

Validation studies of the Edinburgh Postnatal Depression Scale for the antenatal period

Zoltan Kozinszky M.D., Ph.D.¹, Robert B. Dudas M.D., Ph.D., MRCPsych^{2,3}

Reproductive Medicine, Department of Obstetrics and Gynaecology, Karolinska University
Hospital Huddinge, Stockholm, Sweden¹

Department of Psychiatry, University of Cambridge, Addenbrooke's Hospital, Cambridge,
United Kingdom²

Cambridgeshire and Peterborough NHS Foundation Trust, United Kingdom³

Running title: EPDS validation studies in pregnancy

Corresponding author: Zoltan Kozinszky

Reproduktionsmedicin Huddinge, Novum Plan 4, 14186 Huddinge, Sweden

Telefon: +46730720835

e-mail: kozinszky@gmail.com

Abstract

Background: Relatively few studies have focused on the validation of psychometric scales measuring depression during pregnancy. The aim of this review was to critically appraise and review antenatal validation studies of the Edinburgh Postnatal Depression Scale (EPDS).

Methods: A systematic search was performed in MEDLINE, EMBASE, ISI, CINAHL, SCIELO and PsycINFO for the period 1987–2013.

Results: Eleven validation studies met the inclusion criteria. The study design varied between studies. Sensitivity and specificity estimates also varied between 64–100% and 73–100%, respectively. The confidence interval estimates also showed a high degree of variability. Our estimates suggest lower positive predictive values in the general population than those reported in the validation study samples. The sensitivity values in validation studies of the EPDS show fairly large variability, ranging from good to acceptable.

Limitations

Future studies should have larger sample sizes and include both representative and clinical samples and look at the psychometric performance of the EPDS in each trimester.

Conclusions: Due to differences in study design and variation in the cultural/linguistic adaptation, uncertainty remains regarding the comparability of the sensitivity and specificity estimates of different EPDS versions. Future studies should have larger sample sizes, include both representative and clinical samples, and look at the psychometric performance of the EPDS in each trimester. Reporting quality, especially as regards checks to ensure content validity, should be improved.

Highlights

- There is a wide variation across different settings in screening properties in validation studies.

- Cultural variations are likely to account for the majority of this variation.
- Each local version of the EPDS should be validated in its own culture.

Keywords: Edinburgh Postnatal Depression Scale, antepartum, validation, review

Introduction

Antenatal major depressive disorder is a major health problem with a prevalence ranging from 3.1–4.9%. Taken together with the minor forms of the illness the point prevalence rises to 8.5–11.0% during pregnancy (Gaynes et al., 2005, Kuijpers et al., 2001). It is therefore essential that we have a tool to detect this illness effectively in those at risk. The Edinburgh Postnatal Depression Scale (EPDS), a 10-item self-rating scale, was developed by Cox et al. in 1987, originally to detect postnatal depression (Cox et al., 1987). The EPDS was constructed from the Snaith's Irritability, Depression and Anxiety Scale (Snaith et al., 1978) and the Hospital Anxiety and Depression Scale (Zigmond and Snaith, 1983) in addition to items formulated by the constructors (Cox et al., 1987).

Antenatal depression (AND) often remains unnoticed by health professionals, although it could be detected using self-report scales which are economical and do not require extensive rater training. The EPDS is the most widely used screening instrument *perinatally*, and it has also been validated in a non-gestational context showing good psychometric properties amongst men (Matthey et al., 2001), infertile women (Peterson et al., 2006), in the perimenopausal period (Becht et al., 2001), and in women who were not pregnant, apart from its original context, *post partum* (Cox et al., 1996). Without trying to be exhaustive, the question as to whether the EPDS is a uni- or multidimensional scale has also been studied in both perinatal and non-gestational contexts with Pop et al (1992), Töreki et al (2014) and Adouard et al (2005) identifying 2 dimensions, whereas Bergink et al (2011) and de Cock et al (2011) reaching a conclusion of unidimensionality. Before use, a scale devised in another culture, needs to go through cultural validation in order to ensure it is suitable for use in the target population. However, as pointed out by Matthey et al. (Matthey et al., 2006), the EPDS has often been used, without proper validation, during pregnancy for screening out probable

depressive cases with a cut-off defined through a postnatal validation procedure (Kim et al., 2008; Lommatsch et al., 2006; Doornbos et al., 2009; Schultze-Herbrüggen et al., 2007; de Tychey et al., 2005; Otake et al., 2013; Koutra et al., 2013). Furthermore, it seems that different validation sessions are needed for each trimester of the pregnancy (Bergink et al., 2011; Su et al., 2007; Bunevicius et al., 2009a), suggested by different cut-offs found in three validation studies during the course of pregnancy (Bergink et al., 2011; Su et al., 2007; Bunevicius et al., 2009a). Consequently, it could be methodologically incorrect to recruit women for the validation of the EPDS both before and after giving birth (Tran et al., 2011; Hanlon et al., 2008; Werrett and Clifford, 2006; Areias et al., 1996), that is, to validate the scale in a perinatal sample. The EPDS, with its items retained, has been renamed as the Edinburgh Depression Scale (EDS), to take into consideration its new use in the antenatal period.

Recent studies have repeatedly shown that mental health problems during pregnancy (e.g., depression or anxiety) are associated with impaired obstetric and neonatal outcomes, although some studies found no correlation (Choi et al., 2014; Räisänen et al., 2014). Antenatal stress and depression, for example, have been linked to preterm birth and impaired physical health (Dayan et al., 2006; Suri et al., 2007; Rahman et al., 2004), as well as to future emotional, behavioural, and cognitive problems in the child (Talge et al., 2007; Deave et al., 2008; O'Connor et al., 2002). Although we have little direct evidence available so far to prove that the treatment of perinatal depression improves obstetric outcome (Engelstad et al., 2014; Lindqvist et al., 2014) or good quality evidence that depression screening improves depression outcome (Thombs et al., 2014), there is enough evidence to suggest that maternal depression is linked to poor maternal and infant health outcomes that argues in favour of improving routine screening during pregnancy.

Early screening for antenatal depression is recommended by some national professional bodies, such as the UK's National Institute for Health and Clinical Excellence (NICE 2007) ("Antenatal and postnatal mental health: clinical management and service guidance"). Therefore, there is a well-defined and expressed need for an easy-to-administer self-report scale that can be used by healthcare professionals, such as midwives, obstetricians, health visitors, and family doctors.

