The Diagnostic Properties of the Translated Chinese Whooley Questionnaire as A Case-Finding Instrument for Depression Among Chinese Women During and After Pregnancy

Kris Yuet-Wan Lok (✉ krislok@hku.hk)

University of Hong Kong <https://orcid.org/0000-0002-3227-0799>

Charlotte Chow

University of Hong Kong

Shirley W Tan

University of Hong Kong

Robert Smith

University of Hong Kong

Jessie Lin

The Hong Kong Polytechnic University

Choi Wah Kong

United Christian Hospital

Daniel YT Fong

University of Hong Kong

Research

Keywords: NPV, PPV, Diagnostic

Posted Date: October 23rd, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-94450/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Introduction: Rising prevalence and undetected perinatal depression has been described in many countries and report that treating those who are already symptomatic, more effort should be targeted towards screening strategies to identify perinatal depression at the early stage. The Whooley questions is the recommended case finding strategy to aid the identification of perinatal depression. An official Chinese version has not been validated. The aim of this study was to evaluate the diagnostic accuracy and stability of the translated Whooley questionnaire against the gold standard measurement during pregnancy (antenatal) and early after pregnancy (postnatal).

Materials and method: This observational study recruited 131 pregnant women from antenatal clinics in a hospital setting from September 2019 till May 2020 in Hong Kong. We translated the Whooley questionnaire in Chinese and evaluated self-reported responses against an interviewer assessed diagnostic standard (DSM-IV criteria) among 107 women receiving antenatal care at 26-28 weeks gestation. We calculated sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), positive likelihood ratio, negative likelihood ratio and diagnostic odds ratio, with DSM-IV diagnosis taken as the gold standard.

Results: Antenatally, the Chinese Whooley questions had a sensitivity of 79% (95% CI 54.4-93.9), a specificity of 97% (95% CI 90.4-99.3), a positive likelihood ratio of 23.2 (95% CI 7.4-72.1) and a negative likelihood ratio of 0.2 (95% CI 0.1-0.5) in identifying perinatal depression.

Conclusion: Prevalence of depression in pregnancy have increased and screening early remains a significant tool in Western countries. The translated Chinese Whooley questionnaire appears to have acceptable diagnostic accuracy and can be implemented into health services among Chinese population as only requiring two yes/no questions.

Introduction

Depression is the commonest mental health problem and is the major contributor to worldwide disability and global burden of disease. Perinatal depression is defined as depression during pregnancy, around childbirth or within the first year postpartum. It affects up to 20% of pregnant women and is a major public health burden across the globe. When pregnant, women are at a greater vulnerability to having depression. Antenatal depression was found to affect 10.7% of women in a meta-analysis of 21 studies across diverse countries and cultures.¹ In an epidemiological study of 357 Hong Kong Chinese women, 18.9% reported suffering from antenatal depression in the second trimester to 22.1% in the third trimester.² Given that even small reductions in population prevalence have a greater public health benefit than treating those who are already symptomatic, more effort should be targeted towards screening strategies to identify perinatal depression at the early stage.

Perinatal depression is associated with a range of adverse outcomes.³ Evidence suggests an association between depressions experienced during pregnancy (prenatal depression) and adverse neonatal outcomes, poor self-reported health, substance abuse and alcohol abuse, and poor usage of antenatal care services.⁴ Postnatal depression has been shown to have a substantial impact on the mother and her partner, mother-baby interactions⁵, the family⁶ and on the longer term emotional and cognitive development of the baby.⁷ Despite the significant disease burden of perinatal depression, it often goes undetected during the antenatal stage in Hong Kong. Research suggests that women who experience perinatal depressive symptoms use more healthcare resources during pregnancy than do women who are not depressed, which potentially affords providers greater opportunity to screen for and address depressive symptoms during the antenatal period.⁸ Research have found that among women with postpartum depression, over 50% had depression identified either before or during the pregnancy.⁹

