Archives of Gynecology and Obstetrics (2023) 307:1331–1345
https://doi.org/10.1007/s00404-022-06525-0

REVIEW

Predictive validity of the Edinburgh postnatal depression scale and other tools for screening depression in pregnant and postpartum women: a systematic review and meta-analysis

Seong-Hi Park[1] · **Jeung-Im Kim**[2]

Received: 28 December 2021 / Accepted: 10 March 2022 / Published online: 13 April 2022

Abstract

Purpose To compare the predictive validity of the Edinburgh Postnatal Depression Scale (EPDS) and other tools for screening depression in pregnant and postpartum women through a systematic review and meta-analysis.

Methods An electronic search of MEDLINE, EMBASE, CINAHL, and PsycArticles databases was conducted using the following keywords: depression, perinatal-related terms, and EPDS. Quality Assessment of Diagnostic Accuracy Studies-2 was used to assess the risk of bias in diagnostic studies.

Results The search identified 823 articles, of which 17 studies met the inclusion criteria. In 1831 pregnant women from nine studies, pooled sensitivity and specificity of the EPDS were 0.81 and 0.87, respectively, with summary receiver operating characteristic (sROC) curve of 0.90. In 515 postpartum women from six studies, pooled sensitivity, specificity, and sROC were 0.79, 0.92, and 0.90, respectively. We then compared the EPDS with other tools using three or more studies. The sROC curve of the Patient Health Questionnaire-9 was 0.74, which was lower than that (0.86) of the EPDS. The sROC curve of the Beck Depression Inventory and the ten-item Kessler Psychological Distress Scale was 0.91, similar to that of the EPDS (0.90 and 0.87). However, in comparison with the Postpartum Depression Screening Scale (0.98), the sROC curve of the EPDS was 0.54.

Conclusion As a tool specialized for screening depression in pregnant and postpartum women, the EPDS showed excellent performance. Thus, the EPDS can be used in preference to other tools to screen for depression in perinatal women at a primary care setting or a midwifery center.

Keywords Depression · Postpartum depression · Pregnant women · Sensitivity and specificity · Systematic review

Introduction

Depression can occur at any age. Pregnancy and childbirth make some women vulnerable to developing major depressive disorder (MDD) [1]. Although postpartum depression (PPD) is widely known, depression is also common during pregnancy. We call this antepartum depression (APD). Public-health experts are becoming increasingly focused on APD [2]. It is known that 7–20% of pregnant women suffer from a depressive disorder [3]. APD significantly increases the risk for postpartum depression (PDD) [4, 5]. Untreated APD and PPD have multiple potential negative effects on maternal-infant attachment, child development, and can also cause big problems such as infanticide [6–8]. Thus, early and accurate detection of MDD is imperative [8] and it should be required in peri-partum. Fortunately, depression during and after pregnancy is treatable [9]. Its eventual remission rate with early improvement reaches over 80% [10].

There are many screening tools for MDD available to us. The representative tools are the Beck Depression Inventory (BDI), the Center for Epidemiologic Studies for Depression (CES-D), the ten-item Kessler Psychological Distress Scale (K-10), and the Patient Health Questionnaire (PHQ-9), all of which are self-rating tools used for adults [11]. Strong negative emotions such as sadness, loss of interests, and

✉ Jeung-Im Kim
jeungim@sch.ac.kr

[1] School of Nursing, Soonchunhyang University, Cheonan, Korea

[2] School of Nursing, Soonchunhyang University, 31 Soonchunhyang 6-gil, Dongnam-gu, Cheonan 31151, Chungcheongnam-do, Korea

hopelessness are prevalent in all types of depression. But since pregnancy and childbirth are often celebrated events, pregnant women may not be always aware of their depression symptoms [12]. Even if they are aware, they often hide their symptoms because of their compulsion to be a good mother [13]. In addition, symptoms such as fatigue, as well as changes in appetite, and sleep are typical signs of pregnancy and postpartum women, and may be misinterpreted as false negative [14, 15].

Considering this, Cox et al. [15] developed the Edinburgh Postnatal Depression Scale (EPDS) in 1987 to better detect PPD. While other tools such as the BDI and the PHQ-9 asked simple questions about sleep disturbances, the EPDS questions difficulty sleeping in relation to unhappiness, and excluded questions about appetite and fatigue. The EPDS expands its target not only for postpartum women, but also pregnant women [16]. And it is widely used in epidemiologic surveys to measure the prevalence of major depression in perinatal women [17]. Therefore, the EPDS, a perinatal-specific depression screening tool, should be more usefully spent to evaluate depression in perinatal women.

The screening accuracy of the EPDS has been verified through several systematic reviews (SRs) [18–20]. These were SRs to test the performance of the EPDS as a depression screening tool or to find the optimal cutoff scores for the EPDS in pregnant and postpartum women. However, none of the SRs compared the screening accuracy of the EPDS with other depression screening tools or contrasted the screening performance of the EPDS according to APD or PPD.

Thus, the purpose of this study was to compare the screening performance of the EPDS and other depression screening tools, specifically in two ways. First, we compared the predictive validity of the EPDS and other screening tools by dividing the subjects into pregnant, postpartum, and perinatal women (mixed of pregnancy and postpartum). Second, we compared the screening performance of the EPDS with each depression screening tool.

Methods

This study was performed according to the guidelines of the Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy [21] and the Preferred Reporting Items for Systematic Reviews and Meta-Analyses 2020 Statement [22].

Search strategy and literature sources

We systematically searched eligible articles in four electronic databases, MEDLINE, Embase, CINAHL, and PsycArticles, on July 17, 2021. Key search terms were depression, perinatal-related terms, and EPDS. Depression and postpartum were searched based on MeSH terms (both free text and MeSH, exploded). EPDS was searched using its full name and abbreviation. We expanded the search scope using free text searching to search full texts in addition to titles and abstracts. We present the search strategies of all databases in Supplementary Table 1.

Eligibility criteria

Inclusion criteria were: (i) Types of studies, studies that reported diagnostic accuracy such as sensitivity and specificity (e.g., observational studies such as cohort or cross-sectional studies); (ii) types of participants, studies on pregnant and postpartum women over 18 years of age; (iii) indexed tests, *s*tudies that used EPDS-10 items; (iv) comparators, studies on depression screening tools of all types compared to the EPDS (for the meta-analysis, the depression screening tool reported in more than three studies was selected); (iv) gold standard, studies that conducted direct (e.g., the Diagnostic and Statistical Manual of Mental Disorders [DSM], the International Classification of Diseases [ICD]) or structured interviews (e.g., Structured Clinical Interview for DSM [SCID], Mini-International Neuropsychiatric Interview [MINI] etc.) by trained psychiatric professionals using diagnostic criteria for MDD as a gold standard; and (v) types of outcomes, studies with data of true positives (TP), false positives (FP), false negatives (FN), and true negatives (TN). We derived sensitivity, specificity, positive and negative likelihood ratio, diagnostic odds ratio, and sROC curve from these data as the outcome measures.

Exclusion criteria were (i) retrospective studies such as case–control studies; (ii) non-original articles such as reviews, letters, or editorials; (iii) studies using the EPDS to assess risk such as anxiety or suicide; (iv) studies that only presented sensitivity or specificity and did not provide sufficient data to create a two-by-two contingency table; and (v) studies that included subjects with other mental disorders or diseases. However, the language was not limited.

Full text screening and data extraction

After removing duplicate articles, two authors (S-H and J-I) independently selected titles and abstracts for study screening and data extraction to confirm their potential eligibility. If there was a difference in opinion among authors, it was resolved through a consensus-based discussion. We extracted the following information from full texts of selected studies: the year of publication, authors, location, subjects, age, gestational age, gestation or postpartum period in weeks, sample size, gold standard, blinding, cutoff scores of the EPDS and other tools, and outcomes such as TP, FP, FN, and TN.

Quality and risk of bias

The quality of selected studies was assessed using QUADAS-2 (Quality Assessment of Diagnostic Accuracy Studies-2 [23]. QUADAS-2 assesses the risk of bias and applicability with four domains of criteria which include patient selection, index test, reference standard, and flow and timing. Applicability is assessed only in the first three domains. Two independent authors (S-H and J-I) completed this assessment. When there was a disagreement between both authors, through discussion, reached consensus.

Statistical analysis

The meta-analysis was conducted using MetaDiSc 1.4 [25] and the Meta DTA program [26, 27]. Using a bivariate random effect model, the MetaDTA program allowed the evaluation of the accuracy of screening and the heterogeneity across studies [24]. Screening accuracy was evaluated by yielding pooled sensitivity, specificity, positive and negative likelihood ratios (LRs), diagnostic odds ratios (ORs) with 95% of confidence intervals (CIs), and the area under the curve (AUC) of summary receiver operating characteristics (sROC) curve. The area under the curve (AUC) and the index Q* value were analyzed to describe test accuracy. The AUC values were interpreted as follows: AUC of 0.5 = non-informative test; AUC of 0.5–0.7 = low accurate; AUC of 0.7–0.9 = moderate accurate; AUC of 0.9–1 = highly accurate; and AUC of 1 = perfect test [28]. The index Q* value represents the point at which the sensitivity and specificity are equal in the ROC curve, with a value of 1 indicating the accuracy of 100% [29]. The heterogeneity among studies was judged using random effect (RE) correlation.

