



## A comparison of two measures to screen for emotional health difficulties during pregnancy

Stephen Matthey & Anna M. Della Vedova

To cite this article: Stephen Matthey & Anna M. Della Vedova (2018): A comparison of two measures to screen for emotional health difficulties during pregnancy, Journal of Reproductive and Infant Psychology, DOI: [10.1080/02646838.2018.1490497](https://doi.org/10.1080/02646838.2018.1490497)

To link to this article: <https://doi.org/10.1080/02646838.2018.1490497>



Published online: 26 Jul 2018.



Submit your article to this journal [↗](#)



Article views: 16



View Crossmark data [↗](#)



## A comparison of two measures to screen for emotional health difficulties during pregnancy

Stephen Matthey<sup>a</sup> and Anna M. Della Vedova<sup>b</sup>

<sup>a</sup>South Western Sydney Local Health District & University of Sydney (School of Psychology), Sydney, Australia; <sup>b</sup>Clinical and Experimental Sciences Department, Università degli Studi di Brescia, Brescia, Italy

### ABSTRACT

**Objective:** To compare the performance of a generic mood questionnaire (Matthey Generic Mood Questionnaire, MGMQ) against the established Edinburgh Postnatal Depression Scale (EPDS) in perinatal mental health mood screening.

**Background:** Many perinatal clinical services use the EPDS to screen for depression, and some may consider using it to screen for anxiety. A new scale, the MGMQ, is designed to screen for a wide variety of emotions, not just depression or anxiety. It comprises a generic distress question, an impact question, as well as two clinical questions. Its brevity, and categorical scoring format, may also mean it is less susceptible than the EPDS to needing a myriad of different screen-positive scores for women from different cultures and during different perinatal time periods.

**Methods:** Two hundred and ten Italian women in their third trimester of pregnancy completed the EPDS and MGMQ while attending routine antenatal clinic appointments or antenatal classes in the north of Italy, between 2015 and 2016.

**Results:** The Distress and Lower Impact question thresholds showed acceptable receiver operating characteristics with the various EPDS screen positive thresholds. The Higher Impact question threshold, however, had lower than acceptable sensitivity. By contrast, the EPDS was poor at detecting women who on the MGMQ said that they were distressed and significantly bothered by their mood. The possible reasons for the discrepancies in screen-positive status between the two measures are discussed.

**Conclusion:** The MGMQ is a useful tool to aid in screening for a wide range of emotional difficulties in the perinatal period.

### ARTICLE HISTORY

Received 8 November 2017

Accepted 4 April 2018

### KEYWORDS

Perinatal depression;  
distress; screening; measures

It is well established that the presence of symptoms of depression or anxiety in the perinatal period can have negative consequences, not only for the woman and the couple's relationship, but also on pregnancy and birth outcomes (Grigoriadis et al., 2013), the fetus' and infant's development (Della Vedova, 2014; Stein et al., 2014), and the risk of depression in adolescent offspring (Pearson et al., 2013). Thus there is now agreement on the importance of detecting emotional distress as early as possible in perinatal women, preferably starting in the antenatal period (e.g. O'Hara et al., 2012).

Several guidelines or expert opinions recommend the use of the Edinburgh Postnatal Depression Scale (EPDS;<sup>1</sup> Cox, Holden, & Sagovsky, 1987), such as those in America, Canada and Australia (BC Reproductive Mental Health Program & Perinatal Services BC, 2014; beyond-blue, 2011; Committee on Obstetric Practice, 2015). However, the UK guidelines (NICE, National Institute for Clinical Excellence, 2014) recommend the Whooley questions (Whooley, Avins, Miranda, & Browner, 1997), while the Italian guidelines (CeVEAS, 2011) consider that no brief measure has yet been shown to have acceptable psychometric properties.

This diversity of 'expert opinion' regarding the optimal screening instrument shows that there needs to be continuing research in this field, a view recently expressed by Alderdice et al. (2013). Various studies have thus reported on the development of new scales or questions, or the validation of existing scales for use in perinatal screening (e.g. Allison, Wenzel, Kleiman, & Sarwer, 2011; Austin et al., 2010; Beck & Gable, 2000; O'Hara et al., 2012; Somerville et al., 2014). However, many measures are quite long, which makes them impractical within screening contexts. For example, some researchers consider the 10-item EPDS to be too long, and thus have tried to develop shorter versions (e.g. Gollan et al., 2017; Kabir, Sheeder, & Kelly, 2008), while other scales are even longer, such as the short form of the Depression, Anxiety and Stress Scale (DASS: Lovibond & Lovibond, 1995), the Perinatal Depression Inventory (Brodey et al., 2016), and the Perinatal Anxiety Screening Scale (Somerville et al., 2014).

Similarly, most measures aim to screen for just one type of mood difficulty – either just depression, or just anxiety. However, it has been acknowledged by both professionals and women themselves that screening needs to occur for at least both of these mood dimensions (e.g. Coates, De Visser, & Ayers, 2015; Muzik et al., 2000). Combining two multi-item measures to screen for both moods may be impractical in busy clinical settings. In addition, scales using continuous scoring (e.g. the EPDS and most anxiety scales) need to be validated for women from different cultures, and for the ante- and postnatal periods, resulting in a vast array of different validated scores being required (Kozinszky & Dudas, 2015). When a questionnaire has subscales measuring different moods (e.g. depression and anxiety – the EPDS; or these two moods and also stress – the DASS), this then gets even more complicated as the subscales within the one measure have different cut-off scores.

