Received: 16 January 2020 | Revised: 29 May 2020 | Accepted: 3 August 2020

DOI: 10.1002/jclp.23041

RESEARCH ARTICLE

WILEY

Development and validation of a three-item version of the Edinburgh Postnatal Depression Scale

Pablo Martínez[1,2,3,4] | Irene Magaña[1,5] | Paul A. Vöhringer[2,3,6,7] | Viviana Guajardo[2,3,8] | Graciela Rojas[2,3,4,9]

[1]Escuela de Psicología, Facultad de Humanidades, Universidad de Santiago de Chile, Santiago, Chile

[2]Millennium Institute for Research in Depression and Personality, Santiago, Chile

[3]Departamento de Psiquiatría y Salud Mental, Hospital Clínico Universidad de Chile, Santiago, Chile

[4]Millennium Nucleus to Improve the Mental Health of Adolescents and Youths, Imhay, Santiago, Chile

[5]Centro de Estudios Migratorios (CEM), Universidad de Santiago de Chile, Santiago, Chile

[6]Department of Psychiatry, Tufts Medical Center, Boston, MA, USA

[7]Tufts University School of Medicine, Boston, Massachusetts, USA

[8]Servicio de Psiquiatría, Hospital El Pino, Santiago, Chile

[9]Millennium Nucleus of Social Development, Santiago, Chile

Correspondence

Graciela Rojas, Departamento de Psiquiatría y Salud Mental, Hospital Clínico Universidad de Chile, Av La Paz 1003 St, 8431617 Santiago, Chile.
Email: graciela.rojas.castillo@gmail.com

Funding information

Universidad de Santiago de Chile; Ministerio de Salud de Chile (under tender ID 4127-41-LE12); ANID — Millennium Science Initiative / Millennium Institute for Research on Depression and Personality-MIDAP; ANID — Millennium Science Initiative, grant "Millennium Nucleus to Improve the Mental Health of Adolescents and Youths, Imhay"

Abstract

Objective: To develop and validate a brief screening instrument for postpartum depression in resource-constrained primary care settings.

Method: Secondary data analysis of a cohort of 305 mothers ($Mdn_{age} = 26$) attending well-child check-ups in six primary care centers in Santiago, Chile, answered the Edinburgh Postnatal Depression Scale (EPDS), the 36-Item Short Form Health Survey, and the Mini International Neuropsychiatric Interview depression module. A predictive model for postpartum depression was built using logistic and least absolute shrinkage and selection operator regressions, with bootstrap validation.

Results: A three-item version of the EPDS exhibited excellent discriminative capacity (*c* statistic = 0.95) and showed no significant differences versus the full version of the EPDS ($\chi^2(1) = 1.75$, $p = .187$). The best trade-off between sensitivity (92.86%) and specificity (86.70%) was achieved at a cut-off score of 8/9.

Conclusions: The three-item version of the EPDS can save clinicians valuable time, which might potentially improve communication of results to patients.

KEYWORDS

decision support techniques, postpartum depression, primary care, screening, sensitivity and specificity

1 | INTRODUCTION

Worldwide, an estimated eighteen percent of mothers have postpartum depression (Hahn-Holbrook, Cornwell-Hinrichs, & Anaya, 2018), a serious public health issue associated with difficulties in affect regulation and child behavior (Stein et al., 2014), poorer academic performance, cognitive development delays, and depression during adolescence (Liu et al., 2017; Netsi et al., 2018). Effective postpartum depression treatment can benefit marital satisfaction, parental functioning, and child mental health (Cuijpers, Weitz, Karyotaki, Garber, & Andersson, 2015), avoiding huge lifetime costs (Bauer, Knapp, & Parsonage, 2016); therefore, timely treatment must be a priority.

Evidence-based recommendations of the United States Preventive Services Task Force (USPSTF) suggest that primary care screening for postpartum depression can promptly detect mothers with depressive symptoms in need of treatment (O'Connor, Rossom, Henninger, Groom, & Burda, 2016). Moreover, when accompanied by integrated courses of action and organizational supports to prevent cumulative deficits in the continuum of postpartum depression care (e.g., onsite availability of treatment alternatives; Kendig et al., 2017; Olin et al., 2016), screening for this maternal mental health issue may have clinical benefits (O'Connor et al., 2016).

The USPSTF recommends primary care screening for postpartum depression with the Edinburgh Postnatal Depression Scale (EPDS; O'Connor et al., 2016), a 10-item self-report instrument (Cox, Holden, & Sagovsky, 1987), which is one of the most widely used postpartum depression questionnaires, though its application as a screening method remains controversial. For instance, current National Institute for Health and Care Excellence (NICE) guidelines suggests instead the use of the two-item Whooley questions for case-finding followed by the EPDS or the Patient Health Questionnaire-9 item (PHQ-9) for further assessment (National Collaborating Centre for Mental Health [NCCMH UK], 2014).

According to NICE guidelines, the combination of a brief case-finding tool followed by a more formal assessment seems to be the most-cost-effective screening method (NCCMH UK, 2014). However, although simple and highly sensible, the Whooley questions could result in a substantial proportion of false positives (NCCMH UK, 2014), just as the EPDS (Harris, 2016). Consequently, a substantial proportion of mothers who are exposed to lengthy screening methods may be falsely classified as being depressed, and be subjected to unnecessary, costly, and stigmatizing procedures (Harris, 2016). Thus, simple and accurate screening methods for postpartum depression are needed.

