Review

Screening for Postpartum Depression: Barriers, Guidelines, and Measurement Scales

Kathleen Kendall-Tackett *

Posted Date: 7 August 2024

doi: 10.20944/preprints202408.0463.v1

Keywords: postpartum; postnatal; depression; screening; Edinburgh Postnatal Depression Scale; Patient Health Questionnaire

Review

Screening for Postpartum Depression: Barriers, Guidelines, and Measurement Scales

Kathleen Kendall-Tackett

Texas Tech University Health Sciences Center, Amarillo

Abstract: Background: Screening for postpartum depression can lower prevalence and ensure that mothers get adequate treatment and support. Yet, few practitioners screen. The present article is a brief review on barriers to screening and two screening scales that are validated for perinatal women. **Findings:** Despite recommendations from health organizations, most new mothers are not screened for depression. Providers cite lack of time, opening "Pandora's box," and lack of resources for mothers who screen positive as reasons for why they do not screen. The Edinburgh Postnatal Depression Scale and the Patient Health Questionnaires are brief screening scales that have been validated for new mothers and are widely available. **Conclusions:** Screening is the necessary first step to identifying new mothers who are depressed. Practitioners who screen need a clear plan, and to know how to access available community resources, so they will know what to do when a mother screens positive.

Keywords: postpartum; postnatal; depression; screening; Edinburgh Postnatal Depression Scale; Patient Health Questionnaire

---

Screening for Postpartum Depression: Barriers, Guidelines, and Measurement Scales

Screening for depression is the vital first step for identifying and treating mothers affected by it. The US Agency for Healthcare Research and Quality (ARHQ) noted that screening increases the likelihood of postpartum depression being prevented or treated (Myers et al., 2013). A review of 23 studies (N=5,398) found that screening lowered the prevalence of postpartum depression and increased the likelihood of treatment or remission, especially if screening is followed by treatment protocols, care management, and clinical care (O'Connor et al., 2016). Despite this, most new mothers are not screened. This article provides a brief review of recent studies on barriers to screening for postpartum depression, some ways to incorporate screening into clinical practices, and information about the two most widely used screening tools: the Edinburgh Postnatal Depression Scale and the Patient Health Questionnaires.

Screening Guidelines

Health organizations concede that screening is important, but some guidelines are quite vague and not particularly helpful. For example, the American College of Obstetricians and Gynecologists Committee on Obstetric Practice (2023) noted that "routine screening by physicians is important for ensuring appropriate follow-up and treatment" without providing any details about what providers need to do. Busy practitioners are not likely to screen when given such generic advice.

A better example of screening guidelines comes from the UK's National Institute for Health and Clinical Excellence (NICE). NICE specifically recommends who should screen (the primary provider), when it should happen (at the first contact, at 4 to 6 weeks or 3 to 4 months), which professionals should do it (all), and what to ask (National Institute for Health and Care Excellence (NICE), 2020). These guidelines recommend the Edinburgh Postnatal Depression Scale, the Patient Health Questionnaire-9, and the Hospital Anxiety and Depression Scale (HADS).

 doi:10.20944/preprints202408.0463.v1

Barriers to Screening

Healthcare providers cite many reasons for not screening including lack of resources and opening "Pandora's box" (i.e., questions might reveal a situation that providers are unequipped to handle, so it is better to not ask). In Canada, their national health system recommends that providers not screen because potential false positives could overly burden their primary care providers (Lang et al., 2022). That recommendation is troubling. The authors provide no statistics on the number of false positives, and whether they are really a problem. In addition, 23% of Canadian mothers actually have either postpartum depression or anxiety (Statistics Canada, 2019). How will these mothers be identified and treated?

Lack of screening can also be differentially harmful to minority women. In US samples, ethnic minority women are even less likely to be screened for depression. When they are, their rates of depression, anxiety, and PTSD are much higher than they are for White women. The Listening to Mothers in California study included a representative sample of 2,539 women (Declercq et al., 2021). Non-Latina Black women had more depressive symptoms and less use of medication or counseling compared to non-Latina White women. However, screened women were 6 times more likely to receive counseling. The authors concluded that universal screening reduces health inequities and increased rates of treatment, especially for women of color.

