

Review

Anxiety scales used in pregnancy: systematic review

Andrea Sinesi, Margaret Maxwell, Ronan O'Carroll and Helen Cheyne

Background

Anxiety disorders and self-reported symptoms are highly prevalent in pregnancy. Despite their negative impact on maternal and child outcomes, uncertainty remains regarding which symptoms can be considered accurate indicators of antenatal anxiety.

Aims

To examine and synthesise the evidence in relation to the psychometric properties and content of self-report scales used to detect anxiety symptoms in pregnant women.

Method

A systematic search was carried out and the methodological quality of all included studies was assessed. Only those achieving a rating of good or excellent were considered in a synthesis of the best available evidence.

Results

Several anxiety symptoms and domains were identified as promising for screening for general antenatal anxiety and pregnancy-related anxiety, including elevated levels of worry, symptoms of panic, fear of childbirth and excessive worries about the baby's health.

Conclusions

This review contributes to the existing knowledge by identifying a number of anxiety symptoms that can be considered psychometrically robust indicators of antenatal anxiety.

Declaration of interest

None.

Keywords

Pregnancy; anxiety disorders; screening; psychometric properties; pregnancy-specific anxiety.

Copyright and usage

© The Royal College of Psychiatrists 2019. This is an Open Access article, distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives licence (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is unaltered and is properly cited. The written permission of Cambridge University Press must be obtained for commercial re-use or in order to create a derivative work.

Anxiety during pregnancy is estimated to affect between 15 and 23% of women and is associated with increased risk for a range of negative maternal and child outcomes.^{1–3} This has led to growing attention in research^{4,5} and clinical guidelines⁶ over recent years. Antenatal anxiety has been consistently found to be a strong predictor of postnatal anxiety and depression.^{7–11} It has also been linked to adverse birth and child development outcomes, including low birth weight,^{12,13} premature birth^{3,14,15} and detrimental effects on neurodevelopmental, cognitive and behavioural child outcomes.^{4,16,17} Adverse child developmental outcomes found to be associated with antenatal anxiety include, for example, increased risk of language delay,¹⁶ attention-deficit hyperactivity disorder¹⁶ and poorer emotional regulation.¹⁷

Assessing anxiety in pregnancy

The importance of promoting the detection of women experiencing antenatal anxiety has been reflected in recent clinical guidelines. In the UK, the National Institute for Health and Care Excellence (NICE) guidance on perinatal mental health⁶ has for the first time recommended considering use of two screening questions (Generalised Anxiety Disorder scale, GAD-2)¹⁸ for the case-identification of anxiety in pregnant and postnatal women, and the most recent Scottish guidelines have also called for further research in this area.¹⁹ However, the evidence for recommending the GAD-2 is primarily based on its good screening accuracy in the general population,²⁰ with a very limited evidence base in perinatal populations. Although clinical diagnostic interviews are the optimal method of assessment for anxiety disorders, self-report rating scales such as the GAD-2 are often preferred in busy clinical practice and research because of their brevity.²¹

A recent systematic review found that self-reported anxiety symptoms during pregnancy had a pooled prevalence of 22.9% across trimesters.¹ For anxiety disorders based on DSM or ICD diagnostic criteria^{22,23} the overall prevalence was 15.2%. Similar prevalence rates were reported in a number of studies showing that problematic anxiety symptoms affect approximately 15% of women, both in early pregnancy⁵ and in later stages.^{2,24} High levels of self-reported symptoms, as opposed to anxiety disorders, are of relevance as they have also been shown to be associated with negative maternal and child outcomes.^{8,15} In research settings, antenatal anxiety has been measured with a heterogeneity of self-report scales, often in the absence of evidence of their psychometric accuracy in pregnant populations.²⁵

Screening for antenatal anxiety using scales developed for the general population is problematic for various reasons, partly as a result of the unique nature of pregnancy. One of the main concerns relates to the emphasis of many self-report measures of general anxiety on somatic symptoms and their potential confounding role when questions around physical symptoms are used to screen for anxiety during pregnancy.^{26,27} For instance, questions regarding sleep disturbances or palpitations, which are relatively common during pregnancy, may potentially lead to inflated scores. The assessment of antenatal anxiety is further complicated by the fact that anxiety symptoms that women can experience in pregnancy are not limited to the range of anxiety disorders determined by formal diagnostic criteria.^{22,23}

Pregnancy-specific anxiety

The occurrence of pregnancy-specific anxiety has been proposed as a distinct syndrome²⁸ and a number of studies have investigated this unique anxiety type.^{29–31} This emerging construct refers to a

particular anxiety response related to a current pregnancy, which can include fears and worries around labour and delivery, the health of the baby and expected changes in a woman's role.³² There is now good evidence of the clinical distinctiveness of pregnancy-specific anxiety,^{33,34} and some studies indicate that pregnancy-specific anxiety may be a stronger predictor of negative child outcomes than general antenatal anxiety.³⁴ However, women who may be significantly anxious because of pregnancy-related concerns might not meet the diagnostic criteria for a DSM/ICD anxiety disorder and consequently go unrecognised.

Aims

Recent reviews on the psychometric properties of scales to measure perinatal anxiety have highlighted this gap and the lack of anxiety scales with sound psychometric properties for use with pregnant women.^{25,35,36} However, none of these reviews have examined the content of measures with published psychometric data in pregnant populations. Consequently, it remains crucial to establish which symptoms can be considered reliable and valid indicators of maternal antenatal anxiety.

The aim of the present paper was to systematically examine and synthesise both the psychometric properties and content of self-report scales used to assess anxiety in pregnancy in order to identify a core set of anxiety symptoms and anxiety domains with established psychometric properties in pregnant populations. This was achieved by conducting a systematic review of studies reporting at least one psychometric property (i.e. one aspect of reliability or validity) of a self-report measure used to assess antenatal anxiety and by appraising and summarising the best available evidence in the form of a narrative synthesis.

Method

The review was conducted based on guidance for undertaking reviews of clinical tests from the Centre of Reviews and Dissemination³⁷ and COSMIN (CONsensus-based Standards for the selection of health status Measurement INSTRUMENTS) recommendations for systematic reviews of measurement properties,³⁸ and is reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement.³⁹ Ethical approval was not required as the study only involved secondary analysis of anonymised data.

Search strategy and selection criteria

Computerised searches were performed to query the following electronic bibliographic databases: MEDLINE, PsycINFO, and the Cumulative Index to Nursing and Allied Health Literature (CINAHL). The initial objective of the review was to locate primary research articles reporting psychometric properties of self-report rating scales used to assess anxiety symptoms in a pregnant population.

The databases were searched from 1991 up to and including February 2017 and searches were restricted to articles published in peer-reviewed journals and available in English. A combination of four main themes was used in the search. Specifically, the major concepts searched were 'anxiety', 'pregnancy', 'measurement' and 'psychometrics' and search terms included both free text and Medical Subject Headings (MeSH) terms. Major concepts and related synonyms for the four main themes were searched in the title and abstract fields, with several key terms also searched as a major concept within each database (see supplementary Appendix 1 available at <https://doi.org/10.1192/bjo.2018.75>).

Reference lists and citation records of papers included in the review were also inspected for potential inclusion of additional studies. Reports, commentaries, conference proceedings and other grey literature were not searched. Methodological search filters were not applied as there is evidence that, because of the variety of designs used in studies of diagnostic or screening test accuracy, applying methodological filters is likely to result in the omission of a significant number of relevant studies.^{40,41} A predefined list of inclusion and exclusion criteria was applied in relation to type of study, population, construct of interest and type of measurement. A complete list of inclusion and exclusion criteria is provided in the Appendix.

Study selection and data extraction

All articles resulting from the electronic bibliographic database searches were imported into RefWorks and duplicates were removed. Titles and abstracts of articles resulting from the initial search were reviewed to identify potentially relevant studies. When there was an indication that an article may have met the inclusion criteria for the review, the full-text publication was obtained and reviewed. The lead reviewer (A.S.) screened titles and abstracts of all retrieved articles to determine their appropriateness for inclusion in the review. A second reviewer (H.C.) independently screened a sample (10%) of all retrieved articles to establish an index of interrater agreement determined as per cent agreement,⁴² which was 98% for titles and abstracts screened by both reviewers. Discrepancies were discussed and resolved by applying the relevant study eligibility criteria to reach consensus.

The PRISMA flow diagram³⁹ was used to document the different stages of the study selection process (Fig. 1). In relation to data extraction, the full-text article of all studies included in the review was inspected and the full version of the rating scale used was obtained in order to extract information relevant to the review. Data extraction forms and summary tables were developed and piloted on a small number of studies ($n = 6$) identified as eligible for inclusion at an early stage of the review.

For each included study the following information was extracted: (a) author/s, (b) year of publication, (c) country, (d) name of index test, (e) sample size, (f) timing of assessment (expressed as trimester or mean gestational week), (g) construct of interest. For each of the rating scales, we extracted: (a) number of items, (b) type and number of response options (for example Likert scale, dichotomous), (c) time frame assessed (for example past week, past month), (d) score range, (e) total possible score, (f) cut-off score (if available). In order to determine which psychometric properties were evaluated in each study, the COSMIN taxonomy and definitions of measurement properties were used.⁴³ The following psychometric properties were extracted: internal consistency reliability, construct validity, convergent and discriminant validity, structural (i.e. factorial) validity and criterion validity. Definitions of all psychometric properties examined in this review and their corresponding indexes are presented in supplementary Appendix 2.

Quality assessment

An assessment of the methodological quality of each study included in the review was conducted using the COSMIN checklist, specifically developed to evaluate the study quality and risk of bias in systematic reviews of studies on the measurement properties of health measurement instruments.⁴⁴ In this review, five of the nine possible boxes in the checklist were employed as they were considered to be relevant to evaluate the methodological quality of studies assessing the construct of anxiety in pregnancy.