Moreover, though previous efforts have been documented to reduce the full version of the EPDS (Albuquerque et al., 2017; Gollan et al., 2017; Kabir, Sheeder, & Kelly, 2008; Phillips, Charles, Sharpe, & Matthey, 2009), these studies have important limitations. For instance, Albuquerque et al. (2017), Gollan et al. (2017), and Kabir et al. (2008) proposed six-, seven-, and two-item versions of the EPDS, respectively, although evaluation of their performance lacked the use of a formal psychiatric diagnostic assessment to ascertain postpartum depression status. In addition, Phillips et al. (2009), conducted testing of the seven-item version of the EPDS in a highly selected sample of women and finding a substantial proportion of false positives.

The case of Chile, a developing country, might be critical, as major efforts have been made to address the burden of postpartum depression. By the end of the past decade, a universal screening program for postpartum depression in primary care was implemented (Chilean Ministry of Health, 2014), with the EPDS being administered during well-child visits and referral of depressed mothers to the onsite national depression program (Alvarado, Rojas, Minoletti, Alvarado, & Domínguez, 2012). However, despite high levels of screening, a small share of women has access to treatment (~15%; Chilean Department of Health Statistics and Information, 2018b).

In practice, busy primary care clinicians (i.e., a nurse or a physician) who have half an hour to conduct a thorough, comprehensive evaluation of the child's growth and provide guidance to parents and/or caregivers, are also in charge of the administration of the screening for postpartum depression (Chilean Ministry of Health, 2014). In this context, although administering the EPDS is feasible in 5 min, this instrument might be long enough to represent a missed chance for providing tailored information about test results, briefly educate mothers about

postpartum depression, and motivating access to treatment for depressed mothers (Holt, Milgrom, & Gemmill, 2017).

Given the need to develop brief methods for the identification of depressive symptoms during the postpartum period, and bearing in mind their potential value for influencing the postpartum depression treatment cascade and reducing the personal, familial, and social costs associated with this mental disorder (Petrou, Morrel, & Knapp, 2015), abbreviated and easy-to-interpret postpartum depression screening instruments may be a valuable contribution to busy clinicians in resource-constrained primary care settings, boosting these professionals' impact on the promotion of timely access to postpartum depression treatment.

The objective of this study was to develop and validate a brief screening instrument for postpartum depression in resource-constrained primary care settings.

2 | METHODS

2.1 | Research design

This paper reports a secondary analysis of baseline data of a cohort study. The original study was commissioned by the Chilean Ministry of Health to evaluate the prevalence and risk factors of postpartum depression, while also documenting services usage and barriers to access postpartum depression treatment in primary care in Santiago, Chile. Baseline data, which included health, psychological, and social data of postpartum women, were obtained a week after recruitment of participants, with a follow-up assessment 3 months later to evaluate service usage and barriers to access postpartum depression treatment. The present paper used a subset of the baseline data to examine the cross-sectional association of health, psychological, and social factors with the diagnosis of postpartum depression, to further develop a predictive model and risk score instrument for postpartum depression.

2.2 | Participant characteristics

The full sample was composed of 305 mothers aged 18 years and up who took their sons/daughters to 2- and 6-month-old well-child checkups at primary care centers in Santiago agreed to participate voluntarily and informedly, had no intellectual disabilities, and were able to be contacted and evaluated over the phone.

2.3 | Sampling procedures

A two-stage nonprobability sampling procedure was used. First, out of the full set of primary care centers in Santiago (n = 120), we selected those with the largest number of 2-month-old well-child checkups in each of the six metropolitan health services, based on Chilean Ministry of Health data for January–September 2012. This resulted in the inclusion of six primary care centers. At a later stage, participants were consecutively selected in the waiting rooms of each primary care center to reduce sampling bias. During January and February 2013, duly trained research team members approached each potential participant after their well-child visit, explaining the procedures of the study to them, obtaining their informed consent, and performing an initial verification of the eligibility criteria. Participants did not receive any compensation for their participation.

This study was approved by the Ethics Committee of Hospital Clínico Universidad de Chile and by the respective entities of each of the six health services of the Metropolitan region. All participants were duly informed of the details of the study and agreed to participate voluntarily, after signing the previously approved informed consent.

2.4 | Sample size

Based on a study conducted by Alvarado et al. (1992), which evaluated mothers attending primary care centers in Santiago, Chile, reporting a prevalence of 20.7% for postpartum depression, a sample size of at least 300 participants was calculated for the detection of approximately 60 mothers with postpartum depression. According to the number of events per variable, an indicator of sample size for developing predictive models (Pavlou et al., 2015), overfitting risk can be reduced with a predictive model composed of no more than six variables (or six questions in a screening instrument). Given the length of the EPDS (10 items), postpartum depression screening would be greatly simplified with a predictive model of these characteristics.

2.5 | Measures and covariates

Data collection was conducted by clinical members of the research group, trained, and experienced in the use of standardized psychometric instruments and clinical assessments, through a structured phone interview used to collect health, psychological, and social data.

2.5.1 | Primary outcome variable

The main outcome was the presence or absence of a diagnosis of postpartum depression, established through the administration of the depression module of the Mini International Neuropsychiatric Interview (Sheehan et al., 1998), Spanish version (Bobes et al., 1998). The Mini International Neuropsychiatric Interview is a short structured diagnostic interview based on the diagnostic criteria of the fourth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV), published by the American Psychiatric Association (1994). The depression module of this structured diagnostic interview has demonstrated excellent interrater ($\kappa = 1.00$) and test–retest reliability ($\kappa = .87$; Sheehan et al., 1998).

