A systematic review of measures of mental health and emotional wellbeing in parents of

children aged 0-5

Rebecca Webb, Susan Ayers and Camilla Rosan

Author Note

Rebecca Webb & Susan Ayers, Centre for Maternal and Child Health Research, City

University London, EC1V 0HB

Camilla Rosan, Development and Delivery Department, Mental Health Foundation,

Colechurch House, 1 London Bridge Walk, London, SE1 2SX

Correspondence concerning this article should be addressed to: Susan Ayers, Centre for

Maternal and Child Health Research, City University London, EC1V 0HB. Email:

Susan.Ayers@city.ac.uk. Telephone: 0044 (0)207 0405834

Abstract

Background: A significant proportion of women with young children experience mental health problems and recent research suggests fathers may also be affected. This may have a long term negative impact on the child's development with significant costs to society. Appropriate measures are therefore needed to identify parents and children at risk.

Method: This literature review aimed to identify the most reliable, evidence based global measures of mental health for parents of infants from pregnancy to 5 years postpartum (0-5 years). Literature searches were conducted on online databases and hand searches of reference lists were also carried out. Studies were included in the review if they reported information on measures of global psychological distress or wellbeing from 0-5 years postpartum.

Results: A total of 183 studies were included in the review, 19 of which directly examined the psychometric validity of an outcome measure. These studies reported information on 23 outcome measures, 4 of which had been validated in parents of children from 1-5. These were: the General Health Questionnaire (GHQ), the Symptom Checklist (SCL), the Self-Reporting Questionnaire (SRQ) and the Kessler scale (K10/6). Reliability and validity varied across studies.

Limitations: Only a small number of studies included fathers and examined psychometric validity across the entire period of early childhood.

Conclusions: The GHQ was the most frequently validated but results suggest poor reliability and validity. The SRQ and K10/6 were the most promising measures in terms of psychometric properties and clinical utility.

Keywords: Early years; parental wellbeing; measures; psychometrics; reliability; validity

A systematic review of measures of mental health and emotional wellbeing in parents of

children aged 0-5

Pregnancy and the first year after birth are associated with increased risk of affective disorders with research indicating between 10 to 20% of women experience affective disorders during this time (Mann, Gilbody, & Adamson, 2010). Recent research also suggests a significant proportion of men may be affected (Cameron, Sedov, & Tomfohr-Madsen, 2016; Parfitt & Ayers, 2014). Parental mental health problems in the early years can be pre-existing, where existing mental health problems are exacerbated or retriggered during pregnancy or after birth. Alternatively, mental health problems may start during pregnancy or after birth. The most common mental health problems are anxiety, depression, post-traumatic stress disorder (PTSD), and stress-related conditions such as adjustment disorder (Cameron et al., 2016). Severe mental illness, such as puerperal psychosis, is less common but is one of the leading indirect causes of maternal death (Manktelow & Draper, 2015).

Parental mental health problems in the early years are a significant public health concern due the association between such problems and a variety of adverse outcomes for women and their partners. These include parenting distress, poor physical health, financial strain, stressful life events, low social support, and low quality partner relationship (Horwitz, Briggs-Gowan, Storfer-Isser, & Carter, 2007). Children may also be affected. For example, maternal postpartum depression is associated with less sensitive and responsive interactions between the mother and baby (Barry et al., 2015), higher rates of negative emotional expression (Murray, Kempton, Woolgar, & Hooper, 1993), and having a higher likelihood of implementing unpredictable and inconsistent parenting techniques (Beck, Shattuck, Haynie, Crump, & Simons-Morton, 1999). These factors are in turn associated with poor child development (Barry et al., 2015; Fihrer, McMahon, & Taylor, 2009; Grace, Evindar, & Stewart, 2003). Mothers with postpartum depression are also less likely to breastfeed (Dennis &

McQueen, 2009) which in some cases may be related to concerns about the effects of any psychotropic medication they may be taking on their infant (Bonari et al., 2005).

Parental mental health problems are also associated with the infant being more likely to experience mental health problems (Lang & Gartstein, 2017; Wang, 2017). This intergenerational transmission of vulnerability may be due to a number of factors, including epigenetic, environmental and parenting factors. For example, women's mental health during pregnancy may exert specific effects on the developing foetus through neuro-biological foetal programming which may have a long-term effect on the child's development and health. There is increasing evidence for foetal programming effects from animal research and epidemiological research showing associations between anxiety and stress in pregnancy and greater risk of a range of adverse outcomes for the child. For example, longitudinal research suggests infants of women who are anxious during pregnancy are more likely to show fearful or anxious behaviour themselves and be at greater risk of poor development and adverse outcomes such as attention-deficit hyperactivity disorder (Talge, Neal, & Glover, 2007). Poorer emotional and behavioural development may even persist into adolescence, with a study of 7,944 families in England finding that maternal anxiety and depression in pregnancy was associated with the child being twice as likely to have a mental disorder (O'Donnell et al, 2014). However, most children are not affected and the mechanisms underlying which children are affected and in which way are not well understood (Glover, 2016).

Paternal mental illness has been less studied than maternal mental illness. However, there is increasing evidence it may also have a negative impact on the couple's relationship, father-child relationship and child developmental outcomes. For example, paternal stress has been associated with men having more negative perceptions of their marriage and baby (Zelkowitz & Milet, 1997). A meta-analysis of 28 studies of paternal depression found it was significantly associated with a decrease in positive parenting behaviors (such as warmth) and an increase in negative parenting behaviors (such as criticism) (Wilson & Durbin, 2010). Additionally, the early father-child

relationship appeared to influence later childhood outcomes: for example, father's psychological distress has been associated with child behavioral difficulties, emotional difficulties, social functioning and development (Fletcher, Feeman, Garfield, & Vimpani, 2011; Kvalevaag et al., 2013).

In summary, there is substantial evidence suggesting that mental illness for men and women during the early years may have an adverse impact on infant development and child outcomes across a range of domains e.g. cognitive, emotional, motor and social (Milgrom, Ericksen, & Sved-Williams, 2016; Sweeney & MacBeth, 2016). There has been some debate over the importance of timing and the possibility that pregnancy and the first year may represent a sensitive period in infant development, in which the infant may be more susceptible to the effects of mental health difficulties due (in part) to the rapid and substantial neural, cognitive and socio- emotional developments that occur during this time (Sroufe, 2005; Talge et al., 2007). However, there is some evidence that sensitivity may continue up to age 5 with, for example, a study of a cohort of 937 adolescents in Canada finding that those who were initially exposed to maternal depression between the ages of 2 and 5 had a two-fold increased risk of affective disorders in adolescence (Naicker, Wickham & Colman, 2012).