Despite the significance of the problem, no screening tool for perinatal depression are used in Hong Kong and guidance are taken predominantly from Western countries. Screening women at the perinatal stage for possible depression has been identified as important by both healthcare professionals and pregnant women to implement early interventions that would prevent further adverse outcomes occurring at later postnatal stages.^{10,11} In UK, the National Institute for Health and Care Excellence (NICE) produced guidelines on antenatal and postnatal health.¹² These set out recommendations for the detection and treatment of mental health problems during pregnancy and the postnatal period. As part of these guidelines, NICE endorsed a case-finding strategy by recommending the use of two "ultra-brief" questions to aid the identification of perinatal depression (see Appendix 1);

these questions are often referred to as the Whooley questions. NICE have updated their guidelines in which they continue to recommend the use of the Whooley questions during pregnancy and the postnatal period (NICE, 2014) and validation study is currently being conducted³. If someone respond positively to either question a more comprehensive assessment is carried out, to determine whether or not an individual is depressed. The Whooley questions have been validated previously in a small sample of 152 women during pregnancy and the early postnatal period in the UK, with a sensitivity of 100% (95% CI: 77–100%) and a specificity of 68% (95% CI: 58–76%) during pregnancy, with similar estimates during early postnatal period.¹³

Locally, a population-based study of women in Hong Kong completed the Edinburgh Postnatal depression scale (EPDS) in the second and third trimesters and at 6 weeks postpartum and found that EPDS can identify high risk women for postpartum depression and that screening all pregnant women in the second trimester can be a secondary preventive measure.¹⁴ A Chinese version of the EPDS has been tested in Hong Kong and demonstrated good reliability and validity¹⁵, however this is routinely used during postnatal periods and not routinely used during the antenatal stage. A reason for this, could be that EPDS have a larger number of items, relatively lower sensitivity (correctly identifying true cases) and varying cut-off points used across different populations compared to Whooley questions. Diagnostic tests are regarded as providing definitive information about the presence or absence of a disease or condition. By contrast, screening tests such as the Whooley questions places fewer demands on the health care system and more accessible, less invasive, less expensive and less time consuming. Given the known effects of screening during pregnancy in preventing depression, the obvious next step is to determine to which these tests are able to identify the likely presence or absence of depression so appropriate decision making can be encouraged in the health care system in Hong Kong.

There are limited studies examining depression case-finding questions, such as the Whooley questions and the EPDS^{16,17} in Hong Kong. The primary objective of this study to evaluate the diagnostic accuracy of the Whooley questionnaire against the gold standard measurement during pregnancy (antenatal) and early after pregnancy (postnatal). In addition, to identify the stability of positive or negative screening of depression between antenatal stage of pregnancy and postnatal, to estimate if earlier testing optimizes depression screening.

Methods

Participants were recruited from antenatal outpatient clinic from one Hong Kong public hospital. Criteria for selecting the subjects were (1) 18 years of age or older; (2) Cantonese speaking; (3) Hong Kong resident for more than 1 year; (4) singleton pregnancy; (5) no serious medical or obstetrical complications. Women were approached sequentially while attending the antenatal clinic at 26–28 weeks gestation for routine antenatal talk. Convenience sampling was used to recruit all eligible subjects and hence all eligible ones were approached and invited to participate in this study. Women were excluded if they were less than 37 weeks gestation or did not speak Chinese, were not literate, were planning to move or be away from Hong Kong after birth.

Ethical approval was obtained from the Institutional Review Board of the University of Hong Kong-Hospital Authority Hong Kong West Cluster (UW-18-649) and the Kowloon Central/Kowloon East Research Ethics Committee (KC/KE-18-0278).

Screening measures

Index test

Whooley questionnaire comprised two brief case-finding questions recommended in NICE guidelines in United Kingdom. Each question was responded as “yes” or “no”. A “yes” response to either question was considered a positive screen. It was translated in Chinese version by two health care professionals and consensus reached by a consensus meeting. The consensus version was then back translated in English and compared with the original version by KL.