2.5.2 | Candidate predictor variables

The main candidate predictor variables were self-reported depressive symptoms during the postpartum period and health-related quality of life, according to the EPDS and the 36-Item Short Form Health Survey (SF-36), respectively. The EPDS is a brief, 10-Question Self-Report Scale for assessing depressive symptoms during the postpartum period (Cox et al., 1987), validated for use in Chile by Alvarado et al. (1992) and Jadresic, Araya, and Jara (1995). It has demonstrated validity and reliability (split-half reliability = .88, Cronbach's $\alpha = .87$; Cox, 1994). In both Chilean validation studies, a cut-off score equal to or higher than 10 was found to be the best threshold for discriminating the presence or absence of postpartum depression. The EPDS has been widely used in Chile (Jadresic, Nguyen, & Halbreich, 2007). We chose not to include item number 10 of the EPDS ("The thought of harming myself has occurred to me") in the construction of the predictive model because the evaluation of suicidal ideation must be an essential part of maternal mental health programs and should be managed separately from depressive symptoms (Yawn, LaRusso, Bertram, & Bobo, 2015). The SF-36 is a self-report instrument that provides a subjective and generic measure of health-related quality of life. It comprises 36 questions grouped into eight dimensions, with scores ranging from 0 (worst health) to 100 (best health; Ware & Sherbourne, 1992). The SF-36 scales have demonstrated good reliability estimates across diverse patient groups (median Cronbach's $\alpha = .85$) (McHorney, Ware, Lu, & Sherbourne, 1994). It has been validated for use in Chile and has been used to study the health-related quality of life of Chilean mothers with postpartum depression (Olivares-Tirado, 2005; Rojas et al.,

2006). For the secondary analyses herein reported, SF-36 items were preselected in two steps: (1) if a couple of SF-36 and EPDS items had a Spearman correlation coefficient (ρ) > .5, the EPDS item was selected; (2) if one of the items resulting from the first step had ρ > .5, the corresponding item of the SF-12—a reduced version of the SF-36—was selected (Vilagut et al., 2013). If any pairs of items with ρ > .5 remained, the item with the strongest correlation with the overall EPDS score was selected. With this procedure, SF-36 items were reduced to 11.

2.5.3 | Other candidate predictors

Based on a synthesis of systematic evidence about psychiatric disorders in the postpartum period and considering the literature on postpartum depression in Chile (Jadresic et al., 2007; Meltzer-Brody et al., 2018), other candidate predictor variables of a psychological (e.g., history of depression) and social nature (e.g., marital status) were preselected considering those included in the main study. Thus, the following variables were selected: (1) being a single mother (i.e., being single, divorced, or separated); (2) not living with one's partner; (3) having <12 years of schooling; (4) having had an unplanned pregnancy; (5) receiving no help to take care of the baby; (6) having two or more children; and (7) having a history of depression.

2.6 | Statistic and data analyses

The data were inspected to determine whether the proportion of missing data warranted a special procedure (Steyerberg, 2009). The associations between the main outcome and each of the candidate predictor variables were explored, using Student's *t*-test for continuous variables (assuming unequal variances), and the χ^2 test for binary variables. The resulting association measures were differences of means or odds ratio, with their respective confidence intervals (CIs) set at 95%. The candidate predictor variables found to be statistically significant at bivariate analysis ($p < .05$, adjusted for multiple comparisons using the Holm method [Bender & Lange, 2001]), underwent a second selection process in which the variance inflation factor was estimated, eliminating variables with a variance inflation factor ≥ 2 to prevent multicollinearity problems (Thompson, Kim, Aloe, & Becker, 2017).

Afterward, the remaining variables were tested with a multiple logistic regression model with manual backward stepwise selection to find the most parsimonious reduced model according to the Akaike information criterion (Vrieze, 2012), and diagnoses of the regression model were performed to eliminate influential points (Hosmer, Lemeshow, & Sturdivant, 2013). To avoid overfitting and improve the performance of the predictive model for postpartum depression, a least absolute shrinkage and selection operator (LASSO) regression method was employed (Steyerberg, 2009), with cross-validation to determine the most suitable penalization factor, which resulted in some of the regression coefficients shrunken toward zero.

Following the penalization of the predictive model, its internal validity was determined with the bootstrapping technique by drawing 200 random samples with replacement from the original sample, and seeds of random numbers prespecified to ensure the replication of the results. With this technique, statistical values assessing the performance of the predictive model were computed, adjusted for the optimism of the apparent validation in the original sample: (1) Nagelkerke pseudo-R^2, the percentage of explained variation, as a measure of general performance; (2) area under the curve of the receiver operating characteristic (ROC) or *c* statistic, used to evaluate the ability of the predictive model to discriminate between mothers with and without postpartum depression; and, (3) to evaluate the fit between the predicted and the observed probability, the values of the intercept and slope of the recalibrated predictive model and the Hosmer–Lemeshow goodness of fit test were used (Steyerberg, 2009).

Subsequently, an easy-to-interpret, risk score instrument (i.e., the three-item version of the EPDS) for postpartum depression was developed (Steyerberg, 2009), based on the rounding of the regression coefficients penalized using the LASSO method (Table 2, predictive score). To assess potential information loss due to the

rounding of the coefficients in the three-item version of the EPDS, the calibration and discriminative capacity were reevaluated. The ROC curves of the three-item version of the EPDS and the full version of the EPDS (with their 95% CI) were compared, with the existence of statistically significant differences determined by the χ^2 test. To determine the clinical usefulness of the three-item version of the EPDS, statistical values for classification were estimated for a range of cut-off scores. Analyses were assisted with Stata 14.0 (StataCorp, 2015) and R 3.5.1 (R Core Team, 2018), using the glmnet package (Friedman, Hastie, & Tibshirani, 2010).

