



## The diagnostic accuracy of the Edinburgh Postnatal Depression Scale without the self-harm item: Does culture matter?

To the Editor:

We read with keen interest the recent article by [Chen et al. \(2023\)](#), in which the authors evaluated the performance of the Edinburgh Postnatal Depression Scale (EPDS) without the self-harm item, called EPDS-9, compared to the complete EPDS, called EPDS-10. They focused on identifying depression among people who are pregnant or postpartum. The authors concluded that the shortened EPDS-9 performs as well as the EPDS-10, suggesting it as a potential replacement for the full-length EPDS.

Our research partially supports the findings of [Chen et al. \(2023\)](#). Our study sample comprises 1153 pregnant women and 309 postpartum women. These participants were enrolled from 11 healthcare centers located throughout Italy ([Cena et al., 2020](#)). The characteristics of the participants are detailed in a separate publication ([Cena et al., 2020](#); [Cena et al., 2021](#); [Cena et al., 2021a,b](#)). Trained psychologists used unstructured clinical interviews and patient-rated Patient Health Questionnaire-9 (PHQ-9) and EPDS questionnaires to evaluate participants' depression.

Our findings indicate a correlation of 0.998 between EPDS-9 and EPDS-10, observed in both the antepartum and postpartum groups. Only 1% of the participants were negative at EPDS-9 cutoff points of <10 but had a non-zero EPDS item 10 score, and 2% at EPDS-9 cutoff points of <13. Furthermore, EPDS-9 demonstrated excellent accuracy in distinguishing EPDS-10-based depression screening in both perinatal groups, in each of the four commonly used cutoff scores ([Levis et al., 2020](#); [Qiu et al., 2023](#)).

We used the PHQ-9 as a criterion to compare the performance of the EPDS-9 versus EPDS-10, using a cut-off value of 13 (which is indicated as the most appropriate for the detection of major depression in perinatal people [[Levis et al., 2019](#)]). EPDS-9 and EPDS-10 demonstrated comparable sensitivity, specificity, and area under the curve (AUC) performances. In the antepartum group, both the EPDS-9 and EPDS-10 (a) show declining sensitivity with increasing cutoff values, (b) have high specificity across all cutoff values, and (c) have AUC values that suggest they perform reasonably well, though their performance declines with increasing cutoff values. Comparison of AUC values between EPDS-9 and EPDS-10 suggests that there are no significant differences in performance between the two versions of EPDS at cutoff values of 10, 11, and 13. However, there appears to be a significant difference in performance at a cutoff value of 12, with the EPDS-10 performing better. Regarding the postpartum group, although the AUC remains relatively high for both EPDS-9 and EPDS-10 across all cutoff values, the equivalence tests showed a statistically significant difference at all cutoff values (see [Table 1](#)), indicating that there is a significant difference in overall test performance. Specifically, the EPDS-10 outperforms the EPDS-9 at all cutoff values.

<https://doi.org/10.1016/j.jpsychires.2024.05.018>

Received 31 July 2023; Received in revised form 31 January 2024; Accepted 3 May 2024

Available online 4 May 2024

0022-3956/© 2024 Elsevier Ltd. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

We also examined the predictive potential of EPDS-9 for responses to the EPDS self-harm item (item 10). The AUC of EPDS-9 against self-harm responses varied depending on the frequency level, which could be an area for further study. Specifically, EPDS-9's AUC against self-harm above the frequency of "hardly" ranged from 0.716 to 0.826, except for cutoff 13 in the antepartum group, where it dropped significantly to 0.288. This decrease in AUC at the cut-off point of 13 suggests that EPDS-9's ability to predict self-harm responses decreases when this more conservative threshold is used. The AUC against self-harm above the frequency of "sometimes" and "often" ranged, respectively, from 0.712 to 0.826 and from 0.445 to 0.675. These variations emphasize the importance of considering frequency when examining self-harm predictions. [Table 1](#) shows the sensitivity, specificity, and AUC for each cutoff value.

Based on our study, we propose two main findings that support those of [Chen et al.](#) First, EPDS-10 and EPDS-9 are strongly correlated. Second, EPDS-9 exhibited similar sensitivity and specificity in screening major depression among pregnant and postpartum women, compared to full EPDS, across the most commonly used cutoff points.