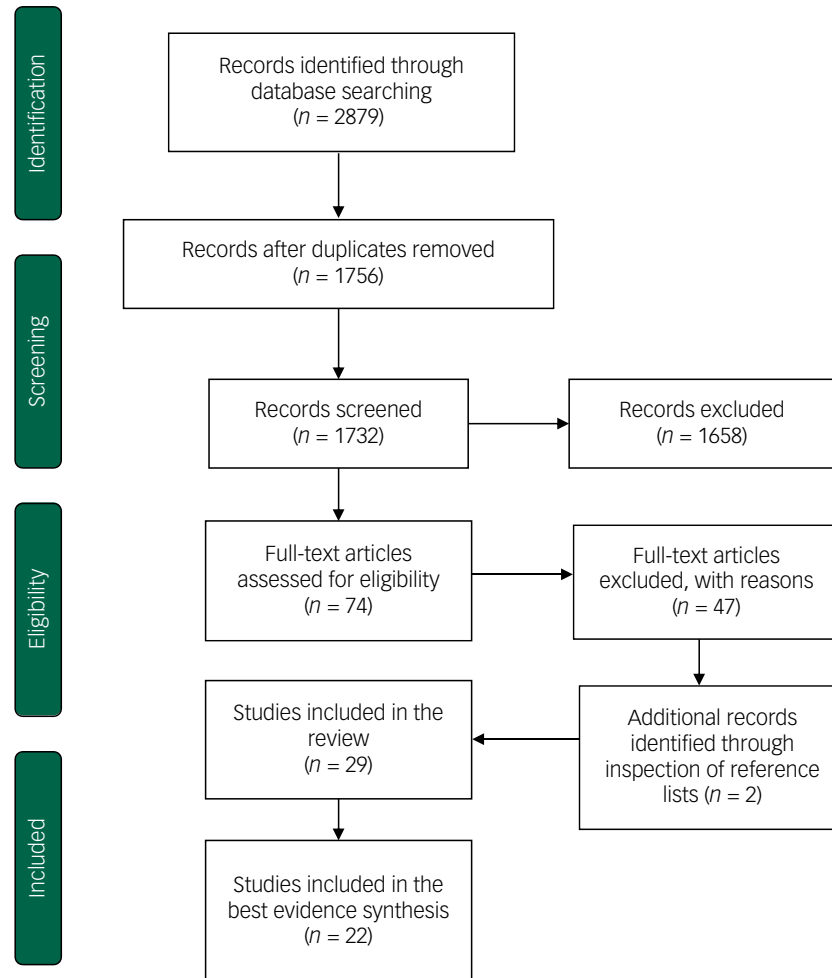


Fig. 1 PRISMA flow diagram of the selection process (based on Moher *et al.*).

Specifically, these were box A (internal reliability), D (content validity), E (structural validity), F (hypotheses testing) and H (criterion validity). Each measurement property is scored on a four-point rating scale as ‘poor’, ‘fair’, ‘good’ or ‘excellent’. An overall score for the methodological quality of a study is determined by using a ‘worse score counts’ system.⁴⁵ The lead reviewer (A.S.) performed the quality assessment for all studies included in the review, with the second reviewer (H.C.) assessing a random sample of studies ($n=5$) to confirm the accuracy of the scoring system. It was decided that only studies that achieved an overall rating of good or excellent were considered in the best-evidence synthesis in order to guarantee the quality of the conclusions reached by the review.

Best-evidence synthesis

The main aim of this review was to examine the psychometric properties and content of anxiety measures used in pregnancy, both at the scale and at the item level, in order to identify specific items (i.e. questions) or anxiety domains with established psychometric properties in this population. A synthesis of the best available evidence is presented for each scale in a narrative form, as the considerable differences across studies in relation to measure used, sample size, time of administration and type of reliability or validity reported precluded a meta-analysis. At the scale level, the psychometric properties discussed above were examined and synthesised. The number of studies, their methodological quality and the consistency of findings were taken into account.

Specifically, the following criteria were used to classify the strength of evidence from one or more studies, based on COSMIN recommendations for quality criteria:⁴⁶ (a) strong evidence: consistent findings in multiple studies of good or excellent methodological quality or in one study of excellent quality, (b) moderate evidence: consistent findings in multiple studies of good or excellent quality, except for one study with contrasting findings, (c) limited evidence: one study of good methodological quality, and (d) unclear or conflicting evidence: contrasting results in multiple studies of good quality. Only items and anxiety domains with moderate or strong evidence of being accurate indicators of anxiety symptoms in pregnancy were considered psychometrically sound in assessing antenatal anxiety.

At the item level, the analysis was primarily based on factor analysis, and specifically on the examination and comparison of coefficients of item loadings on specific anxiety factors for each scale. In psychometrics, the examination of item loadings is recommended in order to determine which items within a scale possess the strongest psychometric properties in terms of their discriminative power,⁴⁷ and can be therefore considered to detect an important aspect of the construct assessed.⁴⁸ Factor analysis is used to reduce variables (i.e. single items) that share common variance into set of clusters (i.e. factors).⁴⁹

In this review, the criteria proposed by Tabachnick & Fidell⁵⁰ and listed as follows were adopted to evaluate the strength of item loading coefficients: (a) 0–0.44, poor; (b) 0.45–0.54, fair; (c) 0.55–0.62, good; (d) 0.63–0.70, very good; (e) >0.70, excellent. Only items that

showed very good or excellent loadings (i.e. 0.63 or above), and for which the strength of evidence from one or multiple studies was moderate or strong according to the criteria discussed above, were considered to be psychometrically sound in measuring anxiety symptoms in pregnancy. When items forming a factor were found to be particularly homogeneous in relation to their content, the entire dimension or domain that the factor represented rather than individual items was selected as a domain identified as psychometrically sound. Secondary indexes that were examined at the item level when factor analysis was not conducted were the correlations between individual items and the remainder of items within a scale (corrected item-total correlations) and item discrimination parameters for analyses based on item-response theory models.

Results

The initial search yielded 2879 citations, which were reduced to 1756 following de-duplication. The titles and abstracts of remaining articles were screened for potentially eligible studies, resulting in 74 publications for which the full-text article was retrieved. At this stage 47 studies were excluded and 2 publications were added from hand searches of reference lists of included studies. This resulted in a final sample of 29 studies included in the review.^{8,11,28,31,33,51–74} The main reasons for excluding studies after retrieving the full text were: (a) no psychometric data available, (b) construct of interest different from inclusion criteria (for example antenatal stress, general mental health), (c) study participants recruited exclusively from high-risk samples. The study selection process is summarised in the PRISMA flowchart (Fig. 1).

The 29 included studies used 9 different scales as index tests to measure antenatal anxiety. The most commonly reported psychometric properties were internal consistency reliability ($n = 27$; 93% of studies), convergent validity ($n = 21$; 72%) and structural validity ($n = 16$; 55%). The characteristics of included studies are presented in Table 1. Included studies showed a considerable degree of heterogeneity in relation to the construct assessed (i.e. general anxiety versus an anxiety disorder versus pregnancy-specific anxiety), gestational age of participants, sample size and type of psychometric properties reported.

As discussed in the Method, a quality assessment of all included studies was performed and only studies achieving a rating of good or excellent in relation to their methodological quality and risk of bias were included in the best-evidence synthesis. Seven studies were given a rating of poor^{59,68,72} or fair^{56,62,65,70} for their methodological quality and were thus not considered in the synthesis. The quality assessment of all 29 studies included in the review is presented in the supplementary Table 1. Further details about the criteria used to rate the methodological quality of all studies included are available from the corresponding author on request.

Best-evidence synthesis

Following an assessment of the methodological quality of all studies, 22 were included in the best-evidence synthesis phase of the review.^{8,11,28,31,33,51–55,57,58,60,61,63,64,66,67,69,71,73,74} This section discusses the findings from these studies through an examination of the psychometric properties of each scale and a critical analysis of the content of their items and anxiety domains found to be psychometrically sound for the assessment of antenatal anxiety. This analysis was carried out accordingly to the criteria discussed in detail in the Method. For clarity of exposition, a synthesis is presented here separately for each scale, whereas the Discussion summarises the general findings of the review.

Edinburgh Postnatal Depression Scale (EPDS) – Anxiety subscale

The EPDS⁷⁵ is a ten-item self-report questionnaire originally developed to screen for postpartum depression, which asks respondents about symptoms of depression experienced in the previous week. Because of the lack of items specific to the postpartum period, the EPDS has also been validated for use with pregnant women.^{76,77} Although the EPDS was developed as a unidimensional measure of depression, it was included in this review because of growing evidence that it contains a separate subscale measuring anxiety rather than depressive symptoms, in both antenatal and postnatal populations.^{78–80}

Six studies included in this review examined the psychometric properties of the EPDS anxiety subscale in pregnant women. All studies except one⁷⁰ achieved an overall methodological quality rating of good^{52,67,71} or excellent^{54,63} and were thus included in the best-evidence synthesis. Four of the five studies examined the factor structure of the EPDS to investigate whether the existence of an anxiety subscale could be confirmed.

Brouwers and colleagues⁵² performed exploratory factor analysis (EFA) of EPDS scores in women in their second trimester of pregnancy. The EFA revealed three components within the EPDS, namely two separate depressive (items 1, 2, 8) and anxiety (items 3, 4, 5) symptoms subscales and a third component consisting only of item 10 ('The thought of harming myself has occurred to me'). However, this third factor was not included in the final factor solution as the authors argued that a single-item loading could not plausibly identify a distinct latent factor.⁵² A two-factor solution, comprising separate depression and anxiety subscales, was therefore proposed. The three items of the anxiety subscale (item 3 'I have blamed myself unnecessarily when things went wrong', item 4 'I have been anxious or worried for no good reason', item 5 'I have felt scared or panicky for no very good reason') were the only ones, among the ten EPDS items, with item loadings on a single factor above the predefined cut-off of 0.63, ranging from 0.68 (item 3) to 0.73 (item 4). An examination of their content appears to indicate that these questions, all loading highly on a single factor, tap important affective and cognitive components of anxiety (for example feeling panicky or worried).

Similar findings were reported by Jomeen & Martin⁶³ in women in their first trimester of pregnancy. EFA resulted in a three-factor solution that included depression and anxiety dimensions, and the same third factor identified by Brouwers and colleagues.⁵² The items loading significantly (>0.63 , range 0.73–0.85) onto the anxiety subscale were entirely consistent (items 3, 4, 5) with those identified in the previous study.⁵² The authors then conducted confirmatory factor analysis (CFA), a more refined data reduction technique than EFA,⁸¹ and tested various predefined factor models including the original unidimensional depression model,⁷⁵ as well as both a two- and a three-factor model identified by Brouwers and colleagues.⁵² Results from the CFA revealed once again a clear superiority of the two-factor solution, thus confirming the previous finding that the EPDS both in early and in mid-pregnancy consistently measures two distinct dimensions of depression and anxiety.