The economic argument for screening and treating parental mental health problems is compelling with evidence that the cost to society of parental mental illness is substantial. A recent economic analysis estimated the cost to UK society of not treating women's perinatal mental health problems is £8.1 billion for every annual cohort of women giving birth. The majority of this cost (72%) was attributable to long term adverse consequences for the child (Bauer, Parsonage, Knapp, Iemmi, & Adelaja, 2014). In Canada, O'Brien, Laporte, and Koren (2009) estimated that approximately 2,593 women discontinued their antidepressants and had a depressive relapse. This resulted in maternal healthcare costs of approximately CA$1 million. Similar results have been found in Australia, where the financial costs associated with maternal postpartum depression were estimated to be AU$61 million in a cohort of 70,997 women (Post and Antenatal Depression

Association, 2012). Furthermore, in the USA, Dagher, McGovern, and Dowd (2014) examined the association between depression in the postpartum period and healthcare expenditure 11 weeks after childbirth and found the mean cost for women with postpartum depression was US$1,046 compared to US$365 for women without depression (2001 prices).

From a public health perspective, a strategy for preventing perinatal mental health problems and treating them effectively therefore has the potential to prevent significant long-term burden of ill-health and problems in parents and children. A critical aspect of this is identifying efficient and acceptable measures of mental health for use with men and women during the early years. Guidelines for screening vary in their recommendations (Milgrom & Gemmill, 2014) but most focus on specific disorders, such as depression. In the UK, national guidelines recommend asking the Whooley questions (Whooley, Avins, Miranda, & Browner, 1997) to identify depression and the General Anxiety Disorder-2 question screen (Spitzer, Kroenke, Williams, & Lowe, 2006) to identify anxiety (NICE, 2014). Other countries use the Edinburgh Postnatal Depression Scale to screen for postpartum depression (Milgrom & Gemmill, 2014). However, this method of focusing on a specific disorder (in this case depression or anxiety) is unlikely to identify women or men experiencing different affective disorders such as PTSD or OCD. This means only a proportion of parents are offered treatment to reduce the possible adverse impact of mental health problems on them and their child.

Another problem is that many of the clinical guidelines for screening do not continue beyond the first year and do not include screening for fathers (NICE, 2014). As we have seen above, there is evidence that children may be affected up to age 5 so screening beyond one year is important on this front as well as the possibility of mental health problems worsening long-term or being recurrent (Brennan et al., 2000; Goodman et al., 2011). Screening fathers is similarly important because of the evidence it may have an adverse impact on fathers, their families and the child (Sweeney & MacBeth, 2016).

Therefore, to improve identification of parents with mental health problems in the early years and reduce the negative impact on child outcomes, a global measure is needed that can be used with mothers and fathers across the early years from 0 to age 5. This literature review aimed to identify reliable, valid and acceptable measures of parental mental health to use with parents during pregnancy and up to 5 years postpartum to identify those with mental health problems. The results of this review will contribute to the wider literature on how best to screen and measure parental wellbeing and outcomes during pregnancy and up to 5 years postpartum.

Method

Identification of Relevant Papers

The literature searches and study selection were conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) guidelines (Moher, Liberati, Tetzlaff, & Altman, 2009). The principal method of identifying studies suitable for the review involved searching the following databases: MEDLINE, PsycINFO, PsychARTICLES, EMBASE, Web of Science, and Scopus. Searches were conducted up to and including February 2016. Search terms formed three categories (the 0-5 early years period, psychological wellbeing or distress and measures) and included, but were not limited to: *natal OR *partum OR parent* AND psychological OR mental OR problem OR disorder OR wellbeing AND measure* OR questionnaire OR scale. Combinations of all search terms and specific syntax used is available in the online supplementary material. In addition, the reference lists of identified articles and existing reviews were hand searched to identify additional relevant papers.

Inclusion and Exclusion Criteria

Studies were included if they were quantitative studies reporting empirical research that examined or included global measures of mental or emotional wellbeing in men and/or women who were expectant or new parents during pregnancy and up to 5 years postpartum. Global measures

were defined as those that examined overall mental health or emotional wellbeing. The 0-5 early

years period was defined as from conception to 5 years postpartum. Studies were included if they

used and reported psychometric or normative data on global measures of mental health or emotional

wellbeing in men and/or women during this period (see Table 1). Studies were excluded if they: (1)

Reported information for measures of specific affective disorders/symptoms such as anxiety or

depression rather than global wellbeing or mental health; (2) If all or some of the sample

measurements were outside the 0-5 early years period; (3) The paper was not available in English;

(4) Non-empirical papers; (5) Dissertation or conference abstracts; (6) Qualitative studies; (7)

Reviews of intervention studies where no information regarding the measure was reported before the

intervention (i.e. at baseline).

Studies using measures originally developed in English that were translated for use in other

countries have been included in the final tables (see online supplementary materials) to provide

information on which languages different measures have been translated into and their validity in

different countries.

Study Selection and Data Extraction

The literature searches provided a total of 24,327 citations, with a further 73 papers identified

through hand searching the reference lists of key papers. After removing duplicates and screening

through abstracts, titles and full texts according to the inclusion and exclusion criteria, 183 studies

remained for inclusion in the review. Most of these studies used a measure of psychological

wellbeing as part of a broader research study but reported some psychometric information. Only 19

of the studies included directly examined the psychometric validity of a measure. Detailed results of

the literature search and screening are shown in Figure 1.

A data extraction sheet (see online supplementary materials) was designed and used to extract

all relevant information from the full text of eligible studies. This included: (1) the measure used; (2)

the language of the measure; (3); the country the study took place in; (4) the amount of participants

in the study; (5) the participant group; (6) participant demographics (including parental gender and

age of children); (7) the design of the study; (8) norming data for the measure; (9); cut off scale of

the measure; (10) how the measure was administered; and (11) any information provided on the

reliability and validity of the measure.

Risk of Bias

Risk of bias was assessed in the 19 studies identified that explored psychometric properties of

measures of parental mental health using the criteria set out by Brink and Louw (2012). This criteria

was developed to help researchers critically appraise reliability and validity studies. It includes 13

items, that can be scored as 'yes' 'no' or 'not applicable', with items that were scored as 'yes' being

assigned a score of 1, and items that were scored as 'no' being assigned a score of 0. These were then

averaged (excluding the answers scored n/a) and multiplied by 100. Studies that scored between 0-33

were labelled as having a high risk of bias, those scoring between 34 – 66 were labelled as having a

medium risk of bias, studies that scored between 67-100 were labelled as having a low risk of bias.

The majority of studies (n = 16) had a low risk of bias, and 3 studies were scored as having a

medium risk of bias. Details of this process are given in the online supplementary material.

<div align="center">

Results

</div>

Study Characteristics

Details of all 183 studies included in the review are given in online supplementary materials.

Nineteen of these studies were validation studies which directly examined the psychometric

properties of measures of global parental mental health and wellbeing. All the other studies included

in the review (n = 164) did not directly examine psychometric properties of a measure of parental

mental health and wellbeing but used one such measure as part of a research study and reported

relevant information such as reliability or means which can be used as norms.

Studies were conducted in 46 different countries, the most common being the USA (n = 27), Norway (n = 18), Australia (n = 18), and the UK (n = 17). All measures evaluated were available in English but as a result of the number of countries that the research was conducted in, norms, reliability and validity from a range of translated versions can also be extracted.

Samples included mothers only, fathers only, and co-parenting pairs. Data were collected in pregnancy, after birth, in the early years, and across the full 0-5 time course. Sample sizes ranged from 14 to 106,935. The majority of studies reported information on the reliability or validity of measures during the perinatal period in mothers. Only a small number of studies were found that included fathers (n = 7 with fathers only; n = 29 with mothers and fathers). Few studies explored validity across the later period of early childhood from age one to five (n = 24).