We presented the sensitivity and specificity of the EPDS and other tools as forest plots and sROC curves. And we categorized subjects into three groups (pregnant, postpartum, and perinatal women) and performed a subgroup analysis. In addition, we analyzed the screening performance for the EPDS and each of the other tools.

Results

Selection process

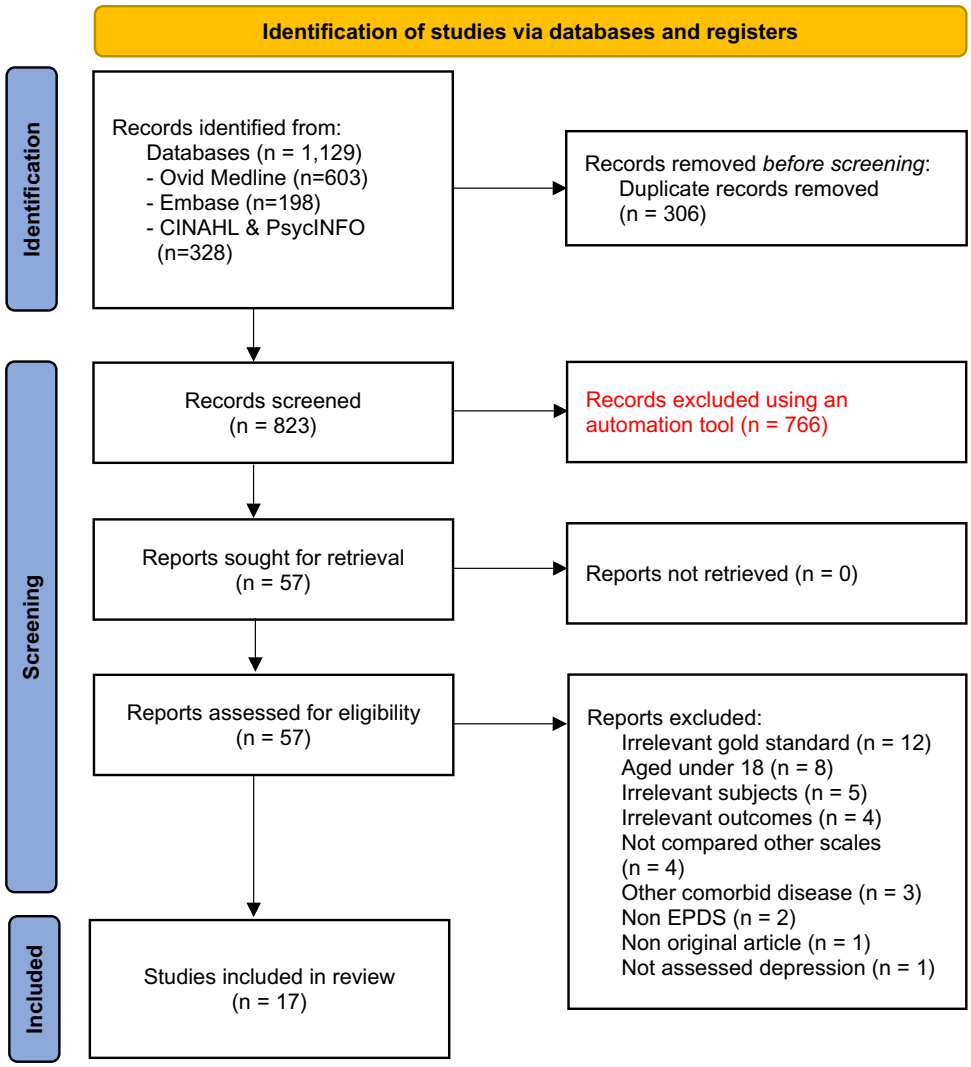
We searched 1,129 articles from the electronic database. Duplicate articles ($k = 306$) were excluded. Inclusion and exclusion criteria were applied to titles and abstracts of 823 articles. When it was difficult to accurately determine the title and abstract, we searched for full texts and checked them. Seventeen studies [30–46] were retained for quantitative synthesis while 806 (97.9%) articles were excluded. The study selection process is detailed in a PRISMA 2020 flow diagram as shown in Fig. 1.

Risk of bias assessment

Among the 17 selected studies, 6 (35.3%) [31, 34, 36, 40–42] were assessed to have a low risk in all domains. In patient selection, eight studies were assessed to have a low risk either randomly [31, 34, 35, 41] or as consecutive samples [30, 36, 40, 42], while one study [44] had a high risk of bias as a convenience sample. Because the EPDS and other tools were all self-reported questionnaires, the risk of bias in the domain of index test was assessed to be low for all studies. Ten studies [31, 33, 34, 36, 40–42, 44–46] were assessed to have a low risk of bias in the reference standard. Of them, nine studies were blinded during the test process. In one study [44], the index test and the reference standard test were performed randomly. In addition, it provided interpretation without information on results of each test. The flow and timing, and applicability concerns of each domain were all assessed to have a low risk (Fig. 2).

Summary included studies

A total of 17 studies analyzed the predictive validity of the EPDS and a total of 2902 women were included. There were nine studies [30, 31, 34–37, 39, 40, 45] on pregnant women. Six studies [32, 40, 42–44, 46] included postpartum women (one [40] of these studies analyzed pregnant and postpartum women separately). Three studies [34, 38, 41] included both pregnant and postpartum women as perinatal women. The average age of women was in their 20 s in 9 studies [30, 31, 33–35, 38, 39, 41, 42] and in their 30 s in 8 studies [32, 36, 37, 40, 43–46]. Selected studies were published in a total of 12 countries. Four studies [38, 40, 43, 46] were from the United States. The United Kingdom [39, 42] and South Africa [33, 35] each had two studies. There were 6 studies [32, 33, 38, 40, 43, 44] with less than 100 women, 6 studies [31, 34, 39, 42, 45, 46] with 100–200 women, and five studies [30, 35–41] with 200 or more women. A total of 11 types of depression screening tools were compared to the EPDS, of which the PHQ-9 [34, 35, 37, 40, 43] and the BDI [31, 36, 38, 45, 46] were reported in 5 studies, and the K10 [35, 39, 42] and the PDSS [43, 44, 46] in 3 studies, respectively. The optimal cut-off score of the EPDS presented in each study ranged from 3/4 to 16. The most common cut-off score was 13, reported in six studies [31, 32, 38–40, 45] (Table 1).