There is a disagreement around whether pregnancy is a risk or a protective factor for mood disorders (Bergink et al., 2011; Cox, Chapman et al., 1996; Töreki et al., 2013; Töreki et al., 2014). Having a cross-culturally valid screening instrument would be very helpful in settling this question.

Investigators using the EPDS to screen for depression should realize that the instrument does not exclusively measure depression but anxiety too (Brouwers et al., 2001, Matthey et al., 2013). This, however, is more of an added benefit rather than a problem, as anxiety disorders, when discovered, can also be effectively treated.

A large review study (Gibson et al., 2009) has previously been published on the validity of the EPDS in both ante- and postnatal depression, however that review could only include three antenatal studies. Since the publication of the Gibson review (Gibson et al., 2009), four more antenatal studies (Töreki et al., 2014; Bergink et al., 2011; Rubertsson et al., 2011; Stewart et al., 2013; Wang et al., 2009) have been published and the conclusions from these highlight the need for making appropriate recommendations in the light of the newly published material.

Aims of the Study

The primary objective of this study was to identify all published antenatal validation studies on the EPDS and to present their design along with the estimated sensitivity, specificity and positive and negative predictive values. An additional aim was to estimate the 95% confidence intervals of the screening values on the basis of the sample sizes and detected depression cases in each study. In addition, we assessed the association between the prevalence of antenatal depression and the estimates of the positive and negative predictive values of the Edinburgh Postnatal Depression Scale and the interrelation between the true and false positive rates of validation studies.

Methods

Search strategy

Previous validation studies of the EPDS during pregnancy in the English-language literature were identified through the MEDLINE, EMBASE, ISI, CINAHL, SCIELO, Cochrane Library and PsycINFO databases from 1987 onwards, when the EPDS was launched, until December 2013, using the following search terms: 'Edinburgh postnatal depression scale' OR 'EPDS' OR 'scale' OR 'tool' OR 'inventory' OR 'questionnaire' AND 'valid?' OR 'screen' OR 'Diagnos?' OR 'semi?structured' OR 'interview' OR 'Valid' OR 'Sensitivity' OR 'Specificity' OR 'predict?' OR 'value' OR 'PPV' OR 'NPV' OR 'likelihood ratio' OR 'ROC' OR 'AUC' AND 'antepartum' OR 'antenatal' OR 'pregnancy' OR 'prenatal' OR 'diagnosis' in the title, keywords, or anywhere in the abstract. All abstracts were read by two reviewers (Z.K. and R.D.) to identify relevant papers. Extensive additional searching was carried out manually, using the reference lists of published papers.

Data quality

The quality of the selected studies was assessed using a grading system based on the York CRD system (Khan et al., 2001). Each study was assigned a grade on the basis of whether the clinicians were appropriately blinded to the EPDS scores, what reference standards were used for diagnosis and the sampling method used.

Data analysis

We tabulated the extracted data and analyzed the screening properties of EPDS for antenatal depression. With the exception of one study (Bergink et al., 2011), relatively small numbers of women were included in the validation studies identified. Also, the studies aimed at choosing a cut-off score that allowed the highest sensitivity value without compromising too much on specificity (Krzanowski and Hand, 2009). The Wilson score method was used for presenting confidence intervals, because diagnosis is a binary variable and leads to a binomial distribution of the EPDS scores (Wallis, 2013).

The positive and negative predictive values (PPV and NPV) of the EPDS and their confidence intervals were also calculated. Unfortunately, there is no available systematic review providing a robust estimate for the point prevalence of major depressive disorder in each trimester in the literature. Prevalence estimates from a meta-analysis would be more representative of the countries in which the included studies were carried out. Consequently, estimation of PPV/NPV is not possible for major depressive episode. For combined depression (major depressive episode and minor depression), re-estimated PPVs and NPVs, assuming 7.4%, 12.8% and 12.0% prevalences of combined depression in the first, second and third trimester, respectively, based on the review of Bennett et al. (Bennett et al., 2004),

were calculated. According to Eberhard-Gran et al. (Eberhard-Gran et al., 2001), predictive values were also estimated at assumed prevalence rates of antenatal depression at 5, 10, 15, 20 and 25%, respectively. Similar to Gibson et al (Gibson et al., 2009), the results were plotted from studies reporting findings for major and combined depression on a scatterplot of true-positive rate (sensitivity) against false-positive rate (1 - specificity), and calculated relevant summary statistics.

All statistics were calculated using Statistical Package for Social Sciences (SPSS), Version 20 (SPSS Inc., Chicago, IL, USA).

Results

Search results

The search strategy identified 11 EPDS validation studies in the antenatal period (Bunevicius et al., 2009a; Su et al., 2007; Felice et al., 2006; Adewuya et al., 2006; Bergink et al., 2011; Adouard et al., 2005; Murray and Cox, 1990; Wang et al., 2009; Rubertsson et al., 2011; Stewart et al., 2013; Töreki et al., 2014), reporting that the EPDS was a valid screening instrument for depressive disorders during pregnancy in 11 different countries. The reference lists in these publications did not include any other validation studies.

Table 1 presents the identified antenatal validation studies. The optimal cut-off for major depression varied from 5.5 to 14.5 (Stewart et al., 2013; Bunevicius et al., 2009a; Su et al., 2007; Felice et al., 2006; Adewuya et al., 2006; Adouard et al., 2005; Murray and Cox, 1990; Wang et al., 2009; Töreki et al., 2014), whereas for combined depression from 4.5 to 13.5 (Stewart et al., 2013; Rubertsson et al., 2011; Felice et al., 2006; Adewuya et al., 2006; Adouard et al., 2005; Murray and Cox, 1990; Töreki et al., 2014). It is of note that only two

studies have evaluated the performance of the EPDS in each trimester (Bergink et al., 2011; Bunevicius et al., 2009a) and they identified different cut-offs. The Taiwanese version of EPDS was validated in the second and third trimester and, again, found different optimal cut-offs (Su et al., 2007). The Maltese version (Felice et al., 2006) was validated among women taken at any time point during the pregnancy, while other researchers validated the EPDS at the end of the first trimester (Töreki et al., 2014), in the first half of the pregnancy (Rubertsson et al., 2011) or in late pregnancy (third trimester) (Adewuya et al., 2006; Adouard et al., 2005; Murray and Cox, 1990; Wang et al., 2009; Stewart et al., 2013).

