INTEGRATIVE REVIEW

DOI: 10.15343/0104-7809.202549e17312025I
ISSN: 1980-3990

Validity evidence of the Edinburgh Postnatal Depression Scale: an integrative review

Renata Costa da Silva[1] **Nayara de Jesus Oliveira**[1] **Joise Magarão Queiroz Silva**[1] **Renata Fernandes do Nascimento Rosa**[1] **Patrícia Figueiredo Marques**[2] **Katia Santana Freitas**[3] **Edméia de Almeida Cardoso Coelho**[1]

[1]Escola de Enfermagem, Universidade Federal da Bahia – UFBA. Salvador/BA, Brasil.
[2]Universidade Federal do Recôncavo – UFRB. Santo Antônio de Jesus/BA, Brasil.
[3]Departamento de Saúde, Universidade Estadual de Feira de Santana – UEFS. Feira de Santana/BA, Brasil.
E-mail: renatalanai@yahoo.com.br

Graphical Abstract

Highlights

- *Robust evidence confirms the validity of the EPDS across different cultural contexts.*
- *The scale demonstrates strong content, construct, criterion, and predictive validity.*
- *Studies show responsiveness to change in clinical interventions.*

Abstract

The prevalence of common mental disorders, including depression, post-traumatic stress disorder, obsessive-compulsive disorder, and a variety of anxiety disorders, represents a major global burden of disease in both high- and low-income countries. This study aimed to identify validity evidence of the Edinburgh Postnatal Depression Scale (EPDS) tested in puerperal women. An integrative review was conducted using the PICO strategy to construct the research question: What validity evidence is available for the Edinburgh Postnatal Depression Scale in puerperal women? The databases searched were PubMed, CINAHL, Embase, Scielo Org., and SCOPUS. Screening was carried out in Rayyan, following the PRISMA 2020 flowchart. Searches were conducted from october to november 2022. Based on inclusion criteria, 17 studies were selected for analytical synthesis. The Edinburgh Postnatal Depression Scale is widely recognized as the gold standard for detecting postpartum depression in women worldwide. This instrument allows the identification of depression risk through different cutoff scores. Understanding how the EPDS functions and its limitations is essential for its effective use across diverse contexts, ensuring accurate screening and early identification of cases requiring specialized follow-up.

Keywords: Postpartum Depression. Mental Health. Women. Validation Study. Psychometrics.

Associate Editor: Edison Barbieri
Mundo Saúde. 2025,49:e17312025
O Mundo da Saúde, São Paulo, SP, Brasil.
https://revistamundodasaude.emnuvens.com.br

Received: 10 march 2025.
Acepted: 01 september 2025.
Published: 30 september 2025.

INTRODUCTION

The perinatal period, defined as the time from conception to one year after birth, is a high-risk stage for the development of mental health disorders. Common mental disorders, including depressive disorders, post-traumatic stress disorder (PTSD), obsessive-compulsive disorder (OCD), and anxiety disorders during the perinatal period, account for a significant proportion of the global disease burden in both high-income and low- and middle-income countries[1,2,3].

Prevalence estimates suggest that up to 17% of the postnatal population experienced depression, with 2 in every 1000 women requiring hospitalization for severe mental illness during the early postpartum period. Perinatal anxiety, bipolar affective disorder (BAD), and PTSD also contribute substantially to maternal morbidity and mortality[4].

Postpartum depression is associated with maternal emotional and psychological recovery after childbirth, as well as other important aspects such as sleep, fatigue, mother–infant bonding, psychosocial support, marital relationships, family dysfunction, and social relationships[5]. Suicide is the leading cause of maternal mortality in the United Kingdom and the second leading cause in the United States[6].

Mental disorders during this critical period are not only associated with increased maternal mortality, suicide, and self-harm[2]; data have shown increased risks of adverse neonatal outcomes, such as fetal growth restriction, postpartum hemorrhage, placental abruption, and stillbirths[7]. Infants exposed to prenatal distress and children with continuous exposure may also face challenges in physical and psychosocial development, including stunting, diarrheal infections, and impaired cognitive development[1,3]. With an estimated one-fourth of children exposed to maternal mental health disorders, timely identification and treatment during the perinatal period are paramount[8].

Postpartum depression (PPD) is a complex construct to assess in practice. Allowing the largest possible number of healthcare professionals to conduct an initial timely evaluation of maternal mental health — while reserving detailed psychiatric assessments only for cases suggestive of PPD — is an appealing approach. Likewise, applied research contexts require rapid and valid instruments[9].

In the late 1980s, Cox *et al.* (1987)[10] argued that an appropriate instrument was needed to assess depressive symptoms after childbirth, since tools available for assessing depression in general populations placed excessive emphasis on somatic symptoms, which could be due to normal physiological adaptations associated with pregnancy. To address this limitation, the authors proposed the Edinburgh Postnatal Depression Scale (EPDS), a simple and widely accepted 10-item screening tool that is easy to complete and does not require specialized knowledge.

According to the World Health Organization (WHO), when a condition is serious, prevalent, and treatable, screening programs should be implemented to identify individuals at high risk[11]. Numerous screening tools have been developed to aid in the early identification and stratification of postpartum depression. Before being applied in clinical practice, validation of such tools in local contexts is essential to ensure their appropriateness for the population and to establish context-specific cutoff points. However, uncertainty regarding the timing of implementation and appropriate thresholds, combined with low acceptability and inconsistent use, means that many barriers to the detection of perinatal mental disorders remain[2,12].