Measures Identified by the Review

The studies in the review used a total of 24 different measures of general psychological health or wellbeing, details of these can be found in Table 2. These are categorised into measures of: (1) current mood; (2) negative symptoms; (3) positive wellbeing; (4) positive and negative symptoms; (5) somatic symptoms; and (6) other measures. Measures that include somatic items (i.e. physical symptoms) are listed separately because these may be confounded by normal symptoms of pregnancy and postpartum. Of the measures identified, only four had been directly evaluated for psychometric properties in parents of children from 0 to 5. These measures and the number of studies that reported their psychometric properties are:

1. General Health Questionnaire (Goldberg & Williams, 1988) 12 item version (GHQ-12, n = 11); 28 item version (GHQ-28, n = 7); 30 item version (GHQ-30, n = 4) and 36 item version (GHQ-36, n = 1).

2. Self-Reporting Questionnaire (SRQ-20, World Health Organisation, 1994, n = 6)

3. (Hopkins) Symptom Check List (Derogatis, Lipman, Rickels, Uhlenhuth, & Covi, 1974) 25 item version (HSCL-25), (n = 1); and 5 item version (HSCL-5), (n = 5)

4. Kessler Psychological Distress Scale (Kessler et al., 2003) 6 item version (K6, n = 3); and 10 item version (K10, n = 2)[1]

Non-validation studies showed that the Brief Symptom Inventory (BSI, Derogatis, 1993) was a widely used measure. The psychometric properties of these five measures are given in Table 3 and outlined in the next section.

A few studies created their own measure with no adequate psychometric evaluation. The absence of psychometric information does not necessarily mean they are invalid measures as many were created for specific populations or for use as very brief assessment tools. Information on these is given in online supplementary materials for researchers who might be interested in using and evaluating these.

Psychometric properties of maternal global mental health outcome measures

General Health Questionnaire

The most extensively validated measure was the GHQ but the results suggest it has poor validity in early years parenting samples. There were discrepancies reported in the structure of the GHQ, with 2 studies finding a 2-factor model to offer a superior fit to the data (Ip & Martin, 2006; Spiteri, Jomeen, & Martin, 2013) suggesting it may not be a unitary measure of psychological wellbeing. Furthermore, Aderibigbe and Gureje (1992) found that only 3 of the 4 subscales of the GHQ made a significant contribution to the discriminatory power of the GHQ. These findings suggest the structural validity of the GHQ is questionable. Despite this, the GHQ appears to have good discriminatory powers (Navarro et al., 2007; Tran et al., 2011), although the EPDS (one of the most cited measures of perinatal depression used in the literature and clinical practice) did

---

[1] One study examined both the K10 and the K6.

outperform the GHQ in two studies (Navarro et al., 2007; Tran et al., 2012). With regards to

reliability of the measure, internal consistency was found to be acceptable to high, however test-

retest reliability was poor. For example, Spiteri found a low correlation ($r = .38$) across all four time

points in which they tested the measure. Due to the poor validity and test-retest reliability of the

GHQ in the early years, this measure cannot be recommended for use with parents during early

years. This is backed up by the conclusions of Prady et al. (2013) who asserted that there is very little

evidence for the GHQ's use in primary care, due to the fact that women from different ethnic groups

respond in very different ways to the items on the questionnaire making the results difficult to

analyse.

Kessler Psychological Distress Scale

The Kessler Psychological Distress Scale was promising. Internal consistency was good to

high ($\alpha = .73$ to $.90$). The K6 and K10 performed well at discriminating diagnosis of major

depression (Tesfaye et al., 2010), panic disorder and social phobia (Spies et al., 2009). The K10 and

K6 yielded high sensitivity and specificity at the optimal cut off point, again indicating good

discriminant validity. However, test-retest reliability was not measured by any of the validity studies,

so no conclusions can be made about its ability to discriminate psychological distress across time

points. Despite this, a non-validation study reported that father's postpartum psychological distress at

one time point correlated with their psychological distress at time three ($r = .48$; Giallo et al., 2014),

suggesting the measure may be able to be successfully used across time points. Furthermore, the K10

is currently being used within the UK NHS children and young people's tier one psychological

services with parents, and is due to be rolled out in tier 2 children and adolescent mental health

services. It will therefore be widely used in NHS settings in the UK. However, only one study

validated this measure and there was no information available for UK samples so it is difficult to

make firm recommendations on this basis.

Self Report Questionnaire

The SRQ was also promising with high discriminant validity with AUCs ranging from 0.74 to 0.826. Furthermore, this measure had high sensitivity and specificity when using the optimal cut off point. The measure had acceptable criterion validity, with a medium correlation between the SRQ and the Comprehensive Psychological Rating Scale ($r = 0.5$; Hanlon et al., 2008). Additionally, one study also found a medium correlation with the EPDS ($r = 0.53$; Abou-Saleh, 1997). No studies measured test-retest reliability so no conclusions can be made about its ability to discriminate psychological distress across time points. However, again only 3 studies validated this measure so it is difficult to make firm recommendations on this basis.

The Hopkins Symptom Checklist

The (Hopkins) Symptom Checklist (HSCL) was only validated by one study. This showed that the HSCL had poorer AUCs than the SRQ and K10, however the specificity and sensitivity of the measure was higher, as was the internal consistency. Non-validation studies showed acceptable to good internal consistency ($\alpha = 0.71$ to 0.87). However, no other studies provided validation information and no studies assessed test-retest reliability. With the lack of data on this measure, it is difficult to provide firm conclusions on the usefulness of this measure in the perinatal period.

The Brief Symptom Inventory

The Brief Symptom Inventory was the fourth most frequently used measure after the GHQ, HSCL and SRQ, with eighteen studies using this measure in early years parenting populations. However, no studies have validated the BSI during this time. Internal consistency was acceptable to high ($\alpha = 0.57$ to 0.99) and the BSI had the highest test-retest reliability of all the measures discussed ($r = .68$ to .91) (see Table 4 for comparisons of the measures). Only one study provided validity data and found that women's scores on the BSI predicted scores on the EPDS and Spielberg Trait Anxiety Inventory showing it is correlated with other measures of the same construct (Cak et al., 2015). The BSI is copyrighted so needs to be purchased for use.

Psychometric properties of paternal global mental health outcome measures

The number of studies that recruited just only fathers was low (n = 7). One of these was a validation study (Tran et al., 2012), and 6 were non-validation studies. Studies were carried out in Asia (n = 1), Europe (n = 1), and Australasia (n = 5). More studies recruited both mothers and fathers (n = 29) and all of these were non-validation studies. The studies that recruited only fathers used the GHQ (n = 4), MHI-5 (n = 2), K6/K10 (n = 1), and HSCL (n = 3).