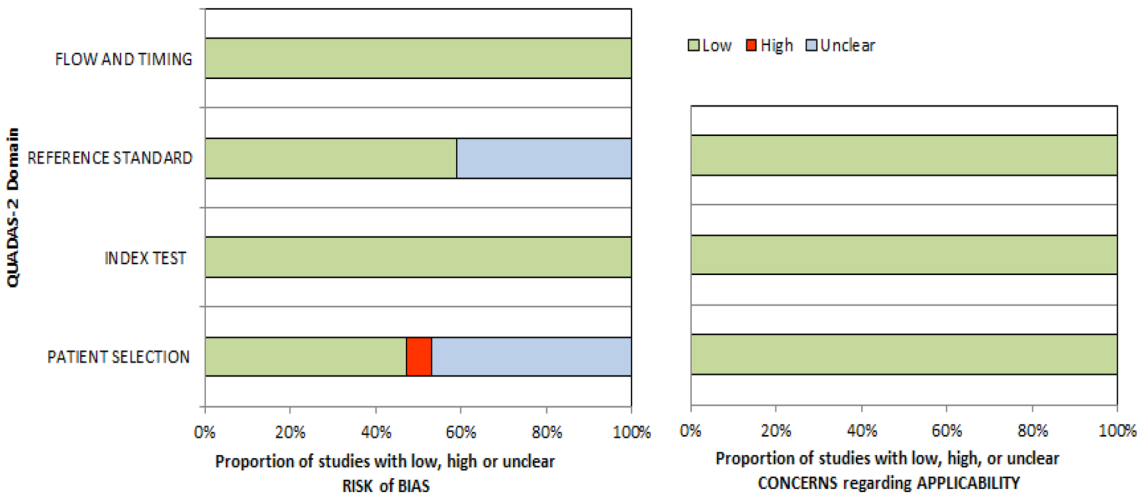


Fig. 1 Flow diagram of article selection

Fig. 2 Quality assessment results of the selected studies by QUADAS-2

Table 1 Characteristics of selected studies

Year	Authors	Location	Subjects	Age	Weeks pregnant/ postpartum	Gold standard	Blind	Total (*n*)	Scales	Cut off	2×2 Table				Value (95% confidence interval)	
											TP	FP	FN	TN	Sensitivity	Specificity
2019	Matthey et al. [30]	Australia	Pregnant women	28.4 ± 5.0	14.2 ± 3.8	MINI	Unknown	252	EPDS	≥ 10	24	6	12	210	0.67 (0.50–0.80)	0.97 (0.94–0.99)
									MGMQ	?	20	13	16	203	0.56 (0.40–0.70)	0.94 (0.90–0.96)
	Naja et al. [31]	Qatar	Pregnant women	28.8 ± 5.0	1-3trimester (2nd 48.4%)	MINI	Yes	128	EPDS	≥ 13	30	9	5	84	0.86 (0.71–0.94)	0.90 (0.83–0.95)
									BDI-II	≥ 19	34	25	1	68	0.97 (0.85–0.99)	0.73 (0.63–0.81)
	Sasaki et al. [32]	Japan	Postpartum women	34.9 ± 4.9	36.6	DSM-5	Unknown	80	EPDS	≥ 13	5	1	4	62	0.56 (0.27–0.81)	0.98 (0.92–1.00)
									CES-D	≥ 21	4	5	5	56	0.44 (0.19–0.73)	0.92 (0.82–0.96)
2018	Chorwe-Sungani et al. [33]	South Africa	Pregnant women	25.8 ± 5.2	27.7 ± 7.9	MINI	Yes	97	EPDS	≥ 10	17	9	8	63	0.68 (0.48–0.83)	0.88 (0.78–0.93)
									HSCL-15	≥ 15	18	5	7	67	0.72 (0.52–0.86)	0.93 (0.85–0.97)
									SRQ	≥ 10	15	2	10	70	0.60 (0.41–0.77)	0.97 (0.90–0.99)
									3-item screener	≥ 1	20	14	5	58	0.80 (0.61–0.91)	0.81 (0.70–0.88)
	Green et al. [34]	Kenya	Pregnant and post-partum women	27.1 ± 5.9	2–3 trimester (4–24)	DSM-5	Yes	193	EPDS	≥ 16	7	51	3	132	0.70 (0.40–0.89)	0.72 (0.65–0.78)
									PHQ-9	≥ 15	7	48	3	135	0.70 (0.40–0.89)	0.74 (0.67–0.80)
									PDEPS	≥ 13	9	18	1	165	0.90 (0.60–0.98)	0.90 (0.85–0.94)
	van Heyningen et al. [35]	South Africa	Pregnant women	26.8 ± 5.9	1–3 trimester (2nd 46.5%)	MINI +	Unknown	376	EPDS	≥ 14	71	56	12	237	0.86 (0.76–0.92)	0.81 (0.76–0.85)
									K10	≥ 15	67	44	16	249	0.81 (0.71–0.88)	0.85 (0.80–0.89)
									PHQ-9	≥ 10	66	53	17	240	0.80 (0.70–0.87)	0.82 (0.77–0.86)
									Whooley	≥ 2	69	44	14	249	0.83 (0.74–0.90)	0.85 (0.80–0.89)
2015	Castro et al. [36]	Brazil	Pregnant women	20–39 (80.0%)	2nd trimester	MINI +	Yes	245	EPDS	≥ 11	35	54	8	150	0.81 (0.67–0.90)	0.74 (0.67–0.79)

Table 1 (continued)

Year	Authors	Location	Subjects	Age	Weeks pregnant/postpartum	Gold standard	Blind	Total (*n*)	Scales	Cut off	2×2 Table				Value (95% confidence interval)	
											TP	FP	FN	TN	Sensitivity	Specificity
									BDI	≥15	32	32	11	172	0.74 (0.60–0.85)	0.84 (0.79–0.89)
2013	Gawlik et al. [37]	Germany	Pregnant women	32.8±4.6	31.8±4.2	SCID	Unknown	273	EPDS	≥12	4	32	1	234	0.80 (0.38–0.96)	0.88 (0.84–0.91)
									PHQ-9	≥10	2	21	3	246	0.40 (0.12–0.77)	0.92 (0.88–0.95)
2012	Tandon et al. [38]	USA	Pregnant and postpartum women	24.4±5.8	31.0±7.3/8.2±3.1	SCID	Unknown	95	EPDS	≥13	22	3	5	65	0.81 (0.63–0.92)	0.96 (0.88–0.98)
									BDI-II	≥20	13	0	14	68	0.48 (0.31–0.66)	1.00 (0.95–1.00)
									CES-D	≥24	19	3	8	65	0.70 (0.52–0.84)	0.96 (0.88–0.98)
2011	Fernandes et al. [39]	UK	Pregnant women	21.5±2.6	35.0±3.0	MINI +	Unknown	194	EPDS	≥13	28	25	0	141	1.00 (0.88–1.00)	0.85 (0.79–0.90)
									K10	≥6	28	31	0	135	1.00 (0.88–1.00)	0.81 (0.75–0.87)
	Flynn et al. [40]	USA	Pregnant women	30.0±8.7	21.0±9.0	DSM-IV	Unknown	81	EPDS	≥13	46	6	12	17	0.79 (0.67–0.88)	0.74 (0.54–0.87)
									PHQ-9	≥10	43	6	15	17	0.74 (0.62–0.84)	0.74 (0.54–0.87)
			Postpartum women	31.0±6.0	12.0±11.0				EPDS	≥13	70	13	6	15	0.92 (0.84–0.96)	0.54 (0.36–0.70)
									PHQ-9	≥10	68	10	8	18	0.89 (0.81–0.95)	0.64 (0.46–0.79)
	Tran et al. [41]	Vietnam	Pregnant and postpartum women	27.0±5.3	At least 28/4–6	SCID	Yes	364	EPDS	≥3/4	76	69	33	186	0.70 (0.61–0.78)	0.73 (0.67–0.78)
									Zung SAS	≥37/38	74	63	35	192	0.68 (0.59–0.76)	0.75 (0.70–0.80)
2010	Tesfaye et al. [42]	UK	Postpartum women	25.3±5.1	10–14	DSM-IV	Yes	100	EPDS	≥7/8	9	20	2	69	0.82 (0.52–0.95)	0.78 (0.68–0.85)
									K10	≥7/8	9	21	2	68	0.82 (0.52–0.95)	0.76 (0.67–0.84)
2008	Hanusa et al. [43]	USA	Postpartum women	30.1±5.8	6–8	SCID	Unknown	29	EPDS	≥10	8	2	5	14	0.62 (0.36–0.82)	0.88 (0.64–0.97)

Table 1 (continued)

Year	Authors	Location	Subjects	Age	Weeks pregnant/postpartum	Gold standard	Blind	Total (*n*)	Scales	Cut off	2×2 Table				Value (95% confidence interval)	
											TP	FP	FN	TN	Sensitivity	Specificity
									PHQ-9	≥10	4	15	9	1	0.31 (0.13–0.58)	0.06 (0.01–0.28)
									PDSS-SF	≥14	12	15	1	1	0.92 (0.67–0.99)	0.06 (0.01–0.28)
	White [44]	New Zealand	Postpartum women	33.0±4.6	7.9±5.1 (2–29)	SCID	Yes	60	EPDS	≥9	14	3	6	37	0.70 (0.48–0.85)	0.92 (0.80–0.97)
									PDSS-SF	≥60	19	9	1	31	0.95 (0.76–0.99)	0.78 (0.62–0.88)
	Su et al. [45]	Taiwan	Pregnant women	32.0±4.0	2–3 trimester (2nd 49.7%)	MINI	Yes	185	EPDS	≥13	19	18	4	144	0.83 (0.63–0.93)	0.89 (0.83–0.93)
									BDI-II	≥11/12	17	28	6	134	0.74 (0.54–0.87)	0.83 (0.76–0.88)
2001	Beck et al. [46]	USA	Postpartum women	31.0±4.8	2–12	SCID	Yes	150	EPDS	≥12	36	1	10	103	0.78 (0.64–0.88)	0.99 (0.95–1.00)
									BDI-II	≥20	26	0	20	104	0.57 (0.42–0.70)	1.00 (0.96–1.00)
									PDSS	≥80	43	2	3	102	0.93 (0.82–0.98)	0.98 (0.93–0.99)

TP true positive, *FP* false positive, *FN* false negative, *TN* true negative, *MINI* mini-international neuropsychiatric interview, *EPDS* Edinburgh postnatal depression Scale, *MGMQ* matthey geriatric mood questionnaire, *BDI* beck depression inventory, *DSM-V* diagnostic and statistical manual of mental disorders-fifth edition *CES-D* center for epidemiologic studies for depression, *PHQ-9* patient health questionnaire-9, *MINI+*MINI plus, *PDEPS* perinatal depression screening, *K-10* ten-item kessler psychological distress scale, *Whooley* whooley questions, *HSCL-15* hopkins symptoms checklist-15, *SRQ* self-reporting questionnaire, *SCID* structured clinical interview for diagnostic and statistical manual of mental disorders, *Zung SAS* Zung self-rating anxiety scale, *DSM-IV* DSM-fourth edition *PDSS-SF* postpartum depression screening scale-short from

Predictive validity of the EPDS versus the other tools by subjects

The EPDS

The predicted validity of the EPDS was analyzed using 17 studies involving 2902 women (Fig. 3). The prevalence of MDD was 26.0%. The sensitivity ranged from 0.56 to 1.00 and the specificity ranged from 0.54 to 0.99. In meta-analysis, pooled sensitivity and specificity were 0.79 (95% CI 0.74–0.84) and 0.88 (95% CI 0.82–0.92), respectively, with RE correlation of -0.506, sROC AUC of 0.89 (SE = 0.02), and Q* value of 0.82 (SE = 0.02).