Screening Scales

The Edinburgh Postnatal Depression Scale (EPDS) and the Patient Health Questionnaires (PHQs) are widely used to screen for postpartum depression and depression in pregnancy. The American Academy of Pediatrics, the US Preventive Services Task Force, and a review in the *American Family Physician* recommends both (Earls & Committee on Psychosocial Aspects of Child and Family Health, 2010; Maurer et al., 2018; Rafferty et al., 2019; Siu & US Preventive Services Task Force (USPSTF), 2016) . However, it is important to note that both measures are *screening scales* and should not be used to diagnose depression—but they frequently are. Both scales have been validated with perinatal samples and have advantages and limitations. All these scales are available for free online.

The Edinburgh Postnatal Depression Scale (EPDS)

The Edinburgh Postnatal Depression Scale (EPDS) is the most widely used postpartum-depression screening tool in the world (Cox et al., 1987). It is a 10-item self-report questionnaire. The EPDS has been validated worldwide and has been translated into many languages. Women are asked to report how they have felt *in the past week*, and each item is scored from 0 to 3. The EPDS was developed to address the limitations of more generic depression measures that do not apply to postpartum women.

Validation of the EPDS

Validation is the process by which researchers determine whether a scale measures what it purports to measure. The new scale is compared to results of established ones known to measure the construct (e.g., depression). If a new scale identifies the same percentage of people as depressed as the established scale, researchers believe that it is a valid measure. The EPDS has been validated against the Center for Epidemiologic Studies-Depression Scale (CES-D), a 20-item diagnostic scale for depression. The CES-D is more detailed and can be used for diagnosis, but the results are consistent with the EPDS (Logsdon et al., 2009). When used in the first 6 weeks, the EPDS performs as well as the Hamilton Rating Scale for Depression-17 and -21, and Beck Depression Inventory when using optimal cutoffs (Myers et al., 2013).

A study validated the Danish version of the EPDS against the DSM-5 and the ICD-10 diagnostic criteria for postpartum depression (Smith-Nielsen et al., 2018). Their sample included 324 women at 2 and 10 months postpartum. They used the Structured Clinical Interview for DSM-5. Eleven was the optimal cutoff for both DSM-5 and ICD-10 criteria.

 doi:10.20944/preprints202408.0463.v1

The EPDS also identifies anxiety. King (2012) used the EPDS as a screening tool for 169 low-income African American women, a group typically underrepresented in postpartum depression studies. She found high rates of depression with her sample. When the cutoff was ≥10, 30% screened positive for depression, and 19% did when the cutoff was ≥13. A confirmatory factor analysis can identify other conditions that a scale measures. With confirmatory factor analysis, she found a three-factor model with depression, anxiety, and anhedonia being the conditions that the EPDS identified (i.e., the best fit for her sample). Shorter versions of the EPDS, such as the EPDS-3, focus specifically on anxiety symptoms (Kabir et al., 2008).

Scale Cutoffs

Twelve is the standard cutoff on the EPDS for depression (Cox, 2019), but higher and lower cutoffs have been used. The cutoff can vary depending on its intended purpose: broad vs narrow screening. Broad screening, with lower cutoffs (EPDS score=9-11), identifies women with milder symptoms and some who are not depressed (i.e., false positives). Narrow screening, with higher cutoffs (EPDS score-12-14+), identifies women with severe depression but misses others with milder symptoms. Which cutoff you use depends on your goals. If you want to identify all women with depression, the lower cutoff is better. Using 9 as a cutoff increased the sensitivity of the scale, which makes it more appropriate for community screening (Dennis, 2004; Dennis et al., 2004). In contrast, a higher cutoff has more specificity, but misses a lot of depressed women. In her study of 594 women, a cutoff of 12/13 at week 1 missed depression in 43% of mothers at 4 weeks, and 53% of mothers at 8 weeks (Dennis, 2004).

In a review of 58 studies (N=15,557), 2,069 women had major depression. Across studies, a cutoff of 11 led to the highest combination of sensitivity and specificity (Levis et al., 2020). The higher cutoff dramatically decreased sensitivity but increased specificity. As expected, a higher cutoff identifies women with more severe symptoms, but lower cutoffs identify the highest percentage of women with mild depression who do not meet diagnostic criteria for major depression. Interestingly, their findings were consistent across countries, but few of the studies were from low- or middle-income countries.