Given these issues, as well as other problems with the EPDS, which include its explicit exclusion of distress if it is considered to be for a good reason (Matthey & Agostini, 2017), a new brief questionnaire has been developed that aims to screen for any type of emotional distress without using separate subscales, and which uses a categorical scoring format – the Matthey Generic Mood Questionnaire (MGMQ: Matthey, Valenti, Souter, & Ross-Hamid, 2013). This study aimed to compare the performance of this MGMQ with the well-established self-report measure, the EPDS.

To make such measurement comparisons, three methodologies have been used in the literature. The first, which is generally considered to be the best, is to use diagnostic status for the relevant mood disorders as the criterion, and to compare the receiver operating characteristics (ROC) of each measure against this. Thus, E Couto et al. (2015) compared the EPDS and Beck Depression Inventory (BDI: Beck, Steer, & Garbin, 1988) against diagnostic status, concluding that the BDI was the preferred instrument for their population due to its better area under the curve value. The second approach has been to use one self-report measure as the criterion against which to judge another self-report measure. Thus, Darwin, McGowan, and Edozien (2016) compared the Whooley

questions against the EPDS as the criterion, and from the resultant ROC concluded that these questions missed around half of the women screening positive on the EPDS, with the implication appearing to be that the EPDS was therefore the preferred screening instrument. The third methodology has been to compare how measures compare with each other, using concordance analyses, but without using diagnostic status or either measure as the criterion. Thus, Affonso, De, Horowitz, and Mayberry (2000) compared the performance of the EPDS with the BDI, and showed that the BDI only detected about half of EPDS screen-positive women, while the EPDS detected around three-quarters of women screening positive on the BDI. They were, however, somewhat inconclusive about the practical implications of this finding. Condon and Corkindale (1997), using the same methodology, found even poorer screen-positive concordance between several symptom-based measures (including the EPDS), and suggested that this was due to the difference in symptoms covered in the different measures, as well as their wording.

The study to be reported investigates the performance of the MGMQ compared to the EPDS using the latter two methodological approaches. Thus, the ROC values for the MGMQ were calculated against the EPDS as the criterion measure, and also analyses were conducted to ascertain how well each of the measures detected the women screening positive on the other measure. Due to funding constraints a diagnostic interview was not administered.

## Method

### Design

The research is a cross-sectional study, assessing depression and anxiety symptoms in Italian women in the third trimester of pregnancy.

### Participants

Women attending public health clinics and hospitals for parents and babies in two medium-size cities in northern Italy were eligible for the study if they were aged 18 or older, in their third trimester of pregnancy and were fluent in Italian (approximately half of the clients were Italian speakers). Three sites across the two cities participated between June 2014 and February 2016. The project received all appropriate Ethical approval, and participants provided informed consent.

### Procedure

Italian-speaking women attending the antenatal clinics for their routine third-trimester visit, and those attending antenatal classes, were approached to participate in the study, with recruitment being undertaken on approximately two days per week across the sites. Those who agreed to participate then completed the relevant measures in the clinic or class.

### Measures

*Background and risk questionnaire.* A questionnaire was created to collect sociodemographic, health, obstetric and psychosocial information.

*Edinburgh Postnatal Depression Scale* (EPDS; Cox et al., 1987). This is a 10-item self-report questionnaire which can be used to screen for possible depression, both during pregnancy and postpartum. Each item is scored from 0 to 3, assessing the woman's mood over the past 7 days. Higher scores reflect poorer mood. Studies have validated its use in various countries and two have validated it postnatally in Italy (Benvenuti, Ferrara, Niccolai, Valoriani, & Cox, 1999; Carpiniello, Pariante, Serri, Costa, & Carta, 1997). Carpiniello et al. reported optimal screen-positive cut-off scores for major depression of 10 or more, while Benvenuti et al. reported that a score of 13 or more detected all six women in their study with major depression. As there has not yet been an Italian validation study during pregnancy, we have therefore used both of these postnatal screen-positive scores to determine screen-positive status for possible depression antenatally.

Several studies have also investigated the scale's ability to screen for anxiety (e.g. Coates, Ayers, & De Visser, 2017), which have often found that items 3, 4 and 5 load on an anxiety subscale (EPDS-3A<sup>1</sup>: Matthey, 2008), although one Italian study found that item 6 also loaded on this subscale (Petrozzi & Gagliardi, 2013). Given that the only work validating this subscale's screen-positive score for an anxiety disorder in a community population reported the optimal cut-off score as being 6 or more (Matthey, 2008), we have used this score to determine screen-positive status for possible anxiety. However, it should be noted that there is evidence that this subscale misses around half of the women with an anxiety disorder (Matthey et al., 2013).

Services frequently use the validated cut-off scores on the EPDS to screen just for possible depression. Given that the EPDS also has the anxiety subscale, as discussed above, some services may decide to screen women for the presence of both of these disorders, by also using the validated cut-off score on this subscale in conjunction with the screen-positive total score for possible depression. We have therefore reported the data to reflect these two screening purposes. This first category is therefore that of screening for just depression (using total scores of 10 or more or 13 more). The second category, which we term 'Distress', uses a total score of 10 or more and/or an anxiety subscale score of 6 or more (lower threshold for distress); or a total score of 13 or more and/or an anxiety subscale score of 6 or more (higher threshold for distress). This combining of the total and anxiety subscale cut-off scores into a distress category has been used in previous research (Della Vedova & Matthey, 2016).