The predictive validity of the EPDS analyzed by subject is as follows. In pregnant women (9 studies including 1831 women), pooled sensitivity and specificity were 0.81 (95% CI 0.75–0.86) and 0.87 (95% CI 0.81–0.91), respectively, with RE correlation of −0.807, sROC AUC of 0.90 (SE = 0.02), and Q* value of 0.83 (SE = 0.02). In postpartum women (6 studies including 515 women), pooled sensitivity and specificity were 0.79 (95% CI 0.67–0.88) and 0.92 (95% CI 0.76–0.98), respectively, with RE correlation of −0.953, sROC AUC of 0.90 (SE = 0.02), and Q* value of 0.83 (SE = 0.02). In perinatal women (3 studies including 652 women), pooled sensitivity and specificity were 0.72 (95% CI 0.64–0.79) and 0.83 (95% CI 0.63–0.93), respectively, with RE correlation of NaN, sROC AUC of 0.74 (SE = 0.04), and Q* value of 0.68 (SE = 0.03).

The other tools

In meta-analysis of all other tools (Fig. 4), pooled sensitivity and specificity were 0.78 (95% CI 0.71–0.84) and 0.86 (95% CI 0.77–0.92), respectively, with RE correlation of −0.304, sROC AUC of 0.87 (SE = 0.02), and Q* value of 0.81 (SE = 0.02).

In pregnant women, pooled sensitivity and specificity were 0.78 (95% CI: 0.70 to 0.84) and 0.86 (95% CI 0.83–0.89), respectively, with RE correlation of −1.000, sROC AUC of 0.89 (SE = 0.01), and Q* value of 0.82 (SE = 0.01). In postpartum women, pooled sensitivity and specificity were 0.80 (95% CI 0.60–0.91 and 0.79 (95% CI 0.33–0.97), respectively, with RE correlation of 0.002, sROC AUC of 0.86 (SE = 0.08), and Q* value of 0.79 (SE = 0.08). In perinatal women, pooled sensitivity and specificity were 0.70 (95% CI 0.58–0.80) and 0.90 (95% CI 0.74–0.97), respectively, with RE correlation of −0.802, sROC AUC of 0.83 (SE = 0.06), and Q* value of 0.76 (SE = 0.06) (Table 2).

Predictive validity of the EPDS versus each other tools

The PHQ-9

There were 5 studies (1054 women) compared the PHQ-9 and the EPDS. Pooled sensitivity and specificity of EPDS were 0.82 (95% CI 0.74–0.89) and 0.79 (95% CI 0.69–0.87), respectively, with RE correlation of −1.000, sROC AUC of 0.86 (SE = 0.03), and Q* value of 0.79 (SE = 0.03). On the other hand, the pooled sensitivity and specificity of the PHQ-9 were 0.72 (95% CI 0.54–0.85) and 0.69 (95% CI 0.40–0.88), respectively, with RE correlation of 0.643, sROC AUC of 0.74 (SE = 0.12), and Q* value of 0.68 (SE = 0.10).

The BDI

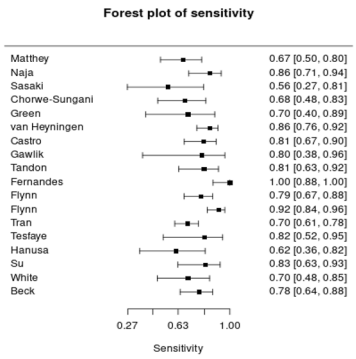
Five studies (8054 women) compared the BDI and the EPDS. The pooled sensitivity and specificity of EPDS were 0.82 (95% CI 0.75–0.87) and 0.92 (0.82–0.97), respectively, with RE correlation of −1.000, sROC AUC of 0.90 (SE = 0.02), and Q* value of 0.83 (SE = 0.02). The pooled sensitivity and specificity of the BDI were 0.73 (95% CI 0.55–0.86) and 0.96 (95% CI 0.71–1.00), respectively, with RE correlation of −0.915, sROC AUC of 0.91 (SE = 0.04), and Q* value of 0.84 (SE = 0.04).

The K-10

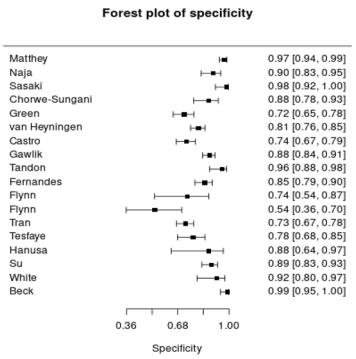
Three studies (670 women) compared the K10 and the EPDS. The pooled sensitivity and specificity of EPDS were 0.92 (95% CI 0.71–0.98) and 0.82 (0.77–0.85), respectively, with RE correlation of 1.000, sROC AUC of 0.87 (SE = 0.04), and Q* value of 0.80 (SE = 0.04). The pooled sensitivity and specificity of the K10 were 0.91 (95% CI 0.70–0.98) and 0.82 (95% CI 0.78–0.86), respectively, with RE correlation of −1.000, sROC AUC of 0.91 (SE = 0.02), and Q* value of 0.84 (SE = 0.02).

The PDSS

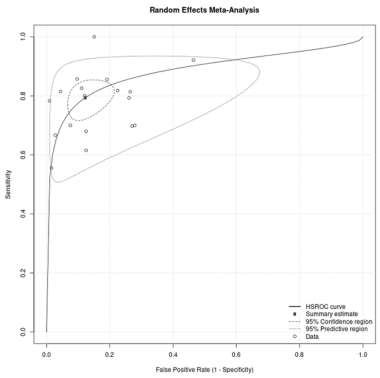
Three studies (239 women) compared the PDSS and the EPDS. The pooled sensitivity and specificity of EPDS were 0.72 (95% CI 0.58–0.82) and 0.95 (0.86–0.99), respectively, with RE correlation of 1.000, sROC AUC of 0.54 (SE = 0.25), and Q value of 0.53 (SE = 0.18). The pooled sensitivity and specificity of the PDSS were 0.94 (95% CI 0.86–0.97) and 0.71 (95% CI 0.09–0.98), respectively, with RE correlation of NaN, sROC AUC of 0.98 (SE = 0.01), and Q* value of 0.94 (SE = 0.02).