2.7 | Ethical approval and informed consent

All procedures performed in this study involving human participants were in accordance with the Ethical Standards of the Institutional and/or National Research Committee of Hospital Clínico Universidad de Chile and by the respective entities of each of the six health services of the Metropolitan region and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants included in the study.

3 | RESULTS

3.1 | Data exploration

The proportion of missing data in the full sample was low, seven cases (2.3%) displayed missing data for the variables "History of depression" ($n = 6$) and "Increased parity" ($n = 1$). Due to the proportion of missing data, the listwise deletion procedure was applied (Steyerberg, 2009); therefore, data analyses considered 298 cases with full information (i.e., estimation sample). In the estimation sample ($Mdn_{age} = 26$, interquartile range = 9), 38.6% of the women were probably depressed (EPDS score ≥ 10) and 21.1% ($n = 63$) were diagnosed with postpartum depression according to the Mini International Neuropsychiatric Interview. The characteristics of the sample for the variables selected and their association with postpartum depression are shown in Table 1. Prevalent conditions in the sample, such as unplanned pregnancy (58.4%) and having two or more children (60.4%), were found to be risk factors for the dependent variable.

3.2 | Development of a predictive model for postpartum depression

Preliminary analyses showed that, from several possible models, the most parsimonious one comprised items 1, 2, 3, 5, and 7 of the EPDS, items 15 and 35 of the SF-36, and having a history of depression (Table 2, naïve regression coefficients). The penalization of regression coefficients using LASSO regression led to including three items of the EPDS in the final predictive model: 2 ("I have looked forward to the future"), 5 ("I have felt frightened or panicky for no good reason"), and 7 ("I have felt so bad that I have had difficulty in sleeping") (Table 2, shrunken regression coefficients).

3.3 | Internal validity of a predictive model for postpartum depression

When comparing changes in the performance statistics in the development sample and the internal validation (bootstrap sampling with replacement), it can be observed that the optimism of the predictive model is negligible, thus indicating good fit (pseudo-R^2), excellent discriminatory capacity (c statistic), perfect calibration values (intercept and slope of the recalibrated model), and adequate goodness of fit (Table 3).

TABLE 1 Health, psychological, and social characteristics of postpartum women, and their association with postpartum depression

	Data		Crude measures of association	
	M	*SD*	*t*	95% CI
Age	27.1	6.6	0.7	[−1.2, 2.5]
EPDS score	9.0	5.5	9.4	[8.2, 10.6]
PCS-36 score	51.0	8.1	−5.0	[−7.6, −2.3]
MCS-36 score	44.0	14.8	−24.8	[−27.8, −21.8]
	n	%	OR	95% CI
Single mother	208	69.8	1.0	[0.5, 2.0]
Living without a partner	116	38.9	2.0	[1.1, 3.7]
Low education	93	31.2	0.9	[0.5, 1.8]
Unplanned pregnancy	174	58.4	1.9	[1.0, 3.6]
Lack of support for child care	115	38.6	1.0	[0.5, 1.8]
Increased parity	180	60.4	1.9	[1.0, 3.6]
History of depression	111	37.3	3.4	[1.8, 6.3]

Note: *N* = 298.

Abbreviations: EPDS, Edinburgh Postnatal Depression Scale; MCS, 36-Item Short Form Health Survey, Mental Component Summary Subscale; PCS, 36-Item Short Form Health Survey, Physical Component Summary Subscale.

TABLE 2 Naïve and shrunken estimates, and predictive risk score

	Regression coefficients		
	Naïve[a]	Shrunken[b]	Predictive risk score[c]
EPDS item 1	1.95	–	–
EPDS item 2	1.27	0.33	3
EPDS item 3	0.68	–	–
EPDS item 5	0.52	0.05	1
EPDS item 7	1.41	0.44	4
SF-36 item 2	−1.23	–	–
SF-36 item 15	1.05	–	–
SF-36 item 35	0.64	–	–
History of depression	0.76	–	–
Shrinkage parameter		0.15	
Effective shrinkage		0–0.31	

Note: *N* = 289. Sample size after deletion of influential observations.

Abbreviations: EPDS, Edinburgh Postnatal Depression Scale; SF-36, 36-Item Short Form Health Survey.

[a]Naïve regression coefficients were calculated by logistic regression using maximum likelihood estimation.

[b]Shrunken regression coefficients were calculated by logistic regression with the least absolute shrinkage and selection operator (LASSO) penalization.

[c]To estimate the predictive risk score, shrunken regression coefficients were rounded to a decimal place and multiplied by 10.

TABLE 3 Performance measures of postpartum depression prediction model in the development and internal validation samples

Performance measures	Development	Internal validation[a]
R^2	51.79%	51.05%
c statistic[b]	0.95	0.94
	[0.92, 0.97]	[0.91, 0.97]
Calibration curve slope	1.00	1.00
Calibration curve intercept	0.00	0.00
Hosmer–Lemeshow goodness-of-fit test[c]	3.21	1.20
	(2, .201)	(2, .547)

Note: $N = 298$. Sample size after deletion of influential observations.

[a]Internal validation with 200 bootstraps resamples with replacement.

[b]*c* statistic and 95% confidence intervals (in square brackets); with 95% bias-corrected confidence interval for internal validation.