A further study included in this review⁶⁷ used the three-item EPDS anxiety subscale (EDS-3A) identified in previous studies to examine its criterion and convergent validity in pregnancy when compared with other anxiety measures. The EDS-3A performed better than both the anxiety subscale of the Hospital Anxiety and Depression Scale (HADS-A⁸²) and the Pregnancy Related Anxiety Questionnaire-Revised (PRAQ-R²⁸) in detecting women with an anxiety disorder as determined by DSM diagnostic criteria. Furthermore, the EDS-3A showed a moderately high correlation

Table 1 General characteristics of studies included in the review

Authors	Year	Participants, <i>n</i>	Gestational age	Country	Index test	Time frame assessed	Target construct
Austin <i>et al</i> ¹¹	2007	748	3rd trimester	Australia	BMWS	'General experience'	Worry
Bayrampour <i>et al</i> ⁵¹	2014	3021	2nd trimester	Canada	STAI (3 six-item short forms)	State (present time); trait (general feelings)	State/trait anxiety
Brouwers <i>et al</i> ⁵²	2001	197	24 weeks	Netherlands	EPDS-A	Previous week	General anxiety
Carmona Monge <i>et al</i> ⁵³	2012	285	Mean 14.1 weeks	Spain	CWS	Present time	Worry during pregnancy
Coates <i>et al</i> ⁵⁴	2016	5551	18 and 32 weeks	UK	EPDS-A	Previous week	General anxiety
Fenaroli & Saita ⁵⁵	2013	522	27–35 weeks	Italy	W-DEQ	Current expectations about childbirth	Fear of childbirth
Garthus-Niegel <i>et al</i> ⁵⁶	2011	1642	32 weeks	Norway	W-DEQ	Current expectations about childbirth	Fear of childbirth
Gourounti <i>et al</i> ⁵⁷	2012	132	11–14 weeks	Greece	CWS	Present time	Worry during pregnancy
Grant <i>et al</i> ⁸	2008	100	35–39 weeks	Australia	STAI (state and trait forms)	State (present time); trait (general feelings)	State/trait anxiety
Green <i>et al</i> ⁵⁸	2003	1207	1st/2nd/3rd trimester	UK	CWS	Present time	Worry during pregnancy
Haines <i>et al</i> ⁵⁹	2015	1410	2nd trimester	Australia	W-DEQ	Current expectations about childbirth	Fear of childbirth
Huizink <i>et al</i> ²⁸	2004	172	1st/2nd/3rd trimester	Netherlands	PRAQ-R	Present time	Pregnancy-related anxiety
Huizink <i>et al</i> ³¹	2016	1144	24 and 34 weeks	Finland	PRAQ-R2	Present time	Pregnancy-related anxiety
Johnson & Slade ⁶⁰	2002	424	3rd trimester	UK	W-DEQ	Current expectations about childbirth	Fear of childbirth
Jomeen & Martin ⁶¹	2004	101	Mean 13.57 weeks	UK	HADS-A	Previous week	General anxiety
Jomeen & Martin ⁶²	2005	129	Mean 13.86 weeks	UK	CWS	Present time	Worry during pregnancy
Jomeen & Martin ⁶³	2005	101	Mean 13.57 weeks	UK	EPDS-A	Previous week	General anxiety
Karimova & Martin ⁶⁴	2003	100	12 and 34 weeks	UK and Uzbekistan	HADS-A	Previous week	General anxiety
Levin ⁶⁵	1991	266	2nd trimester	USA	PAS	Present time	Pregnancy-related anxiety
Marteau & Bekker ⁶⁶	1992	200	Gestational age not reported	UK	STAI (six-item short form)	Present time	State anxiety
Matthey <i>et al</i> ⁶⁷	2013	132	Mean 14.9 weeks	Australia	EPDS-A, HADS-A, PRAQ-R	Various time frames	General and pregnancy-related anxiety
Öhman <i>et al</i> ⁶⁸	2003	200	8–42 weeks	Sweden	CWS	Present time	Worry during pregnancy
Petersen <i>et al</i> ⁶⁹	2009	344	Mean 31.4 weeks	Germany	CWS	Present time	Worry during pregnancy
Simpson <i>et al</i> ⁷⁰	2014	240	1st trimester	Canada	EPDS-A, GAD-7	Previous week	General anxiety
Swalm <i>et al</i> ⁷¹	2010	4706	10–42 weeks	Australia	EPDS-A	Previous week	General anxiety
Tendais <i>et al</i> ⁷²	2014	148	1st/2nd/3rd trimester	Portugal	STAI-S	Present time	State anxiety
Westermeng <i>et al</i> ³³	2015	6004	Mean 19.8 weeks	Netherlands	PRAQ-R	Present time	Pregnancy-related anxiety
Wijma <i>et al</i> ⁷³	1998	196	32 weeks	Sweden	W-DEQ	Current expectations about childbirth	Fear of childbirth
Zhong <i>et al</i> ⁷⁴	2015	946	Mean 9.6 weeks	Peru	GAD-7	Previous two weeks	Generalised anxiety disorder

BMWS, Brief Measure of Worry Severity; STAI, State-Trait Anxiety Inventory; EPDS-A, Edinburgh Postnatal Depression Scale – Anxiety subscale; CWS, Cambridge Worry Scale; W-DEQ, Wijma Delivery Expectancy/Experience Questionnaire; PRAQ-R and PRAQ-R2, Pregnancy-Related Anxiety Questionnaire- Revised; HADS-A, Hospital Anxiety and Depression Scale – Anxiety subscale; PAS, Pregnancy Anxiety Scale; GAD-7, Generalised Anxiety Disorder – 7.

with the HADS-A ($r = 0.68$) and a low to moderate correlation with the PRAQ-R ($r = 0.23$), which may be interpreted as an indication that the three measures tap into different aspects of antenatal anxiety.

Although a potential limitation of the three studies reported above is their relatively small number of participants ($n < 200$), the existence of an anxiety subscale within the EPDS was further confirmed in two subsequent studies with much larger numbers of participants ($n > 4000$). Swalm and colleagues⁷¹ examined the EPDS factor structure in Australian women across the three trimesters of pregnancy. A two-factor solution consisting of anxiety and depression components was found once more to be optimal, accounting for 55% of the score variance (anxiety subscale 29.4%; depression subscale 25.4% of the total variance). Moreover, an analysis of individual item loadings confirmed that items 3, 4 and 5 were the only ones with loadings higher than 0.63 on the anxiety subscale (range 0.75–0.78).

A recent UK population-based study⁵⁴ conducted both EFA and CFA on a large number of participants at two time points (18 and 32 weeks' gestation). Although both EFA and CFA indicated a three-factor model as the best factor solution, this was primarily because of the 'depression' factor that was split into an anhedonia (items 1 and 2) and a depression (items 7–10) factor. Importantly, this was the only study in which item 3 'I have blamed myself unnecessarily when things went wrong' (0.56) did not reach the pre-defined item loading coefficient of 0.63.

In summary, according to the criteria previously discussed to evaluate the strength of evidence in relation to the psychometric properties of reviewed scales, item 3 of the EPDS showed moderate evidence of its psychometric value, and items 4 and 5 demonstrated strong evidence of being psychometrically sound in assessing antenatal anxiety, as their item loadings on the anxiety subscale consistently exceeded the 0.63 cut-off in all reviewed studies.

HADS – Anxiety subscale

The HADS⁸² is a widely popular screening tool⁸³ originally developed to assess anxiety and depression in non-psychiatric patients. This 14-item measure consists of two subscales (anxiety: HADS-A; depression: HADS-D), both comprising seven items and enquiring about feelings over the past week with four response options.⁸² It is particularly important to establish the psychometric properties of the HADS when used in the antenatal period, as a considerable number of studies have used this screening tool to assess anxiety and depression levels in pregnant women, including in recent years.^{5,84}

Three studies included in this review examined psychometric aspects of the HADS in a pregnant population.^{61,64,67} They all achieved a rating of good in relation to their methodological quality. Karimova & Martin⁶⁴ investigated the factor structure of the HADS in the third trimester of pregnancy by conducting EFA of HADS scores in nulliparous women, and a *post hoc* factor analysis revealed a two-factor solution. Specifically, six of the seven HADS-D items loaded higher on one factor and an equal number of HADS-A items loaded higher on a second factor. However, there was significant overlapping of item loadings on the two subscales, with only four HADS-A items (item 3 'I get a sort of frightened feeling as if something awful is going to happen'; item 5 'Worrying thoughts go through my mind'; item 9 'I get a sort of frightened feeling like "butterflies" in the stomach' and item 13 'I get sudden feelings of panic') loading above 0.63 on the anxiety factor. The authors therefore concluded that the seven-item HADS-A and HADS-D subscales do not reliably distinguish between anxiety and depressive symptoms in pregnancy.

Table 2 Item loading coefficients of the Hospital Anxiety and Depression Scale – Anxiety subscale (HADS-A) subscale in Karimova & Martin and Jomeen & Martin

HADS-A items	Karimova & Martin (2003) Anxiety factor (factor 2)	Jomeen & Martin (2004) Anxiety factor (factor 2)
1. I feel tense or wound up	0.18	0.31
3. I get a sort of frightened feeling as if something awful is going to happen	0.67 ^a	0.74 ^a
5. Worrying thoughts go through my mind	0.78 ^a	0.69 ^a
7. I can sit at ease and feel relaxed	0.33	0.07
9. I get a sort of frightened feeling like 'butterflies' in the stomach	0.65 ^a	0.57
11. I feel restless as if I have to be on the move	0.57	0.36
13. I get sudden feelings of panic	0.67 ^a	0.75 ^a

a. Item loadings of 0.63 or above.

A further study was conducted by Jomeen & Martin⁶¹ on women in early pregnancy. Both EFA and CFA revealed a three-factor solution that confirmed that the HADS in pregnancy is not a bi-dimensional measure of anxiety and depression. However, a comparison of individual item loadings of the HADS anxiety subscale in the two studies was carried out in this review to examine psychometric information for each individual item within the HADS anxiety subscale. This is presented in Table 2.