Diagnostic gold standard

To confirm the presence or absence of a current depressive episode, the DSM-IV (Diagnostic and Statistical Manual of Mental Disorders, fourth edition) diagnostic criteria for major depressive disorder were administered in an interview by telephone.¹⁸ Guidance for the administration and interpretation of the criteria was taken from the Structured Clinical Interview for DSM-IV-

Clinical version.¹⁹ Semi-structured questions to identify depressive symptoms were asked in a format suitable for verbal interview²⁰ and have previously been shown to be valid when used over the telephone or face to face.^{21,22} The interview questions and DSM-IV criteria used in our study are available in Appendix 1.

Procedure

The study was conducted in two phases in order to validate the utility of the questions in the antenatal and postnatal periods, as recommended by National Institute for Health and Clinical Excellence.²³ The study is reported according to the STARD (Standards for the Reporting of Diagnostic accuracy studies) statement.²⁴

During the antenatal phase, participants self-completed the questionnaires which consisted of demographic data and the Whooley questionnaire while attending antenatal clinic. A researcher (C.C) was available to answer questions if necessary.

During the postnatal phase, at about five to six weeks postnatally, a copy of the questionnaire with a self-addressed envelope was mailed to the participant and returned within seven days. Non-respondents by seven days were contacted by telephone to remind them completing the questionnaire or to complete the questionnaire by interview over the telephone.

Within 14 days after completion of the Whooley questionnaires during the antenatal phase, a diagnostic interview was conducted by an experience clinician (JL) and researcher (CC) who was blinded to the participants' responses to the Whooley questionnaire before the interview. The researcher (JL and CC) had previous clinical and research experience with the administration of diagnostic interviews. A diagnostic interview was conducted during the postnatal phase based on their preference of time and day on their returned questionnaire. To prevent review bias, the researcher (JL and CC) was blinded to the participants' responses to the Whooley questionnaire before the interview.

Statistical analysis

Baseline characteristics were reported for age, level of education, employment status, marital status, having existing children, history of mental health problems and responses to the individual Whooley questionnaires. Data were reported using proportions, means and standard deviations. We assessed the diagnostic accuracy and utility of the Whooley questionnaire by examining its sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), positive likelihood ratio, negative likelihood ratio and diagnostic odds ratio, with DSM-IV diagnosis taken as the gold standard. Each such measure was accompanied by a 95% confidence interval (CI). Values of these statistics were based on a positive case identified with a positive response to at least one of the two Whooley items. The diagnostic accuracy and utility measures were obtained separately at antenatal and postnatal timepoints.

Stability of positive and negative screening between the two timepoints was assessed according the number of cases that remained stable or changed diagnosis. We anticipate that the stability of the Whooley questionnaire shall be the same to the DSM-IV in changes of cases between the two timepoints. Stability or change in depression status between timepoints was tested by McNemar test.

All analysis was conducted using STATA version 16 (StataCorp. 2019. *Stata Statistical Software: Release 16*. College Station, TX: StataCorp LLC.).²⁵

Results

From September 2019 till May 2020, a total of 131 participants were recruited. One hundred and seven participants completed the Whooley questionnaire and DSM-IV interviews at the antenatal stage. Of those, seventy-eight participants (73%) also completed the Whooley questionnaire and DSM-IV interviews at the postnatal stage (Fig. 1). Baseline characteristics of both groups are shown in Table 1, overall participants that continued in the study at the postnatal follow-up were similar in characteristics to the whole sample at the antenatal baseline stage.

Table 1
Participant characteristics of those who participated at study antenatal stage and postnatal stage.