Study sample/selection

The size of the study sample varied between 60 and 845 women interviewed. In the validation study by Murray and Cox (Murray and Cox, 1990), available mothers were recruited ad hoc by a health care provider and they completed the EPDS before the interview. In some studies (Töreki et al., 2014; Bunevicius et al., 2009a), unselected clinical subjects were recruited in random order or by random number generation during a daily recruitment process (Felice et al., 2006). However, in the French study targeted to validate the EPDS in high-risk pregnancies, women were recruited ad hoc by a health-care provider directing them to a psychiatrist who carried out the questionnaire survey and the interview (Adouard et al., 2005). The selection process was not adequately described in the Taiwanese study, where a research assistant recruited the participants completing the questionnaire before the interview carried out by the psychiatrist (Su et al., 2007). In other studies (Adewuya AO et al., 2013; Rubertsson et al., 2011), the participants were recruited consecutively, whilst all women with high scores and a random sample of women with low scores from a population-based study were included. Pregnant women were recruited as a medium or large, community-based

sample in only two studies (Wang et al., 2009; Bergink et al., 2011) (Table 1). 6 studies did not report the percentage of women who declined participation, and the remaining 5 reported 0-25 percent.

Diagnostic criteria

In an earlier validation study (Murray & Cox, 1990), the EPDS was validated against the Research Diagnostic Criteria (RDC) set for depression, distinguishing between major, minor and intermittent depressive disorder (Spitzer et al., 1975; Spitzer et al., 1978). In most of the later studies (Bunevicius et al., 2009a; Töreki et al., 2014; Su et al., 2007; Adewuya et al., 2006; Adouard et al., 2005; Wang et al., 2009; Rubertsson et al., 2011; Stewart et al., 2013), the Diagnostic and Statistical Manual of Mental Disorders (DSM)-III (American Psychiatric Association, 1980) or DSM-IV (American Psychiatric Association, 1994) was used as gold standard. International Classification of Disease (ICD-10) diagnosis for severe, moderate and mild depressive episode (World Health Organization, 1992) has also been used as a validation instrument (Felice et al., 2006). Some studies included both major and minor depressive disorders in the definition of antenatal depression (Rubertsson et al., 2011; Stewart et al., 2013; Murray and Cox, 1990; Töreki et al., 2014; Adewuya et al., 2006) or minor, intermittent and major depression (Felice et al., 2006), were used. In four studies, a more strict definition, only including major depression, was applied (Wang et al., 2009; Bunevicius et al., 2009a; Adouard et al., 2005; Su et al., 2007). The EPDS was validated according to DSM-IV (American Psychiatric Association, 1994) and ICD-10 criteria (World Health Organization, 1992) in the Netherlands (Bergink et al., 2011) (Table 1). The overall point prevalence of major and combined depression extracted from the validation studies was 5.23% and 11.37%, respectively (not presented in the Tables).

Diagnostic instruments

In all studies, the validation was performed with a clinical interview. In order to establish the diagnosis of major/minor depression, three studies (Adouard et al., 2005; Adewuya et al., 2006; Su et al., 2007) used the Mini International Neuropsychiatric Interview (Sheehan et al., 1998), one study (Murray and Cox, 1990) Goldberg's Standardized Psychiatric Interview (Goldberg, 1972), one study (Felice et al., 2006) the Clinical Interview Schedule (Lewis et al., 1992), one study (Rubertsson et al., 2011) the Primary Care Evaluation of Mental Disorders (PRIME-MD) (Spitzer et al., 1999) and four studies (Stewart et al., 2013; Wang et al., 2009; Töreki et al., 2014; Bunevicius et al., 2009a) the Structured Clinical Interview for DSM-III/IV (SCID; Spitzer et al., 1990; American Psychiatric Association, 1980; American Psychiatric Association, 1994; First et al., 2002). It is of note that there was only one study (Töreki et al., 2014) that clearly stated the diagnostic criteria used for the diagnosis of both major and minor depression. The EPDS was validated against the short version of the Composite International Diagnostic Interview (Robins et al., 1988), a structured diagnostic interview, in the Dutch study (Bergink et al., 2011). In all studies specially trained investigators (psychiatrists or psychologists) conducted the diagnostic interviews, except in the Bergink et al (Bergink et al., 2011) study, in which two third of the participants were interviewed by midwives and the remaining by psychology students who had been specially trained in the diagnosis of antenatal depression.

Only two studies evaluated the reliability of the EPDS with repeated measurements in all three trimesters of pregnancy (Bergink et al., 2011; Bunevicius et al., 2009a) and another one that validated the EPDS in two trimesters (Su et al., 2007), indicating that the EPDS is a reliable instrument for repeated evaluations of depressive symptoms. The high test–retest

reliability of the EPDS was demonstrated throughout one trimester in two studies (Töreki et al., 2014; Wang et al., 2009).

Quality assessment

Table 2 summarizes the results of our quality assessment. Study quality was generally acceptable, with four of the eleven studies achieving grade A (Adewuya et al., 2006; Bergink et al., 2011; Rubertsson et al., 2011; Stewart et al., 2013) indicating an appropriate study design, used in a representative population, and appropriate blinding. The remaining seven studies achieved grade B (Adouard et al., 2005; Bunevicius et al., 2009a; Murray and Cox, 1990; Su et al., 2007; Töreki et al., 2014; Wang et al., 2009) because of having been conducted in a narrow population sample. One study (Felice et al., 2006) was hard to classify.

Sensitivity and specificity

The estimates of sensitivity of the EPDS between studies varied from 70% to 100% for major depression and from 64 to 87% for combined depression. There was also a large variation in specificity from 74% to 97% and from 73% to 96% for major and combined depression, respectively (Table 1). Studies with higher cut-off score for major depression did not always produce lower sensitivity and higher specificity than studies with a lower cut-off, likely reflecting cultural differences. Studies in which a high EPDS cut-off score for combined depression was used (Murray and Cox, 1990; Felice et al., 2006; Rubertsson et al., 2011) generally showed lower sensitivity and higher specificity than the study using a low cut-off score (Töreki et al., 2014) did. Stewart used weighted scores and very low cut-off values and reported acceptable sensitivity and specificity for both major and combined

depression (Stewart et al., 2013). The choice of an appropriate cut-off level was discussed in all studies. Receiver operating characteristic (ROC) curve analyses were carried out to assess the global performance of the EPDS for detecting depression as a diagnosis, as well as to optimize the cut-off points for the EPDS by means of the trapezoidal rule. Inclusion of minor depressive disorders in the definition of postnatal depression tends to lower both sensitivity and specificity (Table 1).