[c]Data for the Hosmer–Lemeshow goodness-of-fit test is χ^2, with degrees of freedom and *p* values in parentheses; the Hosmer–Lemeshow goodness-of-fit test was based on quartiles to ensure the appropriate size of each risk category.

3.4 | Risk score instrument for postpartum depression: The three-item version of the EPDS

A graphical representation of a score table suggests an almost monotonic relationship between a higher three-item EPDS score and a greater likelihood of postpartum depression (Figure S1). Table 4 shows the scoring rules for the three-item version of the EPDS, with a score ranging from 0 to 24. Information loss due to coefficient rounding was negligible, with the three-item version of the EPDS displaying adequate calibration (Hosmer–Lemeshow $\chi^2(2) = 2.57$, $p = .277$), and a discriminative capacity equal to that of the predictive model. A comparison between the ROC curve of the three-item version of the EPDS and that of the full version of the EPDS yielded no statistically significant differences in terms of discriminative capacity ($\chi^2(1) = 1.75$, $p = .187$; Figure 1).

Adjusting for the prevalence of postpartum depression in Chilean primary health care (and in the estimation sample), at a cut-off score of 8/9, the three-item version of the EPDS reached a sensitivity of 92.86%, a specificity of 86.70%, and a positive predictive value of 62.65%, achieving the best trade-off between sensitivity and specificity across a range of cut-off scores (Table 5). As a comparison, according to the Chilean validations of the full version of the EPDS (Alvarado et al., 1992; Jadresic et al., 1995), the best cut-off score for this screening instrument (9/10), achieved a higher sensitivity (100%), but a lower specificity (77.25%) in the estimation sample, thus having a higher rate of false positives.

4 | DISCUSSION

The predictive modeling techniques used led to the development of a three-item version of the EPDS (composed of items 2, 5, and 7 from the full version of the EPDS). After internal validation by bootstrap, the three-item version of the EPDS had negligible optimism, with good performance statistics for detecting postpartum depression. The new instrument, with a score ranging from 0 to 24, attained a discriminative capacity comparable with that of the full version of the EPDS (*c* statistic = 0.95), thus demonstrating its excellent capacity to correctly classify women with and without postpartum depression. The best cut-off score (8/9), displayed a sensitivity of 92.86%, a specificity of 86.70%, and a positive predictive value of 62.65%.

TABLE 4 Scoring rules for the three-item version of the Edinburgh Postnatal Depression Scale

PS-Item	Alternatives	Score
EPDS item 2 "I have looked forward to the future"	As much as I ever did	0
	Rather less than I used to	3
	Definitely less than I used to	6
	Hardly at all	9
EPDS item 5 "I have felt frightened or panicky for no good reason"	Yes, quite a lot	3
	Yes, sometimes	2
	No, not much	1
	No, not at all	0
EPDS item 7 "I have felt so bad that I have had difficulty in sleeping"	Yes, most of the time	12
	Yes, sometimes	8
	Not very often	4
	No, not at all	0

Note: Score ranges from 0 to 24 points. The EPDS items were extracted from Cox et al. (1987).
Abbreviation: EPDS, Edinburgh Postnatal Depression Scale.

The three-item version of the EPDS is an abbreviated version of the most used screening instrument for postpartum depression (O'Connor et al., 2016), and, in contrast to a 10-item questionnaire, it emerges as a more parsimonious alternative. Moreover, its performance compares quite well with validity coefficients found for the full version of the EPDS. For instance, in the Chilean validation studies of the latter, sensitivity and specificity were estimated to be approximately 90% and 80%, respectively (Alvarado et al., 1992; Jadresic et al., 1995); while the English language version of the full EPDS has reported sensitivities and specificities of around 80% and 90%, respectively (O'Connor et al., 2016).

In the literature, factor analyses of the full version of the EPDS have led to the development of six- or seven-item depression subscales (Albuquerque et al., 2017; Gollan et al., 2017; Phillips et al., 2009), and predictive models that have incorporated several sociodemographic and clinical variables (Jiménez-Serrano, Tortajada, & García-Gómez, 2015) or

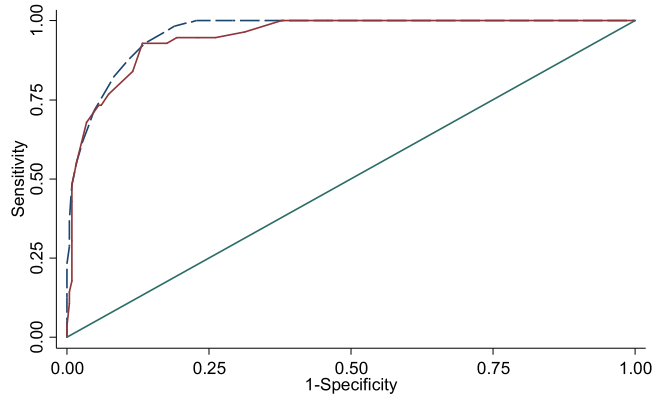

FIGURE 1 Discriminative capacity of the three-item version of the Edinburgh Postnatal Depression Scale (EPDS) versus the full (10-item) version of the EPDS. Lines are receiver operating characteristic (ROC) curves for the full EPDS (blue long dashed line), the three-item version of the EPDS (red continuous line), and the reference line (green diagonal line). The area under the ROC curve (*c* statistic) for the full EPDS was 0.96 (95% confidence interval: 0.94, 0.98), and for the three-item version of the EPDS was 0.95 (95% confidence interval: 0.92, 0.98) [Color figure can be viewed at wileyonlinelibrary.com]