Dennis et al. (2016) compared cutoffs of 9 vs. 12 using a stratified sample of 1,125 Canadian new mothers who were screened at 1 and 16 weeks postpartum. They found that a cutoff of 9 had greater sensitivity and more accurately identified mothers who were likely to be depressed at 16 weeks. They noted that 64 mothers identified at 1 week with depressive symptoms had recovered by 16 weeks. The mothers falsely identified as depressed experienced no harm. The higher cutoff (12/13) led to fewer false positives but was far less sensitive.

Conversely, a practitioner or program with limited resources, who can only treat women with more severe symptoms, may find a higher cutoff to be more appropriate. A review of 23 studies (N=5,398) using the English-language EPDS found that a cutoff of 13 or higher had a sensitivity 0.8 and a specificity of 0.9 for detecting major depression (O'Connor et al., 2016). A random-effects meta-analysis compared the EPDS to the Structured Clinical Interview for DSM (SCID) (Lyubenova et al., 2021). The samples across 29 studies included 7,315 women, 1,017 with SCID scores that indicated major depression. Not surprisingly, a cutoff of 9 identified more women as depressed (28%) than using a higher cutoff. Using a cutoff of 10 identified 22% as depressed, compared to 11% with a cutoff of 13, while only 9% were identified with the SCID. The most stringent cutoff—14—identified the same percentage of women as the SCID. The authors concluded that the EPDS did not accurately estimate prevalence of major depression at lower cutoffs, but did with a higher cutoff.

Advantages and Disadvantages of the EPDS

The EPDS offers several advantages. It is easy to complete. Mothers can answer all the questions in a few minutes. Researchers can justify its use (i.e., most studies use it), which makes it easier to compare results across published articles. The limitations of the EPDS include language and scoring issues.

 doi:10.20944/preprints202408.0463.v1

Language Issues

The EPDS can be difficult to accurately translate because it uses colloquial British English expressions, so literal translation often makes little sense for mothers outside the UK or the Commonwealth (Waqas et al., 2021). A review of 16 studies with mothers in low-to-middle-income countries (LMIC) aggregated findings from 12 countries in 14 languages (Shrestha et al., 2016). The sample included 1,281 mothers. They found that most local-language versions were less precise in identifying depression than the original-English version. Out of 16 studies, only one had a culturally sensitive translation. Most studies simply translated the original language, without cultural adaptation. They found better performance of the EPDS when local expressions were used in the translations. The two studies that made these adaptations almost doubled their sensitivity.

Literal translations were also more difficult for mothers who have low literacy. Shrestha et al. (2016) recommended that translations explicitly consider local terminology for psychological distress and mental health disorders. They recommended that cutoff scores be validated for each culture along with culturally validated standard diagnostic protocol. An inaccurate measure may mean that mothers with depression do not get identified.

"Translation" issues for the EPDS can be a concern even when comparing American vs British English (Cox, 2019). American mothers often find the wording of some questions to be odd (e.g., "things are getting on top of me lately"). Even in the US, using a colloquial British expression, without considering how American mothers describe themselves as being overwhelmed, would not accurately assess how they were feeling. A new US version of the EPDS may address some of these concerns (Moyer et al., 2023).

Scoring Issues

Another challenge is that the scale contains several items that use reverse scoring, which make it easy to mis-score them. Matthey et al. (2013) collected 496 EPDSs from client files and 6 practices in Australia. They asked 22 clinicians to estimate their rate of errors. Clinicians made many more scoring mistakes than they realized. Matthey et al. found that 17% of the forms had at least one error, where clinicians scored answers incorrectly. Most were off by only one point, but that can make the difference between a positive and negative screen for depression.

The Edinburgh Postnatal Depression Scale-3 (EPDS-3)

Shortened versions of the EPDS are often more sensitive and focus particularly on anxiety symptoms (Walker et al., 2015). One study used the anxiety subscale from the EPDS and compared it to the full 10-item scale and the ultra-brief 2-item screener (Kabir et al., 2008). Their sample was 199 14-to-26-year-old participants in an adolescent maternity program. A total of 21% of the mothers met the criteria for depression (EPDS >10). The EPDS-3 had the best performance, with 95% sensitivity and 98% negative predictive value. It identified 16% more mothers as depressed than the full EPDS did. The 3-item screener included the following questions:

I have blamed myself unnecessarily when things went wrong. (item 3)