*Matthey Generic Mood Question* (MGMQ: Matthey et al., 2013). This consists of three questions concerning how the person has been feeling over the past two weeks: (1) (Distress Question): 'Have you felt very stressed, anxious, or unhappy, or found it difficult to cope, for some of the time?' (response options: 'Yes', 'Possibly', 'No'); (2) (Impact question): If 'Yes' or 'Possibly' (to the Distress question), 'How bothered have you been by these feelings?' (response options: 'Not at all'; 'A little bit'; 'Moderately'; 'A lot'); (3) ('Reason for distress' question): If 'A little bit' or more (to the Impact Question), 'What do you think has made you feel this way?'. There is also a fourth question for use within clinical services asking if she/he would like to talk to a health professional ('Wish for referral' question), although this was not included in this study.

In determining the measure's screen-positive threshold Kirkby, Wilson, Calvert, and Draper (2011) recommend surveying expert opinion when ascertaining clinically

meaningful effect sizes on new instruments or interventions. Thus, a convenience sample of 57 experienced Italian mental health clinicians were surveyed as to the responses on the Impact question that would warrant further assessment within a screening context, as well as whether they had any comments about the scale. The majority stated that responses of 'Moderately' or 'A lot' (bothered) on the Impact question would indicate that further assessment was required. Therefore, the response of 'Moderately' bothered constitutes the upper Impact threshold, or clinicians' threshold, for a positive screen on the Impact question, while 'A little' or more is the lower MGMQ Impact threshold. For the Distress question any positive endorsement (i.e. 'Yes' or 'Possibly') was classified as screen-positive.

Matthey et al. (2013) demonstrated in an antenatal English-speaking sample ( $N = 132\text{--}389$ ) where the focus of the study was on anxiety, that the Distress question correctly identified 80% of women meeting diagnostic criteria for an anxiety disorder, 84% of women who scored above the cut-off threshold for possible anxiety on the EPDS-3A, and between 58% and 87% of women scoring high on various general and pregnancy-specific anxiety measures. In a separate study the Distress question detected 89.5% of women screening positive on the Primary Health Questionnaire-2 (PHQ-2: Kroenke, Spitzer, & Williams, 2003), whereas that measure, designed just to screen for depression, only detected between 25% and 58% of women screening positive on the MGMQ (Matthey & Bilbao, 2018).

The measure was translated into Italian, using a standard translation and back-translation procedure (Beaton, Bombardier, Guillemin, & Ferraz, 2000). In addition, apart from some minor changes to the original Italian wording, none of the survey respondents raised any concerns with the concepts being assessed by the MGMQ, and thus it is considered suitable for Italians. This Italian version is thus called the MGMQ-it.

### **Analyses considerations**

*EPDS as criterion.* Given that the MGMQ is screening for both depression and anxiety (and other such moods), the ROC will be determined against the EPDS screen-positive thresholds for depression and/or anxiety (distress).

*Concordance analyses.* Rather than focus on statistical significance, which says nothing about the size of a relationship (Cumming, 2013), the analyses will focus on whether the concordance rates between the measures are clinically meaningfully different, using Rosenthal's (1996) assertion that an absolute percentage difference of about 18 points is equivalent to a medium effect size, or 12 points if the rates are at the extreme ends of the distribution (e.g. 10% or less, or 90% or more).

*Sample size.* A minimum of 15 women screening positive on the EPDS were required for the ROC analyses, based upon previous scale validation work (e.g. Benvenuti et al., 1999; Carpinello et al., 1997; Cox et al., 1987). The likely rate of women scoring 10 or more on the EPDS was around 10–20%, based upon the Italian perinatal literature (Petrozzi & Gagliardi, 2013), and thus a sample size of around 200 participants would be required. This number was also considered to be sufficient to compare the two measures against each other, as previous work on the MGMQ indicates that around a third of women will screen positive to the Distress question.

## Results

### Participants

Approximately 267 eligible women were invited to participate in the study, of which 210 participated and provided complete data (an uptake rate of about 79%). Their demographic characteristics are shown in Table 1.

### Screen-positive rates

Table 2 shows the different rates of women screening positive for each of the EPDS and MGMQ screen positive thresholds, many of which show no meaningful differences between them. Thus, the rates for each of the EPDS thresholds mainly vary between around 9% and 13%, which are all quite similar using Rosenthal's recommendation, although the lower rate for the 13 or more threshold of 3.8% is equivalent to an effect size between small and medium. The MGMQ Higher Impact threshold of 11.9% is therefore similar to most of the EPDS screen-positive rates, while the thresholds for the Lower Impact question and Distress question produce screen positive rates of around 33%, which are therefore meaningfully higher than the EPDS thresholds.