General Health Questionnaire

Tran et al. (2012) validated the GHQ in Vietnamese men whose partners were pregnant, or had recently given birth. Internal consistency was good for the GHQ (Cronbach's a = 0.70). Discriminant validity was also good, with Tran et al. (2012) finding that it outperformed the EPDS and the Zung SAS in terms of discriminant validity (GHQ: AUC = 79.2; EPDS: AUC = 76.7; Zung SAS: AUC = 77.5). When the GHQ was used with fathers it performed better than it did with mothers (Tran et al., 2011). Using a recommended cut off of 0/1 its sensitivity was found to be 75.6% and its specificity was 74.7%). Condon, Boyce and Corkindale (2004) reported that there were no significant differences in GHQ scores from pregnancy to 12 months postpartum suggesting strong test re-test reliability. However, the same pattern was found for the MHI-5 which could suggest test-retest reliability is inflated by the stability of fathers' mood rather than the reliability of the measure. The other two studies that used this measure did not report psychometric properties (Boyce, Condon, Barton, & Corkindale, 2007; Roberts, Bushnell, Collings & Purdie, 2006).

Kessler Psychological Distress Scale.

Only one study reported psychometric information for the K6. Giallo, Cooklin, Wade, D'Esposito, and Nicholson (2014) reported that fathers' scores on the K6 at baseline predicted scores at repeated follow up points; suggesting this measure produces consistent results over time.

Hopkins Symptom Checklist

For the HSCL, Kvalevaag et al. (2014) reported that fathers' scores on the SCL-5 were strongly correlated with their scores on the SCL-25 ($r = 0.92$); suggesting the shortened version is as valid as the longer version.

Mental Health Inventory 5 (Mental Health scale of the SF-36)

As reported above, Condon, Boyce and Corkindale (2004) found that there were no significant differences in MHI-5 scores from pregnancy to 12 months postpartum suggesting good test re-test reliability or strong stability in fathers' mood across time.

Discussion

This review aimed to identify reliable, valid and acceptable measures of parental mental health to use with parents during pregnancy and up to 5 years postpartum. The review found a wide range of measures that are available for measuring global parental mental health and wellbeing. However, very little research directly validated these measures for use with parents in the 0-5 early years. Four measures (General Health Questionnaire, Kessler Psychological Distress Scale, Self Report Questionnaire, Hopkins Symptom Checklist) were validated for use during this time but overall there was fairly limited psychometric data. The majority of studies explored the validity of measures during the perinatal period (0-1 year) in mothers, and only a small number of studies were found that included fathers or explored validity across the entire period of early childhood (0-5 years).

It is important to ensure that there are valid and reliable measures available to enable effective screening of parental mental health and wellbeing in the early years. There is evidence that untreated parental mental health issues can adversely impact the infant and the family (Barry et al., 2015; Kinsella & Monk, 2009; Talge et al., 2007) therefore the ability to screen for these difficulties is important in health care practice. Valid screening tools will enable parental mental health problems to be identified early through universal services, such as maternity and primary care services (Hogg, 2013).

A number of characteristics are important in a screening tool. Firstly, validity and effectiveness are critical if the screening tool is to accurately identify parents who require help. Other characteristics are also important in clinical practice, such as brevity, cost and availability in different languages. In clinical settings it is important that screening measures are brief and easy to complete due to time and funding constraints for both the individual land the healthcare system (Papanicolas & Smith, 2013; Squires, 2011). Similarly, lack of resources may prohibit the use of screening tools that need to be paid for. The availability of questionnaires in different languages to ensure that the measure is suitable for everyone.

The questionnaires identified in this review have advantages and disadvantages. The GHQ is available in 19 languages and has shortened versions (e.g. GHQ-12), making it appealing in the health care context. However due to its poor validity and reliability in this population it cannot be recommended for use in early years parents. The BSI has been widely used (with acceptable internal consistency) but has not been validated for use in this population, is very long (53 items), and has to be purchased. The BSI also includes questions about somatic symptoms, which may be confounded by normal physical symptoms of pregnancy and postpartum. The lack of validity information, length and cost therefore mean it also cannot be recommended for use – particularly in publically funded health services. The HSCL had good internal consistency and discriminant validity with mothers and has been used with fathers. However, it has many versions ranging from 5 to 90 items and many of these need to be purchased. Although it is available in 10 languages for perinatal populations, the multiple versions mean there is limited validation and languages for some of the versions. This, coupled with the length and cost therefore mean it is also unlikely to be suited to use in clinical settings.

The remaining two measures had promising psychometric properties. The SRQ had good discriminant validity and acceptable criterion validity with mothers. It is fairly brief (20 items), free to use and available in five languages. The K10/6 had good internal consistency and good

discriminant validity with mothers and has also been used with fathers. It is brief (6 items), free to use and available in four languages. At the current time these measures therefore seem most appropriate for use with parents of children aged 0-5. However, both these measures have not been widely validated so more research is needed to evaluate them in this population in different cultural contexts.

It is worth also considering newer measures that are available that have not yet been used in early years parenting populations but which may be promising in the future. For example, the CORE-10 is a broad, 10-item measure that includes symptoms (anxiety, depression, insomnia, re-experiencing) and functioning (coping, support) with emerging evidence of good psychometric properties with pregnant women (Coates, Ayers, de Visser, unpublished). This review therefore highlights the need for further validation of the promising measures, or alternatively examination of newer tools to identify psychological distress in the early years.

Limitations

There are a number of limitations to the current review. The first is the sparsity of studies that directly validated measures in the early years. This makes it difficult to draw conclusions or make firm recommendations about the ideal measure to screen for emotional wellbeing in this population. Second, some of the measures that were most validated were developed many years ago. The amount of time since development means older measures are more likely to have been validated in different populations so may be over-represented in reviews such as this one. For example, the GHQ (developed in 1988) has 5 validation studies and 23 non-validation studies reporting psychometric information. In contrast, the Kessler (developed in 2003) has 1 validation study and 4 non-validation studies. As mentioned, there may also be other new measures that are promising but have not yet been evaluated. The amount of psychometric information available is therefore not necessarily a measure of quality and is influenced by time since the measure was developed. This needs to be

considered when drawing conclusions, particularly in relation to newer measures that have not been widely validated.

There are also methodological issues that need considering. A key issue is the potential confounding of measures of maternal psychological wellbeing by normal symptoms of pregnancy and postpartum. The length of time between measures of test-retest reliability also varied between studies and it is possible that the length of time influenced estimates of reliability. In addition, two of the studies were considered to have a medium risk of bias in how the validation was carried out, therefore this should be taken into account before drawing any conclusions.

Finally, although a wide range of screening tools were identified for screening general psychological distress only four of these have been validated in the early years period. More research needs to be done to validate these measures in this population.

Conclusions

The Kessler Psychological Distress Scale and Self Report Questionnaire were both identified as promising global measures of parental mental health from 0 to 5 years in terms of psychometric properties and clinical utility. However, this review also shows that more research is needed to validate available measures of general psychological distress in early years parenting populations. It is possible that other measures of psychological distress, such as the CORE-10, which are not included in this review may also be promising. Therefore, future research should also examine the validity of scales not included here to see if they are valid for use in the early years with both mothers and fathers. Identifying valid and reliable measures will enable effective screening for perinatal distress and potentially prevent the long term negative impact of parental mental health problems.