	Antenatal stage (n = 107)	Postnatal stage (n = 78)
Age in years, mean (SD)	32.4 (4.3)	32.3 (4.2)
Mother is married, n (%)	101 (94.39)	73 (93.59)
Yes	6 (5.61)	5 (6.41)
No		
Location of mother birth, n (%)	62 (57.94)	47 (60.26)
Hong Kong	41 (38.32)	28 (35.90)
Mainland China	4 (3.74)	3 (3.85)
Other		
Number of years living in Hong Kong, n (%)	5 (4.67)	3 (3.85)
< 5 years	41 (38.32)	29 (37.18)
5 to = > 15 years	61 (57.01)	46 (58.97)
since birth		
Level of education, n (%)	6 (5.61)	5 (6.41)
Primary	31 (28.97)	18 (23.08)
Secondary	70 (65.42)	55 (70.51)
University Degree or Above		
Family income in HKD per month, n (%)	12 (11.21)	8 (10.26)
< \$15,000	23 (21.50)	15 (19.23)
\$15,000 to \$29,999	72 (67.29)	55 (70.51)
= > \$30,000		
Anyone else live in your home apart from partner? n (%)	47 (43.93)	35 (44.87)
No	60 (56.07)	43 (55.13)
Yes		
Smoke tobacco during pregnancy, n (%)	106 (99.07)	77 (98.72)
No	1 (0.93)	1 (1.28)
Yes		
Partner currently smoke tobacco, n (%)	89 (83.18)	65 (83.33)
No	18 (16.82)	13 (16.67)
Yes		
Other at home currently smoke tobacco, n (%)	96 (90.57)	71 (91)
No	10 (9.43)	7 (9)
Yes		

	Antenatal stage (n = 107)	Postnatal stage (n = 78)
Employment status during pregnancy, n (%)	76 (71.03)	58 (74.36)
Full-time	9 (8.41)	6 (7.69)
Part-time	1 (0.93)	0 (0)
Casually	21 (19.63)	14 (17.95)
Did not work		
Weeks of paid maternity leave, n (%)	9 (10.71)	6 (9.38)
0	44 (52.38)	29 (45.31)
10	2 (2.38)	2 (3.13)
12	25 (29.76)	23 (35.94)
14	4 (4.76)	4 (6.25)
16		

At the antenatal stage, the DSM-IV identified 19 (18%) with depression. The Whooley questionnaire identified 15 true positive, 85 true negative, 3 false positive and 4 false negative compared to the DSM-IV. Sensitivity for the Whooley questionnaire for at least one positive response to one of the first two items at antenatal stage was 78.9% and a specificity of 96.6%, PPV 83% and NPV 96%, likelihood ratio (positive) 23.2 and likelihood ratio (negative) 0.2 (Table 2). At the postnatal stage, the DSM-IV identified two with depression (7%), The Whooley questionnaire identified 1 true positive, 24 true negative, 4 false positive and 1 false negative compared to the DSM-IV. Sensitivity for the Whooley questionnaire for at least one positive response to one of the first two items at postnatal stage was 76.9% and a specificity of 86.2%, PPV 53% and NPV 95%, likelihood ratio (positive) 5.6 and likelihood ratio (negative) 0.3 (Table 2). The McNemar test suggested that the DSM-IV and Whooley questionnaire remained stable between the antenatal and postnatal stage with no significant differences found ($\chi^2 = 0.05$, $p = 1.00$, $\chi^2 = 1.8$, $p = 0.26$ respectively) (Table 3).

Table 2
2 × 2 Contingency table of diagnostic properties for the Whooley questionnaire.

	DSM-IV positive	DSM-IV negative	Sensitivity (95% CI)	Specificity (95% CI)	Positive predictive value (95% CI)	Negative predictive value (95% CI)	Positive likelihood value (95% CI)	Negative likelihood value (95% CI)	Diagnostic Odds Ratio (95% CI)
Antenatal (n = 107)									
Whooley positive	15	3	78.9 (54.4–93.9)	96.6 (90.4–99.3)	83.3 (58.6–96.4)	95.5 (88.9–98.8)	23.2 (7.4–72.1)	0.2 (0.1–0.5)	106.3(22.6–498.7)
Whooley negative	4	85							
Postnatal (n = 78)									
Whooley positive	10	3	76.9 (46.2–95.0)	86.2 (75.3–93.5)	52.6 (28.9–75.6)	94.9 (85.9–98.9)	5.6 (2.8–10.9)	0.3 (0.1–0.7)	20.74 (5.03–83.94)
Whooley negative	9	56							

Table 3
Contingency table for stability of the DSM-IV and Whooley questionnaire.