TABLE 5 Classification table (%) for the three-item version of the Edinburgh Postnatal Depression Scale

Cut-off score	6/7	7/8	8/9	9/10	10/11
Sensitivity	94.64	92.86	92.86	83.93	76.79
Specificity	80.69	82.40	86.70	88.41	92.70
Positive predictive value	54.08	55.91	62.65	63.51	71.67
Negative predictive value	98.43	97.96	98.06	95.81	94.32
Correctly classified	83.39	84.43	87.89	87.54	89.62

which evaluated the predictive power of biomarkers (Osborne et al., 2016) have been devised. The three-item version of the EPDS reported in this study is at least twice as short and as accurate (or more accurate) as these alternatives, and, importantly, it seems to replicate the commonly found three-factor solution (i.e., "anhedonia": items 1 and 2; "anxiety": items 3, 4, 5, and 6; and, "depression": items 7, 8, 9, and 10) reported by other studies (Coates, Ayers, & de Visser, 2017).

Regarding the clinical relevance of the selected items, it may be said that the three-item version of the EPDS seems to depict an important combination of symptoms for a depressive episode during the postpartum period. For instance, the fact that a new mother may not be able to "[look] forward to the future" (item 2 of the EPDS), is a clear indication of anhedonia—a cardinal symptom of depression. In addition, as noted by the fifth edition of the DSM-5 (American Psychiatric Association, 2013), feeling "frightened or panicky for no good reason" (item 5 of the EPDS), seems to be a common trait of postpartum depression, as mothers with mood episodes may even experience panic attacks. Finally, it seems that the item 7 of the EPDS ("I have felt so bad that I have had difficulty in sleeping"), gives evidence of a clinical picture of depression that is exerting a heavy emotional toll over and above the expected demands of motherhood.

As discussed earlier, the two-item Whooley questions remains a good choice for postpartum depression case-finding in primary care (NCCMH UK, 2014). Its administration, paired with the full version of the EPDS or PHQ-9 for a more thorough assessment, have been identified as one of the most cost-effective screening methods available (NCCMH UK, 2014). Although recommended for its brevity, when the two-item Whooley questions are applied to postpartum women a large proportion of false positives is produced (NCCMH UK, 2014). The three-item version of the EPDS, as short as the two-item Whooley questions, has better overall diagnostic accuracy and substantially reduces the proportion of false positives.

The three-item version of the EPDS emerges as a parsimonious, accurate, easy to administer, and interpret tool, readily available to busy primary care clinicians working in resource-constrained settings. Compared to currently available screening methods, after a brief assessment to detect mothers at risk for postpartum depression, clinicians would have more time available to explore symptoms, provide psychoeducation, and boost patients' treatment motivation, actions that may help to reduce the treatment gap for postpartum depression. For instance, recently Holt et al. (2017) have demonstrated that a brief motivational interviewing intervention during well-child check-ups can increase treatment uptake for postnatal depression and anxiety.

Given the controversies regarding postpartum depression screening with the full version of the EPDS (Harris, 2016), the instrument herein developed and internally validated can be said to address some of these issues thanks to its quick administration and its lower rates of false positives. Moreover, this latter characteristic may reduce the costs to women and the health services of managing incorrectly classified cases (Harris, 2016). For example, based on the validity coefficients reported for the Chilean version of the full EPDS (Alvarado et al., 1992; Jadresic et al., 1995), and according to updated data from the Chilean Department of Health Statistics and Information (2018a), the three-item version of the EPDS would reduce the number of mothers falsely classified as depressed from approximately 7200 to approximately 5500.

Despite promising prospects, the results of the present study have some limitations. First, the study did not involve random sampling and was restricted to Santiago, the capital city of Chile, which may affect the

generalization of results. Second, sample size (298 mothers, of which 63 had postpartum depression) and the number of initial predictor variables (i.e., >20 candidate predictor variables), greatly surpassed the threshold recommended for the ratio of events per variable, and thus be at risk of overfitting (Pavlou et al., 2015). Another major limitation is that the internal validation, which employed the bootstrapping technique, was applied after the final predictive model (penalized with LASSO) was determined, which may not guarantee the consistent selection of predictors and may introduce bias in the calculation of the optimism (i.e., predictive error; Steyerberg, 2009). Thus, a definitive assessment of the test's performance can only be established through external validation, to generalize and apply these findings in real clinical settings. In this regard, futures studies should recruit a random, representative sample of the postpartum women in Chile, with higher sample sizes, and ensuring a balance of potentially confounding variables (e.g., parity, pregnancy planning, or social support) for measurement invariance testing (i.e., equivalence of psychometric properties across groups). Although not strictly a limitation of this study, it must be borne in mind that from a clinical standpoint suicidality assessment should be integral to maternal mental health programs, therefore, the three-item version of the EPDS should be administered including the item 10 of the EPDS to detect those women in need of immediate referral (Yawn et al., 2015).

5 | CONCLUSION

The predictive modeling techniques applied allowed for the reduction of the full (10 items) version of the EPDS to a three-item version of the same instrument, generating a test with two relevant characteristics for postpartum depression screening: simplicity and precision. This abbreviated version of the EPDS saves clinicians valuable time that can be used to potentially improve the communication of test results, which may help to reduce the treatment gap for postpartum depression. Nevertheless, external validation is needed to guarantee its generalization.