References

Abou-Saleh, M. T., & Ghubash, R. (1997). The prevalence of early postpartum psychiatric morbidity in Dubai: A transcultural perspective. *Acta Psychiatrica Scandinavica, 95*(5), 428-432.

Aderibigbe, Y. A., & Gureje, O. (1992). The validity of the 28-item General Health Questionnaire in a Nigerian antenatal clinic. *Soc Psychiatry Psychiatr Epidemiol, 27*(6), 280-283.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: American Psychiatric Association.

Asberg, M., Montgomery, S. A., Perris, C., Schalling, D., & Sedvall, G. (1978). A comprehensive psychopathological rating scale. *Acta Psychiatr Scand Suppl*(271), 5-27.

Barry, T. J., Murray, L., Fearon, R. M., Moutsiana, C., Cooper, P., Goodyer, I. M., . . . Halligan, S. L. (2015). Maternal postnatal depression predicts altered offspring biological stress reactivity in adulthood. *Psychoneuroendocrinology, 52*, 251-260. doi:10.1016/j.psyneuen.2014.12.003

Bauer, A., Parsonage, M., Knapp, M., Iemmi, V., & Adelaja, B. (2014). *The costs of perinatal mental health problems*. London, UK.: Centre for Mental Health and London School of Economics.

Beck, K. H., Shattuck, T., Haynie, D., Crump, A. D., & Simons-Morton, B. (1999). Associations between parent awareness, monitoring, enforcement and adolescent involvement with alcohol. *Health education research, 14*(6), 765-775. doi:10.1093/her/14.6.765

Bonari, L., Koren, G., Einarson, T. R., Jasper, J. D., Taddio, A., & Einarson, A. (2005). Use of antidepressants by pregnant women: evaluation of perception of risk, efficacy of evidence based counseling and determinants of decision making. *Arch Womens Ment Health, 8*(4), 214-220. doi:10.1007/s00737-005-0094-8

Brink, Y., & Louw, Q. A. (2012). Clinical instruments: reliability and validity critical appraisal. *J Eval Clin Pract, 18*(6), 1126-1132. doi:10.1111/j.1365-2753.2011.01707.x

Çak, H. T., Karabekiroğlu, K., Kültür, E. C., Tarakçioğlu, M. C., Kaya, R., Say, G. N., . . . Varol, F. (2015). Relationship between psychiatric symptoms in expectant parents and postpartum depression and infantile colic: A multicenter follow-up study. *Turk Psikiyatri Dergisi, 26*(2), 87-98.

Cameron, E. E., Sedov, I. D., & Tomfohr-Madsen, L. M. (2016). Prevalence of paternal depression in pregnancy and the postpartum: An updated meta-analysis. *J Affect Disord, 206*, 189-203. doi:10.1016/j.jad.2016.07.044

Cox, J. L., Holden, J. M., & Sagovsky, R. (1987). Detection of postnatal depression. Development of the 10-item Edinburgh Postnatal Depression Scale. *Br J Psychiatry, 150*, 782-786.

Dagher, R. K., McGovern, P. M., & Dowd, B. E. (2014). Maternity leave duration and postpartum mental and physical health: implications for leave policies. *J Health Polit Policy Law, 39*(2), 369-416. doi:10.1215/03616878-2416247

Dean, C., Surtees, P. G., & Sashidharan, S. P. (1983). Comparison of research diagnostic systems in an Edinburgh community sample. *Br J Psychiatry, 142*, 247-256.

Dennis, C. L. (2003). Detection, prevention, and treatment of postpartum depression. In D. E. Stewart, E. Robertson, C. L. Dennis, S. L. Grace, & T. Wallington (Eds.), *Postpartum depression: Literature review of risk factors and interventions.* . Toronto: Faculty of Nursing, University of Toronto.

Dennis, C. L., & McQueen, K. (2009). The relationship between infant-feeding outcomes and postpartum depression: a qualitative systematic review. *Pediatrics, 123*(4), e736-751. doi:10.1542/peds.2008-1629

Derogatis, L. R. (1993). *BSI Brief Symptom Inventory: Administration, Scoring, and Procedure Manual* (4th ed.). Minneapolis: National Computer Systems

Derogatis, L. R., Lipman, R. S., Rickels, K., Uhlenhuth, E. H., & Covi, L. (1974). The Hopkins

Symptom Checklist (HSCL): A self-report symptom inventory. *Behavioral Science, 19*(1), 1-

15. doi:10.1002/bs.3830190102

Fihrer, I., McMahon, C. A., & Taylor, A. J. (2009). The impact of postnatal and concurrent maternal

depression on child behaviour during the early school years. *J Affect Disord, 119*(1-3), 116-

123. doi:10.1016/j.jad.2009.03.001

First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. (1992). *Structured Clinical Interview for

DSM-IV-TR Axis I Disorders*. New York: Biometrics Research.

Fletcher, R. J., Feeman, E., Garfield, C., & Vimpani, G. V. (2011). The effects of early paternal

depression on children's development. *Medical Journal of Australia, 195*(11-12), 685-689.

Giallo, R., Cooklin, A., Wade, C., D'Esposito, F., & Nicholson, J. M. (2014). Fathers' postnatal

mental health and child well-being at age five: The mediating role of parenting behavior.

Journal of Family Issues, 35(11), 1543-1562. doi:10.1177/0192513X13477411

Gill, S. C., Butterworth, P., Rodgers, B., & Mackinnon, A. (2007). Validity of the mental health

component scale of the 12-item Short-Form Health Survey (MCS-12) as measure of common

mental disorders in the general population. *Psychiatry Res, 152*(1), 63-71.

doi:10.1016/j.psychres.2006.11.005

Glover, V. (2016). Maternal stress during pregnancy and infant and child outcomes. In A. Wenzel

(Ed.), *The Oxford Handbook of Perinatal Psychology* (pp. 268-283). New York, USA:

Oxford University Press.

Goldberg, D., Cooper, B., Eastwood, M. R., Kedward, H. B., & Shepherd, M. (1970). A standardized

psychiatric interview for use in community surveys. *British journal of preventive & social

medicine, 24*(1), 18-23.

Goldberg, D., & Williams, P. (1988). *A user's guide to the General Health Questionnaire*. Windsor,

UK: NFER-Nelson.

Grace, S. L., Evindar, A., & Stewart, D. E. (2003). The effect of postpartum depression on child cognitive development and behavior: a review and critical analysis of the literature. *Arch Womens Ment Health, 6*(4), 263-274. doi:10.1007/s00737-003-0024-6

Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology, 143*(1), 29-36. doi:10.1148/radiology.143.1.7063747

Hanlon, C., Medhin, G., Alem, A., Araya, M., Abdulahi, A., Hughes, M., . . . Prince, M. (2008). Detecting perinatal common mental disorders in Ethiopia: Validation of the self-reporting questionnaire and Edinburgh Postnatal Depression Scale. *J Affect Disord, 108*(3), 251-262. doi:http://dx.doi.org/10.1016/j.jad.2007.10.023

Hogg, S. (2013). *Prevention in mind. All babies count: Spotlight on Perinatal Mental Health*. NSPCC.