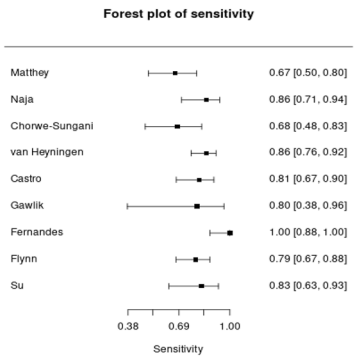
A. Sensitivity in total subjects



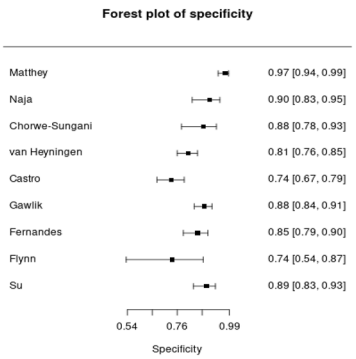
B. Specificity in total subjects



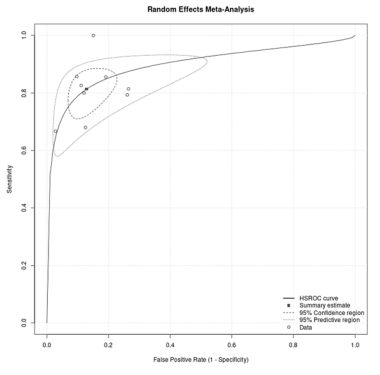
C. sROC curve in total subjects



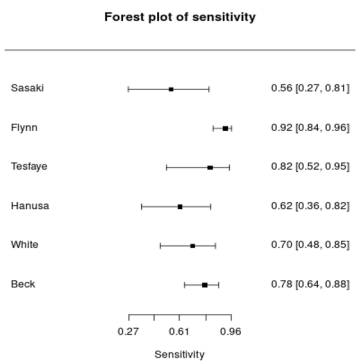
D. Sensitivity in pregnant women



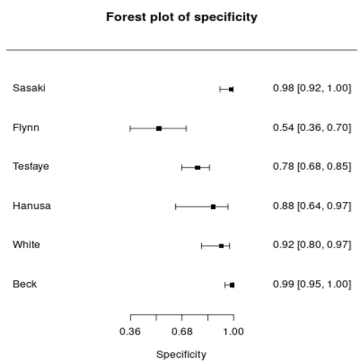
E. Specificity in pregnant women



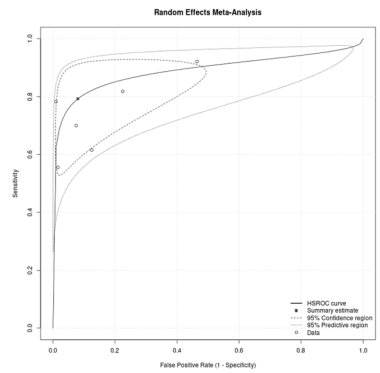
F. sROC curve in pregnant women



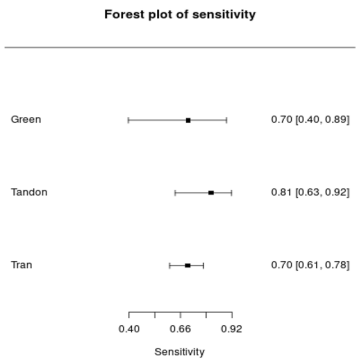
G. Sensitivity in postpartum women



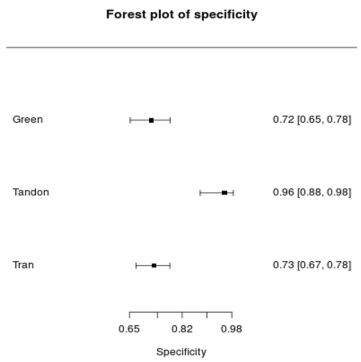
H. Specificity in postpartum women



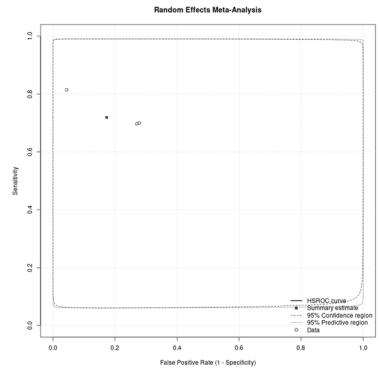
I. sROC curve in postpartum women



J. Sensitivity in perinatal women

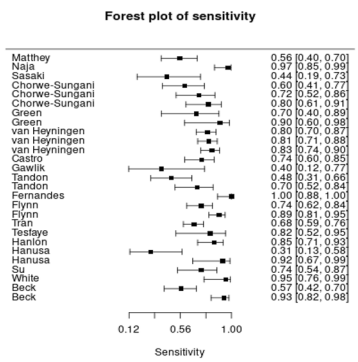


K. Specificity in perinatal women

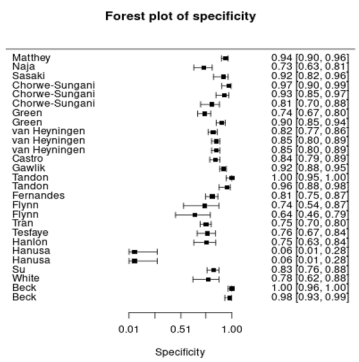


L. sROC curve in perinatal women

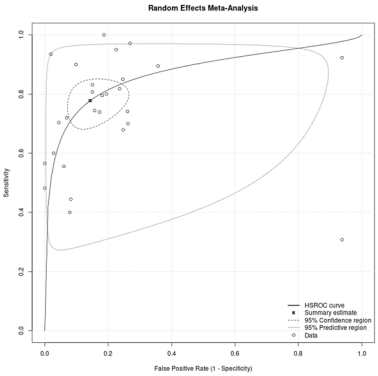
Fig. 3 Predictive validity of the EPDS



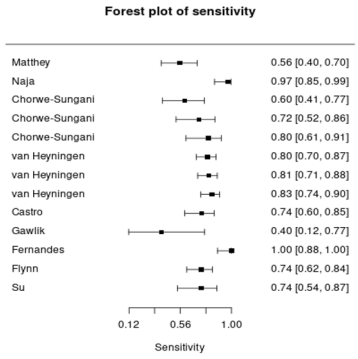
A. Sensitivity in total subjects



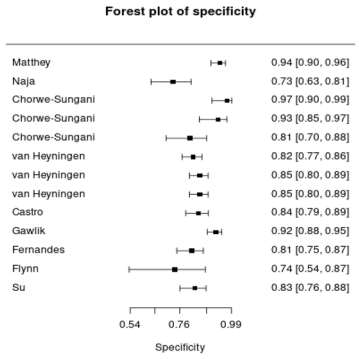
B. Specificity in total subjects



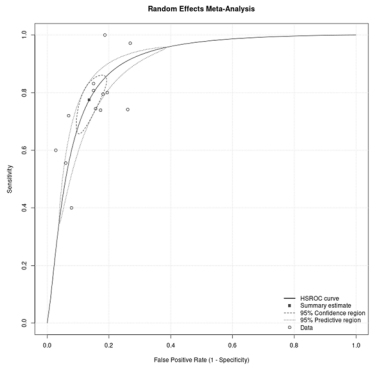
C. sROC curve in total subjects



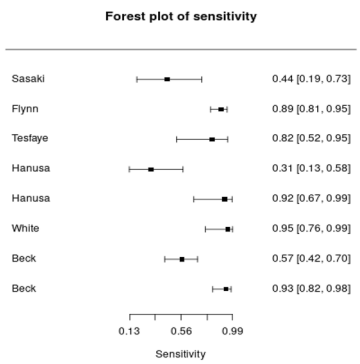
A. Sensitivity in pregnant women



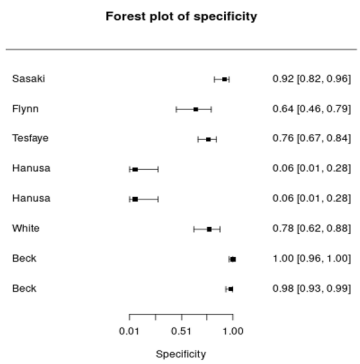
B. Specificity in pregnant women



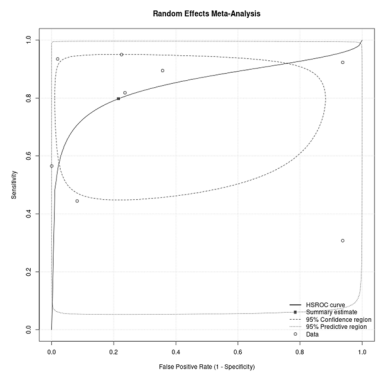
C. sROC curve in pregnant women



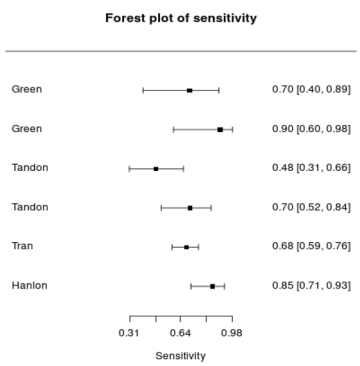
D. Sensitivity in postpartum women



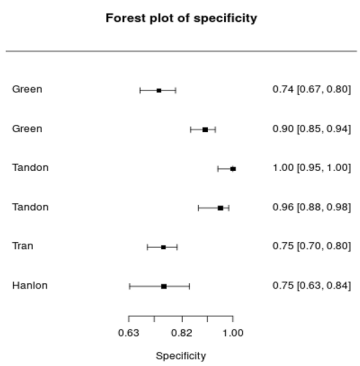
E. Specificity in postpartum women



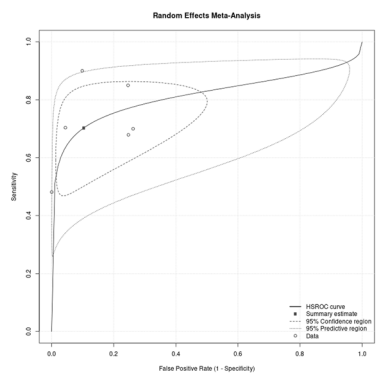
F. sROC curve in postpartum women



G. Sensitivity in perinatal women



H. Specificity in perinatal women



I. sROC curve in perinatal women

Fig. 4 Predictive validity of the other tools

Table 2 Summary results of meta-analysis

Categories		Studies (*k*)	Subjects (*n*)	Prevalence (mean)	Pooled diagnostic test accuracy (95% confidence interval)					RE correlation	Summary ROC curve			
					Sensitivity	Specificity	Positive likelihood ratio	Negative likelihood ratio	Diagnostic odds ratio		AUC	SE	Q*	(SE)
EPDS versus other tools by subjects														
EPDS	Total	17	2902	26.0	0.79 (0.74–0.84)	0.88 (0.82–0.92)	6.48 (4.30–9.76)	0.24 (0.19–0.30)	27.55 (17.01–44.64)	− 0.506	0.89	0.02	0.82	0.02
	Pregnant women	9	1831	23.0	0.81 (0.75–0.86)	0.87 (0.81–0.91)	6.33 (4.42–9.07)	0.21 (0.16–0.28)	29.53 (19.07–45.70)	− 0.807	0.90	0.02	0.83	0.02
	Postpartum women	6	515	34.0	0.79 (0.67–0.88)	0.92 (0.76–0.98)	9.89 (3.30–29.67)	0.23 (0.14–0.36)	43.78 (15.39–124.53)	− 0.953	0.90	0.02	0.83	0.02
	Perinatal women	3	652	21.0	0.72 (0.64–0.79)	0.83 (0.63–0.93)	4.16 (1.76–9.87)	0.34 (0.25–0.47)	12.26 (4.09–36.74)	NaN	0.74	0.04	0.68	0.03
Other tools	Total	17	2902	26.0	0.78 (0.71–0.84)	0.86 (0.77–0.92)	5.44 (3.35–8.83)	0.26 (0.20–0.34)	20.94 (11.40–38.46)	− 0.304	0.87	0.02	0.81	0.02
	Pregnant women	9	1831	23.0	0.78 (0.70–0.84)	0.86 (0.83–0.89)	5.69 (4.71–6.88)	0.26 (0.20–0.35)	21.82 (16.68–28.55)	− 1.000	0.89	0.01	0.82	0.01
	Postpartum women	6	515	34.0	0.80 (0.60–0.91)	0.79 (0.33–0.97)	3.71 (0.75–18.33)	0.26 (0.11–0.63)	14.37 (1.53–135.29)	0.002	0.86	0.08	0.79	0.08
	Perinatal women	3	652	21.0	0.70 (0.58–0.80)	0.90 (0.74–0.97)	6.84 (2.63–17.82)	0.33 (0.24–0.46)	20.63 (7.65–55.61)	− 0.802	0.83	0.06	0.76	0.06
EPDS versus each other tools														
	EPDS	5	1054	36.0	0.82 (0.74–0.89)	0.79 (0.69–0.87)	3.92 (2.70–5.68)	0.22 (0.16–0.32)	17.62 (11.18–27.77)	− 1.000	0.86	0.03	0.79	0.03
	PHQ-9				0.72 (0.54–0.85)	0.69 (0.40–0.88)	2.30 (0.88–5.98)	0.41 (0.18–0.94)	5.67 (0.99–32.61)	0.643	0.74	0.12	0.68	0.10
	EPDS	5	805	23.0	0.82 (0.75–0.87)	0.92 (0.82–0.97)	10.63 (4.28–26.36)	0.20 (0.15–0.27)	53.52 (19.49–146.94)	− 1.000	0.90	0.02	0.83	0.02
	BDI				0.73 (0.55–0.86)	0.96 (0.71–1.00)	16.16 (2.35–110.96)	0.28 (0.17–0.48)	57.24 (10.21–321.12)	− 0.915	0.91	0.04	0.84	0.04
	EPDS	3	670	16.0	0.92 (0.71–0.98)	0.82 (0.77–0.85)	5.01 (3.71–6.75)	0.10 (0.03–0.41)	47.94 (9.65–238.24)	1.000	0.87	0.04	0.80	0.04
	K10				0.91 (0.70–0.98)	0.82 (0.78–0.86)	5.11 (4.13–6.32)	0.11 (0.03–0.40)	45.39 (11.59–177.74)	− 1.000	0.91	0.02	0.84	0.02
	EPDS	3	239	36.0	0.72 (0.58–0.82)	0.95 (0.86–0.99)	16.69 (4.51–61.72)	0.30 (0.19–0.47)	56.34 (11.26–281.83)	1.000	0.54	0.25	0.53	0.18
	PDSS				0.94 (0.86–0.97)	0.71 (0.09–0.98)	3.18 (0.33–30.45)	0.09 (0.03–0.32)	35.42 (1.27–987.99)	NaN	0.98	0.01	0.94	0.02
