

CHAPTER 9

CONCLUSION, LIMITATIONS, AND RECOMMENDATIONS FOR FUTURE RESEARCH

Postpartum depression is a highly prevalent complication of childbirth that often goes unrecognised. It has an impact on the health of the mother, her baby, and on other members of her family. The main objective of this study was to develop, and psychometrically evaluate, the properties of the Afrikaans version of the PDSS. The data generated from this study suggests that the Afrikaans PDSS is an effective screening measure that health practitioners in South Africa can use to identify mothers with PPD.

The Afrikaans PDSS demonstrates good psychometric properties in this study when compared to the English PDSS. Reliability indices for both the PDSS and the Afrikaans PDSS were excellent. Rasch analysis confirmed the presence of subdimensions (known as content scales) in the PDSS, which represents a multidimensional construct of PPD. Examination of fit indices for the English PDSS total screening scale reveal that all the items except the items from the SLP content scale had acceptable fit indices. A similar trend was noted in the analysis of the Afrikaans PDSS where the majority of items that had fit indices beyond the acceptable range of 1.40 were from the SLP content scale. This seems to indicate that these items clearly form a separate construct. These items did not present with fit problems within the SLP content scale.

Results from confirmatory factor analysis and IRT techniques during the development and psychometric testing of the English PDSS by Beck and Gable (2000) demonstrated the presence of seven dimensions and provided support for the construct validity of the PDSS. Beck and Gable (2002) also investigated how well each PDSS item loads on all extracted factors using exploratory factor analysis with their diagnostic sample. The analysis yielded 7 factors with eigenvalues greater than 1.00. Sleep-specific items from the SLP content scale and appetite-specific items from the SLP content scale loaded on separate factors. Item analysis and internal consistency estimates in the diagnostic sample reveal, however, that the reliability of the SLP content scale (.85) was slightly higher when all five items were included in the content scale than if either of the sleep- or appetite-specific items were removed.

Construct validity of the PDSS and Afrikaans PDSS content scales in this study is confirmed by Rasch analysis and Pearson's correlation. Correlation coefficients of the PDSS and Afrikaans PDSS content scales indicate that items correlate higher with the factor it purports to measure. This provides support for the presence of different subscales. Findings from Rasch analysis of fit indices in the PDSS and Afrikaans PDSS content scales, with item fit indices that demonstrate good item performance on the separate dimensions, confirm the unidimensionality of the content scales and provide support for their construct validity. This suggests that both language versions accurately capture and quantify the multidimensional construct of PPD in English and Afrikaans-speaking South African postpartum women.

This discussion therefore focuses on the unidimensionality of the seven content scales. When Rasch analysis was performed with each respective content scale, item fit

MNSQ statistics supported the measurement of a unidimensional construct in each content scale with the exception of two items: Item 28 (I felt that my baby would be better off without me; Ek het gevoel dat dit vir my baba beter sou wees sonder my) in both the English PDSS and the Afrikaans PDSS SUI content scale, and Item 30 (Ek het gevoel asof ek heeltyd aan die gang moes bly) in the Afrikaans PDSS ANX content scale.

Poor fit of Item 28 may be an indication that it was consistently misunderstood by both English and Afrikaans respondents, but considering that the item demonstrated poor fit in both languages, it is more likely that it did not fit the construct of the SUI content scale very well.

Item 30 had fit problems in the total Afrikaans PDSS, borderline fit problems in the total English PDSS, and also demonstrated poor fit in the Afrikaans ANX content scale. In addition to fit problems, Item 30 was also flagged as an item that was found difficult to understand by both English and Afrikaans participants, and presented with DIF in the ANX content scale. This may indicate that the Afrikaans translation was not adequate, that Afrikaans respondents were not familiar with the item content, or that the item's content is not appropriate for this Afrikaans sample.