Horwitz, S. M., Briggs-Gowan, M. J., Storfer-Isser, A., & Carter, A. S. (2007). Prevalence, correlates, and persistence of maternal depression. *J Womens Health (Larchmt), 16*(5), 678-691. doi:10.1089/jwh.2006.0185

Horwitz, S. M., Kelleher, K. J., Stein, R. E., Storfer-Isser, A., Youngstrom, E. A., Park, E. R., . . . Hoagwood, K. E. (2007). Barriers to the identification and management of psychosocial issues in children and maternal depression. *Pediatrics, 119*(1), e208-218. doi:10.1542/peds.2005-1997

Howitt, D., & Cramer, D. (2008). *Introduction to SPSS in Psychology* (4th ed.). Essex, UK: Pearson Education Limited.

Ip, W. Y., & Martin, C. R. (2006). Psychometric properties of the 12-item General Health Questionnaire (GHQ-12) in Chinese women during pregnancy and in the postnatal period. *Psychol Health Med, 11*(1), 60-69. doi:10.1080/13548500500155750

Kaaya, S. F., Fawzi, M. C., Mbwambo, J. K., Lee, B., Msamanga, G. I., & Fawzi, W. (2002).

    Validity of the Hopkins Symptom Checklist-25 amongst HIV-positive pregnant women in

    Tanzania. *Acta Psychiatr Scand, 106*(1), 9-19.

Kessler, R. C., Barker, P. R., Colpe, L. J., Epstein, J. F., Gfroerer, J. C., Hiripi, E., . . . Zaslavsky, A.

    M. (2003). Screening for serious mental illness in the general population. *Arch Gen*

    Psychiatry, 60(2), 184-189.

Kinsella, M. T., & Monk, C. (2009). Impact of maternal stress, depression and anxiety on fetal

    neurobehavioral development. *Clin Obstet Gynecol, 52*(3), 425-440.

    doi:10.1097/GRF.0b013e3181b52df1

Kvalevaag, A. L., Ramchandani, P. G., Hove, O., Assmus, J., Eberhard-Gran, M., & Biringer, E.

    (2013). Paternal Mental Health and Socioemotional and Behavioral Development in Their

    Children. *Pediatrics, 131*(2), e463-e469.

Kvalevaag, A. L., Ramchandani, P. G., Hove, O., Eberhard-Gran, M., Assmus, J., Havik, O. E., . . .

    Biringer, E. (2014). Does paternal mental health in pregnancy predict physically aggressive

    behavior in children? *Eur Child Adolesc Psychiatry, 23*(10), 993-1002. doi:10.1007/s00787-

    014-0587-y

Lang, A.J., Gartstein, M.A. (2017). Intergenerational transmission of traumatization: Theoretical

    framework and implications for prevention. *J Trauma Dissociation*. doi:

    10.1080/15299732.2017.1329773. [Epub ahead of print]

Manktelow, B. M., Smith, L. K., Evans, T. A., Hyman-Taylor, P., Kurinczuk, J. J., Field, D. J.,

    Smith, P. W., Mielewczyk, F.,, & Draper, E. S., on behalf of the MBRRACE-UK

    collaboration. . (2015). *MBRRACE-UK Perinatal Mortality Surveillance Report UK*

    Perinatal Death for Births from January to December 2013 – Supplementary Report.

    Leicester: The Infant Mortality and Morbidity Studies Group, Department of Health

    Sciences,: University of Leicester.

Mann, R., Gilbody, S., & Adamson, J. (2010). Prevalence and incidence of postnatal depression: what can systematic reviews tell us? *Archives of Women's Mental Health, 13*(4), 295-305. doi:10.1007/s00737-010-0162-6

Milgrom, J., Ericksen, J., & Sved-Williams, A. (2016). Impact of parental psychiatric illness on infant development. . In A. L. Sutter-Dallay, N. M. C. Glangeaud-Freudenthal, A. Guedeney, & A. Riecher-Rossler (Eds.), *Joint Care of Parents and Infants in Perinatal Psychiatry* (pp. 47-78). Switzerland: Springer International Publishing.

Milgrom, J., & Gemmill, A.W. (2014). Screening for perinatal depression. *Best Practice and Research in Clinical Obstetrics and Gynaecology, 28*(1), 13-23. doi: 10.1016/j.bpobgyn.2013.08.014.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med, 6*(7), e1000097. doi:10.1371/journal.pmed.1000097

Murray, L., Kempton, C., Woolgar, M., & Hooper, R. (1993). Depressed mothers' speech to their infants and its relation to infant gender and cognitive development. *J Child Psychol Psychiatry, 34*(7), 1083-1101.

Naicker, K., Wickham, M., Colman, I. (2012). Timing of first exposure to maternal depression and adolescent emotional disorder in a national Canadian cohort. PLoS One. 2012;7(3):e33422. doi: 10.1371/journal.pone.0033422.

Navarro, P., Ascaso, C., Garcia-Esteve, L., Aguado, J., Torres, A., & Martin-Santos, R. (2007). Postnatal psychiatric morbidity: a validation study of the GHQ-12 and the EPDS as screening tools. *Gen Hosp Psychiatry, 29*(1), 1-7. doi:10.1016/j.genhosppsych.2006.10.004

NICE. (2014). Antenatal and postnatal mental health: clinical management and service guidance (CG192). Retrieved from https://www.nice.org.uk/guidance/cg192/resources/antenatal-and-postnatal-mental-health-clinical-management-and-service-guidance-35109869806789

O'Brien, L., Laporte, A., & Koren, G. (2009). Estimating the economic costs of antidepressant

discontinuation during pregnancy. *Can J Psychiatry, 54*(6), 399-408.

doi:10.1177/070674370905400607

O'Donnell, K. J., Glover, V., Barker, E. D., & O'Connor, T. G. (2014). The persisting effect of

maternal mood in pregnancy on childhood psychopathology. *Dev Psychopathol, 26*(2), 393-

403. doi: 10.1017/s0954579414000029

Papanicolas, I., & Smith, P. C. (Producer). (2013). Health System Performance Comparison: An

agenda for policy, information and research. *European Observatory on Health Systems and*

Policies Series Retrieved from

http://www.euro.who.int/__data/assets/pdf_file/0009/244836/Health-System-Performance-

Comparison.pdf

Parfitt, Y., & Ayers, S. (2014). Transition to parenthood and mental health in first-time parents.

Infant Ment Health J, 35(3), 263-273. doi:10.1002/imhj.21443

Post and Antenatal Depression Association. (2012, 18/04/2017). The cost of perinatal depression in

Australia FINAL REPORT. Retrieved from

https://www.deloitteaccesseconomics.com.au/uploads/File/PANDA%20Exec%20Summ%20

pdf.pdf

Prady, S. L., Miles, J. N., Pickett, K. E., Fairley, L., Bloor, K., Gilbody, S., . . . Wright, J. (2013).

The psychometric properties of the subscales of the GHQ-28 in a multi-ethnic maternal

sample: results from the Born in Bradford cohort. *BMC Psychiatry, 13*, 55.

doi:10.1186/1471-244x-13-55

Ramchandani, P., Stein, A., Evans, J., & O'Connor, T. G. (2005). Paternal depression in the postnatal

period and child development: a prospective population study. *Lancet, 365*(9478), 2201-

2205. doi:10.1016/s0140-6736(05)66778-5

SIGN. (2012). *Management of perinatal mood disorders*. Edinburgh: SIGN Retrieved from

   http://www.sign.ac.uk/guidelines/fulltext/127/.