Confidence intervals

Confidence intervals of the sensitivity and specificity estimates of the EPDS were not presented in any of the validation studies and could not be calculated in one study (Wang et al., 2009). Our estimates of the 95% confidence intervals for the sensitivity and specificity in each study are presented in Table 1. Sensitivity and specificity values were higher and the confidence intervals narrower for major depression compared to combined depression. Confidence intervals were narrower for specificity than for sensitivity values. Studies that had narrower confidence intervals for sensitivity (Adewuya et al., 2006; Bergink et al., 2011; Felice et al., 2006; Stewart et al., 2013) generally had higher quality grading, reflecting a more representative sample.

Positive and negative predictive values

The positive and negative predictive values (PPVs and NPVs) with confidence intervals of the EPDS total scores for the detection of depression for each validation study are presented in Table 3, except for the Chinese validation (Wang et al., 2009) where they could not be calculated. PPVs ranged between 22% and 75%, whereas NPVs from 92% to 100%.

Higher predictive values can be observed at a higher prevalence of antenatal depression. We also calculated estimates of these values based on the sensitivity and specificity values published in a studies on combined depression, assuming prevalences of 7.4%, 12.8% and 12.0% of combined depression in the first, second and third trimester, respectively, based on the meta-analysis by Bennett et al (Bennett et al., 2004). Our estimates, not very different from the values in the original studies, varied from 28% to 69% for PPVs and from 95% to 98% for NPVs. The differences reflected the possible over-representativity (Felice et al., 2006; Murray and Cox, 1990; Töreki et al., 2014) or under-representativity (Adewuya et al., 2006; Rubertsson et al., 2011) of depressed women in the included studies. From all validation studies, an overall specificity and sensitivity could be extracted for major depression and combined depression in each trimester, and the effect of the changing prevalence of major/combined depression in relation to predictive values could be studied (Eberhard-Gran et al., 2001). Figure 1 illustrates the impact of the prevalence of antenatal depression on the positive predictive value of the EPDS. Figure 2 shows the impact of prevalence on negative predictive value. Figure 3 depicts the true-positive rates (sensitivity) against false-positive rates (1 - specificity) for major and combined depression of the studies reflecting that sensitivity is satisfactorily high with regards to specificity.

Discussion

It is not advisable to use universal cut-off scores, as there can be cultural differences and/or differences in how successful the adaptation of the EPDS was in a given culture, resulting in differences in at what total score the best trade-off between sensitivity and specificity can be found.

There is a painful lack of a systematic review on the prevalence of major depressive disorder in each trimester. Trimester-specific EPDS validation studies are urgently needed in each culture instead of using non-validated cut-offs (Matthey et al., 2001; Matthey et al., 2006). The UK NICE guidelines recommend screening, but there are no trimester-specific cut-offs extracted from ROC-analyses carried out in ante partum validation studies (NICE, 2007). Generally, the EPDS versions validated in various countries are good at identifying major depression alone, with the lowest misclassification rates in studies using community samples (Wang et al., 2009; Bergink et al., 2011), but they perform poorer for the combined or minor depression categories. As in the previous reviews (Gibson et al., 2009; Eberhard-Gran et al., 2001), we found that there were significant differences in study design, population sampled, the timing of testing, language version of the EPDS used, and how cases were diagnosed. Overall, the studies included in this current review have provided evidence for the EPDS as a valid screening tool across different cultures, as visible from the satisfactory ROC analyses results (Area under curves (AUCs) were greater than 0.7) (Hanley and McNeil, 1982). However, there was a wide range of sensitivity, specificity and predictive values at all cut-off points. One possible explanation for this could be differences in content validity; unfortunately, few studies provide sufficiently detailed description to assess this. The high sensitivity of the EPDS is clinically desirable for a screening instrument, as it prevents missing a depressive disorder (Flaherty et al., 1998). High specificity combined with a high negative predictive value also fulfils the requirements for a good screening instrument. It is advisable to choose a cut-off at which the misclassification rate is not too high. Lowering the cut-off (increasing sensitivity) may make sense for established referral pathways (e.g. primary care) where the objective is to maximize detection and referring potential cases on to psychiatry.

As with any screening tool, in selecting a cut-off point, there is a trade-off between the accuracy of ruling in or out a diagnosis. At lower cut-offs, the EPDS performs considerably better overall and generally the studies suggest that it provides 'convincing evidence' (combined depression) or 'strong evidence' (major depression) by which to rule out the diagnosis.

Whilst sensitivity and specificity do not depend on prevalence, PPV and NPV do, and they showed significant variability. There are a number of explanations to be considered. Both methodologies and populations varied greatly between the studies. Samples were drawn from countries with different cultural attitudes to distress and genuine differences in prevalences. Also, different diagnostic interviews and criteria were used.

Although not possible to analyze due to the low number of studies, in a review of the postnatal validation studies (Gibson et al., 2009), it was stated that the language of administration of the EPDS contributed to some differences in the heterogeneity of the screening properties (Gibson et al., 2009). .

As the prevalence of the disorder increases, so too does the PPV. Therefore, studies in which women with AND were over-represented, showed higher PPV values. Bennett et al., in their review, have suggested that the prevalence of depression during pregnancy in the first, second and third trimesters are 7.4%, 12.8% and 12.0%, respectively, in the general population (Bennett et al., 2004). If the true prevalence is close to this value, the PPVs presented in some of the validation studies are misleadingly high (Töreki et al., 2014; Murray and Cox, 1990; Felice et al., 2006) or low (Adewuya et al., 2006). This bias is caused by over- or under-representation of depressed women in the samples. This can be partly explained by the fact that the prevalence of depression is higher in clinical (as opposed to community) settings (Eberhard-Gran et al., 2001).

Whilst the included studies scored well on the quality rating scale there were some general weaknesses. The method of selection of study samples was not always clearly described and in some cases was 'ad hoc', which may bias the results towards including more motivated or psychologically vulnerable women, or under-representing women who were depressed during pregnancy. This method of sampling may under-represent women with depression, chiefly in those cultures, where the prevalence of depression is low. One of the studies (Adouard et al., 2005) was conducted on a narrowly defined population, e.g. women who had experienced obstetric complications, and the results are not likely to be generalizable to a standard community setting.

The relatively wide range of values for sensitivity, specificity and predictive values and their confidence intervals across studies make the interpretation of the utility of this screening instrument more complicated. The heterogeneity in the figures may suggest that the accuracy of the EPDS varies depending on the clinical setting, country, and language of administration.

At present, when the EPDS is used in the general population, it will yield a substantial proportion of false positives, which is costly to service providers, because it generates unnecessary further assessment. Also, the EPDS will miss a considerable number of cases, similar to the majority of other screening tools. The advantage of the EPDS rests in its free availability, ease of administration and general acceptability to women, if given sympathetically. Therefore, if the above caveats are observed, it remains a useful tool in the field of perinatal mental health.