Tool	Antenatal		Postnatal (% of those at antenatal)		McNemar Test (χ^2, p)
DSM-IV	Cases	14	Cases	4 (29)	$\chi^2 = 0.05, p = 1.00$
			Non-Cases	10 (71)	
	Non-Cases	64	Cases	9 (14)	
			Non-Cases	55 (86)	
Whooley	Cases	13	Cases	6 (46)	$\chi^2 = 1.8, p = 0.26$
			Non-Cases	7 (54)	
	Non-Cases	65	Cases	13 (20)	
			Non-Cases	52 (80)	

Discussion

From our study we found that the two translated case finding questions endorsed by the National Institute for Health and Clinical Excellence showed promising diagnostic characteristic among Chinese antenatal population with a strong specificity and moderate sensitivity, suggesting a follow-up questionnaire may not be needed. The Whooley questionnaire among the Chinese population appeared to be stable and so acceptable for use early and later pregnancy and birth.

A recent meta-analysis of 21 studies on depression during pregnancy indicated a prevalence of antenatal depression approximately 10.7%, ranging from 7.4% in the first trimester to 12.8% in the second trimester²⁶. The rate of antenatal depression in individual studies were found to be as high as 24%²⁶. Chinese researchers have published numerous studies examining the rates of depression during pregnancy, depending on the study, these rates varied from 5.5–23.1%.^{27,28} Lee et al, found that the rates of antenatal depression ranged from 18.9% in the second trimester to 22.1% in the first trimester in a cohort of 357 women in Hong Kong.² Our finding was similar to that in the general population in Hong Kong. The study's overall positive predictive value is reported to be 83.3% of screening for depression during antenatal stage and negative predictive value is reported to be 95.5% that truly do not have the condition can therefore be generalize to the population.

Evidence from a Cochrane review suggests that screening or case-finding instruments are a simple, quick and inexpensive method to improve detection and management of depression in non-specialist settings, such as primary care and the general hospital.¹⁶ This approach has important implications for clinical practice. The value of screening or case-finding questionnaires should be considered as a triage tests approach rather than a replacement test to existing methods of assessment²⁴. Triage tests are simple and noninvasive, have no waiting time and do not aim to improve the diagnostic pathway, but instead they reduce the number of patients who need further assessment. The benefit of using the two case finding questions in clinical settings is not necessarily to diagnose perinatal depression, but to reduce the number of women who need extensive clinical assessment or evaluation with much longer questionnaires, such as the Edinburgh Postnatal Depression Scale or the Patient Health Questionnaire-9 by more than 50%.¹³ Overall, the translated Chinese Whooley questions showed high specificity and moderate sensitivity compared to other studies.^{13,29} This is consistent with other studies due to the homogenous sample, which may lead to lowering sensitivity and higher specificity. It has been suggested that using the Whooley questions in primary care or community settings may lead to lowering sensitivity and higher specificity compared to specialist hospital setting due to primary care needing to deal with a more diverse population on a daily basis.³⁰ The specificity of the two case finding questions in our antenatal (96.6%) and postnatal validation study (86.2%) provide further evidence of a simple approach to identify perinatal depression.

The strength of the study was the use of the gold standard comparison so this validation study is confident in the integrity of diagnosis and the follow-up enabled us to examine the stability of the Whooley questions. Limitations of the study was the small

sample size as unanticipated due to the social unrest and the pandemic COVID-19 and we had to stop recruitment which was unavoidable. Other limitations include the homogenous sample as we recruited from one study site and limited to the third trimester and first three postnatal months. Therefore, further research is warranted with a larger sample size involving diverse perinatal populations and longer postnatal follow up. Lastly, the effect of the questions on outcomes of perinatal care warrants evaluation.