Spies, G., Stein, D. J., Roos, A., Faure, S. C., Mostert, J., Seedat, S., & Vythilingum, B. (2009).

   Validity of the Kessler 10 (K-10) in detecting DSM-IV defined mood and anxiety disorders

   among pregnant women. *Arch Womens Ment Health, 12*(2), 69-74. doi:10.1007/s00737-009-

   0050-0

Spiteri, M. C., Jomeen, J., & Martin, C. R. (2013). Reimagining the General Health Questionnaire as

   a measure of emotional wellbeing: A study of postpartum women in Malta. *Women and

   Birth, 26*(4), E105-E111. doi:10.1016/j.wombi.2013.06.002

Spitzer, R. L., Kroenke, K., Williams, J. B., & Lowe, B. (2006). A brief measure for assessing

   generalized anxiety disorder: the GAD-7. *Arch Intern Med, 166*(10), 1092-1097.

   doi:10.1001/archinte.166.10.1092

Squires, D. A. (2011). The U.S. Health System in Perspective: A Comparison of Twelve

   Industrialized Nations. *Issues in International Health Policy, 16*, 1-14.

Sroufe, L. A. (2005). Attachment and development: a prospective, longitudinal study from birth to

   adulthood. *Attach Hum Dev, 7*(4), 349-367. doi:10.1080/14616730500365928

Swallow, B. L., Lindow, S. W., Masson, E. A., & Hay, D. M. (2003). The use of the General Health

   Questionnaire (GHQ-28) to estimate prevalence of psychiatric disorder in early pregnancy.

   Psychology, Health & Medicine, 8(2), 213-217. doi:10.1080/1354850031000087591

Sweeney, S., & MacBeth, A. (2016). The effects of paternal depression on child and adolescent

   outcomes: A systematic review. *J Affect Disord, 205*, 44-59. doi:10.1016/j.jad.2016.05.073

Talge, N. M., Neal, C., & Glover, V. (2007). Antenatal maternal stress and long-term effects on child

   neurodevelopment: how and why? *J Child Psychol Psychiatry, 48*(3-4), 245-261.

   doi:10.1111/j.1469-7610.2006.01714.x

Tesfaye, M., Hanlon, C., Wondimagegn, D., & Alem, A. (2010). Detecting postnatal common

    mental disorders in Addis Ababa, Ethiopia: Validation of the Edinburgh Postnatal Depression

    Scale and Kessler Scales *J Affect Disord* (Vol. 122).

Tran, T. D., Tran, T., & Fisher, J. (2012). Validation of three psychometric instruments for screening

    for perinatal common mental disorders in men in the north of Vietnam. *J Affect Disord,*

    136(1-2), 104-109. doi:10.1016/j.jad.2011.08.012

Tran, T. D., Tran, T., La, B., Lee, D., Rosenthal, D., & Fisher, J. (2011). Screening for perinatal

    common mental disorders in women in the north of Vietnam: A comparison of three

    psychometric instruments. *J Affect Disord, 133*(1-2), 281-293. doi:10.1016/j.jad.2011.03.038

Varescon, I., Leignel, S., Gérard, C., Aubourg, F., & Detilleux, M. (2013). Self-esteem,

    psychological distress, and coping styles in pregnant smokers and non-smokers

    Psychological Reports (Vol. 113).

Wang Y. (2017). Intergenerational Transmission of Depressive Symptoms: The Role of Parental

    Negative Perceptions and Behaviors. *Child Psychiatry Hum Dev*. doi: 10.1007/s10578-017-

    0734-z. [Epub ahead of print]

Whooley, M. A., Avins, A. L., Miranda, J., & Browner, W. S. (1997). Case-finding instruments for

    depression. Two questions are as good as many. *J Gen Intern Med, 12*(7), 439-445.

Wilson, S., & Durbin, C. E. (2010). Effects of paternal depression on fathers' parenting behaviors: a

    meta-analytic review. *Clin Psychol Rev, 30*(2), 167-180. doi:10.1016/j.cpr.2009.10.007

World Health Organisation. (1990). *Composite International Diagnostic Interview*. Geneva: World

    Health Organisation.

World Health Organisation. (1994). *Self-Reporting Questionnaire. .* Geneva: World Health

    Organisation

World Health Organisation. (2008). *Well-being measures in primary health care: the DepCare*

    project. Stockholm, Sweden.

Zelkowitz, P., & Milet, T. H. (1997). Stress and support as related to postpartum paternal mental

    health and perceptions of the infant. *Infant Ment Health J, 18*(4), 424-435.

    doi:10.1002/(SICI)1097-0355(199724)18:4<424::AID-IMHJ8>3.0.CO;2-K

Zigmond, A. S., & Snaith, R. P. (1983). The hospital anxiety and depression scale. *Acta Psychiatr*

    Scand, 67(6), 361-370.

Table 1. Psychometric measures of validity and reliability and their definitions

Psychometric measure	Definition
Internal consistency	Reliability measure. Measures the extent to which the items within a scale are consistent. This is usually calculated by correlating the score of one half of the measure with the other half or using Cronbach's alpha α (which is the average of all possible split-half tests) (Howitt & Cramer, 2008)
Test-retest reliability	Reliability measure. This calculates a measures consistency over time. The larger the correlation, the better the consistency.
Discriminant validity	Refers to the degree to which the scale measures what it intends to. The two main ways of measuring this is through the sensitivity (%) of a measure to pick up known 'cases' and the specificity (%) of a measure to excluded known 'non-cases'. A receiver operating characteristic curve (ROC) analysis can also been performed. The area under this curve (AUC) provides an indication of a scales ability to discriminate between those with and without a diagnosis (Hanley & McNeil, 1982). . It has been suggested that an AUC of above 0.5 suggests the measure is performing above chance (Gill,

	Butterworth, Rodgers, & Mackinnon, 2007; Spies et al., 2009).
Criterion validity	Refers to the extent to which a measure is related to an outcome that it is intended to be related to (i.e. an anxiety measure should correlate strongly with a diagnosis of GAD).