ACKNOWLEDGMENTS

This study was supported by the Universidad de Santiago de Chile; Ministerio de Salud de Chile; the ANID—Millennium Science Initiative /Millennium Institute for Research on Depression and Personality-MIDAP; and the ANID—Millennium Science Initiative, grant "Millennium Nucleus to Improve the Mental Health of Adolescents and Youths, Imhay". This manuscript is based on data used in a previously published report. The authors would like to thank the participants. Scientific editing by Jennifer Cheavens.

ORCID

Pablo Martínez http://orcid.org/0000-0003-4482-7993

REFERENCES

Albuquerque, M. R., Correa, H., Couto, T. C., Santos, W., Romano-Silva, M. A., & Santos, L. M. (2017). A proposal for a new Brazilian six-item version of the Edinburgh Postnatal Depression Scale. *Trends in Psychiatry and Psychotherapy*, *39*, 29–33. https://doi.org/10.1590/2237-6089-2016-0056

Alvarado, R., Rojas, G., Minoletti, A., Alvarado, F., & Domínguez, C. (2012). Depression program in primary health care: The Chilean experience. *International Journal of Mental Health*, *41*, 38–47. https://doi.org/10.2753/IMH0020-7411410103

Alvarado, R., Vera, A., Rojas, M., Olea, E., Monardes, J., & Neves, E. (1992). La Escala de Edimburgo para la detección de cuadros depresivos en el posparto. *Revista de Psiquiatría*, *3-4*, 1177–1181.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: American Psychiatric Association.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: American Psychiatric Association.

Bauer, A., Knapp, M., & Parsonage, M. (2016). Lifetime costs of perinatal anxiety and depression. *Journal of Affective Disorders*, *192*, 83–90. https://doi.org/10.1016/j.jad.2015.12.005

Bender, R., & Lange, S. (2001). Adjusting for multiple testing -when and how? *Journal of Clinical Epidemiology*, *54*, 343–349. https://doi.org/10.1016/S0895-4356(00)00314-0

Bobes, J., González, M. P., Sáiz, P. A., Bousoño, M., Iglesias, C., Bascarán, M. T., & Gibert, J. (1998). La MINI (Mini International Neuropsychiatric Interview): una familia de entrevistas de ayuda diagnóstica en Psiquiatría y Atención Primaria. In M. Gutiérrez, J. Ezcurra, A. González, & P. Pichot (Eds.), *Psiquiatría y otras especialidades médicas* (pp. 401–419). Madrid, Spain: Grupo Aula Médica.

Chilean Department of Health Statistics and Information. (2018a). *Section B: Assessment, administration, and results of screening scales in women: Section B3: administration of the Edinburgh Postnatal Depression Scale to pregnant and postpartum women* [original title in Spanish: Sección B: evaluación, aplicación y resultados de escalas en la mujer: Sección B3: aplicación de Escala de Edimburgo a gestantes y mujeres posparto] [Data]. Retrieved from http://webdeis.minsal.cl/rem2018/?serie=1&rem=24&seccion_id=230&tipo=4&regiones=-1&servicios=0&periodo=2018&mes_inicio=1&mes_final=12

Chilean Department of Health Statistics and Information. (2018b). *Section N: Admissions to primary care/specialty mental health program* [original title in Spanish: Sección N: ingresos al programa de salud mental en APS/especialidad] [Data]. Retrieved from http://webdeis.minsal.cl/rem2018/?serie=1&rem=26&seccion_id=265&tipo=4&tipoReload=4&regiones=-1&regionesReload=-1&servicios=0&serviciosReload=0&periodo=2018&mes_inicio=1&mes_final=12

Chilean Ministry of Health. (2014). *Technical Standard for the supervision of children from 0 to 9 years old* [original title in Spanish: Norma técnica para la supervisión de niños y niñas de 0 a 9 años en la Atención Primaria de Salud. Programa Nacional de Salud de la Infancia] (Technical Standard No.1 66). Retrieved from: http://www.crececontigo.gob.cl/wp-content/uploads/2015/11/Norma-Tecnica-para-la-supervision-de-ninos-y-ninas-de-0-a-9-en-APS.compressed.pdf.

Coates, R., Ayers, S., & de Visser, R. (2017). Factor structure of the Edinburgh Postnatal Depression Scale in a population-based sample. *Psychological Assessment*, *29*, 1016–1027. https://doi.org/10.1037/pas0000397

Cox, J. L. (1994). *Perinatal psychiatry: Use and misuse of the Edinburgh Postnatal Depression Scale (EPDS)*. London, UK: Gaskell.

Cox, J. L., Holden, J. M., & Sagovsky, R. (1987). Detection of postnatal depression: development of the 10-item Edinburgh Postnatal Depression Scale. *British Journal of Psychiatry*, *150*, 782–786. https://doi.org/10.1192/bjp.150.6.782

Cuijpers, P., Weitz, E., Karyotaki, E., Garber, J., & Andersson, G. (2015). The effects of psychological treatment of maternal depression on children and parental functioning: A meta-analysis. *European Child & Adolescent Psychiatry*, *24*, 237–245 . https://doi.org/10.1007/s00787-014-0660-6

Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, *33*, 1–22.