Importantly, however, while the rates are very similar between the EPDS and MGMQ Clinicians' Higher Impact threshold, this is not because most of same women are

**Table 1.** Demographic characteristics of the sample ( $N = 210$ ).

	Mean ( <i>sd</i> )	Range
Mother's age (years)	32.0 (4.7)	18–44
Gestational age (weeks)	33.3 (3.0)	27–39
	%	
Place of birth: Italy	100	
Nulliparous	86.7	
Married or de-facto	99	
Tertiary level of education	52.4	
<i>Health problems in pregnancy</i>		
Maternal	24.8	
Infant	1.4	
Singleton pregnancy	98.6	

**Table 2.** Percentage of the sample screening positive on the various thresholds for each measure ( $N = 210$ ).

Measure threshold	% screening positive ( $N$ )
EPDS 10 or more	10 (21)
EPDS 13 or more	3.8 (8)
EPDS (10+/6+) <sup>1</sup>	12.9 (27)
EPDS (13+/6+) <sup>2</sup>	9.0 (19)
MGMQ Distress question <sup>3</sup>	32.9 (69)
MGMQ Lower Impact threshold <sup>4</sup>	32.9 (69)
MGMQ Higher Impact threshold <sup>5</sup>	11.9 (25)

<sup>1</sup>A total EPDS score of 10 or more, and/or an EPDS-3A score of 6 or more.

<sup>2</sup>A total EPDS score of 13 or more, and/or an EPDS-3A score of 6 or more.

<sup>3</sup>Endorsing on the MGMQ Distress question 'Yes' or 'Possibly'.

<sup>4</sup>Endorsing on the MGMQ Impact question 'A little' or 'Moderately' or 'A lot'.

<sup>5</sup>Endorsing on the MGMQ Impact question 'Moderately' or 'A lot'.

screening positive on both measures. This is true for only about half the women screening positive on each measure, as is clear from the following information.

### EPDS as criterion

Tables 3a and 3b show the ROC of the MGMQ against the various EPDS screen-positive thresholds. As the tables show, the MGMQ Distress and Lower Impact question thresholds detect the majority of women screening positive on the EPDS for possible depression (sensitivity values of 81% and 88%), and a slightly lower percentage for those screening positive on the EPDS for depression and/or anxiety (Distress and Lower Impact question thresholds – sensitivity values of 63% and 70%). Of note, however, are the consistently low positive predictive values (ppvs) (varying between 5% and 28%), indicating that the majority of women screening positive on these two MGMQ questions do not screen positive on the EPDS thresholds. For the MGMQ Higher Impact threshold, the sensitivity values are naturally lower (32–43%), while the ppvs are slightly higher (12–40%).

### Measure concordance

While Tables 3a and 3b show how ‘well’ the MGMQ detects women screening positive on the EPDS (sensitivity values), Table 4 shows how ‘well’ the EPDS detects women screening

**Table 3a.** Receiver operating characteristics of the MGMQ Distress question, using the EPDS as the criterion ( $N = 210$ ).

Criterion threshold	MGMQ Distress question threshold: ‘Yes’ or ‘Possibly’			
	Sensitivity	Specificity	PPV	NPV
EPDS 10+	81	72.5	24.6	97.2
EPDS 13+	87.5	30.7	4.8	98.4
EPDS 10+ and/or EPDS-3A 6+	70.4	72.7	27.5	94.3
EPDS 13+ and/or EPDS-3A 6+	63.2	70.2	17.4	95.0

Note: EPDS 10+ and EPDS 13+ signify ‘10 or more’ and ‘13 or more’ on the Total EPDS score. EPDS-3A 6+ signifies ‘6 or more’ on the EPDS-3A.

**Table 3b.** Receiver operating characteristics of the lower and higher MGMQ Impact question thresholds, using the EPDS as the criterion ( $N = 210$ ).

Criterion threshold		MGMQ Impact question thresholds <sup>1</sup>			
		Sensitivity	Specificity	PPV	NPV
EPDS 10+ }	MGMQ Impact lower threshold	81.0	73.5	25.4	97.2
	MGMQ Impact higher threshold	42.9	91.5	36.0	93.5
EPDS 13+ }	MGMQ Impact lower threshold	87.5	70.3	10.4	99.3
	MGMQ Impact higher threshold	37.5	89.1	12.0	97.3
EPDS 10+ and/or EPDS-3A 6+ }	MGMQ Impact lower threshold	70.4	73.8	28.4	94.4
	MGMQ Impact higher threshold	37.0	91.8	40.0	90.8
EPDS 13+ and/or EPDS-3A 6+ }	MGMQ Impact lower threshold	63.2	71.2	17.9	95.1
	MGMQ Impact higher threshold	31.6	90.1	24.0	93.0

Note: ‘EPDS 10+’ and ‘EPDS 13+’ signify ‘10 or more’ and ‘13 or more’, respectively, on the Total EPDS score. ‘EPDS-3A 6+’ signifies ‘6 or more’ on the EPDS-3A.

<sup>1</sup>Lower threshold: bothered ‘A little’ or more; higher threshold: bothered ‘Moderately’ or more.



**Table 4.** Concordance analyses between EDS and MGMQ screen-positive status.

MGMQ threshold	N	% of MGMQ screen-positive women detected by			
		EPDS 10+	EPDS 13+	EPDS 10+ and/or EPDS-3A 6+	EPDS 13+ and/or EPDS-3A 6+
Distress <sup>1</sup>	69	24.6	10.1	27.5	17.4
Lower Impact threshold <sup>2</sup>	67	25.4	10.4	28.4	17.9
Higher Impact threshold <sup>3</sup>	25	36	12	40	24

Note: EPDS 10+ and EPDS 13+ signify '10 or more' and '13 or more' on the Total EPDS score; EPDS-3A 6+ signifies '6 or more' on the EPDS-3A.