	EPDS	7	1449	20.0	0.73 (0.65–0.80)	0.90 (0.79–0.96)	7.41 (3.46–15.87)	0.30 (0.22–0.39)	25.13 (10.67–59.17)	− 0.424	0.84	0.04	0.77	0.04
	Others				0.70 (0.62–0.77)	0.90 (0.85–0.94)	7.27 (4.76–11.09)	0.33 (0.26–0.43)	21.81 (13.11–36.29)	− 0.552	0.88	0.01	0.81	0.01

RE correlation random effect correlation, *ROC curve* receiver operating characteristic curve, *AUC* area under the curve, SE standard error, *EPDS* Edinburgh postnatal depression scale, *PHQ-9* patient health questionnaire-9, *BDI* beck depression inventory, *PDSS* postpartum depression screening scale, *K-10* ten-item kessler psychological distress scale

Others

Other depression screening tools were reported in 7 studies (1449 women). The pooled sensitivity and specificity of EPDS were 0.73 (95% CI 0.65–0.80) and 0.90 (0.79–0.96), respectively, with RE correlation of -0.424, sROC AUC of 0.84 (SE = 0.04), and Q* value of 0.77 (SE = 0.04). The pooled sensitivity and specificity of other tools were 0.70 (95% CI 0.62–0.77) and 0.90 (95% CI 0.85–0.94), respectively, with RE correlation of −0.552, sROC AUC of 0.88 (SE = 0.01), and Q* value of 0.81 (SE = 0.01).

Discussion

Depression occurs mainly during the postpartum period, but pregnancy also can be an emotional time [1]. In many cases, the depression often sets in while the women are still pregnant, so we prefer the term peri-partum to postpartum [47]. The EPDS was developed as a screening tool to detect PPD [15]. It is also actively used to screen for depression in pregnant women [16]. Therefore, the target of EPDS has been expanded from postpartum women to pregnant women. This was also revealed in the 17 studies included in this review. The studies on postpartum women were published mainly before 2010, the majority of the studies on pregnant women were published recently. The screening accuracy of the EPDS has been proven through several SRs [19, 48]. However, in most cases, only the EPDS was verified [20]. Even if pregnant women were included, no distinction of among depression or depressive order between pregnant and postpartum women was made in the report [18] or adolescents were mixed into subjects [49]. In addition, another SR was limited to one country [50]. Therefore, the objective of this study was to perform the first systematic review and meta-analysis to classify pregnant, postpartum, and perinatal women through literature selected with a standardized method. How the EPDS differed from other depression screening tools was then determined.

First, the quality of the selected studies was assessed using QUADAS-2 [23]. The EPDS is a self-reporting tool. Subjects were women in their 20 s and 30 s. The EPDS as an index test has a quantified scoring system. It can be interpreted that there is no bias in the test execution process. Thus, blinding in reference standards is important when assessing the quality of the literature. Since this study included only studies consisting of unstructured or semi-structured interviews with psychiatrists, there was no bias in the reference standard for correctly classifying the target condition, MDD. Most studies (9 out of 17) were blinded, allowing analysis to be performed using well-designed literature with a relatively low risk of bias.

In the meta-analysis, the EPDS showed a moderate accuracy with an sROC AUC of 0.89 (SE = 0.02) as a tool specialized for perinatal women. When pregnant and postpartum women were analyzed separately, sROC AUC was equal to 0.90 (SE = 0.02) for both women. The pooled sensitivity was slightly higher in pregnant women, while the pooled specificity was slightly higher in postpartum women. The RE correlation had negative values for pregnant women (− 0.807) and postpartum women (− 0.953), and no heterogeneity among studies. These results confirm that the EPDS is a suitable screening tool for not only PPD, but also depression during pregnancy, which was overlooked in the past. However, perinatal women, the sROC AUC was as low accurate as 0.74 relatively. As the number of studies was small, we could not interpret them significantly. The SROC AUC of other tools was 0.87 (SE = 0.03), which showed a low value slightly compared to the EPDS. Therefore, we could confirm that the EPDS is a perinatal-specific depression screening tool. In perinatal women, the sROC AUC was 0.83 and showed low accuracy performance relative to pregnant and postpartum women like the case of the EPDS.