Decision-making is supported by measurement instruments; therefore, ensuring validity evidence of the instruments involved in this assessment process is crucial[12,13]. The data generated by instruments, as well as psychometric studies, are of utmost importance, as recommendations are based on this type of evidence. Thus, advancing with detailed analyses of validity evidence for these instruments addresses a global and growing need for quality measures and aligns with the macro–micro perspective of the pathways that underpin such evidence, in light of psychometric science[13].

Accordingly, this integrative review was developed to identify the validity evidence of the Edinburgh Postnatal Depression Scale (EPDS) tested in puerperal women. This review was deemed necessary to provide an updated synthesis of the psychometric evidence of the EPDS across different contexts.

METHODOLOGY

This study is an integrative review on the validity evidence of the Edinburgh Postnatal Depression Scale (EPDS) translated and adapted for puerperal women. The methodological framework proposed by Whittemore and Knafl (2005)[14] was adopted. The review protocol was registered in PROSPERO under number CRD42024568821.

The guiding question of this integrative review was: "What validity evidence is available for the EPDS in puerperal women?" From this question, the PICO strategy was constructed as follows: P (population/patient) – puerperal women; I (intervention/expo-

Mundo Saúde. 2025,49:e17312025
DOI: 10.15343/0104-7809.202549e17312025I

sure) – validity evidence study of the translated and adapted EPDS assessment instrument; C (comparator group) – not applicable; O (outcome) – high level of evidence.

Once the PICO framework was established, the research question was defined: What levels of validity evidence are available for the EPDS, translated and adapted for puerperal women?

Searches were conducted between October and November 2023, via the CAPES Journals Portal, with institutional access through the Federal University of Bahia and the University of São Paulo, in the following databases: PubMed, CINAHL, Embase, Scielo Org., and SCOPUS.

Inclusion criteria were: articles with samples composed exclusively of puerperal women; studies presenting validity evidence or psychometric properties; no language or time restrictions; full-text and freely available. Exclusion criteria were: review articles, book chapters, dissertations, and theses. The search strategies were developed by the librarian of the School of Nursing, University of São Paulo, and JBI-Brazil. The core search algorithm was structured as follows: ((“NAME OF THE INSTRUMENT”[EPDS])) AND (psychometr* OR valid*)) AND (Postpartum OR Postpartum Period OR “postnatal”)).

Table 1 – Search strategy by database, authorship, 2023.

Data bases	Search strategy	Link	Results
PubMed	((((“Psychiatric Status Rating Scales”[MeSHTerms]) OR “Edinburgh Postnatal Depression Scale”[Title]) AND psychometrics[MeSHTerms]) AND “PostpartumPeriod”[MeSHTerms]) OR postnatal[Title]	https://www-ncbi-nlm-nih.ez10.periodicos.capes.gov.br/pmc?term=((((%E2%80%9CPsychiatric%20Status%20Rating%20Scales%E2%80%9D%5BMeSH%20Terms%5D)%20OR%20%E2%80%9CEdinburgh%20Postnatal%20Depression%20Scale%E2%80%9D%5BTitle%5D)%20AND%20psychometrics%5BMeSH%20Terms%5D)%20AND%20%E2%80%9CPostpartum%20Period%E2%80%9D%5BMeSH%20Terms%5D)%20OR%20postnatal%5BTitle%5D	494
CINAHL	*“Psychiatric Status Rating Scales” OR “Edinburgh Postnatal Depression Scale” AND psychometrics AND “PostpartumPeriod” OR postnatal*	https://www.ebsco. ez10.periodicos.capes.gov.br /pt/search?search=%E2%80%9CPsychiatric+Status+Rating+Scales%E2%80%9D+OR+%E2%80%9CEdinburgh+Postnatal+Depression+Scale%E2%80%9D+AND+psychometrics+AND+%E2%80%9CPostpartum+Period%E2%80%9D+OR+postnatal	14
Embase	(‘psychiatric status rating scales’/exp OR ‘psychiatric status rating scales’ OR ‘edinburgh postnatal depression scale’/exp OR ‘edinburgh postnatal depression scale’) AND psychometrics:ti,ab,kw AND (‘postpartumperiod’:ti,ab,kw OR postnatal:ti,ab,kw)	https://www-embase.ez10.periodicos.capes.gov.br/#advancedSearch/resultspage/history.2/page.1/25.items/orderby.date/source.	18
Scielo.org	Edinburgh Postnatal Depression Scale AND validation	https://search.scielo.org/?q=Escala+de+depress%C3%A3o+p%C3%B3s-parto+de+Edimburgo&lang=pt&count=15&from=0&output=site&sort=&format=summary&fb=&page=1&q=Escala+de+depress%C3%A3o+p%C3%B3s-parto+de+Edimburgo+AND+valida%C3%A7%C3%A3o&lang=pt&page=1	08
SCOPUS	TITLE-ABS-KEY (“Psychiatric Status Rating Scales” OR “Edinburgh Postnatal Depression Scale”) AND TITLE-ABS-KEY (psychometrics) AND TITLE-ABS-KEY (“PostpartumPeriod” OR postnatal)	https://www-scopus.ez10.periodicos.capes.gov.br/results/results.uri?sort=plf-f&src=s&sid=98f81c0a6e-538589b9e71041dc8138d8&sot=a&sdt=a&sl=178&s=TITLE-ABS-KEY+%28%22Psychiatric+Status+Rating+Scales%22+OR+%22Edinburgh+Postnatal+Depression+Scale%22%29+AND+TITLE-ABS-KEY+%28psychometrics%29+AND+TITLE-ABS-KEY%28%22Postpartum+Period%22+OR+postnatal%29&origin=searchadvanced&editSaveSearch=&txGid=72293935e8b524fd8104e054501daaed	24

The stages of study selection and data extraction were conducted by two independent reviewers, with no cases of disagreement. The studies retrieved were exported from the databases into specific files and uploaded to Rayyan Qatar Computing Research Institute (Rayyan QCRI), available at https://rayyan.qcri.org[15], online version, following

the PRISMA 2020 flowchart[16]. Subsequently, the studies were analyzed according to eligibility criteria, and inclusion was defined for the analytical synthesis.