<sup>1</sup>Endorsing on the Distress question 'Yes' or 'Possibly'.

<sup>2</sup>Endorsing on the MGMQ Impact question 'A little' or more.

<sup>3</sup>Endorsing on the MGMQ Impact question 'Moderately' or more.

positive on the MGMQ. As this shows, the lower EPDS distress threshold only detects around 28–40% of women screening positive on the MGMQ, while the higher threshold detects even fewer. The large differences between most of the detection rates between the two measures (e.g. 70.4% detected by the MGMQ lower threshold compared to 28.4% detected by the EPDS lower distress threshold) are clinically meaningful, using Rosenthal's recommendation. It is only for the Higher Impact question that the detection rates between the two measures are similar (around 38%).

Of note also is that for nine of the 27 lower distress threshold EPDS screen-positive women, their reported level of Impact on the MGMQ was just 'a little bothered', while another eight reported no difficulty on the MGMQ Distress question. A similar pattern was found for those women screening positive using the higher EPDS distress threshold.

### Reasons for 'MGMQ screen-positive'

Women gave various reasons on the MGMQ for why they screened positive. Pregnancy-related reasons included family or work difficulties, moving home, anxiety about the birth, concern regarding coping with a new baby and the anticipated changes in their life. Non pregnancy-related concerns included worries over the health of a family member, planning for her own wedding and unspecified external stressors.

### Discussion

Two methodological approaches were taken to investigate the properties of the MGMQ. The first, that of using the EPDS as the criterion measure for distress (for possible depression and/or anxiety), showed that the ROC sensitivity and specificity values of the MGMQ Distress question were, in the main, satisfactory (sensitivity: 63% to 88%; specificity: 31% to 73%), and had ppvs ranging from 5% to 28% (see Table 3a).

The MGMQ Lower Impact threshold (bothered 'a little bit' or more) had similar ROC values. The Higher Impact threshold (bothered 'moderately' or more) had lower sensitivity (32–43%), higher specificity (89–92%) and slightly higher ppvs (12–40%) (see Table 3b).

The second methodological approach was that of inspecting the concordance for screen-positive women between each measure. This showed that whereas the MGMQ Distress and Lower Impact thresholds each detected 63–70% of women screening positive on the EPDS (see Table 3a), the EPDS only detected 10–28% of the women

screening positive on the MGMQ (see Table 4). These figures were no longer different when the MGMQ Higher Impact threshold was used, with both measures then generally detecting about 40% of the high scorers on the other measure, although the higher EPDS threshold of 13 or more resulted in lower detection rates of 12% and 24% (see Tables 3b and 4).

The ROC findings are evidence that the MGMQ compares well to the EPDS in a screening context, with the obtained values being comparable to those reported in some EPDS validation studies (e.g. see Kozinszky & Dudas, 2015; Matthey, Barnett, Kavanagh, & Howie, 2001; Meijer et al., 2014). However, the low ppvs mean that many screen-positive women on the MGMQ would not be high scorers on the EPDS. It would be expected, however, that within a screening setting a woman who says on the MGMQ that she is 'distressed' and that this bothers her 'moderately' (or 'a lot') would take precedence over her low EPDS score when deciding whether to further investigate her mood. The concordance analyses show that the EPDS misses many women screening positive on the MGMQ. If a service used the common EPDS screen-positive threshold of 13 or more, they would, from these data, miss almost all of the women (88%) who state that they are 'moderately' or 'a lot' bothered by their distress. Even the lower EPDS screen-positive threshold would still miss over 60% of these women.

One likely reason for this is that the EPDS screens mainly for depression, with just three items on anxiety, whereas the MGMQ screens for a much broader range of negative emotions. Therefore, the EPDS would not be expected to detect women who are endorsing the MGMQ because of negative moods other than the narrower mood range detected by the EPDS. In addition, as the EPDS uses a symptom-based approach, women will only be detected on this scale if they have enough of the 10 specific symptoms it addresses. If instead their negative mood is characterised by other symptoms (e.g. anger or irritation: Coates et al., 2015), they will score below the threshold on the EPDS despite in fact having a low mood. Also, if her distress is due to worries that she considers to be reasonable – such as being extremely worried about the health of her unborn child – then this too would not necessarily be detected on the EPDS, given that the anxiety questions only assess for worries she experiences 'for no good reason' (e.g. item 4). An examination of the explanations given on the MGMQ for her distress certainly shows that there are often quite reasonable reasons for the woman's distress (e.g. anxiety about the birth). It is also possible that some women complete the EPDS but exclude from consideration negative feelings not related to their pregnancy or postpartum period (e.g. health concerns of others; stress at work, etc.), given that the measure's instructions focus on the fact that the woman has recently given birth or will soon give birth, although this theory would need to be investigated.

Another possible reason why the EPDS misses so many MGMQ screen-positive women is that low scores on the EPDS should not be unquestionably taken to indicate good mental health. Around a third of women who meet criteria for depression score low on the EPDS (Eberhard-Gran, Eskild, Tambs, Opjordsmoen, & Samuelsen, 2001; Kozinszky & Dudas, 2015), while Lydsdottir et al. (2014) have shown that many women scoring low on the EPDS actually have other emotional difficulties, such as bipolar and somatoform disorders, apart from depression.