However, unlike the Japanese sample of [Chen et al.](#), EPDS-9 did not predict the responses of Italian participants to the self-harm item as accurately. We found this discrepancy when comparing the differentiation performance of EPDS-9 versus EPDS-10 using the PHQ-9 as a criterion. Likely, the discrepancy is due to the use of different instruments although as Kessler Psychological Distress Scale (K6) (used by [Chen et al.](#)) and PHQ-9 showed a strong correlation ([Cotton et al., 2021](#)).

It is here important to remember that the EPDS was originally developed in English ([Cox et al., 1987](#)). Consequently, both our study and that of [Chen et al.](#) employed translated versions of the scale. Although both the Japanese and the Italian translations have been validated ([Benvenuti et al., 1999](#); [Okano et al., 1996](#)) and shown to be reliable and valid measures for perinatal depression ([Kubota et al., 2018](#); [Stefana et al., 2023](#)) and have demonstrated a similar factor structure (which includes aspects of anxiety and anhedonia ([Kubota et al., 2014](#); [Stefana et al., 2024](#)), the translation process may contribute to some of the inconsistencies in the data. This highlights a critical issue: the necessity of establishing cross-cultural validity for psychological inventories.

Cultural variations in the subjective experience and expression of affective disorders must be taken into account in clinical assessment ([Kiermaier and Groleau, 2001](#)). They may significantly shape the manifestation of depression symptomatology and impact the openness to answer questions about self-harm, as suggested by numerous studies. Mental health issues such as depression can present differently in various cultures due to differences in social norms, belief systems, and levels of stigma associated with mental health ([Kleinman and Good,](#)

**Table 1**  
Sensitivity, specificity, and AUC values.

	Performance of EPDS-9 against EPDS-10-based screening of depression			
	Cutoff = 10	Cutoff = 11	Cutoff = 12	Cutoff = 13
Antepartum	Sensitivity = 0.962 Specificity = 0.997 AUC = 0.979	Sensitivity = 0.968 Specificity = 1 AUC = 0.985	Sensitivity = 0.968 Specificity = 1 AUC = 0.984	Sensitivity = 0.979 Specificity = 1 AUC = 0.990
Postpartum	Sensitivity = 1 Specificity = 0.996 AUC = 0.998	Sensitivity = 0.971 Specificity = 0.998 AUC = 0.984	Sensitivity = 0.946 Specificity = 1 AUC = 0.973	Sensitivity = 0.961 Specificity = 1 AUC = 0.980
Comparison of performance between EPDS-9 and EPDS-10 against PHQ-9-based screening of depression.				
	Cutoff = 10	Cutoff = 11	Cutoff = 12	Cutoff = 13
Antepartum	EPDS-10 Sensitivity = 0.651 Specificity = 0.929 AUC = 0.790 EPDS-9 Sensitivity = 0.651 Specificity = 0.931 AUC = 0.791 AUC difference = 0.001 <i>p</i> = 0.045 Equivalent = true	EPDS-10 Sensitivity = 0.602 Specificity = 0.950 AUC = 0.776 EPDS-9 Sensitivity = 0.602 Specificity = 0.951 AUC = 0.777 AUC difference = 0.001 <i>p</i> = 0.051 Equivalent = true	EPDS-10 Sensitivity = 0.494 Specificity = 0.969 AUC = 0.732 EPDS-9 Sensitivity = 0.470 Specificity = 0.971 AUC = 0.720 AUC difference = 0.012 <i>p</i> < 0.001 Equivalent = false	EPDS-10 Sensitivity = 0.410 Specificity = 0.984 AUC = 0.697 EPDS-9 Sensitivity = 0.041 Specificity = 0.986 AUC = 0.698 AUC difference = 0.001 <i>p</i> = 0.076 Equivalent = true
Postpartum	EPDS-10 Sensitivity = 1 Specificity = 0.851 AUC = 0.925 EPDS-9 Sensitivity = 0.936 Specificity = 0.840 AUC = 0.888 AUC difference = 0.037 <i>p</i> < 0.001 Equivalent = false	EPDS-10 Sensitivity = 0.894 Specificity = 0.874 AUC = 0.884 EPDS-9 Sensitivity = 0.872 Specificity = 0.874 AUC = 0.873 AUC difference = 0.011 <i>p</i> < 0.001 Equivalent = false	EPDS-10 Sensitivity = 0.851 Specificity = 0.916 AUC = 0.884 EPDS-9 Sensitivity = 0.830 Specificity = 0.920 AUC = 0.875 AUC difference = 0.009 <i>p</i> < 0.001 Equivalent = false	EPDS-10 Sensitivity = 0.787 Specificity = 0.958 AUC = 0.873 EPDS-9 Sensitivity = 0.766 Specificity = 0.958 AUC = 0.862 AUC difference = 0.011 <i>p</i> < 0.001 Equivalent = false
Performance of EPDS-9 against thoughts of self-harm.				
	EPDS-9 cutoff = 10	EPDS-9 cutoff = 11	EPDS-9 cutoff = 12	EPDS-9 cutoff = 13
Antepartum	≥hardly ever Sensitivity = 0.106 Specificity = 0.462 AUC = 0.716 ≥sometimes Sensitivity = 0.538	≥hardly ever Sensitivity = 0.083 Specificity = 0.462 AUC = 0.728 ≥sometimes Sensitivity = 0.538	≥hardly ever Sensitivity = 0.055 Specificity = 0.462 AUC = 0.742 ≥sometimes Sensitivity = 0.538	≥hardly ever Sensitivity = 0.0380 Specificity = 0.538 AUC = 0.288 ≥sometimes Sensitivity = 0.462