In the validation studies discussed, the identified cut-off for combined depression appears to be the most useful for screening purposes. A Norwegian validation study (Berle et al., 2003) recommended the cut-off for minor depression for screening purposes in clinical settings to detect women at risk for developing depression and to identify those who have

minor depression. Identifying women with a psychiatric disorder in pregnancy is of special interest, because they are less likely to attend antenatal clinic appointments as appropriate (Zuckerman et al., 1989), and depression, one of the commonest reasons, is relatively easy and cheap to screen for. A further argument for studying antenatal depression is the significant association of AND with PND (Robertson et al., 2004), which with the treatment of the AND can often be prevented.

Antenatal validation of the EPDS in particular populations in each trimester in every culture is highly recommended before the instrument can be used properly for screening for AND. The message from the antenatal validation studies is clear: antenatal cut-off values are different from the postnatal ones (France: 11.5 antenatal (Adouard et al., 2005;) vs 10.5 in postnatal (Guedeney and Fermanian, 1998), UK: 14.5 antenatal (Murray and Cox, 1990) vs 12.5 postnatal (Cox et al., 1987), Malta: 13.5 antenatal (Felice et al., 2004) vs 11.5 postnatal (Felice et al., 2006), Nigeria: 9.5 antenatal (Adewuya et al., 2006) vs 8.5 postnatal (Uwakwe and Okonkwo, 2003), Hungary: 6.5 antenatal (Töreki et al., 2014) vs 12.5 postnatal (Nagy et al., 2011; or Töreki et al., 2013); Lithuania: 6.5 antenatal (Bunevicius et al., 2009a) vs. 10.5-11.5 postnatal (Bunevicius et al., 2009b); Taiwan: 12.5-13.5 antenatal (Su et al., 2007) vs. 12.5 postnatal (Teng et al., 2005); Mainland China: 9.5 antenatal (Wang et al., 2009) vs. 10.5 postnatal (Lau et al., 2010) and Sweden: 13.5 antenatal (Rubertsson et al., 2011) vs. 11.5 postnatal (Wickberg and Hwang, 1997). In general, the cut-offs are lower in the post- than in the antenatal period. Taking into consideration the trimester in which the validation took place is also important, because the prevalence of depression may fluctuate during the course of pregnancy and, as a consequence, the optimal cut-off of EPDS may also be different.

The difference in the EPDS cut-off scores among the studies may also be partly due to differences in the time of assessment in relation to childbirth (early or late pregnancy, early or late post partum) (Alvarado-Esquivel et al., 2006; Bunevicius et al., 2009a). Su et al. (Su et

al., 2007) stated that they avoided validating the EPDS in the first trimester of the pregnancy due to the possible acute psychological effects of pregnancy; however this was not recognizable in other studies (Bunevicius et al., 2009a; Bergink et al., 2011; Tőreki et al., 2014).

Cultural adaptation is necessary even in countries speaking the same language. E.g. the New-Zealand (Boyce and Todd, 1992) and English version (Cox et al., 1987) or the Spanish (Garcia-Esteve et al., 2003) and Mexican versions (Alvarado-Esquivel et al., 2006) gave different cut-offs in postnatal samples.

There is a trade-off between validating in clinical versus community samples (e.g. Berle et al., 2003 vs. Eberhard-Gran et al., 2001 or Nagy et al., 2011 vs. Tőreki et al., 2013). The second option may provide values more representative to the general population, but it may only be possible to use structured clinical interviews as gold standard for diagnosis (which is more acceptable than self-report symptom measures that are often used in community studies) in smaller, clinical samples.

Higher prevalence of depression in the sample, which can be due to selection bias, can lead to a falsely high PPV, which in turn may influence our decision regarding the recommended cut-off value (Bergink et al., 2011).

Apart from the Hungarian and the Swedish studies (Tőreki et al., 2014; Rubertsson et al., 2011), no other study described how many approached pregnant women declined to participate and whether those who declined were different from participants in terms of sociodemographic characteristics or EPDS scores. Also, they do not report the criteria used for the diagnosis of minor depression. There are discrepancies in what is meant by antenatal depression in terms of criteria (Leight et al., 2010). Establishing international consensus with regard to this would be helpful for future research.

In conclusion, currently, versions of the EPDS validated in an antepartum sample are available in Chinese, Dutch English, French, Hungarian, Lithuanian, Maltese, Nigerian, and Portuguese. It is suggested that the EPDS be validated in every country in each trimester during pregnancy. Also, test-retest reliability indices of the EPDS were not assessed in most of the studies discussed in the current review and the EPDS was only administered on one occasion. However, in the studies that did report figures, short-term (Wang et al., 2009; Töreki et al., 2014) and long-term retest validity was good (Bergink et al., 2011).

Large confidence intervals

When estimating the sensitivity, specificity and predictive values as screening properties of an instrument, narrow confidence intervals would be desirable (Eberhard-Gran et al., 2001). Our estimates reveal variable 95% confidence intervals for the sensitivity, specificity and predictive values in most of the validation studies, which reflect the possibility of the low number of participants, the prevalence of identifiable depression cases and differences in characteristics of the EPDS version applied. It would be advisable to achieve larger samples in future validation studies.

Limitations

This systematic review has confirmed that the screening accuracy of the EPDS in diagnosing depression during pregnancy is satisfactory and that the EPDS can be recommended for use for this purpose. The heterogeneity of included studies prevented progression to meta-analysis and further statistical comparison of the EPDS across settings was not possible.

Declaration of Interests

The authors declare that they have no conflict of interest.

Contributors

Both authors reviewed the literature and drafted the MS. Z.K. performed the statistical analyses in the MS.