I have felt scared and panicky for no very good reason. (item 5)

I have been anxious or worried for no good reason. (item 4)

Kabir et al. suggested that the EPDS-3 is brief enough to be incorporated into well-baby checks and it identified a higher percentage of women as possibly depressed than the full EPDS. Similarly, King (2012) found, with her sample of African American women, that anxiety symptoms were the only significant ones that predicted postpartum depression at 4 to 6 weeks. The anxiety items were self-blame (item 3), anxiety (item 4), scared or panicking (item 5), inability to cope (item 6), and difficulty sleeping (item 7).

 doi:10.20944/preprints202408.0463.v1

Summary

The EPDS is a widely available screening scale that is reliable and valid. There are scoring and language challenges with the EPDS, but these can be overcome. Even with these limitations, the EPDS is an accurate measure of depression in new mothers and is most sensitive with a lower cutoff score.

The Patient Health Questionnaires

The Patient Health Questionnaires (PHQs) have been widely adopted by many health organizations and are available in 2-, 4-, and 9-item versions. The *Diagnostic and Statistical Manual-5* recommends the PHQ-9 as a tool for evaluating the severity of depression in the general population (Sun et al., 2020). The Patient Health Questionnaires (PHQs) are screening tools available for free online, and they are in American English, which is easier for American mothers and may be less colloquial in its phrasing, allowing for easier translation. They are not specific to postpartum depression but have been validated for use with both pregnant and postpartum women. Some find the shorter versions are easier to use in clinical settings. All versions have good reliability and validity.

Patient Health Questionnaire-9 (PHQ-9)

The PHQ-9 is the master scale and includes 9 multiple-choice questions that correspond to the 9 symptoms of major depression in the *Diagnostic and Statistical Manual, 5th Edition* (DSM-5) (American Psychiatric Association, 2013). It has high specificity and sensitivity. Criterion validity was established by comparing it to results from structured mental health professional interviews of 580 patients. Construct validity was assessed by comparing results to the 20-item Short-Form General Health Survey. It was originally validated on a sample of 6,000 patients: 3,000 from primary care and 3,000 from OB/Gyn. The Cronbach's alpha (a measure of internal reliability) was 0.89 for the primary care sample and 0.86 for the OB/Gyn sample. It also had excellent test-retest reliability (Kroenke et al., 2001).

One study compared the Patient Health Questionnaire-9 (PHQ-9) and the depression questions from the Pregnancy Risk Assessment Monitoring Systems (PRAMS-6) (Davis et al., 2013). These measures were validated against the Structured Clinical Interview for Depression (SCID) and the Hamilton Rating Scale for Depression (HAM-D). The PRAMS-6 and PHQ-9 were both effective for screening for postpartum depression and had moderate accuracy when compared with the SCID. The PHQ-9 had slightly better accuracy than the PRAMS-6.

Patient Health Questionnaire-2 (PHQ-2)

The 2-Item Patient Health Questionnaire (PHQ-2) is a shortened version of the PHQ-9, which can be used in all healthcare settings. It has also been validated to use with new mothers (Walker et al., 2015). Both the American Academy of Pediatrics and the American College of Obstetricians and Gynecologists recommend its use (Earls & Committee on Psychosocial Aspects of Child and Family Health, 2010). However, if the score on the PHQ-2 is higher than 3 (possible depression), they recommend screening mothers with the full PHQ-9.

The PHQ-2 assesses frequency of two mental states: anhedonia ("little interest or pleasure in doing things") and depressed mood ("feeling down, depressed or hopeless") during *the past two weeks*. The response categories include Not at All, Several Days, More than Half the Days, and Nearly Every Day. The PHQ-2 was used in the US Centers for Disease Control's Pregnancy Risk Assessment Monitoring Systems (PRAMS) to assess depression in new mothers (Bauman et al., 2020).