The observation that three items of the HADS-A (items 3, 5, 13) are the only ones to reach an item loading above 0.63 on the anxiety factor in both studies is of particular importance. Although the two studies reached the conclusion that the seven-item HADS-A as a whole is not a psychometrically sound measure of anxiety in pregnancy, the three HADS-A items identified here showed a consistent pattern across the two studies, with significantly similar loadings on the anxiety factor. These items would therefore appear to have good psychometric value in assessing specific anxiety symptoms in pregnancy.

A subsequent study⁶⁷ compared the screening performance of the HADS-A with diagnosis of an anxiety disorder according to DSM criteria. The authors found that high anxiety scores on the HADS-A, defined as the top 15% of scores, had poor concordance (34%) with formal diagnosis of an anxiety disorder. The poor concordance with DSM diagnoses seems to confirm the previous findings indicating that the seven-item HADS anxiety subscale as a whole is not a reliable screening tool to assess anxiety in pregnancy. However, based on the evidence provided by the two studies discussed above on the factor structure of the HADS, we conclude that the three identified items represent a shortened version of the HADS-A which, unlike the entire HADS-A, has good evidence of its psychometric properties to measure antenatal anxiety.

State-Trait Anxiety Inventory (STAI)

The STAI⁸⁵ comprises two subscales, each composed of 20 items. It is based on a model of anxiety that distinguishes between state and trait anxiety.⁸⁶ State anxiety refers to the situation-specific, transient component of anxiety. Conversely, trait anxiety reflects a relatively stable personality trait, a dispositional anxiety proneness.⁵⁸ Response options range from one (not at all) to four (very much so) for both the state and trait form, and each scale includes ten anxiety-present (for example 'I am worried') and ten anxiety-absent (for example 'I feel secure') items. The state form asks participants about feelings at the present time, whereas the trait form

enquires about how a respondent generally feels. The STAI has been widely validated in the general population¹¹ and is one of the most common measures used in research to assess anxiety in perinatal women.³⁵

This review located four studies reporting psychometric properties of the STAI in pregnant populations, one of which⁷² was scored poor in relation to its methodological quality. Both the state and trait form of the STAI were used in an Australian study by Grant and colleagues⁸ on women in the third trimester of pregnancy. Internal consistency was found to be high for the full version of the scale, with a Cronbach's alpha (α) of 0.95. A structured diagnostic interview was also used (Mini International Neuropsychiatric Interview)⁸⁷ to identify women meeting DSM-IV diagnostic criteria for an anxiety disorder. The authors found a cut-off score of 40 to yield the highest accuracy in identifying women with a diagnosed anxiety disorder, with a sensitivity of 80.9% and a specificity of 79.7%.⁸ However, they also acknowledged the limited generalisability of the findings because of the relatively small number of participants. The study did not provide any psychometric data at the item level and it was thus not possible to reach conclusions on the psychometric qualities of individual items measuring specific symptoms.

A further study⁶⁶ tested various shortened versions of the STAI-S form to determine the smallest subset of items that preserved high correlations ($r > 0.90$) with the original, 20-item STAI-S. They found that a six-item version produced scores comparable with the full version ($r > 0.94$) while retaining a good level of internal consistency ($\alpha = 0.82$). The six items selected were the ones with the highest correlations with the remaining 19 items of the STAI-S (i.e. corrected item-total correlations). Specifically, the authors identified three anxiety-present and three anxiety-absent items, corresponding to the following emotional states: calm, tense, upset, relaxed, content and worried. This is a significant finding, as it identifies a number of symptoms (i.e. feeling tense, upset or worried) that correlate highly with the 20-item STAI-S total score, providing an initial indication that these anxiety-present symptoms may be considered relatively accurate indicators of problematic anxiety in pregnancy.

This was confirmed in a further study by Bayrampour and colleagues⁵¹ that examined the psychometric properties of three six-item shortened versions of the STAI-S when compared with the full state form. The three short versions are the ones discussed above⁶⁶ and two other versions developed in non-perinatal populations. The six-item version by Marteau & Bekker⁶⁶ had the highest correlation with the sum score of the full form ($r = 0.94$). Furthermore, confirmatory factor analysis was conducted and the version by Marteau & Bekker⁶⁶ was found once more to consistently have the best values for all fit indexes considered, with the three anxiety-present items (i.e. feeling tense/upset/worried) all found to have coefficient item loadings above 0.63, a further indication of their psychometric soundness.

In sum, the three items from the STAI-S short form discussed above were identified in two studies of good methodological quality^{51,66} as potentially reliable indicators of anxiety symptoms during pregnancy.

GAD-7

The GAD-7¹⁸ was developed in 2006 as a brief screening measure for generalised anxiety disorder. Its original psychometric validation study, in a large number of primary care patients indicated very good screening accuracy in identifying people with a diagnosis of generalised anxiety disorder.¹⁸ The scale consists of seven items asking respondents about some of the core generalised anxiety disorder symptoms (for example excessive or persistent worry, trouble relaxing) experienced in the previous 2 weeks. As previously

discussed, the first two questions of the GAD-7 (GAD-2) have been recently recommended by NICE as a brief screening measure for anxiety in perinatal women.⁶

Only two studies examining the measurement properties of the GAD-7 in a pregnant population were identified by this review,^{70,74} and only one⁷⁴ achieved a satisfactory rating for its methodological quality. Importantly, this was one of the few included studies that performed assessment of a scale against a gold-standard clinical interview, the Composite International Diagnostic Interview,⁸⁸ to determine the criterion validity of the scale. In this antenatal sample at a cut-off score of seven or above, notably different from the cut-off of ten identified in the general population, the measure yielded moderately good sensitivity (73%) and specificity (67%).⁷⁴ Internal consistency was close to excellent ($\alpha = 0.89$).

Both EFA and CFA were conducted, and confirmed the unidimensional structure (i.e. a single factor) of the GAD-7 previously found in the general population.¹⁸ The results of the factor analysis indicated that the seven items loaded on a single factor with item loadings all exceeding 0.63. In order to identify which items provided the most accurate screening performance we thus examined the item discrimination parameters, which are based on item-response theory and indicate how well individual items differentiate between different levels of the target condition among respondents.⁸⁹ Two items showed considerably higher discrimination parameter estimates than the remaining ones. These were item 3 'Worrying too much about different things' (2.05) and item 2 'Not being able to stop or control worrying' (2.04), which clearly tap into the experience of pervasive or persistent worry typical of generalised anxiety disorder. All other items exhibited substantially lower discrimination parameter estimates. Considering that this study was of excellent methodological quality, the two identified items have consequently strong evidence of their psychometric value in the antenatal period.

Brief Measure of Worry Severity (BMWS)

A single study¹¹ was located reporting psychometric data of the BMWS⁹⁰ in pregnant women. Self-report scales assessing the construct of 'worry' were included in this review (Appendix 1) as worry is a core clinical feature of generalised anxiety disorder.^{22,91} A number of studies indicate that generalised anxiety disorder is the most common anxiety disorder in pregnancy^{1,72} and for this reason worry can be hypothesised to be an important dimension of the construct of antenatal anxiety. The BMWS was developed as a unidimensional measure of the functional impact and severity of worry.⁹⁰ It includes eight items assessing different aspects of worry. Respondents are asked to rate their general or usual experience of worrying, with four verbally anchored response options (not true at all – definitely true).⁹⁰

Austin *et al* aimed to determine whether the construct of worry as measured by the BMWS, defined as 'dysfunctional trait cognitive anxiety', was a significant predictor of postnatal depression.¹¹ Internal consistency was very good ($\alpha = 0.89$) and the BMWS also showed good convergent validity with the STAI trait ($r = 0.71$). Although psychometric properties of the scale at the item level were not reported, there was evidence that the construct of worry as measured by the BMWS is a reliable indicator of antenatal anxiety. First, the BMWS was found to have good construct validity in these pregnant participants, as it showed significant correlations with a number of other variables linked to a current episode of anxiety and depression.¹¹ Moreover, it was a better predictor of postnatal depression than the STAI-S after controlling for possible confounding factors. As the literature indicates that antenatal anxiety is a predictor of postnatal depression,^{7,8,10} it appears than

the BMWS taps into a core component of antenatal anxiety considering its good predictive validity.

Consequently, the construct of worry has strong evidence of being psychometrically robust according to the criteria used in this review (i.e. consistent findings in multiple studies of good or excellent methodological quality) as it was also identified as psychometrically sound in other studies previously discussed in this synthesis.

Cambridge Worry Scale (CWS)

The CWS is a 16-item measure assessing the extent and content of women's worries during pregnancy.⁵⁸ The 16 items in the CWS enquire both about worries specific to pregnancy, such as 'The possibility of miscarriage', 'The possibility of something being wrong with the baby' or 'Giving birth', and more general concerns including 'Money problems' and 'Your relationship with your family and friends'. Items are scored on a six-point Likert-type scale with verbally described anchors ranging from zero (not a worry) to five (major worry) and referring to the present time.⁵⁸

Six studies examining psychometric aspects of the CWS in a pregnant population were included in this review, four of which are considered here. The other two studies were rated as poor⁶⁸ or fair⁶² for their methodological quality.

Green and colleagues⁵⁸ were the first to investigate the structural validity (i.e. the factor structure) of the CWS. A longitudinal design was used in a large number ($n = 1207$) of British women completing the CWS at gestational weeks 16, 22 and 35. The authors analysed scores at these three time points by means of principal component analysis (PCA), a form of exploratory factor analysis. The PCA revealed a four-factor structure, consisting of the following factors: (a) socio-medical aspects of having a baby, (b) socio-economic issues, (c) health of mother and baby, and (d) relationships with partner, family and friends. This four-factor solution was subsequently replicated in all the other studies examined in this synthesis.^{53,57,69} This can be considered robust evidence of factorial stability of the CWS in different populations and stages of pregnancy.

The convergent validity of the CWS was examined by comparing it with STAI state and trait scores^{57,58,69} and with the anxiety subscale of the Symptom Checklist-90⁹² by Carmona Monge and colleagues.⁵³ Two of the four CWS subscales were found to have the highest correlations with state anxiety (STAI-S) scores across studies. These were the 'socio-medical' and the 'health of mother and baby' factors. For the purpose of this review, we specifically focused on these two factors, both because of their higher correlations with state anxiety and because the content of items in these subscales appears to reflect worries more closely related to pregnancy. Thus, an examination of individual item loadings for these two factors was carried out.