Subgroup analyses of other tools were performed only when there were three or more studies. However, due to the small number of studies, it was not possible to be analyzed by subjects. As a result of meta-analysis of all studies reporting other tools, the sROC AUC was 0.87 (SE = 0.02), which was similar to that of the EPDS (0.89). However, there were differences in subgroup analysis for each tool. Compared with five studies (1,054 women), the sROC AUC of the PHQ-9 was 0.74 (SE = 0.12), which was lower than that of EPDS 0.86 (SE = 0.03). SR by Wang et al. [51] has also compared the EPDS and the PHQ-9. However, quantitative meta-analysis was not performed. Only median AUC was presented (EPDS: 0.88; PHQ-9: 0.86. Only one in every five studies showed the PHQ-9 differed from the EPDS: Gawlik et al. [37] in sensitivity and Hausa et al. [43] in specificity. Therefore, although the screening performance of the EPDS in perinatal women is rather good, by only these results, we did not interpret as that the PHQ-9 had lower screening performance than the EPDS.

On the other hand, in comparison with the PDSS (0.98), the sROC of the EPDS was very low accuracy at 0.54. Both tools are perinatal-specific depression screening tools. The EPDS was developed in 1987 and the PDSS was made in 2000 [52]. In all individual studies, the PDSS displayed higher sensitivity, and the EPDS had higher specificity. So, the pooled sensitivity was relatively high in PDSS (0.94) while the pooled specificity was high in EPDS (0.95). When interpreting only these results, it was considered that they could be complementary tools. First, if women with suspected depression are selected through the PDSS, and then the normal range of woman is correctly excluded by the EPDS, it seems that there will be a synergistic effect to

increase the accuracy of depression screening in perinatal women. However, in this case, there were only three studies and the RE correlation value was not calculated in PDSS. In EPDS, there was a heterogeneity between studies as a positive value (1.000). In addition, the PDSS includes both a long and short version. Thus, interpretation of this tool was withheld. For the BDI, sROC AUC was 0.91, which was similar to that of EPDS (0.90). RE correlation was negative for both. There was no heterogeneity between studies. In the case of other tools, eight depression screening tools such as the CES-D and Zung Self-Rating Anxiety Scale were included and analyzed.

Limitation

This study also has limitations. First, the cut-off score of the EPDS was the optimal score suggested by each study was applied. In most studies, the recommended cut-off score for the EPDS was 13, which was within the optimal cut-off score range of 9–13 in Fellmech's SR [53]. However, in some studies, exceptional cut-off scores such as 3/4 [41] and 16 [34] were used. Second, in the meta-analysis results of 17 studies on pregnant and postpartum women with subgroup analysis using the BDI, the RE correlation indicating heterogeneity between studies had negative values. However, in subgroup analysis with other tools such as the PDSS, the RE correlation indicating heterogeneity between studies had positive values. In the present meta-analysis, the cause of heterogeneity could not be identified due to a small number of studies. Third, this meta-analysis did not include studies on women under the age of 18 with high-risk pregnancies as subjects because young pregnant and postpartum women with age under 18 should be considered as a different group of subjects to estimate the sensitivity and specificity of EPDS for screening postpartum depression.

Conclusion

This study demonstrates that EPDS is an excellent depression screening tool for pregnant and postpartum women aged over 18 years based on 17 well-designed studies with a low risk of bias. In this study, the sROC AUC of the EPDS was 0.9, indicating moderate accuracy. In a primary care setting or a midwifery center, the EPDS is a very good screening tool that can be applied before other tools when trying to screen for depression in perinatal women. Among other tools for screening depression, the BDI has the most similar screening accuracy to the EPDS. The PHQ-9 and the K10 can also be used.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00404-022-06525-0.

Author contributions All authors (SHP and JIK) contributed to the study's conception and design, material preparation, data collection and analysis. Advanced analysis was performed by SHP. All authors have written the first draft of the manuscript and commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding This work was supported by the Soonchunhyang University Research Fund. And this study was accomplished relating to searching causes of preterm birth by a grant by a grant (NRF 2019R1F1A104579912) of the Basic Science Research Program through the National Research Foundation (NRF) funded by the Ministry of Education, Republic of Korea. The funder had no further role in the conduct of the research.

Data availability Yes.

Declarations

Conflict of interest The authors have no conflicts of interest or any competing interests to declare.

References

1. Biaggi A, Conroy S, Pawlby S, Pariante CM (2016) Identifying the women at risk of antenatal anxiety and depression: a systematic review. J Affect Disord 191:62–77. https://doi.org/10.1016/j.jad.2015.11.014
2. Hay DF, Pawlby S, Waters CS, Sharp D (2008) Antepartum and postpartum exposure to maternal depression: different effects on different adolescent outcomes. J Child Psychol Psychiatry 49:1079–1088. https://doi.org/10.1111/j.1469-7610.2008.01959.x
3. Kingston D, Kehler H, Austin MP et al (2018) Trajectories of maternal depressive symptoms during pregnancy and the first 12 months postpartum and child externalizing and internalizing behavior at three years. PLoS One 13(4):e0195365. https://doi.org/10.1371/journal.pone.0195365
4. Guintivano J, Manuck T, Meltzer-Brody S (2018) Predictors of postpartum depression: a comprehensive review of the last decade of evidence. Clin Obstet Gynecol 61(3):591–603. https://doi.org/10.1097/GRF.0000000000000368
5. Leigh B, Milgrom J (2008) Risk factors for antenatal depression, postnatal depression and parenting stress. BMC Psychiatry 8(1):24. https://doi.org/10.1186/1471-244X-8-24
6. ACOG committee on practice bulletins-obstetrics (2018) ACOG Practice Bulletin: Clinical management guidelines for obstetrician-gynecologists number 92, April 2008 (replaces practice bulletin number 87, November 2007) use of psychiatric medications during pregnancy and lactation. Obstet Gynecol 111:1001–1020. https://doi.org/10.1097/AOG.0b013e31816fd910 (**Reaffirmed Obstet Gynecol 131:185. Reaffirmed 2020**)
7. Stewart DE (2011) Clinical practice depression during pregnancy. N Engl J Med 365(17):1605–1611. https://doi.org/10.1056/NEJMcp1102730
8. Fitelson E, Kim S, Baker AS, Leight K (2010) Treatment of postpartum depression: clinical, psychological and pharmacological options. Int J Womens Health 3:1–14. https://doi.org/10.2147/IJWH.S6938
9. Horowitz JA, Goodman JH (2005) Identifying and treating postpartum depression. JOGNN 34(2):264–273. https://doi.org/10.1177/0884217505274583