**Table 1 (continued)**

	Performance of EPDS-9 against EPDS-10-based screening of depression			
	Cutoff = 10	Cutoff = 11	Cutoff = 12	Cutoff = 13
	Specificity = 0.894 AUC = 0.716 ≥often Sensitivity = 0.000 Specificity = 0.889 AUC = 0.445	Specificity = 0.917 AUC = 0.728 ≥often Sensitivity = 0.000 Specificity = 0.912 AUC = 0.456	Specificity = 0.945 AUC = 0.742 ≥often Sensitivity = 0.000 Specificity = 0.939 AUC = 0.470	Specificity = 0.962 AUC = 0.712 ≥often Sensitivity = 0.000 Specificity = 0.958 AUC = 0.479
Postpartum	≥hardly ever Sensitivity = 0.266 Specificity = 0.125 AUC = 0.805 ≥sometimes Sensitivity = 0.875 Specificity = 0.734 AUC = 0.805 ≥often Sensitivity = 0.500 Specificity = 0.720 AUC = 0.610	≥hardly ever Sensitivity = 0.223 Specificity = 0.125 AUC = 0.826 ≥sometimes Sensitivity = 0.875 Specificity = 0.777 AUC = 0.826 ≥often Sensitivity = 0.500 Specificity = 0.762 AUC = 0.631	≥hardly ever Sensitivity = 0.183 Specificity = 0.375 AUC = 0.721 ≥sometimes Sensitivity = 0.625 Specificity = 0.817 AUC = 0.721 ≥often Sensitivity = 0.500 Specificity = 0.808 AUC = 0.654	≥hardly ever Sensitivity = 0.014 Specificity = 0.375 AUC = 0.743 ≥sometimes Sensitivity = 0.625 Specificity = 0.860 AUC = 0.743 ≥often Sensitivity = 0.500 Specificity = 0.850 AUC = 0.675

1985). In some societies, such as the Chinese one, psychological symptoms may be expressed more somatically, which may influence the detection of depressive symptoms through tools such as the EPDS (Ryder et al., 2008).

Concerning self-harm and suicidal ideation, cultural factors can significantly influence the willingness to disclose such experiences. For example, some cultures may have high levels of stigma associated with mental health conditions or self-harm behaviors, making individuals less likely to report these experiences openly (Chu et al., 2010). Additionally, cultures like, for example, the Chinese one prioritize collective identity over individualism may see a higher level of self-stigma, resulting in a lower level of openness about mental health struggles, including self-harm (Yang et al., 2007).