In relation to the 'socio-medical' subscale, one item ('Giving birth') was found to load above the predefined criterion of 0.63 in all studies, thus demonstrating strong evidence of its psychometric properties in assessing a major worry in pregnancy. Another three items showed moderate strength of evidence as they loaded above 0.63 on the 'socio-medical' subscale in all studies apart from one. Specifically, 'Internal examinations' had an item loading coefficient of 0.61 in Gourounti and colleagues,⁵⁷ but item loadings above 0.63 in all the other studies; 'Going to hospital' (0.68–0.79), apart from Gourounti and colleagues⁵⁷ (0.47); and 'Coping with the new baby' (0.65–0.68), except for the study by Petersen and colleagues,⁶⁹ in which its loading was 0.58.

An inspection of the second factor examined, 'Health of mother and baby', indicated two further items with loadings >0.63 in all the studies, namely 'The possibility of miscarriage', which ranged

between 0.75⁵⁸ and 0.85⁵³, and 'The possibility of something being wrong with the baby' (range 0.65–0.83^{53,58}). The other two items included in this subscale, 'Own health' and 'Health of someone else close', consistently loaded below the pre-defined cut-off.

In summary, three items of the CWS ('Giving birth', 'The possibility of miscarriage', 'The possibility of something being wrong with the baby') demonstrated strong evidence of their psychometric properties. Three further items ('Internal examinations', 'Going to hospital' 'Coping with the new baby') showed a moderate strength of evidence of their psychometric value in pregnancy.

Wijma Delivery Expectancy/Experience Questionnaire (W-DEQ – Version A)

The W-DEQ⁷³ was developed in the late nineties to assess the construct of fear of childbirth. Within the research literature on pregnancy-specific anxiety, fear of childbirth or tokophobia has emerged as a central dimension of pregnancy-specific anxiety.^{5,34,93} The W-DEQ Version A⁷³ includes 33 items enquiring about thoughts and feelings relating to the approaching childbirth, with six response options ranging from 'not at all' to 'extremely'.

Five studies included in the present review reported psychometric information on the W-DEQ in an antenatal population,^{55,56,59,60,73} and three studies achieved a good or excellent methodological quality rating.^{55,60,73} In the original development study of the W-DEQ,⁷³ internal consistency of the measure was excellent ($\alpha = 0.93$). The authors also provided good evidence of the face and construct validity of the W-DEQ, with all items formulated based on the clinical experience of the first two authors and incorporating women's input in the wording of items. The W-DEQ showed higher correlations with other anxiety measures than with extraversion or depression measures. However, these correlations were only moderate (STAI-T: $r = 0.54$; S-R Inventory of anxiousness: $r = 0.52$), thus showing a degree of conceptual overlap but also a sufficient level of variance left to indicate that the W-DEQ measures other than anxiety as a dispositional trait.⁷³

At the item level, item-total correlations were ranked and the authors examined the ten items with the highest ranking. Two domains of fear of childbirth, 'Negative feelings towards childbirth' and 'Fear of labour and delivery', were identified among the items more strongly correlated with the sum score, thus suggesting a stronger relation with the overall construct of fear of childbirth. As single items composing the W-DEQ are very specific to a given feeling or cognitive appraisal, we considered it appropriate to focus on domains of fear of childbirth rather than individual items.

Two other studies^{55,60} included in this synthesis conducted factor analysis of W-DEQ scores and found four distinct dimensions of the construct of fear of childbirth as measured by the scale. Johnson & Slade⁶⁰ named the four identified domains Fear, Lack of positive anticipation, Isolation and Riskiness. The latter two refer to feelings of isolation related to childbirth and to the extent to which women anticipate risks for the child during delivery. Fenaroli & Saita⁵⁵ also found a four-factor structure of the W-DEQ, and although the four domains were named with slightly different labels than those used by Johnson & Slade⁶⁰, the four factors were considerably similar and had a high degree of conceptual overlap. In this best-evidence synthesis two dimensions of pregnancy-specific anxiety, namely Fear of labour and delivery and Negative feelings towards childbirth (corresponding to Lack of positive anticipation in Fenaroli & Saita⁵⁵), were thus found to exhibit strong evidence of being psychometrically sound in assessing this specific aspect of antenatal anxiety. A third dimension (Fear for baby's health) showed moderate strength of evidence as, although it was identified in two studies,^{55,60} contrasting results were found in another study.⁷³

PRAQ-R and PRAQ-R2

This pregnancy-specific anxiety measure is composed of ten items assessing various manifestations of anxiety related to a current pregnancy. Each item asks about feelings at the present time and has five response options ranging from 'never' to 'very often'. Its original version (PRAQ)⁹⁴ consisted of 58 items and was developed based on previous anxiety measures.

The first study testing the psychometric properties of the PRAQ was carried out by Huizink and colleagues²⁸ who initially tested a revised, 34-item version (PRAQ-R²⁸) of the original PRAQ on 230 nulliparous women. The authors' aim was to examine the factorial structure of the PRAQ-R and test the hypothesis that pregnancy-specific anxiety could be differentiated from general anxiety by comparing STAI and PRAQ-R scores. They found that only between 8 and 27% of the PRAQ-R variance was accounted for by the index of general anxiety at different time points during pregnancy, with no linear association found between the two measures. This was interpreted as evidence of the distinctiveness of the pregnancy-specific anxiety construct²⁸ and highlighted once more that measures of general anxiety cannot be accurately used to identify women experiencing fears and worries specific to pregnancy.

The authors initially conducted EFA and removed a number of items because of high error variance, resulting in a final version comprising ten items (PRAQ-R). A subsequent CFA revealed that a solution with three factors provided the best fit to the data, with the three identified factors labelled by the researchers 'Fear of giving birth' (three items), 'Fear of bearing a physically or mentally handicapped child' (four items) and 'Concern about one's appearance' (three items). All individual items loaded on one of the factors above the cut-off of 0.63, except for one item (0.50), 'I am worried about not being able to control myself during labour and fear that I will scream'. Similarly to the approach used for the W-DEQ and discussed above, we considered the whole factors rather than individual items making up a given factor.

Two further studies^{31,33} included here tested the measurement properties of the PRAQ-R, and both replicated the previous finding of a three-factor structure of the PRAQ-R by means of CFA. As the original participants of the ten-item PRAQ-R were exclusively composed of nulliparous women, Westerneng and colleagues³³ aimed to test the factorial stability of the three-factor solution of the PRAQ-R²⁸ on a large ($n > 6000$) data-set of both nulliparous and parous women. This involved the deletion of item 8 'I am anxious about the delivery because I have never experienced one before', as it was not suitable for use with women who had already experienced childbirth. CFA confirmed the same three-factor structure of the original ten-item PRAQ-R with good indexes of fit to the data for both nulliparous and parous women.

Three factors were also found in a recent study³¹ that replaced item eight of the original PRAQ-R with the more generic 'I am anxious about the delivery' in order to preserve a ten-item scale while making it appropriate for all pregnant women irrespective of parity (PRAQ-R2)³¹. All item loadings were once more above 0.63 (range: 0.70–0.93) except for two items, 'I am worried about not being able to control myself during labour and fear that I will scream', similarly to Huizink and colleagues,²⁸ and 'I sometimes think that our child will be in poor health or will be prone to illnesses'.

In summary, across the three studies examined here^{28,31,33} eight items from the PRAQ-R were found to consistently have high loadings on one of three factors (i.e. pregnancy-specific anxiety domains). These three pregnancy-specific anxiety domains, namely 'Fear of

giving birth', 'Fear of bearing a physically or mentally handicapped child' and 'Concern about one's appearance', were all identified in studies of good or excellent methodological quality, thus providing strong evidence of being accurate indicators of pregnancy-specific anxiety.

Discussion

There are several important findings to this study. First, this review has identified a number of anxiety items and domains from existing self-report scales with demonstrated psychometric value when used to assess symptoms of anxiety in pregnant women. To the best of our knowledge, this is the first study to analyse the content of self-report anxiety measures used in the antenatal period and provide recommendations for the accurate assessment of maternal antenatal anxiety based on a systematic synthesis of published psychometric data.

A second, significant finding of this paper is that it highlights the scarcity of studies reporting psychometric properties of scales employed to measure anxiety in pregnancy. A considerable number of studies using self-report scales to assess antenatal anxiety were not included in this review as no measurement properties of the scale used were reported. It would appear that in most cases researchers have selected a given anxiety measure only based on its widespread use and good psychometric properties in the general population.³⁶ However, assuming that the measurement properties of a psychological scale developed for the general population are preserved in pregnancy is incorrect for various reasons discussed earlier in this paper (i.e. undue emphasis on somatic symptoms, lack of validated cut-off scores and norms for pregnant populations, role of pregnancy-specific anxiety).

A further limitation of the literature is that only a dearth of studies located by this review ($n = 5$)^{8,67,70,72,74} validated a measure against a reference 'gold' standard such as a structured diagnostic interview. Testing a scale against a reference standard provides evidence of the screening accuracy of a measure, also referred to as its criterion validity, arguably the single most important aspect of psychometric validation of a scale.⁴⁸

Perhaps even more surprisingly, only two studies^{70,74} were identified that examined the psychometric properties of the GAD-7 in a pregnant population, and only one⁷⁴ was found to have satisfactory methodological quality. As previously reported, the GAD-2 (i.e. the initial two questions of the GAD-7) is the measure currently recommended by NICE in the UK to screen for anxiety in pregnant women, followed by administration of the GAD-7 if a woman scores three or higher on the GAD-2.⁶ The only methodologically robust study providing psychometric information on the GAD-7 in a pregnant population⁷⁴ was also somewhat limited by focusing exclusively on the screening accuracy of the GAD-7 for generalised anxiety disorder, without providing any evidence of its screening ability for other anxiety disorders in pregnancy. Furthermore, sub-analyses to assess the screening ability of the GAD-2 as opposed to the full GAD-7 were not conducted, thus leaving unanswered the question of whether the GAD-2 can be used as an ultra-brief screening scale for problematic anxiety symptoms in pregnancy, as per recent guidelines.⁶

Key best-evidence findings

Eight self-report measures were considered in the synthesis of the best available evidence presented above. One further scale located by this review (Pregnancy Anxiety Scale⁶⁵) was not examined at the best-evidence stage as the single study reporting its psychometric properties was rated poor for its methodological quality.⁶⁵

The key findings regarding anxiety items and domains identified as accurate indicators of antenatal anxiety, as discussed in the Results, are summarised here. A complete list of all the identified anxiety items and domains is also presented in the supplementary Table 2. Furthermore, a table summarising all the correlations between scales included in the review is available in supplementary Table 3.