10. Szegedi A, Jansen WT, van Willigenburg AP, van der Meulen E, Stassen HH, Thase ME (2009) Early improvement in the first 2 weeks as a predictor of treatment outcome in patients with major depressive disorder: a meta-analysis including 6562 patients. J Clin Psychiatry 70(3):344–353. https://doi.org/10.4088/JCP.07m03780
11. Ren Y, Yang H, Browning C, Thomas S, Liu M (2015) Performance of screening tools in detecting major depressive disorder among patients with coronary heart disease: a systematic review. Med Sci Monit 21:646–653. https://doi.org/10.12659/MSM.892537
12. Buist A, Bilszta J, Barnett B et al (2005) Recognition and management of perinatal depression in general practice: a survey of GPs and postnatal women. Aust Fam Phys 34:787–790
13. Callister LC, Beckstrand RL, Corbett C (2011) Postpartum depression and help-seeking behaviors in immigrant hispanic women. JOGNN 40(4):440–449
14. ACOG (2018) ACOG committee opinion no. 757 summary: screening for perinatal depression. Obstet Gynecol 132:1314–1316. https://doi.org/10.1097/AOG.0000000000002928
15. Cox JL, Holden JM, Sagovsky R (1987) Detection of postnatal depression. development of the 10-item Edinburgh postnatal depression scale. Br J Psychiatry 150:782–786. https://doi.org/10.1192/bjp.150.6.782
16. Rachel E (2015) Perinatal mental health: the Edinburgh postnatal depression scale, Manual 2nd edn. In: Cox J, Holden JM, Henshaw CM (eds) RCPsych, vol 25. Primary Health Care, London, p 12
17. Lyubenova A et al (2021) Depression prevalence based on the Edinburgh postnatal depression scale compared to structured clinical interview for DSM disorders classification: systematic review and individual participant data meta-analysis. Int J Methods Psychiatr Res 30(1):1–13. https://doi.org/10.1002/mpr.1860
18. Levis B, Yan XW, He C, Sun Y, Benedetti A, Thombs BD (2019) Comparison of depression prevalence estimates in meta-analyses based on screening tools and rating scales versus diagnostic interviews: a meta-research review. BMC Med 17(1):65–65. https://doi.org/10.1186/s12916-019-1297-6
19. Levis B, Negeri A, Sun Y et al (2020) Accuracy of the Edinburgh postnatal depression scale (EPDS) for screening to detect major depression among pregnant and postpartum women: systematic review and meta-analysis of individual participant data. BMJ 371:m4022. https://doi.org/10.1136/bmj.m4022
20. Thombs BD, Benedetti A, Kloda LA et al (2015) Diagnostic accuracy of the Edinburgh postnatal depression scale (EPDS) for detecting major depression in pregnant and postnatal women: protocol for a systematic review and individual patient data meta-analyses. BMJ Open 5(10):e009742. https://doi.org/10.1136/bmjopen-2015-009742
21. Bossuyt PM, Davenport C, Deeks JJ et al. (2013) Cochrane handbook for systematic reviews of diagnostic test accuracy. The cochrane training. https://training.cochrane.org/resource/cochrane-handbook-systematic-reviews-diagnostic-test-accuracy. Accessed 15 May 2021
22. Page MJ, McKenzie JE, Bossuyt PM et al (2021) The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 372:n71. https://doi.org/10.1136/bmj.n71
23. Whiting PF, Rutjes AW, Westwood ME et al (2011) QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 155:529–536. https://doi.org/10.7326/0003-4819-155-8-201110180-00009
24. Negeri ZF, Shaikh M, Beyene J (2018) Bivariate random-effects meta-analysis models for diagnostic test accuracy studies using arcsine-based transformations. Biom J 60:827–844. https://doi.org/10.1002/bimj.201700101
25. Zamora J, Abraira V, Muriel A et al (2006) Meta-DiSc: a software for meta-analysis of test accuracy data. BMC Med Res Methodol 6(1):31–31. https://doi.org/10.1186/1471-2288-6-31
26. Patel A, Cooper N, Freeman S et al (2021) Graphical enhancements to summary receiver operating characteristic plots to facilitate the analysis and reporting of meta-analysis of diagnostic test accuracy data. Res Synth Methods 12:34–44. https://doi.org/10.1002/jrsm.1439
27. Freeman SC, Kerby CR, Patel A et al (2019) Development of an interactive web-based tool to conduct and interrogate meta-analysis of diagnostic test accuracy studies: MetaDTA. BMC Med Res Methodol 19:81
28. Greiner M, Pfeiffer D, Smith RD (2000) Principles and practical application of the receiver-operating characteristic analysis for diagnostic tests. Prev Vet Med 45:23–41. https://doi.org/10.1016/S0167-5877(00)00115-X
29. Walter SD (2002) Properties of the summary receiver operating characteristic (SROC) curve for diagnostic test data. Stat Med 21:1237–1256
30. Matthey S, Souter K, Valenti B, Ross-Hamid C (2019) Validation of the MGMQ in screening for emotional difficulties in women during pregnancy. J Affect Disord 256:156–163. https://doi.org/10.1016/j.jad.2019.05.037
31. Naja S, Al Kubaisi N, Chehab M, Al Dahshan A, Abuhashem N, Bougmiza I (2019) Psychometric properties of the Arabic version of EPDS and BDI-II as a screening tool for antenatal depression: evidence from Qatar. BMJ Open 9(9):e030365. https://doi.org/10.1136/bmjopen-2019-030365
32. Sasaki Y, Baba T, Oyama R, Fukumoto K, Haba G, Sasaki M (2019) Re-evaluation of the Edinburgh postnatal depression scale as screening for post-partum depression in Iwate prefecture, Japan. J Obstet Gynaecol Res 45(9):1876–1883. https://doi.org/10.1111/jog.14042
33. Chorwe-Sungani G, Chipps J (2018) Validity and utility of instruments for screening of depression in women attending antenatal clinics in Blantyre district in Malawi. S Afr Fam Pract 60:114–120. https://doi.org/10.4102/safp.v60i4.4888
34. Green EP, Tuli H, Kwobah E, Menya D, Chesire I, Schmidt C (2018) Developing and validating a perinatal depression screening tool in Kenya blending Western criteria with local idioms: a mixed methods study. J Affect Disord 228:49–59. https://doi.org/10.1016/j.jad.2017.11.027
35. van Heyningen T, Honikman S, Tomlinson M, Field S, Myer L (2018) Comparison of mental health screening tools for detecting antenatal depression and anxiety disorders in South African women. PLoS One [Electronic Res] 13(4):e0193697. https://doi.org/10.1371/journal.pone.0193697
36. Castro CT, Martins Brancaglion MY, Nogueira CM, Bergo PA, Duarte GF, Nicolato R, Lopes PAR, Vitor LH, Correa H (2015) What is the best tool for screening antenatal depression? J Affect Disord 178:12–17. https://doi.org/10.1016/j.jad.2015.02.003
37. Gawlik S, Waldeier L, Muller M, Szabo A, Sohn C, Reck C (2013) Subclinical depressive symptoms during pregnancy and birth outcome–a pilot study in a healthy German sample. Arch Womens Ment Health 16(2):93–100. https://doi.org/10.1007/s00737-012-0320-0
38. Tandon SD, Cluxton-Keller F, Leis J, Le HN, Perry DFA (2012) Comparison of three screening tools to identify perinatal depression among low-income African American women. J Affect Disord 136(1–2):155–162. https://doi.org/10.1016/j.jad.2011.07.014
39. Fernandes MC, Srinivasan K, Stein AL, Menezes G, Sumithra R, Ramchandani PG (2011) Assessing prenatal depression in the rural developing world: a comparison of two screening measures. Arch Womens Ment Health 14(3):209–216. https://doi.org/10.1007/s00737-010-0190-2

40. Flynn HA, Sexton M, Ratliff S, Porter K, Zivin K (2011) Comparative performance of the Edinburgh postnatal depression scale and the patient health questionnaire-9 in pregnant and postpartum women seeking psychiatric services. Psychiatry Res 187(1–2):130–134. https://doi.org/10.1016/j.psychres.2010.10.022
41. Tran TD, 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. https://doi.org/10.1016/j.jad.2011.03.038
42. 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 122(1–2):102–108. https://doi.org/10.1016/j.jad.2009.06.020
43. Hanusa BH, Scholle SH, Haskett RF, Spadaro K, Wisner KL (2008) Screening for depression in the postpartum period: a comparison of three instruments. J Womens Health 17(4):585–596. https://doi.org/10.1016/j.jmwh.2008.07.005
44. White G. (2008) A comparison of the postpartum depression screening scale (PDSS) with the Edinburgh postnatal depression scale (EPDS). New Zealand college of midwives journal 10(39):28–32. https://link.gale.com/apps/doc/A206107529/HRCA?u=anon~2d3f0f29&sid=googleScholar&xid=3b00c4bb. Accessed 7 Aug 2021
45. Su KP, Chiu TH, Huang CL, Ho M, Lee CC, Wu PL, Lin CY, Liau CH, Liao CC, Chiu WC, Pariante CM (2007) Different cut-off points for different trimesters? the use of Edinburgh postnatal depression scale and beck depression inventory to screen for depression in pregnant Taiwanese women. Gen Hosp Psychiatry 29(5):436–441. https://doi.org/10.1016/j.genhosppsych.2007.05.005
46. Beck CT, Gable RK (2001) Comparative analysis of the performance of the postpartum depression screening scale with two other depression instruments. Nurs Res 50(4):242–250. https://doi.org/10.1097/00006199-200107000-00008
47. Payne JL (2016) Recent advances and controversies in peripartum depression. Curr Obstet Gynecol Rep 5(3):250–256. https://doi.org/10.1007/s13669-016-0167-x
48. O'Connor E, Rossom RC, Henninger M et al. (2016) Screening for depression in adults: an updated systematic evidence review for the US preventive services task force. evidence synthesis No. 128. AHRQ publication No. 14-05208-EF-1. Rockville, MD: Agency for healthcare research and quality. http://lps3.www.ncbi.nlm.nih.gov.libproxy.snu.ac.kr/books/NBK349027/?report=reader. Accessed 13 Nov 2021
49. Nelson HD, Cantor A, Pappas M, Weeks C (2020) Screening for anxiety in adolescent and adult women: a systematic review for the women's preventive services initiative. Ann Intern Med 173(1):29–41. https://doi.org/10.7326/M20-0579
50. Reuland D, Cherrington A, Watkins GS, Bradford DW, Blanco RA, Gaynes BN (2009) Diagnostic accuracy of Spanish language depression-screening instruments. Ann Fam Med 7(5):455–462. https://doi.org/10.1370/afm.981
51. Wang L, Kroenke K, Stump TM, Monahan PO (2021) Screening for perinatal depression with the patient health questionnaire depression scale (PHQ-9): a systematic review and meta-analysis. Gen Hosp Psychiatry 68:74–78. https://doi.org/10.1016/j.genhosppsych.2020.12.007
52. Beck CT, Gable RK (2000) Postpartum depression screening scale: development and psychometric testing. Nurs Res 49(5):272–282. https://doi.org/10.1097/00006199-200009000-00006
53. Fellmeth G, Harrison S, Opondo C, Nair M, Kurinczuk JJ, Alderdice F (2021) Validated screening tools to identify common mental disorders in perinatal and postpartum women in India: a systematic review and meta-analysis. BMC Psychiatry 21(1):200–209. https://doi.org/10.1186/s12888-021-03190-6