A study of postpartum women from Kenya found that the PHQ-2 did not detect depression as well as other measures (Larsen et al., 2023). This study included 3,605 women and used several scales to assess the prevalence of moderate-to-severe depression: the Center for Epidemiologic Studies Depression Scale (CESD-10), the Edinburgh Postnatal Depression Scale (EPDS), the Patient Health Questionnaire-9 and -2 (PHQ-9 and PHQ-2). The PHQ-2 had the poorest performance in detecting depression compared to the other scales.

 doi:10.20944/preprints202408.0463.v1

Another study compared the Edinburgh Postnatal Depression Scale to the PHQ-2 for their accuracy in detecting depression in 252 pregnant women from Australia (Slavin et al., 2020). They also found that the PHQ-2 had low sensitivity (i.e., failed to detect women who were depressed) but high specificity using cutoffs of both 2 or 3. The authors tested women at 4 points during pregnancy and postpartum and concluded that the PHQ-2 missed an unacceptably high number of women with major depression. One possible reason is that the PHQ-2 does not include questions about anxiety, which researchers who developed the EPDS-3 discovered were critical. The PHQ-4 addresses this limitation.

Patient Health Questionnaire-4 (PHQ-4)

The PHQ-4 combines the PHQ-2 with questions from the Generalized Anxiety Disorder-2 scale (GAD-2), a 2-item measure of anxiety (Kroenke et al., 2009). The PHQ-4 uses the same response categories as the PHQ-2. A score of 3 or more on the first two questions indicates anxiety symptoms, and a score of 3 or more on the PHQ-2 on questions 3 and 4 indicates depressive symptoms (Wicke et al., 2022). The scale accurately predicts generalized anxiety disorder, panic disorder, and social anxiety disorder but not posttraumatic stress disorder. A study of 6,874 women from 64 countries examined the impact of giving birth during COVID-19 (Basu et al., 2021). The PHQ-4 yielded results consistent with other measures of PTSD and worry and indicated that 31% exceeded the threshold for depression or anxiety.

The Listening to Mothers in California survey (N=2,539) found that the PHQ-4 was a valid measure for all mothers in their study including racial/ethnic minorities (Declercq et al., 2021). The PHQ-4 was also used to measure depression and anxiety in an 11-year longitudinal study of 20 women who developed pregnancy and lactation-related osteoporosis (Gehlen et al., 2019).

A Spanish study tested the psychometric properties and validation of the PHQ-4 with a sample of 845 pregnant women recruited from public hospitals (Rodriguez-Munoz et al., 2020). Exploratory and confirmatory factor analysis found that the PHQ-4 adequately assessed depression and anxiety during pregnancy. They concluded that it was a reliable and valid instrument for use with pregnant women.

A study from rural Pakistan used three short screening tools to screen for depression and anxiety: the PHQ-4, the 4-item Hamilton Depression Rating Scale, and the Community Informant Detection Tool (Waqas et al., 2021). All were effective but the PHQ-4 had "excellent psychometric properties to screen for perinatal depression," with a sensitivity of 93.4% and specificity of 91.7%.

In the US, a study examined the reliability and validity of the PHQ-4 for pregnant women (Barrera et al., 2021). Their sample included 1,148 pregnant women of diverse cultural background who were English (n=587) or Spanish (n=561) speaking. Using confirmatory factor analysis, they found that the PHQ-4 showed good reliability and validity for pregnant women as a screen for anxiety and depression.

Summary

Like the EPDS, the PHQs have solid track records as assessment tools for new mothers. The PHQ-9 and PHQ-4 have better accuracy. The PHQ-2 can be used if its limitations are recognized and addressed.

Conclusions

Screening is a necessary first step to treatment. Mothers cannot be treated if practitioners do not know they are depressed. Health organizations recognize the importance of screening and recommend it—but most new mothers are never screened. Unfortunately, some health guidelines are vague, and practitioners do not know what to do when mothers screen positive. Barriers to screening include time constraints, not having clearly identified steps of what to do when there is a positive screen, and fear of opening "Pandora's box." When administrators mandate screening, but do not follow up to make sure that it was done, practitioners are unlikely to do it. Organizations need

 doi:10.20944/preprints202408.0463.v1

to make sure that providers have "buy in" and understand why it is relevant to their clinical practice. Electronic prompts in medical records are one way to increase screening in health settings.

Many programs administer the EPDS because it is so widely used. However, practitioners should consider whether this scale is appropriate for mothers in their communities. Providers should also be aware of possible scoring mistakes and how to safeguard against them. Low cutoffs on the EPDS will identify more depressed mothers. However, some communities with scarce resources might opt for a more stringent cutoff allowing them to triage mothers with severe depression. The EPDS-3 has good reliability and validity and identifies anxiety symptoms.