Items assessing excessive, generalised worry were found to be psychometrically sound in the antenatal period in the EPDS, HADS-A, BMWS, GAD-7 and STAI-S. Overall, there was strong evidence of the psychometric robustness of items measuring the domain of worry, with consistent findings in multiple studies of good or excellent quality. Since excessive worry is essentially a cognitive symptom, it could be argued that it is less susceptible to the physical and physiological changes of pregnancy, and it remains thus a good indicator of problematic anxiety in pregnancy as it is in the general population.

A second anxiety domain that showed good evidence of its psychometric soundness in pregnant populations concerned items tapping into symptoms of fear or panic. Feelings of fear are another important component of different anxiety disorders, including panic disorder, agoraphobia, social anxiety disorder and specific phobia.^{22,95} In this review, items assessing the fear/panic domain were identified as psychometrically sound for use in pregnancy in various scales, including the HADS-A, the EPDS and several pregnancy-specific anxiety scales.

Other specific symptoms identified by this review showed moderate evidence of their screening ability in the assessment of antenatal anxiety. These included being excessively self-critical (EPDS, item 3), feeling upset (STAI-S, item 6) and the experience of nervous or motor tension (STAI-S, item 3). Although these symptoms may not appear to be specific to anxiety disorders, these findings are in line with the well-established tripartite model of anxiety and depression. This model postulates that depressive and anxiety disorders share a common component of general emotional distress, and the symptoms above can be categorised as manifestations of general distress, which can be present in both depressive and anxiety symptomatology.⁹⁶

In relation to anxiety symptoms specifically related to pregnancy, fear of childbirth was shown to be a good indicator of pregnancy-specific anxiety. Specifically, pregnancy-specific anxiety symptoms of fear related to giving birth exhibited strong evidence of their psychometric value in the W-DEQ (several items) and the PRAQ-R (two items related to fear of childbirth).

Items assessing persistent worries specifically related to pregnancy also showed good psychometric properties in the CWS, the W-DEQ and the PRAQ-R. The worries with the strongest evidence to support their screening accuracy related to concerns regarding the health or safety of the baby and the possibility of miscarriage. Other worries, including being in hospital and worrying about future parenting showed only moderate evidence of their screening value (see supplementary Table 2). It may be argued that most women are likely to experience some degree of concern regarding these aspects of pregnancy, but that in women experiencing clinical levels of anxiety these worries may be more intense or persistent (i.e. higher severity or frequency).

Strengths and limitations

The present review has a number of strengths. Only studies with good or excellent methodological quality as determined by the COSMIN checklist⁴⁵ were included in the best-evidence synthesis, thus guaranteeing that the conclusions reached were only based on the strongest evidence available. We also used a comprehensive search strategy that was devised to locate studies testing the

psychometric properties of both general anxiety scales and pregnancy-specific anxiety measures, unlike previous reviews that were focused mostly or exclusively on general anxiety or pregnancy-specific anxiety scales.^{35,36} A second reviewer independently checked a sample of studies, both in the initial phase of screening of titles and abstract and for the quality assessment of included studies, as per best practice recommendations for systematic reviews.³⁷ The review was reported according to the PRISMA reporting guidelines³⁹ (see supplementary Table 4).

Several limitations also have to be acknowledged. Searches were limited to research articles in English and restricted to publications from 1991 onwards, this being the year when the first pregnancy-specific anxiety scale was developed. The generalisability of the review findings may also be somewhat limited by the fact that we did not include studies from countries with substantial cultural differences compared with the UK (i.e. Asian and African countries) for which cultural equivalence of psychological symptoms cannot be assumed.^{97,98}

Implications and future directions

The accurate identification of women experiencing high levels of anxiety symptoms in pregnancy is important and deserves clinical attention for several reasons. Whereas postnatal depression has been the focus of most research in perinatal mental health in the past decades,^{21,24} there is now a substantial body of research indicating that anxiety in pregnant women is common and is associated with increased risk for negative maternal and child outcomes.^{3,32,99} In the UK, the Royal College of General Practitioners has identified perinatal mental health as a clinical priority¹⁰⁰ and a recent report from the London School of Economics has estimated the costs of neglecting perinatal mental health problems in the UK to be a striking figure of £8.1 billion for every annual cohort of women, with approximately three-quarters of this cost related to the adverse long-term impact on children.¹⁰¹

Among the range of perinatal mental health problems that women can experience, anxiety disorders have the highest prevalence.¹ Consequently, a number of authors in recent years have advocated the use of a brief scale for the universal screening of antenatal anxiety.^{36,102} To the best of our knowledge, no anxiety scales have been developed that are specific to the antenatal period and take into account both general and pregnancy-specific anxiety symptoms. Most studies have used measures of general anxiety, but the clinical importance of including screening for pregnancy-specific anxiety symptoms is supported by studies indicating that pregnancy-specific anxiety may be a better predictor of adverse birth and child development outcomes than general anxiety during pregnancy.^{34,103}

Future research is needed to conduct robust psychometric studies of existing measures in sufficiently large samples and ideally including validation against a reference standard. The development of a new anxiety scale specifically constructed for use in pregnancy and that takes into account both general anxiety and symptoms of pregnancy-specific anxiety would also be highly desirable.

In sum, despite the research literature on prevalence, risk factors and treatment of antenatal anxiety having decisively grown in recent years,^{1,8,104} this review clearly points out how evidence regarding the screening performance of anxiety scales for use in pregnancy, including the one currently recommended by NICE, remains insufficient. The lack of measures with a sufficient evidence base constitutes a considerable barrier to the identification of pregnant women experiencing problematic anxiety symptoms, the initial step if they are to be offered the appropriate support or treatment. This is, in turn, an important missed opportunity for early prevention of negative health outcomes for women and their children. This review improves the current understanding of anxiety symptomatology

in pregnant women and may contribute to provide the theoretical basis for the development of a psychometrically robust screening scale for antenatal anxiety.

Andrea Sinesi, MSc, PhD Student, Nursing, Midwifery and Allied Health Professions Research Unit, University of Stirling, UK; **Margaret Maxwell**, PhD, Professor of Health Services and Mental Health Research and Director, Nursing, Midwifery and Allied Health Professions Research Unit, University of Stirling, UK; **Ronan O'Carroll**, MPhil, PhD, Professor of Psychology, Department of Psychology, University of Stirling, UK; **Helen Cheyne**, RGN, RM, PhD, Professor of Maternal and Child Health and Deputy Director, Nursing, Midwifery and Allied Health Professions Research Unit, University of Stirling, UK

Correspondence: Andrea Sinesi, Nursing, Midwifery and Allied Health Professions Research Unit (NMAHP RU), Unit 13 Scion House, Stirling University Innovation Park, University of Stirling, Stirling FK9 4NK, UK. Email: andrea.sinesi@stir.ac.uk

First received 2 Aug 2018, final revision 2 Nov 2018, accepted 9 Nov 2018

Funding

A.S. was supported by a Doctoral Training Fellowship from the Chief Scientist Office of the Scottish Government Health Directorates (DTF/15/03).

Acknowledgements

We thank Alex Pollock (Nursing, Midwifery and Allied Health Professions Research Unit, Glasgow Caledonian University) for her contribution to the design of the study.

Supplementary material

Supplementary material is available online at <https://doi.org/10.1192/bjo.2018.75>.

Appendix

Inclusion and exclusion criteria

	Inclusion criteria	Exclusion criteria
Type of study	Primary research articles reporting at least one psychometric property of a self-report rating scale used to assess anxiety symptoms in pregnancy. Published in a peer-reviewed journal in English in or after 1991.	Studies conducted in countries with substantial cultural differences with the UK (i.e. African and Asian countries) for which cultural equivalence cannot be assumed. Qualitative studies on the experience of anxiety symptoms during pregnancy.
Population	Pregnant or perinatal participants (for perinatal samples, subgroup analyses of psychometric properties of the measure available for the subsample of pregnant women).	Sample composed exclusively of women with high-risk pregnancies, because of obstetric complications (for example pre-eclampsia, ectopic pregnancy) or high psychosocial risk.
Construct of interest	A specific anxiety disorder, as determined by DSM-5 or ICD-10 diagnostic criteria. General 'anxiety' or 'worry' Pregnancy-related anxiety, as defined by Huizink and colleagues. ²⁸ Fear of childbirth, as this is deemed to be a relevant component of pregnancy-specific anxiety.	Any other construct, as for example general mental health, mental disorders other than anxiety during pregnancy, stress or postnatal anxiety.
Type of measurement	All studies with published psychometric data using self-report rating scales to assess anxiety symptoms in pregnancy and meeting the inclusion criteria for type of study, population, and construct of interest detailed above were eligible for inclusion in the review.	Anxiety symptoms assessed exclusively with methods other than a self-report rating scale (for example open questions by a health professional, physiological measures of anxiety). Scale designed to be completed by a health professional after observation (i.e. not self-report). Scale not developed to generate a total score or single-item measures.