In the development of the English PDSS (Beck & Gable, 2000), using data from the diagnostic sample, the deletion of a particular item did not increase the reliability of the respective content scales, with only one exception – Item 28. Deleting this item increased the reliability of the content scale (.90). Item analysis and internal consistency estimates revealed that Item 28 had the lowest correlation (.39) of the five items that comprise the SUI content scale, but the reliability of the scale still remained sufficiently high when the

item was not deleted ($\alpha = .86$). Clinical experts judged the content of Item 28 to be a good fit with the operational definition of the SUI content scale (Beck & Gable, 2001b). Item 30 had moderate correlations with the diagnostic sample (.55) and the reliability of the ANX content scale (.80) did not improve if this item was deleted (.78 if Item 30 was deleted). Results from confirmatory factor analysis of the PDSS using the development sample (Beck & Gable, 2000) revealed that both Item 28 and Item 30 (along with all other PDSS items) had a good fit with the hypothesized model (.75 and .66 respectively).

Item 2 (I got anxious over even the littlest things that concerned my baby; Die geringste dingetjie wat met my baba te doen het, het my angstig gemaak) had large DIF in the ANX content scale, borderline fit problems in the English PDSS ANX content scale, and was flagged by seven participants (five English participants and two Afrikaans participants) as an item that was difficult to understand. This suggests that the English version of this item was not well understood by the English participants of this sample. In future studies, it may be explored whether the term ‘anxious’ is appropriate for the broader English-speaking population in South Africa, particularly because some South Africans refer to the term ‘nerves’ to describe anxiety. Therefore, Item 2 may not be considered appropriate for all English-speaking South Africans.

Although no fit problems were evident for Item 25 (I had a difficult time making even a simple decision; Ek het dit moeilik gevind om die eenvoudigste besluit te neem), this item did present with large DIF in the total PDSS and less DIF in the PDSS MNT content scale. The DIF in the content scale was relatively small ($z = 2.00$) and could be due to measurement error or sample idiosyncrasies. Item 25 was not marked as an item that was

difficult to fully understand. The performance of this item may, however, need to be monitored in future studies.

In summary, both the English PDSS and the Afrikaans PDSS performed fairly well in these English and Afrikaans South African samples. The items that were identified as most problematic were Item 2, Item 25, Item 28, and Item 30. Other items that may require minor revision of their Afrikaans translation, although they did not present with fit or DIF problems, are Item 9, Item 23, and Item 31. Furthermore, together with the problematic items listed above, the following items need to be monitored in future studies: Item 24, Item 25, Item 32, and Item 34.

Analysis of the 5-point Likert response categories provided evidence for meaningful score interpretations. The rating scale analysis showed that responses to the different categories are separated sufficiently so as not to warrant combining some of the response categories. The response data therefore defines the dimension well, with higher responses corresponding to higher levels of agreement with the construct being measured.

A limitation of this study is that the reading level of the Afrikaans participants was not established prior to completing the Afrikaans PDSS. All participants were asked, however, to indicate whether there were items that they did not understand. This was done in order to account for comprehensibility of the PDSS in English participants and of the Afrikaans PDSS in Afrikaans participants. During the screening process some items were identified that a number of mothers found difficult to fully understand. Both the Afrikaans and the original English PDSS should therefore be used with caution in South African women who are not proficient in either of these languages. Future studies may consider

using larger samples where the literacy or education level of respondents can function as strata which can be compared for item location shift.

The translation methodology employed in this study was chosen to help ensure adequate semantic translation of the PDSS into Afrikaans. Analysis revealed, however, that DIF was present in a number of items. Some significant differences were evident between the English and Afrikaans samples, which may account for some of the DIF found. It is also possible that DIF could be attributed to the translation, possible cultural differences in the verbal expression of emotional symptoms, or differences in the manifestation of depression symptoms.

Allalouf, Hambleton, and Sireci (as cited in Allalouf, 2003, p. 56) found different causes for DIF in different item types like analogies, sentence completion, and reading comprehension in verbal reasoning items. Cultural relevance was found to contribute to DIF in reading comprehension items. In order to eliminate the possibility that DIF is due to cultural relevance, a complementary study using only English and Afrikaans speaking participants from the same cultural group should be conducted. This may be an area of future research.

