Malay by a group of schoolteachers who are bilingual (Malay and English). Another group consisting of primary care doctors, all well versed in both languages then back translated the Malay version using the usual back translation technique (17). The final version was later pilot tested among 20 postpartum women from a health centre (Bakar Bata Health Centre) in Alor Setar, Kedah, a state North West of Peninsular Malaysia. It was found to be acceptable and no revision was required. In addition, the instrument was also reviewed by two psychiatrists, two physicians and a general practitioner to ensure satisfactory face, content and semantic equivalence of the translated instrument as outlined by Flaherty et al (18).

Stage II

61 Malay women, between 4-12 weeks postpartum attending the same health centre between July and August 2002 were approached in the second stage. These women were initially given the following questionnaires:

1. Questionnaires on socio-demographic details
2. The Malay version of the Edinburgh Postnatal Depression Scale (EPDS) (19)

3. The Malay version of the Beck Depression Inventory II

All of them were subsequently interviewed using a semi-structured interview (Clinical Interview Schedule – CIS) (20) and the 17-item Hamilton Rating Scale for Depression (HRSD-17) (21). Diagnoses were made based on the Tenth Edition of the International Classification of Disease (22).

Stage III

In the third stage, we recruited 354 Malay women between 4 – 12 weeks post delivery from selected health centres within the Kota Setar and Pendang districts in Kedah through systematic random sampling. The centres involved were 3 large health centres (Bakar Bata, Jalan Putra and Pendang) and 2 smaller health centres (Sungai Tiang and Kubor Panjang). Participating women were required to complete both the Malay versions of EPDS (19) and BDI -II.

Statistical analyses

All the results were analyzed using the Statistical Package for Social Sciences version 10.05 (23)

Results

Preliminary assessment

Preliminary appraisal of the face, content and semantic equivalence of the Malay version of BDI-II was accomplished in stage 1 (details as above).

Concurrent validity and reliability

Evaluations of these measures were completed in stage II. At this stage, the mean age of the

Table I: The Receiver Operating Characteristics (ROC) of the Malay version of the BDI-II using ICD-10 criteria for depressive episodes

Score	Sens (%)	Spec. (%)	PPV (%)	NPV (%)	MR (%)	Effi. (%)
6.5	100	88.89	53.85	100	9.84	90.16
7.5	100	94.44	70	100	4.92	95.08
8.5	100	96.30	77.78	100	3.28	96.72
9.5	100	98.15	87.50	100	1.64	98.36
10.5	85.71	98.15	85.71	98.15	3.28	96.72
11.5	71.43	100	83.33	96.36	4.92	95.08
12.5	71.43	100	100	96.43	3.28	96.72
13.5	57.14	100	100	94.74	4.92	95.08
14.5	57.14	100	100	94.74	4.92	91.80

Spec=specificity; Sens=sensitivity; PPV=positive predictive value; NPV=negative predictive value; MR=misclassification rate; Effi=efficiency

participants was 26.62 years (standard deviation = 5.51) and they were between 30 to 84 days post delivery (mean = 55.11 days; standard deviation= 15.84). Their overall mean BDI-II scores was 4.43 (standard deviation = 5.45). The median BDI-II score was notably higher among the depressed women (median= 16; inter quartile range = 11 – 22) compared to their non-depressed counterparts (median = 2.5; inter quartile range = 0 – 5) and this difference was statistically significant. (Mann Whitney U: 2 tailed p value < 0.01).

The internal consistency of the Malay version of the BDI-II was 0.89. The Spearman correlations between their BDI-II scores and, their EPDS and 17 items Hamilton Rating Scale for Depression scores were 0.72 and 0.75 respectively. The unequal length Spearman Brown split half reliability of the instrument was 0.84. 11.48% of the participants were diagnosed to have significant depressive episodes (4 mild; 3 moderate) based on ICD-10 criteria. The concurrent validity of the Malay version of BDI-II was evaluated at several cut-off scores against ICD-10 diagnoses of a depressive episode (Table I and Figure I). The Receiver Operating Characteristics (ROC) of the Malay version of BDI-II is shown in Table II. Using the optimum threshold value of 9.5 (cut-off score 9 / 10), the Malay version of BDI-II was able to detect depression of varying severity with a sensitivity of 100%, specificity of 98.15%, positive predictive value of 87.50 %, negative

predictive value of 100 %, efficiency of 98.36% and misclassification rate of 1.64%. The area under the ROC curve was calculated to be 0.995.