The MGMQ does, however, miss some EPDS screen-positive women (10–40%). A possible reason for this is that a high score on the EPDS does not necessarily mean that the woman is significantly affected by her feelings, given that the EPDS does not have a question directly

addressing this issue. If despite scoring high on the EPDS the woman was not concerned with her mood, she would not therefore endorse that she was bothered 'moderately' or 'a lot' on the MGMQ. In addition, around half of all women screening positive on the EPDS have transient distress (Matthey & Ross-Hamid, 2012), and many of these women anticipate that they will soon be feeling better (Matthey, 2016). It is possible that such women who screen positive on the single administration of the EPDS, while also then screening positive on the MGMQ Distress question would not then screen positive on the Impact question.

While the screen-positive overlap between the two measures is quite poor, this phenomenon has also been found in other studies comparing different mood measures (e.g. Codon & Corkindale, 1997; Affonso et al., 2000). For most of the comparisons, however, it is evident that the MGMQ detects meaningfully more of the women screening positive on the EPDS than vice versa.

With respect to the practical aspects of the MGMQ, the fact that there are only three positive categorical responses on the MGMQ's Impact question may be beneficial. Measures such as the EPDS that have a continuous scoring format mean that there are a multitude of different validated cut-off scores across genders, cultures, and perinatal time periods, making them extremely difficult to use correctly within clinical settings (Matthey & Agostini, 2017; Woody, Ferrari, Siskind, Whiteford, & Harris, 2017). To date, the Australian and Italian cultures have the same clinician-determined screen-positive threshold on the MGMQ Impact question, and further work is being undertaken to see if this also applies to other cultures. In addition, the inclusion of the 'Reason for distress' and 'Wish for referral' questions on the MGMQ are likely to be clinically useful when used in a routine screening context, although this would need to be determined given that Darwin et al. (2016) have shown that in some situations a 'wish for referral' type question can impede the open communication of feelings by clients.

There are several limitations within this antenatal study. Firstly, the thresholds for the EPDS have been based upon the validated cut-off scores from two Italian postnatal studies, and an English-speaking study for the anxiety component of the EPDS. Secondly, the EPDS and MGMQ screen for mood over different time periods (1 week and 2 weeks, respectively), which could account for some of the lack of concordance, although this issue also applies to studies validating the EPDS against diagnostic status. Thirdly, all the women completed the EPDS first, followed, after some non-mood questions, by the MGMQ. There could therefore be some order effect influencing the data, although similar findings were obtained in the original MGMQ study where the order was in the opposite direction (Matthey et al., 2013). Fourthly, no diagnostic status comparison between the two measures was undertaken.

## Conclusion

The data indicate that the MGMQ has merit when screening for a broader range of emotions than those detected by the EPDS. The initial Distress question detects nearly all women who would have screened positive on the EPDS depression thresholds, and most that would have screened positive on the broader EPDS distress thresholds. Which of the two MGMQ Impact thresholds would be ideal within screening settings is, however, debatable. While most clinicians surveyed considered that the Higher Impact response of 'Moderately (bothered)' would be the threshold of choice, and this was the same in another cultural group (Australian), this naturally has poorer sensitivity values

against the EPDS than the lower threshold of 'A little bit (bothered)'. When the two measures are compared side by side, however, the EPDS fails to detect the majority of women screening positive on the MGMQ questions. This may be of concern, especially when the many limitations of the EPDS are considered, as recently discussed by Matthey and Agostini (2017).

Given that many services however have the EPDS firmly embedded within their clinical practice, and it is recognised that this on its own misses many women with high levels of anxiety (Austin, 2017), we would suggest that the trialling of the brief MGMQ, together with the EPDS, may be beneficial in improving routine perinatal mental health screening.

## Notes

1. Changed for consistency purposes from the 'EDS-3A', with permission from the author.

## Disclosure statement

No potential conflict of interest was reported by the authors.

## References

- Affonso, D. D., De, A. K., Horowitz, J. A., & Mayberry, L. J. (2000). An international study exploring levels of postpartum depressive symptomatology. *Journal of Psychosomatic Research*, 49, 207–216.
- Alderdice, F., Ayers, S., Darwin, Z., Green, J., Jomeen, J., Kenyon, S., & Savage-McGlynn, E. (2013). Measuring psychological health in the perinatal period: Workshop consensus statement, 19 March 2013. *Journal of Reproductive and Infant Psychology*, 31(5), 431–438.
- Allison, K. C., Wenzel, A., Kleiman, K., & Sarwer, D. B. (2011). Development of a brief measure of postpartum distress. *Journal of Women's Health*, 20, 617–623.
- Austin, M.-P. (2017, October). *The 2017 Australian clinical practice guidelines for mental health care in the perinatal period*. Paper presented at the Australasian Marce Society for Perinatal Mental Health 2017 Conference. Brisbane, Australia.
- Austin, M. P. V., Hadzi-Pavlovic, D., Priest, S. R., Reilly, N., Wilhelm, K., Saint, K., & Parker, G. (2010). Depressive and anxiety disorders in the postpartum period: How prevalent are they and can we improve their detection?. *Archives of Women's Mental Health*, 13(5), 395–401.
- BC Reproductive Mental Health Program & Perinatal Services BC. (2014). British Columbia Best Practice Guidelines for Mental Health Disorders in the Perinatal Period. Retrieved from <http://www.perinatalservicesbc.ca/Documents/Guidelines-Standards/Maternal/MentalHealthDisordersGuideline.pdf>.
- Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*, 25, 3186–3191.
- Beck, A. T., Steer, R. A., & Garbin, M. G. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review*, 8, 77–100.
- Beck, C. T., & Gable, R. K. (2000). Postpartum Depression Screening Scale: Development and psychometric testing. *Nursing Research*, 49, 272–282.
- Benvenuti, P., Ferrara, M., Niccolai, C., Valoriani, V., & Cox, J. L. (1999). The Edinburgh postnatal depression scale: Validation for an Italian sample. *Journal of Affective Disorders*, 53, 137–141.
- beyondblue. (2011). *Clinical Practice Guidelines for depression and related disorders – Anxiety, bipolar disorder and puerperal psychosis – In the perinatal period. A guideline for primary care health professionals*. Melbourne: Author.
- Brodey, B., Goodman, S. H., Baldasaro, R. E., Brooks-DeWeese, A., Wilson, M., Brodey, I. S. B., & Doyle, N. M. (2016). Development of the Perinatal Depression Inventory (PDI)-14 using item response