Van de Vijver and Hambleton (as cited in Allalouf, 2003, p. 56) recommend that a panel of psychologists with knowledge about potential causes of DIF and linguistic experts be involved in the translation and revision process. Future studies may attempt to improve the Afrikaans translation of the PDSS by changing the wording of the items with DIF in the target language and then retesting for DIF to determine if a different translation resulted in non-DIF or lower DIF items. Synonym questionnaires may also help to determine whether

some words that appeared in items with DIF were well understood, such as anxious, failure, irritable, angstig, misluk, konsentreer, and geïrriteerd.

The Inconsistent Responding Index (INC) of the PDSS is a useful basic measure of a respondent's consistency in completing the PDSS, providing an indication of response validity. The INC score is derived from ten pairs of PDSS items for which ratings are typically very similar. The INC score is a count of the amount of item pairs for which there is a discrepancy of more than one point in the respondent's rating. When there is an INC score of four (in other words, four pairs of items showed inconsistent responding), then there is an 85% likelihood that the PDSS items were not completed in a way that consistently reflected the screening scale's content. An INC score of five results in a 94% likelihood, while an INC score of six results in a 97% likelihood that the PDSS items were not completed consistently. Beck and Gable (2000) recommend that the examiner regard an INC score of four or more as an indication that the respondent did not complete the PDSS consistently and that the PDSS cannot be interpreted accurately. Inconsistent responding may be due to a respondent misreading an item, marking a response other than the one intended, an inability to maintain sufficient focus during the time taken to complete the PDSS, or it may reflect a misunderstanding of the test instructions or of the item content. Consulting the respondent about the discrepancy in scores may provide some additional information about the comprehensibility of the items. A total of 16 participants (seven Afrikaans and nine English) had INC scores of 4 or more. The researcher did not follow up with these participants to determine the reason for the discrepancy in their responses. These participants were also not eliminated from the sample. In this study, items that participants marked as difficult to understand did not contribute greatly to high INC scores.

A useful screening questionnaire should be able to correctly identify, with consistent reliability, the illness it purports to measure. The screening questionnaire should also be able to correctly identify those persons who do not have the illness. In order to assess a screening questionnaire's accuracy (such as measures of sensitivity, specificity, and positive predictive value), it is necessary to know whether the respondent has been diagnosed with the illness or not. This may be achieved by means of comparison to a reference standard, such as a diagnostic interview. At the time of data collection, this study was not funded. Due to limited resources a decision was made not to use the DSM-IV diagnostic interview as the gold standard to determine how many women would have been diagnosed with major PPD, minor PPD, and with no depression. Obtaining a sample of sufficient size from across South Africa who would be willing to participate through detailed psychiatric interviews in the early postpartum period was thought to be an insurmountable effort for one person to undertake. The researcher rather sought to determine how well the PDSS correlated with two other brief self-report screening measures which were easily administered, and could be completed online.

Not using a diagnostic interview, like the Structured Clinical Interview for DSM (SCID), was a major limitation of this study which resulted in verification bias. Firstly because no comparison could be made between the scores of the three screening measures used in this study with the diagnostic classification of membership to the non-depressed, or to the minor or major depressed groups to determine the accuracy of the measures in an English and Afrikaans South African population. And secondly, it was not possible to determine whether some participants' scores on the screening scales were influenced by comorbidity of other disorders, like eating disorders for example, which, according to

Nishizono-Maher et al (as cited in Nishizono-Maher et al., 2004, p.189) is not uncommon in persons with very high scores on the EPDS. Verification bias often occurs in clinical research, frequently as a consequence of either resource limitations or ethical dilemmas, or both (Hanusa, Scholle, Haskett, Spadaro, & Wisner, 2008). It is recommended that future studies establish the screening accuracy of the PDSS and the Afrikaans PDSS in South African postpartum women by comparing the screening outcomes to the SCID.