Conclusion

Overall, the translated Chinese Whooley questionnaire appears to have acceptable diagnostic accuracy and can be implemented into health services among Chinese population as only requiring two yes/no questions. The next stage is to research approaches using the Whooley questions to identify, treat and/or prevent depression in Chinese pregnant women and new mothers.

Abbreviations

DSM-IV: Diagnostic Statistical Manual

EPDS: Edinburgh Postnatal depression scale

NICE: National Institute of Clinical Excellence

NPV: Negative predictive value

PPV: Positive predictive value

STARD: Standards for the Reporting of Diagnostic accuracy studies

Declarations

Funding

The project was supported by Seed Fund for Basic Research, the University of Hong Kong (Reference: 201711159208).

Author information

Affiliations

School of Nursing, The University of Hong Kong, Hong Kong, 4/F, William M.W. Mong Block, 21 Sassoon Road, Pok Fu Lam, Hong Kong

Kris Yuet Wan Lok, Robert Smith & Daniel Yee Tak Fong

Contributions

KYWL: design and conceptualized the study, interpretation of data and writing the manuscript; **CC, ST, JL** participated in acquisition of the data; **RS** contributed to analysis of data, drafting and participated in interpretation of data; **CWK** and **DYT** supervised the conception and protocol of study and participated in interpretation of data. All authors read and approved the final manuscript.

Ethics declarations

Ethics approval and consent to participate

Ethical approval was obtained from Institutional Review Board of the Hong Kong/University of Hong Kong-Hospital Authority Hong Kong West Cluster (UW-18-649) and the Kowloon Central/Kowloon East Research Ethics Committee (KC/KE-18-0278). Written informed consent was obtained for each participant.

Consent for publication

Not applicable

Competing interests

The authors declare that they have no competing interests.

Availability of supporting data

The datasets generated and/or analysed during the current study are not publicly available because it contain personal data but are available from the corresponding author on reasonable request.

Acknowledgements

We are grateful for the time given by our participants to contribute to this research. Funding for this study was provided by University of Hong Kong Seed Fund for Basic Research.

References

1. Schatz DB, Hsiao MC, Liu CY. Antenatal Depression in East Asia: A Review of the Literature. *Psychiatry Investig.* 2012;9(2):111-118.
2. Lee DT, Chan SS, Sahota DS, Yip AS, Tsui M, Chung TK. A prevalence study of antenatal depression among Chinese women. *J Affect Disord.* 2004;82(1):93-99.
3. Littlewood E, Ali S, Ansell P, et al. Identification of depression in women during pregnancy and the early postnatal period using the Whooley questions and the Edinburgh Postnatal Depression Scale: protocol for the Born and Bred in Yorkshire: PeriNatal Depression Diagnostic Accuracy (BaBY PaNDA) study. *BMJ Open.* 2016;6(6):e011223.
4. Marcus SM. Depression during pregnancy: rates, risks and consequences—Motherisk Update 2008. *Can J Clin Pharmacol.* 2009;16(1):e15-22.
5. Murray L, FioriCowley A, Hooper R, Cooper P. The impact of postnatal depression and associated adversity on early mother-infant interactions and later infant outcome. *Child Dev.* 1996;67(5):2512-2526.
6. Lovestone S, Kumar R. Postnatal Psychiatric-Illness - the Impact on Partners. *Br J Psychiatry.* 1993;163:210-216.
7. Murray L, Sinclair D, Cooper P, Ducournau P, Turner P. The socioemotional development of 5-year-old children of postnatally depressed mothers. *J Child Psychol Psyc.* 1999;40(8):1259-1271.
8. Andersson L, Sundstrom-Poromaa I, Wulff M, Astrom M, Bixo M. Implications of antenatal depression and anxiety for obstetric outcome. *Obstet Gynecol.* 2004;104(3):467-476.
9. Dietz PM, Williams SB, Callaghan WM, Bachman DJ, Whitlock EP, Hornbrook MC. Clinically identified maternal depression before, during, and after pregnancies ending in live births. *Am J Psychiatry.* 2007;164(10):1515-1520.
10. Eastwood J, Ogbo FA, Hendry A, Noble J, Page A, Early Years Research G. The Impact of Antenatal Depression on Perinatal Outcomes in Australian Women. *PLoS One.* 2017;12(1):e0169907.
11. Littlewood E, Ali S, Dyson L, et al. In: *Identifying perinatal depression with case-finding instruments: a mixed-methods study (BaBY PaNDA - Born and Bred in Yorkshire PeriNatal Depression Diagnostic Accuracy)*. Southampton (UK)2018.
12. Excellence NloHaC. *Antenatal and postnatal mental health: clinical management and service guidance (Clinical guidelines 192)*. 2020.
13. Mann R, Adamson J, Gilbody SM. Diagnostic accuracy of case-finding questions to identify perinatal depression. *CMAJ.* 2012;184(8):E424-430.
14. Lau Y, Wong DF, Chan KS. The utility of screening for perinatal depression in the second trimester among Chinese: a three-wave prospective longitudinal study. *Archives of women's mental health.* 2010;13(2):153-164.
15. Lee DT, Yip SK, Chiu HF, et al. Detecting postnatal depression in Chinese women. Validation of the Chinese version of the Edinburgh Postnatal Depression Scale. *Br J Psychiatry.* 1998;172:433-437.