- theory: A comparison of the BDI II, EPDS, PDI, and PHQ-9. *Archives of Women's Mental Health*, 19, 307–316.
- Carpiniello, B., Pariante, C. M., Serri, F., Costa, G., & Carta, M. G. (1997). Validation of the Edinburgh Postnatal Depression Scale in Italy. *Journal of Psychosomatic Obstetrics and Gynaecology*, 18, 280–285.
- CeVEAS (2011). Linea Guida 20. Gravidanza Fisiologica. Aggiornamento 2011 Ministero della Salute: Italia. (Guideline 20. Normal pregnancy, 2011 updates. Ministry of Health of Italy). Retrieved from [http://www.salute.gov.it/portale/documentazione/p6\\_2\\_2\\_1.jsp?lingua=italiano&id=1436](http://www.salute.gov.it/portale/documentazione/p6_2_2_1.jsp?lingua=italiano&id=1436).
- Coates, R., Ayers, S., & De Visser, R. (2017). Factor structure of the Edinburgh Postnatal Depression Scale in a population-based sample. *Psychological Assessment*, 29, 1016–1027.
- Coates, R., De Visser, R., & Ayers, S. (2015). Not identifying with postnatal depression: A qualitative study of women's postnatal symptoms of distress and need for support. *Journal of Psychosomatic Obstetrics & Gynecology*, 36, 114–121.
- Committee on Obstetric Practice. (2015). Screening for perinatal depression. Committee Opinion No. 630. American College of Obstetricians & Gynecologists. *Obstetrics & Gynecology*, 125, 1268–1271.
- Condon, J. T., & Corkindale, C. J. (1997). The assessment of depression in the postnatal period: A comparison of four self-report questionnaires. *Australian and New Zealand Journal of Psychiatry*, 31, 353–359.
- Cox, J., Holden, J., & Sagovsky, R. (1987). Detection of postnatal depression: Development of the 10 item Edinburgh Postnatal Depression Scale. *British Journal of Psychiatry*, 150, 782–786.
- Cumming, G. (2013). The new statistics: A how-to guide. *Australian Psychologist*, 48, 161–170.
- Darwin, Z., McGowan, L., & Edozien, L. C. (2016). Identification of women at risk of depression in pregnancy: Using women's accounts to understand the poor specificity of the Whooley and Arroll case finding questions in clinical practice. *Archives of Women's Mental Health*, 19, 41–49.
- Della Vedova, A. M. (2014). Maternal psychological state and infant's temperament at three months. *Journal of Reproductive and Infant Psychology*, 32, 520–534.
- Della Vedova, A. M., & Matthey, S. (2016). The relative risks, and the likelihoods, of becoming postnatally distressed in the presence of common psychosocial risks: A study with Italian-speaking mothers. *International Journal of Mental Health Promotion*, 18, 276–290.
- E Couto, T. C., Brancaglion, M. Y. M., Cardoso, M. N., Protzner, A. B., Garcia, F. D., Nicolato, R., & Corrêa, H. (2015). What is the best tool for screening antenatal depression?. *Journal of Affective Disorders*, 178, 12–17.
- Eberhard-Gran, M., Eskild, A., Tambs, K., Opjordsmoen, S., & Samuelsen, S. O. (2001). Review of validation studies of the Edinburgh Postnatal Depression Scale. *Acta Psychiatrica Scandinavica*, 104, 243–249.
- Gollan, J. K., Wisniewski, S. R., Luther, J. F., Eng, H. F., Dills, J. L., Sit, D., ... Wisner, K. L. (2017). Generating an efficient version of the Edinburgh Postnatal Depression Scale in an urban obstetrical population. *Journal of Affective Disorders*, 208, 615–620.
- Grigoriadis, S., VonderPorten, E. H., Mamisashvili, L., Tomlinson, G., Dennis, C. L., Koren, G., ... Martinovic, J. (2013). The impact of maternal depression during pregnancy on perinatal outcomes: A systematic review and meta-analysis. *Journal of Clinical Psychiatry*, 74, e321–e341.
- Kabir, K., Sheeder, J., & Kelly, L. S. (2008). Identifying postpartum depression: Are 3 questions as good as 10?. *Pediatrics*, 122, e696–e702.
- Kirkby, H. M., Wilson, S., Calvert, M., & Draper, H. (2011). Using e-mail recruitment and an online questionnaire to establish effect size: A worked example. *BMC Medical Research Methodology*, 11, 89.
- Kozinszky, Z., & Dudas, R. B. (2015). Validation studies of the Edinburgh Postnatal Depression Scale for the antenatal period. *Journal of Affective Disorders*, 176, 95–105.
- Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2003). The Patient Health Questionnaire-2: Validity of a two-item depression screener. *Medical Care*, 41, 1284–1292.
- Lovibond, S. H., & Lovibond, P. F. (1995). *Manual for the Depression Anxiety Stress Scales*. University of New South Wales, Australia: Psychology Foundation Monograph.