The three instruments that were selected for use in this study were considered suitable instruments for identifying major PPD. Beck and Gable (2001a; 2001c) report that the ability of the PDSS to detect PPD is comparable to the ability of the Structured Clinical Interview for DSM-IV Axis 1 Disorders (SCID) to detect major depressive disorder. The PDSS, while still a relatively new measure, has also been translated into a number of languages in recent years. At present English and Spanish versions are readily available. It has also been used in different socioeconomic and ethnic groups. The PDSS is easily administered and, although not as brief as the EPDS, can be administered in five to ten minutes. The PDSS Short Form comprises the first seven items of the complete 35 item PDSS. This version may be completed in as little as one to two minutes and yields a total score which is comparable to the full PDSS total score as it provides an index of the general severity of the mother's PPD symptoms. The Short Form does not allow scoring of the PDSS dimensions (symptom content scales) but gives an indication whether the mother is in need of formal psychiatric evaluation. An advantage that the 35 item PDSS offers over the seven item PDSS Short Form, the EPDS, and the QIDS-SR16 is that it allows investigators and clinicians to identify the dimensions of PPD in which mothers have elevated symptoms, like emotional lability, anxiety and insecurity, and loss of self. This is

particularly helpful to clinicians and therapists for determining specific areas that need to be addressed and for selecting the more suitable treatment options. The PDSS is copyrighted and must be purchased for use in clinical as well as research settings.

The QIDS-SR16 is a relatively new, but increasingly used, brief 16-item self-report measure (Rush et al., 2003; Trivedi, Rush, Ibrahim, et al., 2004). The QIDS-SR16 has demonstrated favourable psychometric profiles across paediatric, adult, and elderly populations with a major depressive disorder. The QIDS is regarded as an effective screening measure for depression in a variety of settings as it is available in both clinician and self-rated versions, it is not time-consuming, and it is administered easily (Rush et al., 2005; Rush et al., 2006; Trivedi, Rush, Crismon, et al., 2004). Although it has not been used extensively in PPD studies, Bernstein et al. (2008) regard the QIDS-SR16 a useful measure in the assessment of PPD as it screens for symptoms of agitation and restlessness as well as decision-making and concentration ability which they found to be symptomatically different between postpartum depression and other depression samples. The IDS, which is a longer format and comprises all 16 items of the QIDS, has demonstrated excellent sensitivity, good specificity and moderate PPV in English and Spanish speaking postpartum women (Yonkers et al., 2001).

The EPDS has been used extensively in postpartum studies across the world, been translated and validated in other cultures, and has been found to be a reliable and valid measure for the detection of PPD (e.g. Navarro et al., 2007). The EPDS has been found to be a valid screening instrument for PPD in low income, socially disadvantaged urban South African women (Lawrie et al., 1998) and an isiXhosa version of the EPDS has shown reliable scores in isiXhosa-speaking postpartum women in South Africa (De Bruin et al.,

2004). Hanusa et al (2008) report that the advantages of using the EPDS is its brevity, ease of administration, it can be used free of charge by investigators, it has many translated versions and has been used with several different socioeconomic and ethnic groups.

Results from this study of continuous screening scores for the PDSS, the EPDS, and the QIDS-SR16 indicated that all three measures were highly correlated and effective in their ability to place women in order of increasing risk for a major depressive disorder. The QIDS-SR16 was translated into Afrikaans for the purpose of this study, but the effectiveness of the Afrikaans translation was not established. Correlation statistics with the QIDS-SR16 must therefore be interpreted cautiously.

This study had several limitations that affect its generalizability. The sample was relatively homogenous (85% white, 89% married, 67% with tertiary education) and the screening measures were not administered randomly. The effect of all the participants completing the questionnaires in the same order may have impacted on internal validity.

All postpartum women who met the research criteria were encouraged to participate in the study. Contact was also established with antenatal women at antenatal classes and if they were willing to participate in the study then they were contacted after their babies were born. There were, however, many participants that were referred from a variety of health practitioners, some of whom suspected the presence of PPD. Many mothers were also eager to participate and be screened for symptoms of PPD after reading articles the researcher wrote about PPD in popular baby magazines. The means used to recruit mothers for participation was another limitation of this study. The participants were asked to volunteer and were not selected randomly which may have resulted in selection bias and accounted

for the higher prevalence rate (47.9%) of major PPD as assessed by the PDSS. This prevalence rate can therefore not be generalised across a similar South African sample.