The Patient Health Questionnaires are excellent choices for screening perinatal women that health organizations recommend. The PHQs are easy-to-use scales that many prefer over the EPDS because the language is less colloquial. The scales with the highest reliability are the PHQ-9 and the PHQ-4. The PHQ-2 is useful, but its reliability and validity are lower. The PHQ-4 also includes a brief measure for anxiety adapted from the Generalized Anxiety Scale-2.

Screening Is the First Step. What Happens Next?

Screening should lead to action. According to the US Preventive Services Task Force, screening for depression should take place in a context where there are adequate systems in place to ensure accurate diagnosis, effective treatment, and appropriate follow-up (O'Connor et al., 2016). Yeaton-Massey and Herrero (2019) concur and recommend education, support, and timely referral to resources. In practical terms, this means figuring out what to do if a mother screens positive.

Unfortunately, some practitioners engage in magical thinking about depression. They believe that including a screening scale in their assessments "fixes" depression. It doesn't. To really be effective, practitioners need a plan and to know what happens when mothers screen positive. Screening without follow up is useless.

As you contemplate possible screening with your patient population, I urge you to find resources in your community that can help depressed mothers. If a mothers screens positive, you will know what to do and where to send them. Catching depression early can stop its escalation into more severe symptoms and can save mothers and their families years of misery. It is very much worth the effort.

References

American College of Obstetricians and Gynecologists. (2023). Treatment and management of mental health conditions during pregnancy and postpartum. *5*(June). https://www.acog.org/clinical/clinical-guidance/clinical-practice-guideline/articles/2023/06/treatment-and-management-of-mental-health-conditions-during-pregnancy-and-postpartum

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual-V*. American Psychiatric Association.

Barrera, A., Moh, Y. S., Nichols, A., & Le, H.-N. (2021). The factor reliability and convergent validity of the Patient Health Questionnaire-4 among an international sample of pregnant women. *Journal of Women's Health (Larchmont)*, *30*(4), 525-532. https://doi.org/10.1089/jwh.2020.8320

Basu, A., Kim, H. H., Basaldua, R., Choi, K. W., Charron, L., Kelsall, N., Hernandez-Diaz, S., Wyszynski, D. F., & Koenen, K. C. (2021). A cross-national study of factors associated with women's perinatal mental health and wellbeing during the COVID-19 pandemic. *PLoS One*, *16*(4). https://doi.org/https://doi.org/10.1371/journal.pone.0249780

Bauman, B. L., Ko, J. Y., Cox, S., D'Angelo, D. V., Warner, L., Folger, S., Tevendale, H. D., Coy, K. C., Harrison, L., & Barfield, W. D. (2020). Postpartum depressive symptoms and provider discussions about perinatal depression: United States 2018. *Morbidity & Mortality Weekly Report*, *69*(19), 575-581.

Cox, J. (2019). Thirty years with the Edinburgh Postnatal Depression Scale: Voices from the past and recommendations for the future. *British Journal of Psychiatry*, *214*, 127-129. https://doi.org/10.1192/bjp.2018.245

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

Davis, K., Pearlstein, T., Stuart, S., O'Hara, M. W., & Zlotnick, C. (2013). Analysis of brief screening tools from the detection of postpartum depression: Comparisons of the PRAMS 6-item instrument, PHQ-9, and structured interviews. *Archives of Women's Mental Health*, *16*, 271-277.

Declercq, E., Feinberg, E., & Belanoff, C. (2021). Racial inequities in the course of treating perinatal mental health challenges: Results from Listening to Mothers in California. *Birth*, *49*, 132-140. https://doi.org/10.1111/birt.12584

Dennis, C.-L. (2004). Can we identify mothers at risk for postpartum depression in the immediate postpartum period using the Edinburgh Postnatal Depression Scale? *Journal of Affective Disorders*, *78*, 163-169.

Dennis, C.-L., Janssen, P. A., & Singer, J. (2004). Identifying women at-risk for postpartum depression in the immediate postpartum period. *Acta Psychiatrica Scandanavica*, *110*, 338-346.