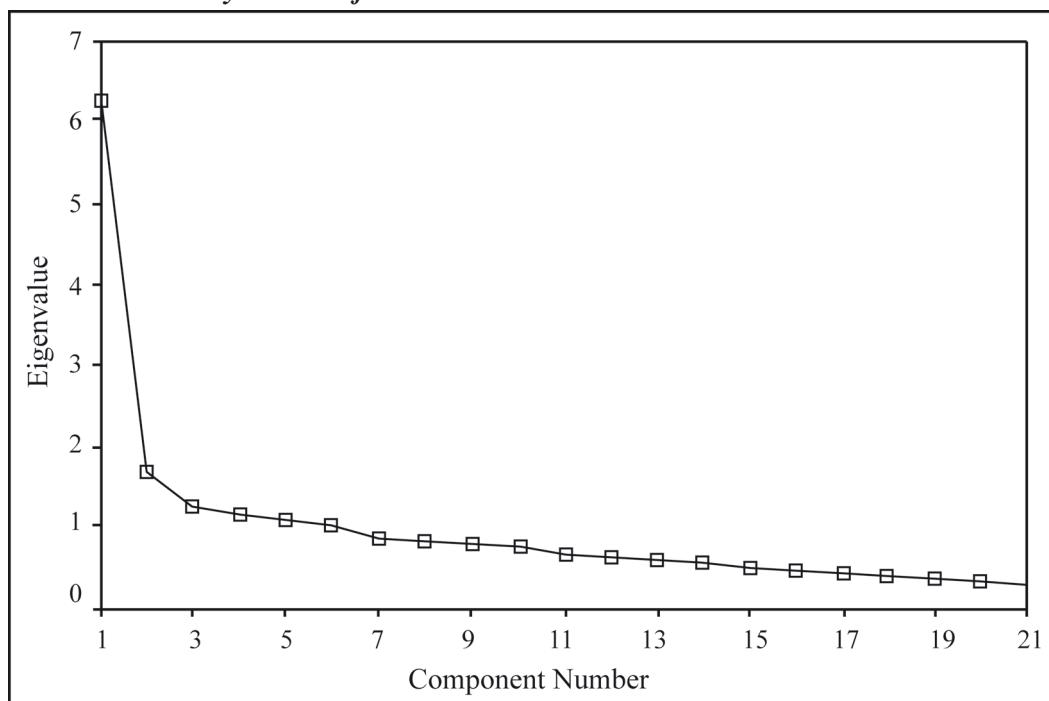
Factor structure

This was evaluated in the final stage (stage 3) of the study. The mean age of the participants in this stage was of 29.56 years (standard deviation = 5.81) and they were between 37 to 65 days post delivery (mean = 41.53 days; standard deviation= 9.40). Their mean overall scores on the BDI-II and EPDS were 6.17 (standard deviation= 6.44) and 5.82 (standard deviation = 4.30) respectively. The Spearman correlation between the latter two scores was 0.74.

Since the Bartlett's test of sphericity for the data collected at this stage was significant ($p < 0.01$) and the Kaiser-Meyer-Olkin measure of sampling adequacy for the 21 items scale BDI-II was 0.87, which Kaiser (24) reported as marvellous, we proceeded to factor analysis. Principle component analysis with varimax rotation provided the best and most interpretable solution. Examination of the initial statistics revealed 3 factors could be extracted which accounted for 43.56% of the variance (factor 1= 16.70%, factor 2= 14.66% and factor 3= 12.20%). The scree plot graphically displayed the eigen values of each factor and suggested that there was one predominant factor (figure II).

Based on analyses, two matrices were produced: a component matrix and a rotated

Figure 2: *Scree plot of the eigen values of the rotated factors from the 21 items of the Malay version of the BDI-II.*



component matrix. The latter had notably fewer complex variables and simpler structures (table II). The three factor model forced the affective symptoms onto the first factor comprising of 8 items (5, 8, 10, 11, 12, 13, 17 and 18) with factor loadings ranging from 0.37 to 0.66; somatic symptoms on to the second factor with another 8 items (1, 2, 4, 15, 16, 19, 20 and 21) with loadings ranging from 0.41 to 0.75; and cognitive symptoms on to the third factor comprising of 5 items (3, 6, 7, 9, 14) with factor loadings ranging from 0.51 to 0.75. The factor correlation matrices (table III) indicated that these factors were moderately correlated. Internal consistencies (Cronbach's alphas) for each of the subscales namely affective, somatic and cognitive were 0.80, 0.75 and 0.68 respectively.

Discussion

Previous studies in the Western populations have suggested that BDI might not be a suitable

instrument for childbearing women (12, 25). O'Hara, Neunaber, and Zekoski (25) for example, reported a certain percentage of depressed women were missed with the BDI because of its limited sensitivity in detecting minor depression. In another study with a sample of 62 primiparous women and a comparison group of 36 nulliparous women, Huffman et al (26) concluded that the total BDI score over represented depressive symptoms because of the confusing influence exerted by normal physiological changes associated with the postpartum period. Concurringly, Whiffen (13) pointed out that poor performance of the BDI may be attributed to the high rate of minor depression among this group of women, to which the BDI was particularly insensitive. Combining the above limitations with the introduction of the Edinburgh Postnatal Depression Scale (EPDS) (5), BDI has somewhat lost its popularity in postpartum research.