16. Hewitt C, Gilbody S, Brealey S, et al. Methods to identify postnatal depression in primary care: an integrated evidence synthesis and value of information analysis. *Health Technol Assess.* 2009;13(36):1-145, 147-230.
17. Brealey SD, Hewitt C, Green JM, Morrell J, Gilbody S. Screening for postnatal depression - is it acceptable to women and healthcare professionals? A systematic review and meta-synthesis. *J Reprod Infant Psychol.* 2010;28(4):328-344.
18. American Psychological Association. *Diagnostic and statistical manual of mental disorders.* 4th ed 1994.
19. First M, Spitzer R, Gibbons M. *Structured clinical interview for DSM-IV - clinical version (SCID-CV)(user's guide and interview).* 1997.
20. Williams JW, Noel PH, Cordes JA, Ramirez G, Pignone M. Is this patient clinically depressed? *Jama-Journal of the American Medical Association.* 2002;287(9):1160-1170.
21. Simon GE, Revicki D, Vonkorff M. Telephone Assessment of Depression Severity. *J Psychiatr Res.* 1993;27(3):247-252.
22. Rohde P, Lewinsohn PM, Seeley JR. Comparability of telephone and face-to-face interviews in assessing axis I and II disorders. *Am J Psychiatry.* 1997;154(11):1593-1598.
23. National Institute for Health and Clinical Excellence. Antenatal and postnatal mental health: clinical management and service guidance. 2020; www.nice.org.uk/guidance/cg192.
24. Bossuyt PM, Reitsma JB, Bruns DE, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *Croat Med J.* 2003;44(5):635-638.
25. *Stata statistical software: release 16.0* [computer program]. College Station, Tx: StataCorp LP; 2020.
26. Bennett HA, Einarson A, Taddio A, Koren G, Einarson TR. Prevalence of depression during pregnancy: systematic review. *Obstet Gynecol.* 2004;103(4):698-709.
27. Qiu JY, Xiao ZP. Reproductive-related depression in women: A review of two decades of research in China. *Asia-Pac Psychiat.* 2010;2(1):19-25.
28. Klainin P, Arthur DG. Postpartum depression in Asian cultures: a literature review. *Int J Nurs Stud.* 2009;46(10):1355-1373.
29. Bosanquet K, Bailey D, Gilbody S, et al. Diagnostic accuracy of the Whooley questions for the identification of depression: a diagnostic meta-analysis. *BMJ Open.* 2015;5(12):e008913.
30. Gjerdingen D, Crow S, McGovern P, Miner M, Center B. Postpartum Depression Screening at Well-Child Visits: Validity of a 2-Question Screen and the PHQ-9. *Ann Fam Med.* 2009;7(1):63-70.

Figures

Figure 1

Participant flow chart of the study