Table 2. Summary of measures and their content, use and language availability

Measure	Abbreviation	No. of Items	What the scale measures	Time frame of response	No. used measure	No. Psychometric	Countries measure used in	Language of measure
Measures of current mood								
Multiple Affect Adjective Checklist ©	MAACL	132	State and trait positive and negative affect	Not reported	1	1	South America	Spanish
Positive and Negative Affect Scale	PANAS	20	Positive affect, negative affect	Right now OR past week	3	2	Australia USA	English
Profile of Mood States ©	POMS	65	Anger, confusion, depression, fatigue, tension, vigour	Right now OR past week	5	2	Australia Japan USA	English Japanese
University of Wales Institute of Science & Technology Mood Adjective Checklist	UMACL UWIST	24 48	UMACL: Energetic arousal, tense arousal, hedonic tone, UWIST has additional items and	Right now	1	1	Poland	Polish

includes anger

Measures of negative symptoms								
Adult Wellbeing Scale		18	Depression, anxiety, irritability (outward and inward directed)	Last few days	1	0	UK	English
Crown-Crisp Experiential Index	CCEI	48	Depression, anxiety, phobia, obsessional, somatic anxiety, hysterical personality		1	0	UK	English
General Health Questionnaire ©	GHQ-6 GHQ-12 GHQ-28 GHQ-30	6 12 28 30	GHQ12: General psychopathology; GHQ28: Depression, anxiety/insomnia, somatic, social dysfunction; GHQ30: Excludes items related to physical illness	Last 7 days	65	24	Angola, Australia, Belgium, Brazil, Chile, China, Denmark, Finland, France, Hong Kong, India, Iran, Italy, Japan, Malta, Nepal, Netherlands, New	Arabic, Chilean, Chinese, Danish, Dutch, English, French, Iranian, Italian, Japanese, Maltese, Mirpuri, Nepalese, Nigerian, Norwegian, Polish,

							Zealand, Norway, Nigeria, Poland, Saudi Arabia, Spain, UK, USA, Vietnam	Spanish, Urdu, Vietnamese
Hopkins Symptom Check List Symptom Checklist	HSCL Or SCL	58 25 15 5 8	SCL58: Depression, anxiety, somatisation, obsessive-compulsive, interpersonal sensitivity; SCL25: Depression, anxiety. SCL15: subscale of the; SCL25 which measures depression; SCL5: five items	Last week	26	6	Australia, Bosnia & Herzegovina, Canada, Egypt, France, Germany, Italy, Israel, Netherlands, Norway, Spain, Switzerland, Tanzania, USA	Arabic, Bosnian, English, Dutch, French, German, Italian, Norwegian, Spanish, Swiss.
Kessler Psychological Distress Scale	K6 K10	6 10	General distress as an indication of serious mental illness	30 days	8	4	Australia, Japan, New Zealand, South Africa	English, Isi, Japanese, Zulu

Marcé Clinical Checklist		53	Clinician-administered scale to identify psychological disorders	Not applicable	1	0	New Zealand	English
Mental Health Inventory	MHI-5	5	MHI-5: General distress	Last month	5	2	Australia, USA	English, Spanish
	MHI-18*	18	MHI-18: Depression, anxiety, behavioural control, positive affect					
Self-Reporting Questionnaire	SRQ-20	20	General distress. Developed as a screening tool to identify probable cases of affective disorders	Last 30 days	21	6	Bangladesh, Brazil, Ethiopia, Ghana, Pakistan, UK, United Arab Emirates	Amharic, Arabic, English, Pakistani, Twi
Measures of positive wellbeing								
WHO Well-being Index 5	WHO-5	5	Positive wellbeing, physical energy, interest in life	Two weeks	1	0	USA	English
Warwick-Edinburgh	WEMW	7*	SWEMWBS (7 items)	Two weeks	1	1	USA	English

Mental Well Being Scale	BS	14	general positive wellbeing. WEMWBS (14 items) subjective wellbeing, positive mental health					
Measures of positive and negative symptoms								
Psychological General Wellbeing Inventory	PGWBI	22	Positive wellbeing, anxiety, depressed mood, self-control, general health and vitality	Past month	2	1	Sweden, USA	English, Swedish
Wellbeing Questionnaire	W-BQ	12	Positive wellbeing, negative wellbeing, energy		1	0	UK	English
Measures that include somatic symptoms								
Brief Symptom Inventory ©	BSI	53	Depression, anxiety, somatisation, hostility,	Last 7 days	18	10	Finland, Netherlands, Portugal, Turkey,	Dutch, Finnish, English, Portuguese,

			obsessive-compulsive,				USA	Turkish
			interpersonal					
			sensitivity, phobia,					
			paranoia, psychoticism					
Kellner Symptom	KSQ	92	Depression, anxiety,	Past week	1	0	USA	English
Questionnaire			anger-hostility, somatic					
			symptoms					
Primary Care Mental	PRIME-		Depression, anxiety,	Two weeks	3	0	Sweden, USA	NR
Disorders Screening	MD		somatic symptoms,					
Questionnaire			substance use,					
			eating/weight change					
Somatic & Psychological	SPHERE	12	SPHERE-12:	Past few weeks	1	0	New Zealand	English
Health Report		34*	Psychological health;					
			physical symptoms and					
			fatigue					
Other								
Prenatal Psychosocial	PPP		Stress, support from		1	1	Switzerland	NR

Profile			partner, support from others				
Short Form Quality of Life Scale ©	SF-12 SF-36	12 36	Mental health-related quality of life (MCS), physical health-related quality of life (PCS)	16	0	Australia, Canada, China, Denmark, New Zealand, Sweden, Taiwan, UK, USA	Chinese, Danish, English, Swedish, Taiwanese
Chinese Health Questionnaire	CHQ	12 30	Non-UK measure	3	2	China	Chinese
Shona Symptom Questionnaire	SSQ		Non-UK measure	2	0	Zimbabwe	NR

Table 3. Summary of psychometric properties and norms of widely used and evaluated measures

Measure	No. of studies reporting psychometric properties	No validation studies	AUC	Sensitivity	Specificity	Alpha (α)	Test-retest (r)	No. of studies in late perinatal period	No. of studies with fathers	Population based means
GHQ	23	5	0.72-0.904	74.7-84.6	56.5-92.1	.64-.95	.24-.49	6	4	**UK:** T1 mean 12.09 (SD 5.73); T2 mean 11.16 (SD 5.46) (Van Bussell et al. 2006); **India**: Pregnant women 3rd trimester mean 3.30 (SD = .15) (Bhat et al., 2015); **Norway:** Mothers: T1 mean 22.0; T2 mean 17.2; T3 mean 16.7; Fathers: T1 mean 16.4; T2 mean 17.9; T3 mean 15.9. Percentage defined to have psychological distress, mothers: T1: 37%, T2: 21%, T3: 19%, fathers: T1: 13%, T2: 11%, T3: 11% (Skari

										et al. 2002).
Kessler	4	1	0.86-0.91	50-84.6	54-82.7	.73-.90	.48	2	1	**Australia:** Mothers T1 mean 3.20 (SD=3.01); T2 mean 2.95 (SD = 3.10) (Giallo et al., 2014); **New Zealand:** First time parents mean 13.4 (SD=5). Subsequent parents mean 13.2 (SD=3.8) (McKenzie & Carter, 2013)
SRQ	6	1	0.74-0.826	59.2-83	80-85.4	.78-.85	NR	1	0	**Ethiopia:** Median score (25th centile, 75th centile): in pregnancy 2 (0, 4); two months postpartum 0 (0, 2); prevalence of CMD: 12% in pregnancy, 4.6% in postpartum period (Senturk et al., 2012)
HSCL	6	1	0.495-0.797	88	89	.80-.93	NR	3	3	**Norway:** Mothers mean 1.42 (Kvalevaag et al. 2014)
BSI	10	0	NR	NR	NR	.57 -.99	.68-.91	4	0	**Netherlands**: Mean 0.26 during pregnancy (N = 4848)

NOTES: * non-validation studies only. NR = not reported