The methodological quality assessment of the studies was performed in two stages. In Stage 1, an adapted protocol comprising seven criteria based on the COSMIN 2018 Risk of Bias checklist[17] was used. Each study was analyzed according to the following protocol criteria, C1 – Is there a clear definition of the construct to be measured? C2 – Is the origin of the construct clear? (Is there a theory, a conceptual model, or disease framework used, or was a clear rationale presented to define the construct to be measured?) C3 – Was a clear definition of the context in which the instrument will be used presented? C4 – Were the procedures for validity evidence conducted in a population representative of the target population for which the instrument is proposed? C5 – Was the original development study mentioned, along with clarification of the population and context for which it was developed? C6 – Were the procedures for validating the instrument presented? C7 – Were the procedures for analyzing the instrument's reliability presented? Each criterion was rated as Excellent, Good, Fair, or Poor.

All studies were recorded in a Microsoft Excel® spreadsheet, analyzed according to the criteria (C1–C7) above, and finally classified as: Adequate (clearly described), Acceptable (partially described), Doubtful (unclear/not described), and Not applicable (not related to the study objective).

Nevertheless, all studies meeting eligibility criteria were retained regardless of methodological quality, in order to minimize bias in conducting this analysis, since the objective of this review was to investigate the validity evidence of the measurement instrument in question. This process complemented the analysis of the sufficiency of such evidence in each study, leading to the subsequent stage.

In Stage 2 of the evaluation of validity evidence studies, the criteria for assessing the sufficiency of available evidence were defined[17].

From the final studies included in the analytical synthesis, data were extracted regarding the instrument, target population, stages of validity evidence, and psychometric properties assessed. Data from the studies were extracted and recorded in a Microsoft Excel® spreadsheet developed by the lead author. Results were presented in tables and figures, according to the most appropriate representation, and discussed in light of the literature.

The ethical aspects of the study were respected in accordance with the principles of good research practices, ensuring the integrity and transparency of the process.

RESULTS

The selection flow was followed, and 558 potentially relevant studies were identified, of which 55 were excluded due to duplication. A total of 503 articles proceeded to title and abstract screening. Of these, 461 were excluded, leaving 42 for full-text reading and eligibility analysis. After complete assessment of the articles, 25 studies were excluded for not addressing the proposed topic within their samples, resulting in 17 studies included in the analytical synthesis.

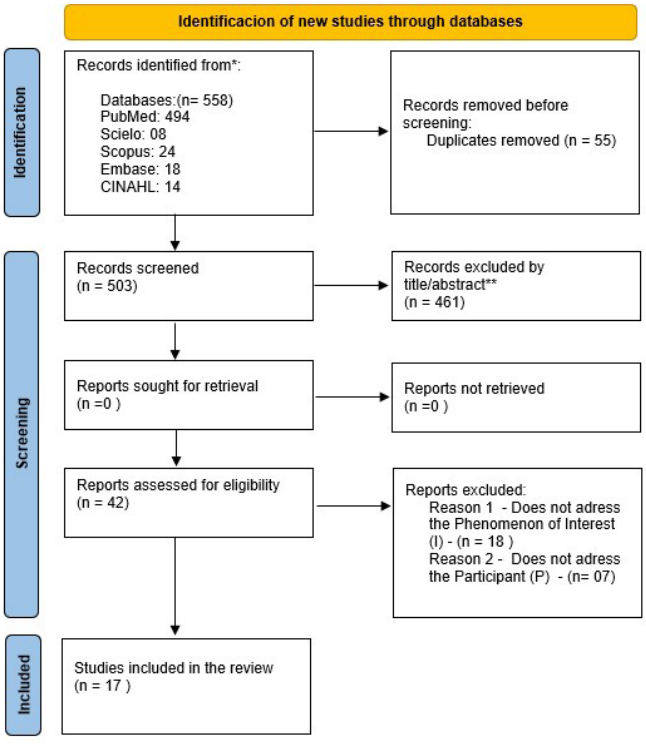

Figure 1 – Flowchart of the integrative review process adapted from the PRISMA Statement, 2024.

The analyzed data were organized into two stages: the first involved the methodological analysis of each study, and the second consisted of the analysis of the steps undertaken in each study to gather validity evidence of the instrument in question.

The methodological quality stage classifies each study according to the procedures addressed. Although more general, this analysis already contributes to the second stage, in which the steps of each proposed evidence were examined in greater detail. Figure 1 provides an illustrative classification of the studies according to the seven criteria proposed in the protocol adopted for this study.

Table 2 – List of selected studies according to methodological quality based on the COSMIN protocol, authorship, 2024.