References

- Dennis CL, Falah-Hassani K, Shiri R. Prevalence of antenatal and postnatal anxiety: systematic review and meta-analysis. *Br J Psychiatry* 2017; **210**: 315–23.
- Heron J, O'Connor TG, Evans J, Golding J, Glover V. The course of anxiety and depression through pregnancy and the postpartum in a community sample. *J Affect Disord* 2004; **80**: 65–73.
- Dunkel-Schetter C, Lobel M. Pregnancy and birth: a multilevel analysis of stress and birth weight. In *Handbook of Health Psychology 2* (eds T Revenson, A Baum, J Singer): 427–53. Lawrence Erlbaum, 2011.
- Stein A, Pearson RM, Goodman SH, Rapa E, Rahman A, McCallum M, et al. Effects of perinatal mental disorders on the fetus and child. *Lancet* 2014; **384**: 1800–19.
- Rubertsson C, Hellström J, Cross M, Sydsjö G. Anxiety in early pregnancy: prevalence and contributing factors. *Arch Womens Ment Health* 2014; **17**: 221–8.
- National Institute for Health and Care Excellence. *Antenatal and Postnatal Mental Health: Clinical Management and Service Guidance (CG192)*. NICE, 2014.
- Milgrom J, Gemmil A, Bilszta JL, Hayes B, Barnett B, Brooks J, et al. Antenatal risk factors for postnatal depression: a large prospective study. *J Affect Disord* 2008; **108**: 147–57.
- Grant K, McMahon C, Austin MP. Maternal anxiety during the transition to parenthood: a prospective study. *J Affect Disord* 2008; **108**: 101–11.
- Verreault N, DaCosta D, Marchand A, Ireland K, Dritsa M, Khalife S. Rates and risk factors associated with depressive symptoms during pregnancy and with postpartum onset. *J Psychosom Obstet Gynaecol* 2014; **35**: 84–91.
- Sutter-Dallay AL, Giaccone-Marcasche V, Glatigny-Dallay E, Verdoux H. Women with anxiety disorders during pregnancy are at increased risk of intense postnatal depressive symptoms: a prospective survey of the MATQUID cohort. *Eur Psychiatr* 2004; **19**: 459–63.
- Austin MP, Tully L, Parker G. Examining the relationship between antenatal anxiety and postnatal depression. *J Affect Disord* 2007; **101**: 169–74.
- Field T, Diego M, Hernandez-Reif M, Figueroa B, Deeds O, Ascencio A, et al. Comorbid depression and anxiety effects on pregnancy and neonatal outcome. *Infant Behav Dev* 2010; **33**: 23–9.
- Diego MA, Jones NA, Field T, Hernandez-Reif M, Schanberg S, Kuhn C, et al. Maternal psychological distress, prenatal cortisol, and fetal weight. *Psychosom Med* 2006; **68**: 747–53.
- Berle JO, Mykletun A, Daltveit AK, Rasmussen S, Holsten F, Dahl AA. Neonatal outcomes in offspring of women with anxiety and depression during pregnancy: a linkage study from The Nord-Trøndelag Health Study (HUNT) and Medical Birth Registry of Norway. *Arch Womens Ment Health* 2005; **8**: 181–9.
- Ding XX, Wu YL, Xu SJ, Zhu RP, Jia XM, Zhang SF, et al. Maternal anxiety during pregnancy and adverse birth outcomes: a systematic review and meta-analysis of prospective cohort studies. *J Affect Disord* 2014; **159**: 103–10.
- Talge NM, Neal C, Glover V. Antenatal maternal stress and long-term effects on child neurodevelopment: how and why? *J Child Psychol Psychiatry* 2007; **48**: 245–61.
- Van Der Bergh BR, Mulder EJ, Mennes M, Glover V. Antenatal maternal anxiety and stress and the neurobehavioural development of the fetus and child: links and possible mechanisms. A review. *Neurosci Biobehav Rev* 2005; **29**: 237–58.
- Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med* 2006; **166**: 1092–7.

- 19 Scottish Intercollegiate Guidelines Network (SIGN). *Management of Perinatal Mood Disorders (127)*. SIGN, 2012.
- 20 National Institute for Health and Care Excellence. *Common Mental Health Disorders: Identification and Pathways to Care (CG123)*. NICE, 2011.
- 21 Austin MP. Antenatal screening and early intervention for 'perinatal' distress, depression and anxiety: where to from here? *Arch Womens Ment Health* 2004; **7**: 1–6.
- 22 American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders (5th edn) (DSM-5)*. American Psychiatric Publishing, 2013.
- 23 World Health Organization (WHO). *The ICD-10 Classification of Mental and Behavioural Disorders: Clinical Descriptions and Diagnostic Guidelines*. WHO, 1992.
- 24 Goodman JH, Chenausky KL, Freeman MP. Anxiety Disorders during pregnancy: a systematic review. *J Clin Psychiatry* 2014; **75**: 1153–84.
- 25 Evans K, Spiby H, Morrell CJ. A psychometric systematic review of self-report instruments to identify anxiety in pregnancy. *J Adv Nurs* 2015; **71**: 1986–2001.
- 26 Swallow BL, Lindow SW, Masson EA, Hay DM. The use of the General Health Questionnaire (GHQ-28) to estimate prevalence of psychiatric disorder in early pregnancy. *Psychol Health Med* 2003; **8**: 213–7.
- 27 Marchesi C, Ampollini P, Paraggio C, Giaraconi G, Ossola P, DePanfilis C, et al. Risk factors for panic disorder in pregnancy: a cohort study. *J Affect Disord* 2014; **156**: 134–8.
- 28 Huizink AC, Mulder EJ, Robles De Medina PG, Visser GH, Buitelaar JK. Is pregnancy anxiety a distinctive syndrome? *Early Hum Dev* 2004; **79**: 81–91.
- 29 Phillips J, Sharpe L, Matthey S, Charles M. Maternally focussed worry. *Arch Womens Ment Health* 2009; **12**: 409–18.
- 30 Poikkeus P, Saisto T, Unkila-Kallio L, Punamaki RL, Repokari L, Vilksa S, et al. Fear of childbirth and pregnancy-related anxiety in women conceiving with assisted reproduction. *Obstet Gynecol* 2006; **108**: 70–6.
- 31 Huizink AC, Delforterie MJ, Scheinim NM, Tolvanen M, Karlsson L, Karlsson H. Adaptation of pregnancy anxiety questionnaire—revised for all pregnant women regardless of parity: PRAQ-R2. *Arch Womens Ment Health* 2016; **19**: 125–32.
- 32 Dunkel Schetter C, Tanner L. Anxiety, depression and stress in pregnancy: implications for mothers, children, research, and practice. *Curr Opin Psychiatry* 2012; **25**: 141–48.
- 33 Westerneng M, de Cock P, Spelten ER, Honig A, Hutton EK. Factorial invariance of pregnancy-specific anxiety dimensions across nulliparous and parous pregnant women. *J Health Psychol* 2015; **20**: 164–72.
- 34 Blackmore ER, Gustafsson H, Gilchrist M, Wyman C, O'Connor TG. Pregnancy-related anxiety: evidence of distinct clinical significance from a prospective longitudinal study. *J Affect Disord* 2016; **197**: 251–8.
- 35 Meades R, Ayers S. Anxiety measures validated in perinatal populations: a systematic review. *J Affect Disord* 2011; **133**: 1–15.
- 36 Brunton RJ, Dryer R, Saliba A, Kohlhoff J. Pregnancy anxiety: a systematic review of current scales. *J Affect Disord* 2015; **176**: 24–34.
- 37 Centre for Reviews and Dissemination (CRD). *Systematic Reviews: CRD's Guidance for Undertaking Reviews in Health Care*. CRD, 2009.
- 38 Terwee CB. *Protocol for Systematic Reviews of Measurement Properties*. COSMIN, 2011.
- 39 Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med* 2009; **151**: 264–70.
- 40 Leeflang MM, Scholten RJ, Rutjes AW, Reitsma JB, Bossuyt PM. Use of methodological search filters to identify diagnostic accuracy studies can lead to the omission of relevant studies. *J Clin Epidemiol* 2006; **59**: 234–40.
- 41 Whiting P, Westwood M, Beynon R, Burke M, Sterne JA, Glanville J. Inclusion of methodological filters in searches for diagnostic test accuracy studies misses relevant studies. *J Clin Epidemiol* 2011; **64**: 602–7.
- 42 McHugh ML. Interrater reliability: the kappa statistic. *Biochem Med* 2012; **22**: 276–82.
- 43 Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol* 2010; **63**: 737–45.
- 44 Mokkink LB, Terwee CB, Knol DL, Stratford PW, Alonso J, Patrick DL, et al. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: a clarification of its content. *BMC Med Res Methodol* 2010; **10**: 1–8.
- 45 Terwee CB, Mokkink LB, Knol DL, Ostelo RW, Bouter LM, de Vet HC. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Qual Life Res* 2012; **21**: 651–7.
- 46 Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007; **60**: 34–42.
- 47 Streiner DL, Norman GR, Cairney J. *Health Measurement Scales: A Practical Guide to Their Development and Use (5th edn)*. Oxford University Press, 2014.
- 48 DeVellis RF. *Scale Development, Theory and Applications (3rd edn)*. SAGE Publication, 2012.
- 49 Bartholomew D, Knotts M, Moustaki I. *Latent Variable Models and Factor Analysis: A Unified Approach (3rd edn)*. John Wiley & Sons, 2011.
- 50 Tabachnick BG, Fidell LS. *Using Multivariate Statistics (5th edn)*. Pearson Education Inc., 2007.
- 51 Bayrampour H, McDonald S, Fung T, Tough S. Reliability and validity of three shortened versions of the State Anxiety Inventory scale during the perinatal period. *J Psychosom Obstet Gynecol* 2014; **35**: 101–7.
- 52 Brouwers EP, van Baar AL, Pop VJ. Does the Edinburgh Postnatal Depression Scale measure anxiety? *J Psychosom Res* 2001; **51**: 659–63.
- 53 Carmona Monge FJ, Peñacoba-Puente C, Marín Morales D, Carretero Abellán I. Factor structure, validity and reliability of the Spanish version of the Cambridge Worry Scale. *Midwifery* 2012; **28**: 112–9.
- 54 Coates R, Ayers S, Visser R. Factor structure of the Edinburgh Postnatal Depression Scale in a population-based sample. *Psychol Assess* 2017; **29**: 1016–27.
- 55 Fenaroli V, Saita E. Fear of childbirth: a contribution to the validation of the Italian version of the Wijma Delivery Expectancy/Experience Questionnaire (WDEQ). *TPM Test Psychom Methodol Appl Psychol* 2013; **20**: 131–54.
- 56 Garthus-Niegel S, Størksen HT, Torgersen L, Von Soest T, Eberhard-Gran M. The Wijma Delivery Expectancy/Experience Questionnaire - a factor analytic study. *J Psychosom Obstet Gynaecol* 2011; **32**: 160–3.
- 57 Gourounti K, Lykeridou K, Taskou C, Kafetsios K, Sandall J. A survey of worries of pregnant women: reliability and validity of the Greek version of the Cambridge Worry Scale. *Midwifery* 2012; **28**: 746–53.
- 58 Green JM, Kafetsios K, Statham HE, Snowdon CM. Factor structure, validity and reliability of the Cambridge Worry Scale in a pregnant population. *J Health Psychol* 2003; **8**: 753–64.
- 59 Haines HM, Pallant JF, Fenwick J, Gamble J, Creedy DK, Toohill J, et al. Identifying women who are afraid of giving birth: a comparison of the fear of birth scale with the WDEQ-A in a large Australian cohort. *Sex Reprod Healthc* 2015; **6**: 204–10.
- 60 Johnson R, Slade P. Does fear of childbirth during pregnancy predict emergency caesarean section? *BJOG An Int J Obstet Gynaecol* 2002; **109**: 1213–21.
- 61 Jomeen J, Martin C. Is the hospital anxiety and depression scale (HADS) a reliable screening tool in early pregnancy? *Psychol Health* 2004; **19**: 787–800.
- 62 Jomeen J, Martin CR. The factor structure of the Cambridge Worry Scale in early pregnancy. *J Prenat Perinat Psychol Heal* 2005; **20**: 25–48.
- 63 Jomeen J, Martin CR. Confirmation of an occluded anxiety component within the Edinburgh Postnatal Depression Scale (EPDS) during early pregnancy. *J Reprod Infant Psychol* 2005; **23**: 143–54.
- 64 Karimova G, Martin C. A psychometric evaluation of the Hospital Anxiety and Depression Scale during pregnancy. *Psychol Health Med* 2003; **8**: 89–103.
- 65 Levin JS. The factor structure of the pregnancy anxiety scale. *J Health Soc Behav* 1991; **32**: 368–81.
- 66 Marteau TM, Bekker H. The development of a six-item short-form of the state scale of the Spielberger State-Trait Anxiety Inventory (STAI). *Br J Clin Psychol* 1992; **31**: 301–306.
- 67 Matthey S, Valenti B, Souler K, Ross-Hamid C. Comparison of four self-report measures and a generic mood question to screen for anxiety during pregnancy in English-speaking women. *J Affect Disord* 2013; **148**: 347–51.
- 68 Öhman SG, Grunewald C, Waldenström U. Women's worries during pregnancy: testing the Cambridge Worry Scale on 200 Swedish women. *Scand J Caring Sci* 2003; **17**: 148–52.
- 69 Petersen JJ, Paulitsch MA, Guethlin C, Gensichen J, Jahn A. A survey on worries of pregnant women—testing the German version of the Cambridge worry scale. *BMC Public Health* 2009; **9**: 490–98.
- 70 Simpson W, Glazer M, Michalski N, Steiner M, Frey BN. Comparative efficacy of the generalized anxiety disorder 7-item scale and the Edinburgh Postnatal Depression Scale as screening tools for generalized anxiety disorder in pregnancy and the postpartum period. *Can J Psychiatry* 2014; **59**: 434–40.
- 71 Swalm D, Brooks J, Doherty D, Nathan E, Jacques A. Using the Edinburgh Postnatal Depression Scale to screen for perinatal anxiety. *Arch Womens Ment Health* 2010; **13**: 515–22.
- 72 Tendais I, Costa R, Conde A, Figueiredo B. Screening for depression and anxiety disorders from pregnancy to postpartum with the EPDS and STAI. *Span J Psychol* 2014; **17**: 1–9.