Therefore, it is crucial to keep cultural factors in mind when interpreting the effectiveness of measures such as EPDS-9 and EPDS-10 in different cultures and perinatal populations (pregnant versus postpartum people). The variance between Chen et al.'s and our samples in terms of the predictive precision of EPDS-9 for self-harm responses underscores the need for culturally sensitive approaches in the detection of depression. More research is needed to understand the specific cultural factors at play in the various phases of the perinatal process and adapt the instruments accordingly to improve their validity and reliability.

Lastly, Chen et al.'s suggestion to omit the self-harm item in order to help avoid confusion and potential psychological distress brought to the responders should be considered with caution. Overreliance on item 10 can surely lead to a strain on resources due to mandatory follow-up assessments, but when psychological assessment is done well it is always therapeutic to some degree. Furthermore, as we explained before, in certain cultures (and more generally, in certain people) a 0 score on item 10 does not mean a 0 risk of suicide.

In conclusion, although EPDS-9 shows a performance similar to that of EPDS-10 in the screening of major depression, we recommend the use of the full EPDS. The variance in predictive accuracy between different population samples highlights the need for future research to further validate EPDS-9 in specific cultures and perinatal populations.

**CRedit authorship contribution statement**

**Alberto Stefana:** Conceptualization, Formal analysis, Writing – original draft, Writing – review & editing. **Loredana Cena:** Investigation, Project administration, Writing – review & editing. **Alice Trainini:** Investigation, Writing – review & editing. **Gabriella Palumbo:** Writing – review & editing. **Antonella Gigantesco:** Writing – review & editing. **Fiorino Mirabella:** Writing – review & editing.

**Declaration of competing interest**

The authors have no conflicts of interest to declare. All co-authors have seen and agree with the contents of the manuscript and there is no financial interest to report.