Autores	C1	C2	C3	C4	C5	C6	C7
Skodová *et al.* 2021[14]	AD	AD	AD	AC	AD	AD	AD
Boran *et al.* 2020[15]	AD	AD	AD	AD	AD	AD	AD
Greena *et al.* 2018[16]	AD	AD	AC	AC	DV	AD	AD
Smith-Nielsen *et al.* 2018[17]	AD	AD	AC	AD	AD	AD	AD
Albuquerque *et al.* 2018[18]	AD	AD	AC	AC	DV	AD	AD
Syam *et al.* 2021[19]	AD	AD	AC	AC	DV	AD	DV
Hartley *et al.* 2014[20]	AD	AD	AC	AD	AD	AD	AD
Toreki *et al.* 2014[21]	AD	AD	AD	AD	DV	AC	AD
Kheirabalde *et al.* 2012[22]	AD	AD	AD	AD	AD	AD	AD
Lee King *et al.* 2012[23]	AD	AD	AC	AD	AD	AD	DV
Reichenheim *et al.* 2011[24]	AD	AD	AD	AD	AC	AD	DV
Montazeri *et al.* 2007[25]	AD	AD	AD	AD	AD	AD	AD
Santos *et al.* 2007[26]	AD	AD	AD	AD	DV	AC	AD
Santos *et al.* 2007[27]	AD	AD	AD	AD	AD	AC	DV
Jardri *et al.* 2006[28]	AD	AD	AC	AC	AD	AC	DV
Werrett *et al.* 2006[29]	AD	AD	AD	AD	AD	AC	AD
Clifford *et al.* 1999[30]	AD	AD	AD	AD	AD	AC	DV

AD = Adequate; **AC** = Acceptable; **DV** = Doubtful
Source: Mokkink *et al.* (2018)[17].

The methodological limitations of the studies, with only 17% classified as "Adequate" according to COSMIN, directly affect the validity and applicability of the EPDS, with implications for the accuracy of results and a greater risk of bias due to failure to follow methodological criteria[17].

The construct definition was consistently addressed, being rated as "Adequate" and "Acceptable" in 100% of the studies. In 58% of the studies, the testing context of the instrument in question was clearly reported, and in 64% of the studies, a representative population for which the use of the instrument was proposed was employed, with reference to the original instrument and clarification of the population for which it had been developed. Regarding psychometric procedures, all studies presented descriptions of validity evidence of the instrument; indeed, this was an eligibility criterion for inclusion in the analytical synthesis, with 64% presenting procedures for reliability analysis of the instrument.

Regardless of the classification achieved in this stage of analysis, no study was excluded, since this stage was not intended as a detailed analysis of validity evidence. Thus, all 17 studies proceeded to the next stage of analysis. Validity is the central concept of psychometrics and is related to the interpretability of test scores, as indicated by AERA, APA, and NCME (2014)[18], with the quality of a test being directly related to its validity evidence.

Given the diversity and complexity of each type of source of validity evidence and the combination of various qualitative and quantitative techniques applied at each stage — which have been reported in the literature over the years — it becomes important to provide clarity to the contemporary model, enabling accurate identification in analyses of the steps undertaken in each study and the results achieved according to the objectives established herein.

It was therefore decided to explicitly establish the criteria and evaluation indicators for each of the sources of validity evidence analyzed in the retrieved studies. Table 3 presents these concepts in detail.

Table 3 – Characteristics of the publications, authorship, 2024.

Study	Objective	Sample	Validity Evidence	Dimensionality	Location
Skodová *et al*. (2021)[14]	Examine the factor structure and psychometric properties of the Slovak version of the EPDS	577	EFA, CFA, Alfa Cronbach	Three-dimensional	Slovakia
Boran *et al*. (2020)[15]	Determine the factor structure of the EPDS using evidence-based analytical techniques	1.614	EFA, CFA	Unidimensional	Turkey
Greena *et al*. (2018)[16]	Validate the EPDS and PHQ-9 in rural Kenya	193	Content validity, Cronbach's Alpha, Test–retest, Sensitivity, Specificity	Not assessed	Kenya
Smith-Nielsen *et al*. (2018)[17]	Validate the Danish EPDS against a diagnosis of depression according to DSM-5 and ICD-10	324	EFA, Sensitivity, Specificity, PPV, NPV, ROC curve	Three-dimensional	Denmark
Albuquerque *et al*. (2018)[18]	Verify and compare the metrics of two different 6-item EPDS subscales	3.891	EFA, External validity, Cronbach's Alpha, Sensitivity, Specificity, PPV, NPV, ROC curve	Bidimensional	Amazon and Northeastern Brazil
Syam *et al*. (2021)[19]	Confirm the factor model of the Indonesian EPDS version and test factor consistency in puerperal women	616	EFA, CFA	Three-dimensional	Indonesia
Hartley *et al*. (2014)[20]	Analyze the factor structure of the EPDS among Hispanic mothers in the United States	220	EFA, Cronbach's Alpha	Bidimensional	USA
Toreki *et al*. (2014)[21]	Assess the validity of the EPDS for postpartum depression screening in Hungary	266	EFA, Cross-cultural adaptation, Cronbach's Alpha, Test–retest, Sensitivity, Specificity, ROC curve	Bidimensional	Hungary
Kheirabalde *et al*. (2012)[22]	Assess the psychometric properties and diagnostic accuracy of the EPDS in a sample of Iranian women	2.762	External validity, Cronbach's Alpha, Sensitivity, Specificity, ROC curve, Cross-cultural adaptation	Bidimensional	Iran
Lee King *et al*. (2012)[23]	Assess the underlying structure of the EPDS using a model comparison approach in confirmatory factor analysis	169	CFA	Three-dimensional	USA
Reichenheim *et al*. (2011)[24]	Examine whether raw scores adequately represent factor scores based on latent models	811	EFA, CFA, TRI	Three-dimensional	Rio de Janeiro, Brazil
Montazeri *et al*. (2007)[25]	Translate and test the reliability and validity of the EPDS in Iran	100	Cross-cultural adaptation, Test–retest, EFA, Cronbach's Alpha	Three-dimensional	Iran
Santos *et al*. (2007)[26]	Assess the EPDS for screening and diagnosis of postpartum depression	378	External validity, Sensitivity, Specificity, ROC curve	Not assessed	Pelotas, Brazil
Santos *et al*. (2007)[27]	Compare the accuracy of two instruments for postpartum depression screening	378	External validity, Sensitivity, Specificity, ROC curve	Not assessed	Pelotas, Brazil