- 73 Wijma K, Wijma B, Zar M. Psychometric aspects of the W-DEQ; a new questionnaire for the measurement of fear of childbirth. *J Psychosom Obstet Gynecol* 1998; **19**: 84–97.
- 74 Zhong QY, Gelaye B, Zaslavsky AM, Fann JR, Rondon MB, Sanchez SE, et al. Diagnostic validity of the generalized anxiety disorder-7 (GAD-7) among pregnant women. *PLoS ONE* 2015; **10**: 1–17.
- 75 Cox J, Holden J, Sagovsky R. Detection of postnatal depression. Development of the 10-item Edinburgh Postnatal Depression Scale. *Br J Psychiatry* 1987; **150**: 782–6.
- 76 Murray D, Cox JL. Screening for depression during pregnancy with the Edinburgh Depression Scale (EDDS). *J Reprod Infant Psychol* 1990; **8**: 99–107.
- 77 Green J, Murray D. The use of the Edinburgh Postnatal Depression Scale in research to explore the relationship between antenatal and postnatal dysphoria. In *Perinatal Psychiatry: Use and Misuse of the Edinburgh Postnatal Depression Scale* (eds J Cox and J Holden): 180–98. Gaskell, 1994.
- 78 Pop VJ, Komprou IH, van Son MJ. Characteristics of the Edinburgh Post Natal Depression Scale in The Netherlands. *J Affect Disord* 1992; **26**: 105–10.
- 79 Ross LE, Gilbert Evans SE, Sellers EM, Romach MK. Measurement issues in postpartum depression, pt 1: anxiety as a feature of postpartum depression. *Arch Womens Ment Health* 2003; **6**: 51–7.
- 80 Matthey S, Fisher J, Rowe H. Using the Edinburgh postnatal depression scale to screen for anxiety disorders: conceptual and methodological considerations. *J Affect Disord* 2013; **146**: 224–30.
- 81 Child D. *The Essentials of Factor Analysis (3rd edn)*. Continuum International Publishing Group, 2006.
- 82 Zigmond AS, Snaith RP. The Hospital Anxiety and Depression Scale. *Acta Psychiatr Scand* 1983; **67**: 361–70.
- 83 Cosco TD, Doyle F, Ward M, McGee H. Latent structure of the Hospital Anxiety and Depression Scale: a 10-year systematic review. *J Psychosom Res* 2012; **72**: 180–4.
- 84 Owen DJ, Wood L, Tomenson B, Creed F, Neilson JP. Social stress predicts preterm birth in twin pregnancies. *J Psychosom Obstet Gynaecol* 2017; **38**: 63–72.
- 85 Spielberger CD, Gorsuch RL, Lushene R. *Manual for the State-Trait Anxiety Inventory*. Consulting Psychologists Press, 1970.
- 86 Spielberger CD, Gorsuch RL, Lushene R, Vagge PR, Jacobs GA. *Manual for the State-Trait Anxiety Inventory*. Consulting Psychologists Press, 1983.
- 87 Sheehan DV, Lecrubier Y, Harnett-Sheehan K, Amorim P, Janavs J, Weiller E, et al. The Mini International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview. *J Clin Psychiatry* 1998; **59**: 22–33.
- 88 Kessler RC, Üstün TB. The World Mental Health (WMH) Survey Initiative version of the World Health Organization (WHO) Composite International Diagnostic Interview (CIDI). *Int J Methods Psychiatr Res* 2004; **13**: 93–121.
- 89 Li Y, Baser R. Using R and WinBUGS to fit a generalized partial credit model for developing and evaluating patient-reported outcomes assessments. *Stat Med* 2012; **31**: 2010–26.
- 90 Gladstone GL, Parker GB, Mitchell PB, Malhi GS, Wilhelm KA, Austin MP. A Brief Measure of Worry Severity (BMWS): personality and clinical correlates of severe worriers. *J Anxiety Disord* 2005; **19**: 877–92.
- 91 Cuijpers P, Sijbrandij M, Koole S, Huibers M, Berking M, Andersson G. Psychological treatment of generalized anxiety disorder: a meta-analysis. *Clin Psychol Rev* 2014; **34**: 130–40.
- 92 Derogatis L. *SCL-90-R. Administration, Scoring and Procedures Manual for the Revised Version of the SCL-90*. John Hopkins University Press, 1977.
- 93 Heimstad R, Dahloe R, Laache I, Skogvoll E, Schei B. Fear of childbirth and history of abuse: implications for pregnancy and delivery. *Acta Obstet Gynecol Scand* 2006; **85**: 435–40.
- 94 Van den Bergh B. The influence of maternal emotions during pregnancy on fetal and neonatal behavior. *J Prenat Perinat Psychol Health* 1991; **5**: 119–30.
- 95 Craske M, Rauch S, Ursano R, Prenoveau J, Pine D, Zinbarg R. What is an anxiety disorder? *Depress Anxiety* 2009; **26**: 1066–85.
- 96 Clark LA, Watson D. Tripartite model of anxiety and depression: psychometric evidence and taxonomic implications. *J Abnorm Psychol* 1991; **100**: 316–36.
- 97 Gunay EY, Gul A. Reliability and validity of the Cambridge Worry Scale in pregnant Turkish women. *Midwifery* 2015; **31**: 359–64.
- 98 Takegata M, Haruna M, Matsuzaki M, Shiraishi M, Murayama R, Okano T, et al. Translation and validation of the Japanese version of the Wijma Delivery Expectancy/Experience Questionnaire version A. *Nurs Health Sci* 2013; **15**: 326–32.
- 99 Gavin N, Meltzer-Brody S, Glover V, Gaynes BN. Is population-based identification of perinatal depression and anxiety desirable?: A public health perspective on the Perinatal Depression Care Continuum. In *Identifying Perinatal Depression and Anxiety* (eds J Milgrom and AW Gemmill): 11–31. John Wiley & Sons, 2015.
- 100 Royal College of General Practitioners. *Perinatal Mental Health is a Clinical Priority for the College*. RCGP, 2017 (<http://www.rcgp.org.uk/about-us/news/2017/february/perinatal-mental-health-is-a-clinical-priority-for-the-college-says-rcgp.aspx>).
- 101 Bauer A, Parsonage M, Knapp M, Iemmi V, Adelaya B. *Costs of Perinatal Mental Health Problems*. London School of Economics and Political Science, 2014.
- 102 Biaggi A, Conroy S, Pawlby S, Pariante CM. Identifying the women at risk of antenatal anxiety and depression: a systematic review. *J Affect Disord* 2016; **191**: 62–77.
- 103 Davis EP, Sandman CA. Prenatal psychobiological predictors of anxiety risk in preadolescent children. *Psychoneuroendocrinology* 2012; **37**: 1224–33.
- 104 Marchesi C, Ossola P, Amerio A, Daniel BD, Tonna M, De Panfilis C. Clinical management of perinatal anxiety disorders: a systematic review. *J Affect Disord* 2016; **190**: 543–50.