**References**

- Benvenuti, P., Ferrara, M., Nicolai, C., Valoriani, V., Cox, J.L., 1999. The Edinburgh Postnatal depression Scale: validation for an Italian sample. *J. Affect. Disord.* 53 (2), 137–141.
- Cena, L., Gigantesco, A., Mirabella, F., Palumbo, G., Camoni, L., Trainini, A., Stefana, A., 2021a. Prevalence of comorbid anxiety and depressive symptomatology in the third trimester of pregnancy: analysing its association with sociodemographic, obstetric, and mental health features. *J. Affect. Disord.* 295, 1398–1406.
- Cena, L., Gigantesco, A., Mirabella, F., Palumbo, G., Trainini, A., Stefana, A., 2021b. Prevalence of maternal postnatal anxiety and its association with demographic and socioeconomic factors: a multicentre study in Italy. *Front. Psychiatr.* 12, 737666.
- Cena, L., Mirabella, F., Palumbo, G., Gigantesco, A., Trainini, A., Stefana, A., 2020. Prevalence of maternal antenatal anxiety and its association with demographic and socioeconomic factors: a multicentre study in Italy. *Eur. Psychiatr.* 63 (1), e84.
- Cena, L., Mirabella, F., Palumbo, G., Gigantesco, A., Trainini, A., Stefana, A., 2021. Prevalence of maternal antenatal and postnatal depression and their association with sociodemographic and socioeconomic factors: a multicentre study in Italy. *J. Affect. Disord.* 279, 217–221.
- Cena, L., Palumbo, G., Mirabella, F., Gigantesco, A., Stefana, A., Trainini, A., Tralli, N., Imbasciati, A., 2020. Perspectives on early screening and prompt intervention to identify and treat maternal perinatal mental health. Protocol for a prospective multicenter study in Italy. *Front. Psychol.* 11, 365.
- Chen, C., Okubo, R., Okawa, S., Nakagawa, S., Tabuchi, T., 2023. The diagnostic accuracy of the Edinburgh Postnatal Depression Scale without the self-harm item. *J. Psychiatr. Res.* 165, 70–76.
- Chu, J.P., Goldblum, P., Floyd, R., Bongar, B., 2010. The cultural theory and model of suicide. *Appl. Prev. Psychol.* 14, 25–40.
- Cotton, S.M., Mensink, J., Filia, K., et al., 2021. The psychometric characteristics of the Kessler Psychological Distress Scale (K6) in help-seeking youth: what do you miss when using it as an outcome measure? *Psychiatr. Res.* 305, 114182.
- Cox, J.L., Holden, J.M., Sagovsky, R., 1987. Detection of postnatal depression: development of the 10-item Edinburgh postnatal depression Scale. *Br. J. Psychiatry* 150 (6), 782–786.
- Kleinman, A., Good, B. (Eds.), 1985. *Culture and Depression: Studies in the Anthropology and Cross-Cultural Psychiatry of Affect and Disorder*. University of California Press.
- Kubota, C., Inada, T., Nakamura, Y., Shiino, T., Ando, M., Aleksic, B., et al., 2018. Stable factor structure of the Edinburgh Postnatal Depression Scale during the whole peripartum period: results from a Japanese prospective cohort study. *Sci. Rep.* 8 (1), 17659.
- Kubota, C., Okada, T., Aleksic, B., Nakamura, Y., Kunimoto, S., Morikawa, M., Shiino, T., Tamaji, A., Ohoka, H., Banno, N., Morita, T., 2014. Factor structure of the Japanese version of the Edinburgh Postnatal Depression Scale in the postpartum period. *PLoS one* 9 (8), e103941.
- Levis, B., Benedetti, A., Thombs, B.D., DEPRESSion Screening Data (DEPRESSD) Collaboration, 2019. Accuracy of Patient Health Questionnaire-9 (PHQ-9) for screening to detect major depression: individual participant data meta-analysis. *BMJ* 365, 11476.
- Levis, B., Negeri, Z., Sun, Y., Benedetti, A., Thombs, B.D., DEPRESSion Screening Data (DEPRESSD) EPDS Group, 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.
- Okano, T.M.M., Masuji, F., Tamaki, R., Nomura, J., Miyaoko, H., 1996. Validation and reliability of Japanese version of the EPDS. *Archives of Psychiatric Diagnostics and Clinical Evaluation* 7, 525–533.
- Qiu, X., Wu, Y., Sun, Y., et al., 2023. Individual participant data meta-analysis to compare EPDS accuracy to detect major depression with and without the self-harm item. *Sci. Rep.* 13, 4026.
- Ryder, A.G., Yang, J., Zhu, X., et al., 2008. The cultural shaping of depression: somatic symptoms in China, psychological symptoms in North America? *J. Abnorm. Psychol.* 117, 300–313.
- Stefana, A., Cena, L., Trainini, A., Palumbo, G., Gigantesco, A., Mirabella, F., 2024. Screening for prenatal and postnatal maternal depression: comparative performance of the Edinburgh Postnatal Depression Scale and Patient Health Questionnaire-9. *Ann. Istituto Super.* 60 (1).
- Stefana, A., Langfus, J.A., Palumbo, G., Cena, L., Trainini, A., Gigantesco, A., Mirabella, F., 2023. Comparing the factor structures and reliabilities of the EPDS and the PHQ-9 for screening antepartum and postpartum depression: a multigroup confirmatory factor analysis. *Arch. Womens Ment. Health* 26 (5), 659–668.
- Yang, L.H., Kleinman, A., Link, B.G., Phelan, J.C., Lee, S., Good, B., 2007. Culture and stigma: adding moral experience to stigma theory. *Soc. Sci. Med.* 64, 1524–1535.

Alberto Stefana\*

Department of Brain and Behavioral Sciences, University of Pavia, Pavia, Italy

Loredana Cena, Alice Trainini

Department of Clinical and Experimental Sciences, Section of Neuroscience, Observatory of Perinatal Clinical Psychology, University of Brescia, Brescia, Italy

Gabriella Palumbo, Antonella Gigantesco<sup>1</sup>, Fiorino Mirabella<sup>1</sup>  
Center for Behavioural Sciences and Mental Health, National Institute of Health, Rome, Italy

\* Corresponding author. Department of Brain and Behavioral Sciences, University of Pavia, Via Forlanini 6, 27100, Pavia, Italy.  
E-mail address: [alberto.stefana@unipv.it](mailto:alberto.stefana@unipv.it) (A. Stefana).

<sup>1</sup> These authors equally contributed to this work and should be considered co-last authors.