to be continued...

...continuation - Table 3.

Study	Objective	Sample	Validity Evidence	Dimensionality	Location
Jardri *et al.* (2006)[28]	Validate the use of the EPDS in the early postpartum period and identify markers for risk of postnatal depression	815	External validity, Sensitivity, Specificity, PPV, NPV, ROC curve	Not assessed	France
Werrett *et al.* (2006)[29]	Validate a Punjabi translation of the EPDS	24	Cronbach's Alpha, Sensitivity, Specificity, PPV, Cross-cultural adaptation	Not assessed	India
Clifford *et al.* (1999)[30]	Develop and conduct a preliminary validation of the EPDS for use in the Punjabi-speaking community	98	Cross-cultural adaptation, EFA	Unidimensional	India

EPDS: Edinburgh Postnatal Depression Scale; PPD: Postpartum Depression; EFA: Exploratory Factor Analysis; CFA: Confirmatory Factor Analysis; PPV: Positive Predictive Value; NPV: Negative Predictive Value; ROC: Receiver Operating Characteristic; IRT: Item Response Theory; Cross-cultural adaptation: Adaptation of the instrument to cultural/linguistic contexto.

Low quality of source data compromises the reliability of any subsequent analysis. If the studies serving as the basis for evaluation are not robust, conclusions about the EPDS's validity—its ability to measure what it is intended to measure—become less reliable.

DISCUSSION

Since its conception, the EPDS has been adapted for use in several countries and has become the most widely used instrument for initial screening of postpartum depression (PPD)[19]. The EPDS has been extensively examined, and numerous studies have evaluated its psychometric properties. Several studies focused on its dimensional structure, with at least fourteen comprising sample sizes above 150 individuals[20-34], which allows for the use of robust multivariate techniques.

Although Cox *et al.* (1987)[10] originally proposed the EPDS as a unidimensional measurement tool and this has been supported by some authors[18,22] the majority of factor analyses have shown that the EPDS is better defined through multifactor structures, whether by two[25-28] or three factors[20,21,24,29,30,31].

The methodological quality analysis revealed critical issues, particularly regarding the definition of the construct to be measured and its origin. What becomes evident is the gap between the recommendations and actual practices concerning the concept of validity evidence, as recommended since 1999 by the Standards[35] and consolidated by the same institutions in the 2014 Standards.

Cross-cultural adaptation of the EPDS is a robust field of study and demonstrates that cultural differences can significantly affect the instrument's validity. A more detailed analysis reveals important variations in its psychometric properties, such as cutoff score, factor structure, and even the way depression is expressed[13,36].

Content validity was addressed in only two studies, which were limited to translation of the instrument. Content validity is the assessment of how accurately and comprehensively the elements of a construct or attributes representative of a target population in a specific context are measured[37]. This type of validation is carried out by a committee of experts in the research topic and methodology used, including cross-cultural adaptation, content validity index, and agreement index[38]. Cross-cultural validity was not adequately evaluated in any of the included studies.

Cross-cultural adaptation is of paramount importance in validation studies, especially when a research instrument such as a questionnaire or scale is translated from one language into another for use in a new cultural context. It ensures that the instrument is not only linguistically understandable but also culturally equivalent and relevant to the target population. This process avoids cultural bias, guarantees validity and reliability, maintains conceptual equivalence, and improves acceptance and comprehension. Validating an instrument in a new cultural context requires a rigorous process that ensures multiple forms of equivalence. Ignoring this step may lead to inaccurate, invalid, and, in some cases, even harmful research results[38,39].

No study described the response process. Questioning respondents from the intended population about their performance strategies or responses to specific items could yield evidence that enriches construct definition[40].

The types of validity evidence most frequently explored by authors were analyses of internal structure, followed by relationships with other variables.

According to the criteria established for this study, the most recurrent procedures were factor analyses (exploratory and confirmatory) and reliability analyses through internal consistency, using Cronbach's alpha and test–retest methods.

It is worth highlighting that only two studies described the criteria for defining/calculating the study sample. Regarding evidence of relations with other variables, the most predominant were predictive analyses using ROC curves, to identify the sensitivity and specificity of the instrument, as well as the cutoff score for the specific population and version studied.

Another commonly used criterion was convergent validity, assessed by correlations with another instrument and score comparisons between groups. The sensitivity and specificity of the EPDS depend on the cutoff score applied[41]. A detailed analysis of the 17 studies concluded that a cutoff score of 11 or higher maximizes the combined sensitivity and specificity of the EPDS.