References

1. Adouard, F., Glangeaud-Freudenthal, N.M., Golse, B., 2005. Validation of the Edinburgh postnatal depression scale (EPDS) in a sample of women with high-risk pregnancies in France. *Arch. Womens. Ment. Health.* 8:89–95.
2. Adewuya, A.O., Ola, B.A., Dada, A.O., Fasoto, O.O., 2006. Validation of the Edinburgh Postnatal Depression Scale as a screening tool for depression in late pregnancy among Nigerian women. *J. Psychosom. Obstet. Gynaecol.* 27:267–272.
3. Alvarado-Esquivel, C., Sifuentes-Alvarez, A., Salas-Martinez, C., Martínez-García, S., 2006. Validation of the Edinburgh Postpartum Depression Scale in a population of puerperal women in Mexico. *Clin. Pract. Epidemiol. Ment. Health.* 29;2:33.
4. Antenatal and Postnatal Mental Health. Clinical management and service guidance. Clinical Guideline 45. London:National Institute for Clinical Excellence (NICE), 2007.
5. American Psychiatric Association. Diagnostic and statistical manual of mental disorders, 3rd edn. Washington, DC: American Psychiatric Press, 1980.
6. American Psychiatric Association. Diagnostic and statistical manual of mental disorders, 4th ed. Washington, DC: American Psychiatric Association; 1994.
7. Areias, M.E., Kumar, R., Barros, H., Figueiredo, E., 1996. Comparative incidence of depression in women and men, during pregnancy and after childbirth. Validation of the Edinburgh Postnatal Depression Scale in Portuguese mothers. *Brit. J. Psychiatry.* 169:30–35.
8. Becht, M.C., van Erp, C.F., Teeuwisse, T.M., van Heck, G.L., van Son, M.J., Pop, V.J., 2001. Measuring depression in women around menopausal age: Towards a validation of the Edinburgh depression scale. *J. Affect. Disord.* 63:209–213.

9. Beck, A.T., Ward, C.H., Mendelson, M., Mock, J., Erbaugh, J., 1961. An inventory for measuring depression. *Arch. Gen. Psychiatry*. 4:561–571.
10. Bennett, H.A., Einarson, A., Taddio, A., Koren, G., Einarson, T.R., 2004. Prevalence of depression during pregnancy: Systematic review. *Obstet. Gynecol.* 103:698–709.
11. Bergink, V., Kooistra, L., Lambregtse-van den Berg, M.P., Wijnen, H., Bunevicius, R., van Baar, A., Pop, V., 2011. Validation of the Edinburgh Depression Scale during pregnancy. *J. Psychosom. Res.* 70:385–389.
12. Berle, J.Ø., Aarre, T.F., Mykletun, A., Dahl, A.A., 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.* 76:151–156.
13. Boyce, P.M., Todd, A.L., 1992. Increased risk of postnatal depression after emergency caesarean section. *Med. J. Aust.* 157:172–174.
14. Brouwers, E.P., van Baar, A.L., Pop, V.J., 2001. Does the Edinburgh Postnatal Depression Scale measure anxiety? *J. Psychosom. Res.* 51:659–663.
15. Bunevicius, A., Kusminskas, L., Pop, V.J., Pedersen, C.A., Bunevicius, R., 2009a. Screening for antenatal depression with the Edinburgh Depression Scale. *J. Psychosom. Obstet. Gynaecol.* 30:238–243.
16. Bunevicius, A., Kusminskas, L., Bunevicius, R., 2009b. Validation of the Lithuanian version of the Edinburgh Postnatal Depression Scale. *Medicina (Kaunas)*, 45:544–548.
17. Choi, S.K., Park, Y.G., Park, I.Y., Ko, H.S., Shin, J.C., 2014. Impact of antenatal depression on perinatal outcomes and postpartum depression in Korean women. *J. Res. Med. Sci.* 19:807–812.

18. Cox, J.L., Holden, J.M., Sagovsky, R., 1987. Detection of postnatal depression. Development of the 10-item Edinburgh Postnatal Depression Scale. *Brit. J. Psychiatry.* 150:782–786.
19. Cox, J.L., Chapman, G., Murray, D., Jones, P., 1996. Validation of the Edinburgh Postnatal Depression Scale (EPDS) in non-postnatal women. *J. Affect. Disord.* 39:185–189.
20. Dayan, J., Creveuil, C., Marks, M.N., Conroy, S., Herlicoviez, M., Dreyfus, M., Tordjman, S., 2006. Prenatal depression, prenatal anxiety, and spontaneous preterm birth: a prospective cohort study among women with early and regular care. *Psychosom. Med.* 68:938–946.
21. Deave, T., Heron, J., Evans, J., Emond, A., 2008. The impact of maternal depression in pregnancy on early child development. *BJOG.* 115:1043–1051.
22. De Cock E.S., Emons W.H., Nefs G., Pop V.J., Pouwer F., 2011. Dimensionality and scale properties of the Edinburgh Depression Scale (EDS) in patients with type 2 diabetes mellitus: the DiaDDzoB study. *BMC Psychiatry.* 11:141.
23. Doornbos, B., Dijck-Brouwer, D.A., Kema, I.P., Tanke, M.A., van Goor, S.A., Muskiet, F.A., Korf, J., 2009. The development of peripartum depressive symptoms is associated with gene polymorphisms of MAOA, 5-HTT and COMT. *Prog. Neuropsychopharmacol. Biol. Psychiatry.* 33:1250–1254.
24. de Tyche, C., Spitz, E., Briançon, S., Lighezzolo, J., Girvan, F., Rosati, A., Thockler, A., Vincent, S., 2005. Pre- and postnatal depression and coping: a comparative approach. *J. Affect. Disord.* 85:323–326.
25. Eberhard-Gran, M., Eskild, A., Tambs, K., Opjordsmoen, S., Samuelsen, S.O., 2001. Review of validation studies of the Edinburgh Postnatal Depression Scale. *Acta. Psych. Scand.* 104:243–239.

26. Engelstad, H.J., Roghair, R.D., Calarge, C.A., Colaizy, T.T., Stuart, S., Haskell, S.E., 2014. Perinatal outcomes of pregnancies complicated by maternal depression with or without selective serotonin reuptake inhibitor therapy. *Neonatology*. 105:149–154.
27. Felice, E., Saliba, J., Grech, V., Cox, J., 2006. Validation of the Maltese version of the Edinburgh Postnatal Depression Scale. *Arch. Womens. Ment. Health*. 9:75–80.
28. First, M.B., Spitzer, R.L., Gibbon, M., Williams, J.B.W., 2002. Structured clinical interview for DSM-IV-TR Axis I Disorders, research version. In: Biometrics Research, Patient Edition. New York State Psychiatric Institute, New York.
29. Flaherty, J.A., Gaviria, F.M., Pathak, D., Mitchell, T., Wintrob, R., Richman, J.A., 1998. Developing instruments for cross-cultural psychiatry research. *J. Nerv. Ment. Dis.* 176,257–263.
30. 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.
31. Gaynes, B., Gavin, N., Meltzer-Brody, S., Lohr, K.N., Swinson, T., Gartlehner, G., Brody, S., Miller, W.C., 2005. Perinatal depression: prevalence, screening accuracy, and screening outcomes. Evidence Report/Technology Assessment No. 119. (Prepared by the RTI-University of North Carolina Evidence-based Practice Center, under Contract No. 290-02-0016.) AHRQ Publication No. 05-E006-2. Rockville, MD: Agency for Healthcare Research and Quality.
32. Gibson, J, McKenzie-McHarg, K., Shakespeare, J., Price, J., & Gray, R. (2009) A systematic review of studies validating the Edinburgh Postnatal Depression Scale in antepartum and postpartum women. *Acta. Psych. Scand.* 119:350–364.
33. Goldberg, D. The detection of psychiatric illness by questionnaire. Maudsley monograph 21. Oxford: Oxford University Press, 1972.