- Lydsdottir, L. B., Howard, L. M., Olafsdottir, H., Thome, M., Tyrfinngsson, P., & Sigurdsson, J. F. (2014). The mental health characteristics of pregnant women with depressive symptoms identified by the Edinburgh Postnatal Depression Scale. *Journal of Clinical Psychiatry*, 75, 393–398.
- Matthey, S. (2008). Using the Edinburgh Postnatal Depression Scale to screen for anxiety disorders. *Depression & Anxiety*, 25, 926–931.
- Matthey, S. (2016). Differentiating between Transient and Enduring distress on the Edinburgh Depression Scale within screening contexts. *Journal of Affective Disorders*, 196, 252–258.
- Matthey, S., & Agostini, F. (2017). Using the Edinburgh Postnatal Depression Scale for women and men - some cautionary thoughts. *Archives of Women's Mental Health*, 20, 345–354.
- Matthey, S., Barnett, B. E. W., Kavanagh, D. J., & Howie, P. (2001). Validation of the Edinburgh Postnatal Depression Scale for men, and comparison of item endorsement with their partners. *Journal of Affective Disorders*, 64, 175–183.
- Matthey, S., & Bilbao, F. (2018). A comparison of the PHQ-2 and MGMQ for screening for emotional health difficulties during pregnancy. *Journal of Affective Disorders*, 234, 174–179.
- Matthey, S., & Ross-Hamid, C. (2012). Repeat testing on the Edinburgh Depression Scale and the HADS-A in pregnancy: Differentiating between transient and enduring distress. *Journal of Affective Disorders*, 141, 213–221.
- Matthey, S., Valenti, B., Souter, K., & Ross-Hamid, C. (2013). Comparison of four self-report measures, and a generic mood question, to screen for anxiety during pregnancy in English-speaking women. *Journal of Affective Disorders*, 148, 347–351. (+ corrigendum: Matthey, S. (2014). *Journal of Affective Disorders*, 155, 307)
- Meijer, J. L., Beijers, C., Van Pampus, M. G., Verbeek, T., Stolk, R. P., Milgrom, J., ... Burger, H. (2014). Predictive accuracy of Edinburgh Postnatal Depression Scale assessment during pregnancy for the risk of developing postpartum depressive symptoms: A prospective cohort study. *BJOG*, 121, 1604–1610.
- Muzik, M., Klier, C. M., Rosenblum, K. L., Holzinger, A., Umek, W., & Katschnig, H. (2000). Are commonly used self-report inventories suitable for screening postpartum depression and anxiety disorders?. *Acta Psychiatrica Scandinavica*, 102, 71–73.
- NICE - National Institute for Clinical Excellence. (2014). *Antenatal and postnatal mental health: Clinical Management and Service Guidance Guidelines*. NICE clinical guideline CG192. London, England: National Institute for Health and Clinical Excellence.
- O'Hara, M. W., Stuart, S., Watson, D., Dietz, P. M., Farr, S. L., & D'Angelo, D. (2012). Brief scales to detect postpartum depression and anxiety symptoms. *Journal of Women's Health*, 21, 1237–1243.
- Pearson, R. M., Evans, J., Kounali, D., Lewis, G., Heron, J., Ramchandani, P. G., ... Stein, A. (2013). Maternal depression during pregnancy and the postnatal period: Risks and possible mechanisms for offspring depression at age 18 years. *JAMA Psychiatry*, 70, 1312–1319.
- Petrozzi, A., & Gagliardi, L. (2013). Anxious and depressive components of Edinburgh Postnatal Depression Scale in maternal postpartum psychological problems. *Journal of Perinatal Medicine*, 41, 343–348.
- Rosenthal, J. A. (1996). Qualitative descriptors of strength of association and effect size. *Journal of Social Service Research*, 21, 37–59.
- Somerville, S., Dedman, K., Hagan, R., Oxnam, E., Wettinger, M., Byrne, S., & Page, A. C. (2014). The Perinatal Anxiety Screening Scale: Development and preliminary validation. *Archives of Women's Mental Health*, 17, 443–454.
- Stein, A., Pearson, R. M., Goodman, S. H., Rapa, E., Rahman, A., McCallum, M., ... Pariente, C. M. (2014). Effects of perinatal mental disorders on the fetus and child. *The Lancet*, 384(9956), 1800–1819.
- Whooley, M. A., Avins, A. L., Miranda, J., & Browner, W. S. (1997). Case finding instruments for depression: Two questions as good as many. *Journal of General Internal Medicine*, 12, 439–445.
- Woody, C. A., Ferrari, A. J., Siskind, D. J., Whiteford, H. A., & Harris, M. G. (2017). A systematic review and meta-regression of the prevalence and incidence of perinatal depression. *Journal of Affective Disorders*, 219, 86–92.