In this context, the widespread use of Cronbach's alpha for internal consistency analysis stands out, often referred to as reliability, even though contemporary psychometrics points out numerous limitations in its use[42] limitations that continue to be ignored in practice, despite long-standing concerns.

Cronbach's alpha relies on a strict assumption called tau-equivalence, which presumes that all items on the scale measure the same construct with equal strength (i.e., identical factor loadings). This assumption rarely holds in practice. As a result, Cronbach's alpha tends to underestimate true internal consistency when items have different loadings or when the factor structure is more complex[42].

Furthermore, the alpha value is highly influenced by the number of items on the scale. An instrument with more items can artificially present a higher alpha, even if inter-item correlations are low. This can lead to erroneous conclusions about scale quality[13,42].

To overcome these limitations, methods that do not depend on the tau-equivalence assumption and provide a more accurate assessment of internal consistency are recommended, such as McDonald's Omega coefficient and Item Response Theory (IRT)[13].

No study aimed to analyze validity evidence related to testing consequences. The importance of studies addressing the effects of psychometric test use and the elements contributing to individual and social consequences is recognized; however, there remains debate in the literature about this source of evidence, and studies with this focus are scarce[43].

Although the EPDS has broad evidence of content and construct validity, the literature still lacks studies exploring validity related to the consequences of its application. It is crucial that future research evaluate the impact of EPDS use on clinical decision-making and maternal and neonatal outcomes. Investigating whether routine application of the scale increases referral rates for treatment and consequently improves maternal and infant health outcomes is an essential step in strengthening the instrument's clinical relevance[1].

The EPDS is a valuable tool, but its application must be guided by validation evidence in each cultural context. For healthcare professionals, the following recommendations can optimize its use in clinical practice: adjustment of cutoff scores, contextualization of results, consideration of the sociocultural context, and use of the EPDS as a starting point for action. By following these recommendations, healthcare professionals can use the EPDS as an effective yet mindful tool that truly contributes to early diagnosis and improved maternal and infant health outcomes.

CONCLUSION

The Edinburgh Postnatal Depression Scale (EPDS) is the most validated and widely used tool for postpartum depression worldwide. It is a self-report questionnaire validated only for screening PPD and includes different validated cutoff scores.

Despite some methodological shortcomings, such as inadequate sample sizes, inappropriate multivariate models, and/or failure to properly model the categorical nature of items, the reviewed literature shows more consistency than otherwise.

It is important to pay attention to the different versions available in the literature of the same instrument, sometimes in shorter versions with fewer items, sometimes applied during pregnancy, as well as its mode of administration. Understanding the instrument's characteristics can strongly support decision-making for its use in the target population, ensuring comprehension of the steps undertaken and results achieved.

Implications for Future Research

The literature review on EPDS validation demonstrates progress in the field but also highlights

critical gaps that must be addressed to ensure the EPDS remains a relevant and clinically useful screening tool. The following research directions are suggested to advance knowledge on cross-cultural validation and psychometrics in the context of postpartum depression: more rigorous methodological approaches, IRT, factor structure analysis in different cultures, and investigation of testing consequences. By addressing these issues, researchers will be able to provide strong evidence on the clinical value and utility of the EPDS, reinforcing its role as an indispensable tool in maternal health.

CRediT author statement

Conceptualization: Silva RC; Coelho EAC; Freitas KS. Methodology: Silva RC; Coelho EAC; Freitas KS; Marques PF. Validation: Silva RC; Coelho EAC; Freitas KS. Statistical analysis: Marques PF; Oliveira NJ; Silva JMQ; Rosa RFN. Formal analysis: Oliveira NJ; Silva JMQ; Rosa RFN. Investigation: Silva RC; Oliveira NJ; Silva JMQ; Rosa RFN. Resources: Silva RC; Coelho EAC; Freitas KS; Marques PF. Writing – original draft preparation: Silva RC; Coelho EAC; Freitas KS. Writing – review and editing: Oliveira NJ; Silva JMQ; Rosa RFN. Visualization: Marques PF; Oliveira NJ; Silva JMQ; Rosa RFN. Supervision: Silva RC; Coelho EAC; Freitas KS. Project administration: Silva RC; Coelho EAC; Freitas KS.

All authors have read and agreed to the published version of the manuscript.

Funding

The authors did not receive funding for the development of the present research.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the workreported in this paper.