Dennis, C.-L., Merry, L., Stewart, D., & Gagnon, A. J. (2016). Prevalence, continuation, and identification of postpartum depressive symptomatology among refugee, asylum-seeking, non-refugee immigrant, and Canadian-born women: Results from a prospective cohort study. *Archive of Women's Mental Health*, *19*, 959-967. https://doi.org/ 10.1007/s00737-016-0633-5

Earls, M. F., & Committee on Psychosocial Aspects of Child and Family Health. (2010). Clinical report incorporating recognition and management of perinatal and postpartum depression into pediatric practice. *Pediatrics*. https://doi.org/10.1542/peds.2010-2348

Gehlen, M., Lazarescu, A. D., Hinz, C., Schwarz-Eywill, M., Pfeifer, M., Balasingam, S., & Maier, A. (2019). Long-term outcome of patients with pregnancy and lactation-associated osteoporosis (PLO) with a particular focus on quality of life. *Clinical Rheumatology*, *38*(12), 3575-3583. https://doi.org/10.1007/s10067-019-04758-0

Kabir, K., Sheeder, J., & Kelly, L. S. (2008). Identifying postpartum depression: Are 3 questions as good as 10? *Pediatrics*, *122*, e696-e702.

King, P. A. (2012). Replicability of structural models of the Edinburgh Postnatal Depression Scale (EPDS) in a community sample of postpartum African American women with low socioeconomic status. *Archives of Women's Mental Health*, *15*, 77-86.

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine*, *16*, 606-613.

Kroenke, K., Spitzer, R. L., Williams, J. B. W., & Lowe, B. (2009). An ultra-brief screening scale for anxiety and depression: The PHQ-4. *Psychosomatics*, *50*, 613-621.

Lang, E., Colquhoun, H., LeBlanc, J. C., Riva, J. J., Moore, A., Traversy, G., Wilson, B., Grad, R., & Canadian Task Force on Preventive Health Care. (2022). Recommendation on instrument-based screening for depression during pregnancy and the postpartum period. *Canadian Medical Association Journal*, *194*(28), E981-E-989. https://doi.org/10.1503/cmaj.220290

Larsen, A., Pintye, J., Obhiambo, B., Mwongeli, N., Marwa, M. M., Watoyi, S., Kinuthia, J., Abuna, F., Gomez, L., Dettinger, J., Bhat, A., & John-Stewart, G. J. (2023). Comparing depression screening tools (CESD-10, EPDS, PHQ-9, and PHQ-2) for diagnostic performance and epidemiologic associationis among postpartum Kenyan women: Implications for research and practice. *Journal of Affective Disorders*, *324*, 637-644. https://doi.org/10.1016/j.jad.2022.12.101

Levis, B., Negeri, Z., Sun, Y., Benedetti, A., & Thombs, B. D. (2020). Accuracy of the Edinburgh Postnatal Depression Scale (EPDS) to detect major depression among pregnant and postpartum women: A systematic review and meta-analysis of individual participant data. *British Medical Journal*, *371*, m4022. https://doi.org/10.1136/bmj.m4022

Logsdon, M. C., Wisner, K. L., & Hanusa, B. H. (2009). Does maternal role functioning improve with antidepressant treatment in women with postpartum depression. *Journal of Women's Health*, *18*(1), 85-90.

Lyubenova, A., Neupane, D., Levis, B., Wu, Y., Sun, Y., He, C., & Thombs, B. D. (2021). Depression prevalence based on the Edinburgh Postnatal Depression Scale compared to Structured Clinical Interview for DSM Disorders classfication: Systematic review and individual participant data meta-analysis. *International Journal of Methods in Psychiatric Research*, *30*, e1860. https://doi.org/10.1002/mpr.1860

Matthey, S., Lee, C., Crncec, R., & Trapolini, T. (2013). Errors in scoring the Edinburgh Postnatal Depression Scale. *Archives of Women's Mental Health*, *16*, 117-122.

Maurer, D. M., Raymond, T. J., & Davis, B. N. (2018). Depression: Screening and diagnosis. *American Family Physician*, *98*(8), 508-515.

Moyer, S. W., Kinser, P. A., Nunziatio, J. D., Holmes, C. M., & Salisbury, A. L. (2023). Development of the Edinburgh Postnatal Depression Scale-United States: An updated perinatal mental health screening tool using a respectful care and trauma-informed approach. *Journal of Women's Health*. https://doi.org/10.1089/jwh.2023.0141

Myers, E. R., Aubuchon-Endsley, N., Bastian, L. A., Gierisch, J. M., Kemper, A. R., & Swamy, G. K. (2013). *Efficacy and safety of screening for postpartum depression* Agency for Healthcare Research and Quality.