It was not determined whether there were significant differences in the psychometrics of the PDSS when participants completed the questionnaires online compared to those participants who completed the questionnaires in the presence of the researcher using pen-paper administration. This may be investigated in a future study.

Baker and Oswalt (2008) screened women from a rural community for PPD and found a significant relationship between race, ethnicity and depression rates. The homogeneity of the sample in this study, however, did not make it possible to determine the influence that these particular maternal demographic characteristics may have had on the rate of PPD. Future studies with concentrated South African ethnic groups may provide a clearer indication of the factors associated with PPD within particular groups in the culturally diverse South African population.

The risk factors for PPD in this South African sample are generally consistent with those reported in other studies. Multiple regression analysis revealed eleven risk factors that were significantly associated with screening positive for major PPD using the PDSS. These were a history of psychiatric illness – depression in particular, antenatal depression in recent pregnancy, postpartum blues, lack of support from the baby's father, lack of support from friends, life stress, infant temperament, difficulty conceiving, feeling negative or ambivalent about expecting this baby, fearful of childbirth, and concern about health related issues regarding the infant, like colic, sleeping and feeding problems, and allergies. When the life stress variables were correlated with the participants' scores on the PDSS, the

following variables were found to have a strong relationship to a positive screen for major PPD: moving house, partner changing jobs, partner's job loss, financial concerns, marital problems, family problems, and to a lesser degree, getting married within the past two years.

Mothers with a previous diagnosis of PPD and those who expressed dissatisfaction with the care received during labour and delivery were slightly more likely to have a positive score of major PPD. These variable were, however, not statistically significant when multiple regression analysis was employed.

Health practitioners who have contact with antenatal as well as postpartum women should be made aware of the risk factors for PPD as well as the symptoms of PPD in an effort to care for the mother's well-being in a comprehensive manner as opposed to focussing only on obstetric factors. Health practitioners should be particularly vigilant about those women who have an existing depression or a history of depression. Antenatal women who become depressed during their pregnancies, or are negative or ambivalent about becoming pregnant, or who seem especially anxious about the delivery of their babies should be closely monitored by their health practitioners for symptoms of PPD after delivery. Women need to be educated about postpartum blues and encouraged to seek advice from their health practitioner if their symptoms of postpartum blues persist for longer than two weeks. New mothers should be asked whether they are feeling isolated since the birth of their baby, and whether their partners provide them with adequate support. It is not uncommon for a new mother to complain that her infant sleeps poorly, or that she struggles with feeding her infant, or that her infant is difficult to settle, or suffers from colic or other health-related concerns. These concerns should not be taken lightly.

These mothers may need emotional support as well as support with their infants presenting problems.

The significant predictors of PPD that were identified in this study may serve as red flags to health practitioners that a woman is at risk of developing PPD. Health practitioners should monitor these women carefully so that specific interventions can be implemented before PPD takes its toll on the mother, her infant, and her family.

Very few obstetricians, family practitioners and midwives in South Africa screen postpartum women routinely for symptoms of PPD. If routine screening is not practiced an effort should be made for the routine distribution of information to pregnant and postpartum women about PPD, what the symptoms are, where they may be screened for symptoms of PPD, and where they may obtain treatment. Creating awareness about PPD may potentially encourage women to discuss their symptoms with their health practitioners.

In conclusion, it is evident that PPD is an illness which touches the lives of many South African women. It is likely that many of these women would suffer in silence if their symptoms were not recognised. Screening for symptoms of PPD by using a screening scale like the PDSS or the EPDS will assist in improving the PPD diagnostic rate. It is imperative though, that when screening is done for PPD, that efficient systems are in place to provide treatment and follow-up for mothers with positive results.