REFERENCES

1. Sultan P, Ando K, Elkhateb R, et al. Assessment of Patient-Reported Outcome Measures for Maternal Postpartum Depression Using the Consensus-Based Standards for the Selection of Health Measurement Instruments Guideline: A Systematic Review. JAMA Netw Open [Internet]. 2022 [citado 2025 ago 29];5(6):e2214885. Available from: https://doi.org/10.1001/jamanetworkopen.2022.14885
2. Howard LM, Khalifeh H. Perinatal mental health: a review of progress and challenges. World Psychiatry. 2020; 19(3):313-327. doi:10.1002/wps.20769
3. Gelaye B, et al. Epidemiology of maternal depression, risk factors, and child outcomes in low-income and middle-income countries. Lancet Psychiatry. 2016;3(10):973-982. doi:10.1016/S2215-0366(16)30284-X.
4. Dennis CL, et al. Prevalence of antenatal and postnatal anxiety: systematic review and meta-analysis. Br J Psychiatry. 2017;210(5):315-323. doi:10.1192/bjp.bp.116.187179
5. Maternal, Newborn and Infant Clinical Outcome Review Programme. Saving Lives, Improving Mothers' Care: Lessons Learned to Inform Maternity Care From the UK and Ireland Confidential Enquiries Into Maternal Deaths and Morbidity 2017-19. Mothers and Babies: Reducing Risk Through Audits and Confidential Enquiries Across the UK; 2021. Available from: https://www.npeu.ox.ac.uk/assets/downloads/mbrrace-uk/reports/maternal-report-2021/MBRRACE-UK_Maternal_Report_2021_-FINAL-_WEB_VERSION.pdf.
6. Sultan P, Ando K, Elkhateb R, et al. Assessment of Patient-ReportedOutcomeMeasures for Maternal Postpartum Depression Usingthe Consensus-Based Standards for theSelection of Health Measurement Instruments Guideline: A Systematic Review. JAMA Netw Open. 2022;5(6):E2214885. doi:10.1001/jamanetworkopen.2022.14885
7. McAllister-Williams RH, Arango C, Blier P, et al. The identification, assessment and management of difficult-to-treat depression: an international consensus statement. J Affect Disord. 2020;267:264-282. doi:10.1016/j.jad.2020.02.023.
8. Abel KM, et al. Prevalence of maternal mental illness among children and adolescents in the UK between 2005 and 2017: a national retrospective cohort analysis. Lancet Public Health. 2019;4(6):291-300. doi:10.1016/S2468-2667(19)30059-3.
9. Montazeri A, Torkan B, Omidvari S. The Edinburgh Postnatal Depression Scale (EPDS): translation and validationstudy of theIranianversion. BMC Psychiatry. 2007; 7: 11. doi:10.1186/1471-244X-7-11.
10. Cox JL, Holden JM, Sagovsky R. Detection of postnatal depression. Development of the 10-item Edinburgh postnatal depression scale. Br J Psychiatry. 1987;150(6):782–6.doi:10.1192/bjp.150.6.782.
11. World Health Organization. WHO recommendations on maternal and newborn care for a positive postnatal experience: web annexes. Geneva: World Health Organization; 2022. Available from: https://apps.who.int/iris/handle/10665/352612.
12. Sambrook MS, Cairns L, Pullen LSW, Opondo C, Fellmeth G, Alderdice F. Validated tools to identify common mental disorders in the perinatal period: A systematic review of systematic reviews. J Affect Disord.2022;298:634-643.doi:10.1016/j.jad.2021.11.011.
13. Ferreira CR, Sanches MFTV, Matos GPNBR, RebustiniF, Domingues MARC. Instruments for evaluating social support networks for Brazilian elderlypeople: a scoping review. Concilium. 2024; 24(7): 432–451.doi:10.53660/CLM-3223-24G24.
14. Whittemore R, Knafl K. The integrative review: update methodology. J Adv Nurs. 2005;52(5):546-553. doi:10.1111/j.1365-2648.2005.03621.x.
15. Mourad Ouzzani, Hossam Hammady, Zbys Fedorowicz e Ahmed Elmagarmid. Rayyan — um aplicativo web e móvel para revisões sistemáticas. Revisões Sistemáticas (2016) 5:210, doi: 10.1186/s13643-016-0384-4.
16. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71. doi:10.1136/bmj.n71.
17. Mokkink LB et al. COSMIN Risk of Bias checklist for systematic reviews of Patient Reported Outcome Measures. Qual. Life Res. 2018; 27(5): 1171-79. doi:10.1007/s11136-017-1765-4.
18. American Educational Research Association, American Psychological Association, National Council on Measurement in Education. Standards for educational and psychological testing. Washington, DC: Author; 1999.
19. Smith-Nielsen J, Matthey S, Lange T, et al. Validation of the Edinburgh Postnatal Depression Scale against both DSM-5 and ICD-10 diagnostic criteria for depression. BMC Psychiatry. 2018;18:393. doi:10.1186/s12888-018-1965-7.
20. Boran P, Waqas A, Aşkan ÖÖ et al. Screening of postpartum depression among new mothers in Istanbul: a psychometricevaluation of