34. 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.
35. Hanley, J.A. McNeil, B.J., 1982. The Meaning and Use of the Area under a Receiver Operating Characteristic (ROC) Curve. *Radiology*. 143:29–36.
36. Hanlon, C., Medhin, G., Alem, A., Araya, M., Abdulahi, A., Hughes, M., Tesfaye, M., Wondimagegn, D., Patel, V., Prince, M., 2008. Detecting perinatal common mental disorders in Ethiopia: validation of the self-reporting questionnaire and Edinburgh Postnatal Depression Scale. *J. Affect. Disord.* 108:251–262.
37. Khan, K.S., ter Riet, G., Popay, J., Nixon, J., Kleijnen, J. Study quality assessment. In Khan, K.S., ter Riet, G., Glanville, J., Snowden, A.J., Kleijnen, J., eds. *Undertaking systematic reviews of research on effectiveness: CRD's guidance for carrying out or commissioning reviews*. 2nd ed. CRD Report 4. University of York. 2001. www.york.ac.uk/inst/crd/crdreports.htm.
38. Kim, J.J., Gordon, T.E., La Porte, L.M., Adams, M., Kuendig, J.M., Silver, R.K., 2008. The utility of maternal depression screening in the third trimester. *Am. J. Obstet. Gyn.* 199:509.e1-5.
39. Koutra, K., Chatzi, L., Bagkeris, M., Vassilaki, M., Bitsios, P., & Kogevinas, M., 2013. Antenatal and postnatal maternal mental health as determinants of infant neurodevelopment at 18 months of age in a mother-child cohort (Rhea Study) in Crete, Greece. *Soc. Psychiatry. Psychiatr. Epidemiol.* 48:1335–1345.
40. Krzanowski, W.J., Hand, D.J. *ROC Curves for Continuous Data*. Chapman, & Hall/CRC Press, 2009.

41. Kuijpers, J.L., Vader, H.L., Drexhage, H.A., Wiersinga, W.M., van Son, M.J., Pop, V.J., 2001. Thyroid peroxidase antibodies during gestation are a marker for subsequent depression postpartum. *Eur. J. Endocrinol.* 145:579–584.
42. Lau, Y., Wong, Y., Yin, L., Chan, K.S., Guo, X., 2010. Validation of the Mainland Chinese version of the Edinburgh Postnatal Depression Scale in Chengdu mothers. *Int J. Nurs. Stud.* 47:1139–1151.
43. Leight, K.L., Fitelson, E.M., Weston, C.A., Wisner, K.L., 2010. Childbirth and mental disorders. *Int. Rev. Psychiatry.* 22:453–471.
44. Lewis, L., Pelosi, A., Araya, R., Dunn, G., 1992. Measuring psychiatric disorder in the community: a standardized assessment for use by lay interviewers. *Psychol. Med.* 22:465–486.
45. Lindqvist, P.G., Nasiell, J., Gustafsson, L.L., Nordstrom, L., 2014. Selective serotonin reuptake inhibitor use during pregnancy increases the risk of postpartum hemorrhage and anemia: a hospital-based cohort study. *J. Thromb. Haemost.* 12:1986–1992.
46. Lommatzsch, M., Hornyk, K., Zingler, C., Schuff-Werner, P., Höppner, J., Virchow, J.C., 2006. Maternal serum concentrations of BDNF and depression in the perinatal period. *Psychoneuroendocrinology.* 31:388–394.
47. Matthey, S., Barnett, B., Kavanagh, D.J., Howie, P., 2001. Validation of the Edinburgh Postnatal Depression Scale for men, and comparison of item endorsement with their partners. *J. Affect. Disord.* 64:175–184.
48. Matthey, S., Henshaw, C., Elliott, S., Barnett, B., 2006. Variability in use of cut-off scores and formats on the Edinburgh Postnatal Depression Scale: implications for clinical and research practice. *Arch. Womens. Ment. Health.* 9:309–315.

49. Matthey, S., Fisher, J., Rowe, H., 2013. Using the Edinburgh postnatal depression scale to screen for anxiety disorders: conceptual and methodological considerations. *J. Affect. Disord.* 146:224–230.
50. Murray, D., Cox, J.L., 1990. Screening for depression during pregnancy with the Edinburgh Depression Scale (EPDS). *J. Reprod. Infant. Psychol.* 8:99–107.
51. Nagy, E., Molnar, P., Pal, A., Orvos, H., 2011. Prevalence rates and socioeconomic characteristics of post-partum depression in Hungary. *Psychiatry. Res.* 185:113–120.
52. O'Connor, T.G., Heron, J., Golding, J., Beveridge, M., Glover, V., 2002. Maternal antenatal anxiety and children's behavioural/emotional problems at 4 years. Report from the Avon Longitudinal Study of Parents and Children. *Brit. J. Psychiatry.* 180:502–508.
53. Otake, Y., Nakajima, S., Uno, A., Kato, S., Sasaki, S., Yoshioka, E., Ikeno, T., Kishi, R., 2014. Association between maternal antenatal depression and infant development: a hospital-based prospective cohort study. *Environ. Health. Prev. Med.* 19:30–45.
54. Peterson, B.D., Newton, C.R., Rosen, K.H., Skaggs, G.E., 2006. The relationship between coping and depression in men and women referred for in vitro fertilization. *Fertil. Steril.* 85:802–804.
55. Pop, V.J., Komproe, I.H., van Son, M.J., 1992. Characteristics of the Edinburgh Post Natal Depression Scale in the Netherlands. *J. Affect. Disord.* ;26:105–110.
56. Rahman, A., Iqbal, Z., Bunn, J., Lovel, H., Harrington, R., 2004. Impact of maternal depression on infant nutritional status and illness: a cohort study. *Arch. Gen. Psychiatry.* 61:946–952.
57. Räisänen, S., Lehto, S.M., Nielsen, H.S., Gissler, M., Kramer, M.R., Heinonen, S., 2014. Risk factors for and perinatal outcomes of major depression during pregnancy: a population-based analysis during 2002-2010 in Finland. *BMJ. Open.* 4:e004883.