National Institute for Health and Care Excellence (NICE). (2020). Antenatal and postnatal mental health: Clinical management and service guidance. https://www.nice.org.uk/guidance/cg192

O'Connor, E., Rossom, R. C., Henniger, M., Groom, H. C., & Burda, B. U. (2016). Primary care screening for and treatment of depression in pregnant and postpartum women: Evidence report and systematic review for the US Preventive Services Task Force. *JAMA*, *315*.

 doi:10.20944/preprints202408.0463.v1

Rafferty, J., Mattson, G., Earls, M. F., Yogman, M. W., Committee on Psychosocial Aspects of Child and Family Health, Gambon, T. B., Lavin, A., & Wissow, L. S. (2019). Incorporating recognition and management of perinatal depression into pediatric practice. *Pediatrics*, *143*(1), e20183260. https://doi.org/10.1542/peds.2018-3260

Rodriguez-Munoz, M. D. L. F., Ruiz-Segovia, N., Soto-Balbuena, C., Le, H.-N., Olivares-Crespo, M. E., & Izquierdo-Mendez, N. (2020). The psychometric properties of the Patient Health Questionnaire-4 for pregnant women. *Internationa Journal of Environmental Research in Public Health*, *17*, 7583. https://doi.org/10.3390/ijerph17207583

Shrestha, S. D., Pradhan, R., Tran, T. D., Gualano, R. C., & Fisher, J. R. W. (2016). Reliability and validity of the Edinburgh Postnatal Depression Scale (EPDS) for detecting perinatal common mental disorders (PCMDs) among women in low-and lower-middle-income countries: A systematic review. *BMC Pregnancy and Childbirth*, *16*.

Siu, A. L., & US Preventive Services Task Force (USPSTF). (2016). Screening for depression in adults: US Preventive Services Task Force recommendations. *JAMA*, *315*(4), 380-387. https://doi.org/10.1001/jama.2015.18392

Slavin, V., Creedy, D. K., & Gamble, J. (2020). Comparison of screening accuracy of the Patient Health Questionnaire-2 using two case-identification methods during pregnancy and postpartum. *BMC Pregnancy and Childbirth*, *20*, 211. https://doi.org/10.1186/s12884-020-02891-2

Smith-Nielsen, J., Matthey, S., Lange, T., & Vaever, M. S. (2018). Validation of the Edinburgh Postnatal Depression Scale against both DSM-5 and ICD-10 diagnostic criteria for depression. *BMC Psychiatry*, *18*, 393. https://doi.org/10.1186/s12888-018-1965-7

Statistics Canada. (2019). Maternal mental health in Canada, 2018/2019. *The Daily*. https://www150.statcan.gc.ca/n1/daily-quotidien/190624/dq190624b-eng.htm

Sun, Y., Fu, Z., Bo, Q., Mao, Z., Ma, X., & Wang, C. (2020). The reliability and validity of PHQ-9 in patients with major depressive disorder in psychiatric hospitals. *BMC Psychiatry*, *20*.

Walker, E. R., Gao, J., & Xie, B. (2015). Postpartum psychosocial and behavioral health: A systematic review of Self-Administered Scales validated for postpartum women in the United States. *Women's Health Issues*, *25*, 586-600.

Waqas, A., Malik, A., Atif, N., Nisar, A., Nazir, H., Sikander, S., & Rahman, A. (2021). Scalable screening and treatment response monitoring for perinatal depression in low- and middle-income countries. *International Journal of Environmental Research and Public Health*, *18*(13), 6693. https://doi.org/10.3390/ijerph18136693

Wicke, F. S., Krakau, L., Lowe, B., Beutel, M. E., & Brahler, E. (2022). Update of the standardization of the Patient Health Questionnaire-4 PHQ-4) in the general population. *Journal of Affective Disorders*, *312*, 310-314. https://doi.org/10.1016/j.jad.2022.06.054

Yeaton-Massey, A., & Herrero, T. (2019). Recognizing maternal mental health disorders: Beyond postpartum depression. *Current Opinions in Obstetrics & Gynecology*, *31*(2), 116-119. https://doi.org/10.1097/GCO.0000000000000524