theTurkish Edinburgh Postnatal Depression Scale. BMC Res Notes. 2020;13:355.doi:10.1186/s13104-020-05196-x.
21. Greena EP, Tuli H, Kwobah E, Menya D, Chesire I, Schmidt C. Developing and validating a perinatal depression screening tool in Kenya blending Western criteriawith local idioms: A mixedmethodsstudy. J Affect Disord. 2018; 228:49-59.doi:10.1016/j.jad.2017.11.027.
22. Smith-Nielsen J, Matthey S, Lange T, et al. Validation of the Edinburgh Postnatal Depression Scale against both DSM-5 and ICD-10 diagnostic criteria for depression. BMC Psychiatry. 2018;18:393. doi:10.1186/s12888-018-1965-7.
23. Albuquerque MR, Corrêa H, Couto TC, Santos W, et al. A proposal for a new Brazilian six-item version of the Edinburgh Postnatal Depression Scale. Trends Psychiatry Psychother. 2017;39(1):29-33. doi:10.1590/2237-6089-2016-0056. doi.10.1590/2237-6089-2016-0056.
24. Syam M, Qasim E, Kadrianti et al. Factor structure of the Edinburgh postnatal depression scale Indonesian Version. Med Clin Pract.2021;4:100238. doi:10.1016/j.mcpsp.2021.100238
25. Hartley CM, Barroso N, Rey Y, Pettit JW, Bagner DM. Factor Structure and Psychometric Properties of English and Spanish Versions of the Edinburgh Postnatal Depression Scale AmongHispanicWomen in a Primary Care Setting. J. Clin. Psychol. 2014;70:1240-50. doi:10.1002/jclp.22101.
26. Toreki A, Andó B, Dudas RB, Dweik D, Janka Z, Kozinszky Z, Keresztúri A. Validation of the Edinburgh Postnatal Depression Scale as a screening tool for postpartum depression in a clinical sample in Hungary. Midwifery. 2014; 30(8):911-8. doi: 10.1016/j.midw.2014.02.008.
27. Kheirabadi GR, Maracy MR, Akbaripour S, Masaeli N. Psychometric properties and diagnostic accuracy of the Edinburgh Postnatal Depression Scale in a sample of Iranian women. Iran J Med Sci. 2012;37(1):32-38.Avaliable from: https://pubmed.ncbi.nlm.nih.gov/23115428/
28. Lee King PA. Replicability of structural models of the Edinburgh Postnatal Depression Scale (EPDS) in a community sample of postpartum African American womenwithlowsocioeconomic status. Arch Womens Ment Health. 2012;15:77–86. doi: 10.1007/s00737-012-0260-8.
29. Reichenheim ME, Moraes CL, Oliveira AS, Lobato G. Revisiting the dimensional structure of the Edinburgh Postnatal Depression Scale (EPDS): empirical evidence for a general factor. BMC Med Res Methodol. 2011;11:93.doi:10.1186/1471-2288-11-93
30. Montazeri A, Torkan B, Omidvari S. The Edinburgh Postnatal Depression Scale (EPDS): translation and validationstudy of theIranianversion. BMC Psychiatry. 2007; 7: 11. doi:10.1186/1471-244X-7-11.
31. Santos IS, Matijasevich A, Tavares BF, Barros AJ, Botelho IP, Lapolli C, et al. Validation of the Edinburgh Postnatal Depression Scale (EPDS) in a sample of mothers from the 2004 Pelotas Birth Cohort Study. Cad Saude Publica. 2007;23(11):2577-2588. doi:10.1590/s0102-311x2007001100005.
32. Santos IS, Matijasevich A, Tavares BF, et al. Comparing validity of Edinburgh scale and SRQ20 in screening for post-partum depression. Clin Pract Epidemiol Ment Health. 2007;3:18. doi:10.1186/1745-0179-3-18
33. Jardri R, Pelta J, Maron M, Thomas P, Delion P, Codaccioni X, et al. Predictive validation study of the Edinburgh Postnatal Depression Scale in the first week after delivery and risk analysis for postnatal depression. J Affect Disord. 2006;93(1-3):169-176. doi:10.1016/j.jad.2006.03.00
34. Werrett J, Clifford C. Validation of the Punjabi version of the Edinburgh Postnatal Depression Scale (EPDS). Int J Nurs Stud. 2006;43(2):227-236. doi:10.1016/j.ijnurstu.2004.12.007.
35. American Educational Research Association, American Psychological Association, National Council on Measurement in Education. Standards for educational and psychological testing. Washington, DC: Author; 2014.
36. Clifford C, Day A, Cox J, Werrett J. A cross-cultural analysis of the use of the Edinburgh Post-Natal Depression Scale (EPDS) in health visiting practice. J Adv Nurs. 1999;30(3):655-664. doi:10.1046/j.1365-2648.1999.01115.
37. MacDermid JC. ICF Linking and Cognitive Interviewing Are Complementary Methods for Optimizing Content Validity of Outcome Measures: An Integrated Methods Review. Front Rehabilit Sci [Internet]. 2021 [citado 2025 ago 22];2:702596. doi: doi.10.3389/fresc.2021.702596
38. Echevarria-Guanilo ME, Gonçalves N, Romaniski PJ. Psychometric properties of measurement instruments: conceptual basis and evaluation methods - part II. Texto Contexto Enferm [Internet]. 2019 [citado 2025 ago 22];28:e20170311. doi: 10.1590/1980-265X-tce-2017-0311
39. American Educational Research Association, American Psychological Association, National Council on Measurement in Education. Standards for educational and psychological testing. Washington, DC: Author; 2014.
40. Skodová Z, Bánovčinová Ľ, Urbanová E, Grendár M, Bašková M. Factor Structure of the Edinburgh Postnatal Depression Scale in a Sample of Postpartum Slovak Women. Int J Environ Res Public Health. 2021;18(12):6298. doi:10.3390/ijerph18126298.
41. Levis B, Negeri Z, Sun Y, Benedetti A, Thombs BD, DEPRESsion Screening Data (DEPRESSD) EPDS Group. 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. 2020;370:m4022. doi:10.1136/bmj.m4022.
42. Trizano-Hermosilla I, Alvarado JM. Best alternatives to Cronbach's alpha reliability in realistic conditions: congeneric and asymmetrical measurements. Front Psychol. 2016;7:769. doi:10.3389/fpsyg.2016.00769
43. Iliescu D, Greiff S. On Consequential Validity. European Journal of Psychological Assessment [Internet]. 2021 May 1 [cited 2025 Aug 22];37(3):163–6. Available from: https://doi.org/10.1027/1015-5759/a000664

How to cite this article: Silva, R.C., Oliveira, M.J., Silva, J.M.Q., Rosa, R.F.N., Marques, P.F., Freitas, K.S., Coelho, E.A.C. (2025). Validity evidence of the Edinburgh Postnatal Depression Scale: an integrative review. *O Mundo Da Saúde*, 49. https://doi.org/10.15343/0104-7809.202549e17312025I. Mundo Saúde. 2025,49:e17312025.