However, contemporary data has revived the interests on the use of this instrument among recently delivered mothers. For example, in a sample of 145

Table II: Pattern matrix from the factor analysis of the 21 items of the Malay version of the BDI-II.

	Items	Factor		
		1	2	3
11	Agitation	0.66		
10	Crying	0.62		
17	Irritability	0.61		
8	Self criticalness	0.61		
5	Guilty feelings	0.59		
12	Loss of interest	0.57		
13	Indecisiveness	0.55		
18	Change in appetite	0.37		
20	Tiredness		0.75	
15	Loss of energy		0.75	
4	Loss of pleasure		0.67	
2	Pessimism		0.48	
21	Loss of libido		0.46	
16	Change in sleeping pattern		0.42	
19	Concentration difficulty		0.41	
1	Sadness		0.41	
6	Punishment feelings			0.75
14	Worthlessness			0.61
7	Slef dislike			0.56
9	Suicidal thoughts or wishes			0.55
3	Past failure			0.51

Extraction Method: Principal Component Analyses.

Rotation Method: Varimax with Kaiser Normalization

(Rotation converged in 6 iterations; Loadings below 0.30 are suppressed)

Table II: Component correlation matrix from the factor analysis of the 21 items of the Malay version of the BDI-II

Factor	1	2	3
1	0.67	0.56	0.49
2	-0.31	0.81	-0.50
3	-0.68	0.19	0.71

Extraction Method: Principal Component Analyses; Rotation Method: Keiser Normalization.

Chinese women, Lee et al (14) demonstrated a much better psychometric performance of BDI in detecting postpartum depression. According to the authors, a more carefully chosen cut-off score using the ROC curve was crucial since their data indicated that the sensitivity and predictive value of the instrument would be very different had they chosen a cut-off score that is merely 2 points higher (sensitivity of 82% and positive predictive value of 50% at cut-off score of 10/11; compared to 65% and 52 % respectively at the cut-off score of 12/13) (14). Furthermore, the fact that Asian women are more likely to present their depression with somatic symptoms may also render BDI (with its somatic components) particularly suitable for their Chinese subjects (14).

BDI-II is the most recent upgraded version of the BDI (15). So far data on the psychometric performance and accuracy of the instrument in screening for postpartum depression have yet to be reported in the literature (8). Preliminary data from this study suggested that BDI-II can be a useful screening instrument for postpartum depression. Its psychometric performance is quite impressive both in terms of reliability and validity.

The internal consistency and split half reliability were relatively high (Cronbach alpha = 0.89; unequal length Spearman Brown = 0.84) indicating a high level of homogeneity among items in the scale. The instrument also showed good discriminant and concurrent validity. Significant difference between the BDI-II scores in the depressed group compared to their non-depressed counterparts (Mann Whitney U: 2 tailed p value < 0.01); and good correlations between the scores on the instrument and both the Malay version of EPDS (Spearman's rho = 0.72) and the 17-item HRSD (Spearman's rho = 0.75) were evidence in support of the validity of the Malay version of BDI-II.

Along the same line, this instrument had demonstrated remarkable specificity and sensitivity in detecting depression among these women. At an optimal cut-off point of 9/10, the instrument was able to detect 100% of those with postpartum

depression with a positive predictive value of approximately 89.15%. Such figures are as impressive as those of EPDS conducted among the same target population (19). Perhaps, the good performance of the instrument may also be attributed to the well chosen optimal cut-off points as suggested by Lee et al (14) and the fact that Malays have the same tendency to somatize their depressive symptoms as the other Asian women (27)

Finally factor analytic examination of our data suggested that a three factor model of the symptom presentation provided the most adequate account of the BDI-II symptom data. The three identifiable subgroups were the affective, cognitive and somatic dimensions. Confirmation of the reliabilities of the subscales was provided by their high internal consistencies. Although in the original validation by Beck and Steer (15) the somatic and affective components merged into one dimension and the cognitive component remained as a separate dimension, the same three factor model has been demonstrated in more recent studies of the same nature (28).

Conclusion

Our findings indicated that this version of the BDI-II is psychometrically strong and appropriate for use in assessing depressive symptomatology among Malay postpartum women in Kedah, North West of Peninsular Malaysia.

Correspondence :

Dr. Wan Mohd. Rushidi Wan Mahmud,
213, Kilometer 7, Jalan Kuala Kedah,
06600 Alor Setar, Kedah Darul Aman.
rushidi@kb.usm.my; shakinah@tm.net.my

References

1. O'Hara M, Swain AM. Rates and risk factors of postpartum depression: A meta-analysis. *International Review of Psychiatry* 1996; **8**: 37 – 54.