58. Robertson, E., Grace, S., Wallington, T., Stewart, D.E., 2004. Antenatal risk factors for postpartum depression: a synthesis of recent literature. *Gen. Hosp. Psychiatry.* 26:289–295.
59. Robins, L.N., Wing, J., Wittchen, H.U., Helzer, J.E., Babor, T.F., Burke, J., Farmer, A., MD; Jablenski. A., Pickens., R., Regier., D.A., Sartorius, N., Towle, L.H., 1988. The Composite International Diagnostic Interview. An epidemiologic instrument suitable for use in conjunction with different diagnostic systems and in different cultures. *Arch. Gen. Psychiatry.* 45:1069–1077.
60. Rubertsson, C., Börjesson, K., Berglund, A., Josefsson, A., Sydsjö, G., 2011. The Swedish validation of Edinburgh Postnatal Depression Scale (EPDS) during pregnancy. *Nord. J. Psychiatry.* 65:414–418.
61. Schulte-Herbrüggen, O., Litzke, J., Hornyk, K., Zingler, C., Höppner, J., Virchow, J.C., Hellweg, R., Lommatzsch, M., 2007. Maternal nerve growth factor serum levels in the perinatal period. *J. Reprod. Immunol.* 74:170–173.
62. Sheehan, D.V., Lecrubier, Y., Harnett-Sheehan, K., Amorim, P., Janavs, J., Weiller, E., Hergueta, T., Baker, R., Dunbar, G.C., 1998. The Mini International Neuropsychiatric Interview (M.I.N.I.): The Development and Validation of a Structured Diagnostic Psychiatric Interview. *J. Clin. Psychiatry.* 59: 22–33.
63. Snaith, R.P., Constantopoulos, A.A., Jardine, M.Y., McGuffin. P., 1978. A clinical scale for the self-assessment of irritability. *Brit. J. Psychiatry.* 132:164–171.
64. Spitzer, R.L., Endicott, J., Robins, E. Research Diagnostic Criteria Instrument no. 58. New York: New York State Psychiatric Institute, 1975.
65. Spitzer, R.L., Williams, J.B.W., Gibbon, M., First, M.B. User's guide for the structured clinical interview for DSM-III-R. Washington, D.C.: American Psychiatric Press; 1990.

66. Spitzer, R.L., Kroenke, K., Williams, J.B., 1999. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. *JAMA*. 282:1737–1744.
67. Stewart, R.C., Umar, E., Tomenson. B., Creed, F., 2013. Validation of screening tools for antenatal depression in Malawi-A comparison of the Edinburgh Postnatal Depression Scale and Self Reporting Questionnaire. *J. Affect. Disord.* 150:1041–1047.
68. Su, K.P., Chiu, T.H., Huang, C.L., Ho, M., Lee, C.C., Wu, P.L., Lin, C.Y., Liao, C.H., Liao, C.C., Chiu, W.C., Pariante, C.M., 2007. Different cutoff points for different trimesters? The use of Edinburgh Postnatal Depression Scale and Beck Depression Inventory to screen for depression in pregnant Taiwanese women. *Gen. Hosp. Psychiatry*. 29:436–441.
69. Suri, R., Altshuler, L., Hellemann, G., Burt, V.K., Aquino, A., Mintz, J., 2007. Effects of antenatal depression and antidepressant treatment on gestational age at birth and risk of preterm birth. *Am. J. Psych.* 164:1206–1213.
70. Talge, N.M., Neal, C., Glover, V., 2007. Antenatal maternal stress and long-term effects on child neurodevelopment: how and why? *J. Child. Psychol. Psychiatry*. 48:245–261.
71. Teng, H.W., Hsu, C.S., Shih, S.M., Lu, M.L., Pan, J.J., Shen, W.W., 2005. Screening postpartum depression with the Taiwanese version of the Edinburgh Postnatal Depression scale. *Compr. Psychiatry*. 46:261–265.
72. Thombs, B.D., Stewart, D.E., 2014. Depression screening in pregnancy and postpartum: how close are we? *J. Psychosom. Res.* 77:244-245.
73. Tőreki, A., Andó, B., Keresztúri, A., Sikovanyecz, J., Dudas, R.B., Janka, Z., Kozinszky, Z, Pál, A., 2013. The Edinburgh Postnatal Depression Scale: translation and antepartum validation for a Hungarian sample. *Midwifery*. 29:308–315.

74. Tőreki, A., Andó, B., Dudas, R.B., Dweik, D., Janka, Z., Kozinszky, Z., Keresztúri, A., 2014. The Edinburgh Postnatal Depression Scale: translation and antepartum validation for a Hungarian sample. *Midwifery*. 30:911–918.
75. Tran, T.D., Tran, T., La, B., Lee, D., Rosenthal, D., Fisher, J., 2011. Screening for perinatal common mental disorders in women in the north of Vietnam: a comparison of three psychometric instruments. *J. Affect. Disord.* 133:281–293.
76. Uwakwe, R., Okonkwo, J.E., 2003. Affective (depressive) morbidity in puerperal Nigerian women: validation of the Edinburgh Postnatal Depression Scale. *Acta Psych Scand*, 107:251–259.
77. Wallis, S.A., 2013. Binomial confidence intervals and contingency tests: mathematical fundamentals and the evaluation of alternative methods. *J. Quantit. Linguistics*. 20:178–208.
78. Wang, Y., Guo, X., Lau, Y., Chan, K.S., Yin, L., Chen, J., 2009. Psychometric evaluation of the Mainland Chinese version of the Edinburgh Postnatal Depression Scale. *Int. J. Nurs. Stud.* 46:813–823.
79. Werrett, J., Clifford, C., 2006. Validation of the Punjabi version of the Edinburgh postnatal depression scale (EPDS). *Int. J. Nurs. Stud.* 43:227–236.
80. Wickberg, B., Hwang, C.P., 1997. Screening for postnatal depression in a population-based Swedish sample. *Acta. Psych. Scand.* 95:62–66.
81. World Health Organization. The ICD-10 classification of mental and behavioural disorders. Clinical descriptions and diagnostic guidelines. Geneva: WHO, 1992.
82. Zigmond, A.S., Snaith, R.P., 1983. The Hospital Anxiety and Depression Scale. *Acta. Psych. Scand.* 67:361–370.

83. Zuckerman. B., Amaro, H., Bauchner, H., Cabral, H., 1989. Depressive symptoms during pregnancy: relationship to poor health behaviors. *Am. J. Obstet. Gyn.* 160:1107–1111.