2. Beck CT. Teetering the edge: A substantive theory of postpartum depression. *Nursing Research* 1993; **41**: 272-282.
3. Holden J. The role of health visitors in postnatal depression. *International Review of Psychiatry* 1996; **8**: 79 – 86.
4. Beck CT, Gable RK. Comparative analyses of the performance of the Postpartum Depression Screening Scale with two other depression instruments. *Nursing Research*. 1999; **50**: 242-250.
5. Cox JL, Holden JM, and Sagovsky R. Detection of postpartum depression - development of the 10-item Edinburgh Postnatal Depression Scale. *British Journal of Psychiatry* 1987; **50**: 782 – 786.
6. Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J. An inventory for measuring depression. *Archives of General Psychiatry* 1961; **4**: 561 – 571.
7. Radloff L. The CESD scale: A self-report depression scale for research in general population. *Applied Psychological Measurement* 1977; **1**: 385 – 401.
8. Beck CT, Gable RK. Postpartum Depression Screening Scale: Development and psychometric testing. *Nursing Research* 2000; **49 (5)**: 272 – 282.
9. Beck AT, Steer RA. *Beck Depression Inventory*. The Psychological Cooperation, San Antonio, TX, 1993.
10. Beck AT, Rush AJ, Shaw BF, Emery G. *Cognitive therapy for depression*. New York: Guildford, 1979.
11. Beck AT, Steer RA, Brown BK. Psychometric properties of Beck Depression Inventory: Twenty five years of evaluation. *Clinical Psychology Review* 1988; **8**: 77 – 100.
12. Harris B, Huckle P, Thomas R, Johns S, Fung H. The use of rating scales to identify post-natal depression. *British Journal of Psychiatry* 1989; **154**: 813 – 817.
13. Whiffen VE. Screening for postpartum depression: A methodological note. *Journal of Clinical Psychology* 1988; **44**: 367 – 371.
14. Lee DTS, Yip ASK, Chiu HFK, Leung TYS, Chung TKH. Screening for postnatal depression: Are specific instruments mandatory? *Journal Affective Disorder* 2001; **63**: 233 – 238.
15. Beck AT, Steer RA, Brown GK. *Manual for Beck Depression Inventory-II*. The Psychological Corporation, San Antonio TX, 1996.
16. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders (Fourth Edition)*. American Psychiatric Association, Washington. 1994.
17. Brislin RW. Back-translation for cross-cultural research. *Journal Cross Cultural Psychology* 1970; **1**: 185 – 216. .
18. Flaherty JA, Graviria M, Pathak D, Mitchell T, Wintrob R, Richman JA, Birz. Developing instruments for cross cultural research. *Journal Nervous Mental Disorder* 1988; **176(5)**: 257 – 263.
19. Rushidi WMWM, Amir A, Mahmood NM. Revalidation study of the Edinburgh Postnatal Depression Scale (EPDS) among Malay postpartum women attending the Bakar Bata Health Centre in Alor Setar, Kedah, North West of Peninsular Malaysia. *Malaysian Journal of Medical Sciences* 2003; **10 (2)**: 71 – 75.
20. Goldberg DP, Cooper B, Eastwood MR, Kedward HB, Shephard MA. Standardized psychiatric interview for use in community surveys. *British Journal of Preventive Social Medicine* 1970; **24**: 18 – 23.
21. Hamilton M. A rating scale for depression. *Journal Neurology Neurosurgery Psychiatry* 1960; **23**: 56 – 62.
22. *The International Classification of Disease (ICD-10) Classification of Mental and Behavioural Disorders. Clinical Descriptions and Diagnostic Guidelines*. World Health Organization, Geneva, 1992.
23. SPSS Inc. *Statistical Package for Social Science (SPSS) 10.0 for Windows*. Chicago. United States, 1999.
24. Kaiser HF. An index of factorial simplicity. *Psychometrika* 1974; **35**: 31 - 36.
25. O'Hara M.W., Neunaber D.J., Zekoski E.M. Prospective study of postpartum depression: Prevalence, course and predictive factors. *Journal of Abnormal Psychology* 1984; **93**: 158 – 171.
26. Huffman LC, Lamour M, Bryan YE and Pederson FA. Depressive symptomatology during pregnancy and the postpartum period: Is the Beck Depression Inventory applicable? *Journal Reproductive Infant Psychology* 1990; **8**: 87 – 89.
27. Razali SM. Somatic presentations of depressed Malay patients: A comparison with non-somatic presentations. *Malaysian Journal of Psychiatry* 1994; **2**: 47 – 59.
28. Buckley TC, Parker JD, Heggie BS. A psychometric evaluation of the BDI-II in treatment seeking substance abusers. *Journal of Substance Abuse Treatment* 2001; **20**: 197 – 204.