Gollan, J. K., Wisniewski, S. R., Luther, J. F., Eng, H. F., Dills, J. L., Sit, D., … Wisner, K. L. (2017). Generating an efficient version of the Edinburgh Postnatal Depression Scale in an urban obstetrical population. *Journal of Affective Disorders*, *208*, 615–620. https://doi.org/10.1016/j.jad.2016.10.013

Hahn-Holbrook, J., Cornwell-Hinrichs, T., & Anaya, I. (2018). Economic and health predictors of national postpartum depression prevalence: A systematic review, meta-analysis, and meta-regression of 291 studies from 56 countries. *Frontiers in Psychiatry*, *8*, 248. https://doi.org/10.3389/fpsyt.2017.00248

Harris, L. (2016). Screening for perinatal depression: A missed opportunity. *The Lancet*, *387*, 505. https://doi.org/10.1016/S0140-6736(16)00265-8

Holt, C., Milgrom, J., & Gemmill, A. W. (2017). Improving help-seeking for postnatal depression and anxiety: A cluster randomized controlled trial of motivational interviewing. *Archives of Women's Mental Health*, *20*, 791–801. https://doi.org/10.1007/s00737-017-0767-0

Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression* (3rd ed.). Hoboken, NJ: John Wiley & Sons.

Jadresic, E., Araya, R., & Jara, C. (1995). Validation of the Edinburgh Postnatal Depression Scale (EPDS) in Chilean postpartum women. *Journal of Psychosomatic Obstetrics and Gynaecoloy*, *16*, 187–191. https://doi.org/10.3109/01674829509024468

Jadresic, E., Nguyen, D. N., & Halbreich, U. (2007). What does Chilean research tell us about postpartum depression (PPD)? *Journal of Affective Disorders*, *102*, 237–243. https://doi.org/10.1016/j.jad.2006.09.032

Jiménez-Serrano, S., Tortajada, S., & García-Gómez, J. M. (2015). A mobile health application to predict postpartum depression based on machine learning. *Telemedicine Journal and e-Health*, *21*, 567–574. https://doi.org/10.1089/tmj.2014.0113

Kabir, K., Sheeder, J., & Kelly, L. S. (2008). Identifying postpartum depression: Are 3 questions as good as 10? *Pediatrics*, *122*, e696–e702. https://doi.org/10.1542/peds.2007-1759

Kendig, S., Keats, J. P., Hoffman, M. C., Kay, L. B., Miller, E. S., Moore Simas, T. A., … Lemieux, L. A. (2017). Consensus bundle on maternal mental health: Perinatal depression and anxiety. *Journal of Obstetric, Gynecologic, and Neonatal Nursing*, *46*, 272–281. https://doi.org/10.1016/j.jogn.2017.01.001

Liu, Y., Kaaya, S., Chai, J., McCoy, D. C., Surkan, P. J., Black, M. M., … Fawzi, M. C. (2017). Maternal depressive symptoms and early childhood cognitive development: A meta-analysis. *Psychological Medicine*, *47*, 680–689. https://doi.org/10.1017/S003329171600283X

McHorney, C. A., Ware, J. E., Lu, J. F. R., & Sherbourne, C. D. (1994). The MOS 36-Item Short-Form Health Survey (SF-36): III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. *Medical Care*, *32*, 40–66.

Meltzer-Brody, S., Howard, L. M., Bergink, V., Vigod, S., Jones, I., & Munk-Olsen, T. (2018). Postpartum psychiatric disorders. *Nature Reviews Disease Primers*, *4*, 18022. https://doi.org/10.1038/nrdp.2018.22

National Collaborating Centre for Mental Health. (2014). *Antenatal and postnatal mental health: Clinical management and service guidance*. Leicester, UK: British Psychological Society.

Netsi, E., Pearson, R. M., Murray, L., Cooper, P., Craske, M. G., & Stein, A. (2018). Association of persistent and severe postnatal depression with child outcomes. *JAMA Psychiatry*, *75*, 247–253. https://doi.org/10.1001/jamapsychiatry.2017.4363

O'Connor, E., Rossom, R. C., Henninger, M., Groom, H. C., & Burda, B. U. (2016). Primary care screening for and treatment of depression in pregnant and postpartum women: Evidence report and systematic review for the US Preventive Services Task Force. *Journal of the American Medical Association*, *315*, 388–406. https://doi.org/10.1001/jama.2015.18948

Olin, S.-C. S., Kerker, B., Stein, R. E. K., Weiss, D., Whitmyre, E. D., Hoagwood, K., & Horwitz, S. M. (2016). Can postpartum depression be managed in pediatric primary care? *Journal of Women's Health*, *25*, 381–390. https://doi.org/10.1089/jwh.2015.5438

Olivares-Tirado, P. (2005). *Perfil del estado de salud de beneficiarios de ISAPRES: informe preliminar*. Santiago, Chile: Superintendencia de Salud.

Osborne, L., Clive, M., Kimmel, M., Gispen, F., Guintivano, J., Brown, T., … Kaminsky, Z. (2016). Replication of epigenetic postpartum depression biomarkers and variation with hormone levels. *Neuropsychopharmacology*, *41*, 1648–1658. https://doi.org/10.1038/npp.2015.333

Pavlou, M., Ambler, G., Seaman, S. R., Guttman, O., Elliot, P., King, M., & Omar, R. Z. (2015). How to develop a more accurate risk prediction model when there are few events. *BMJ*, *351*, h3868. https://doi.org/10.1136/bmj.h3868

Petrou, S., Morrel, C. J., & Knapp, M. (2015). An overview of health economic aspects of perinatal depression. In J. Milgrom & A. W. Gemmill (Eds.), *Identifying perinatal depression and anxiety: evidence-based practice in screening, psychosocial assessment, and management* (pp. 228–239). West Sussex, UK: Wiley-Blackwell.

Phillips, J., Charles, M., Sharpe, L., & Matthey, S. (2009). Validation of the subscales of the Edinburgh Postnatal Depression Scale in a sample of women with unsettled infants. *Journal of Affective Disorders*, *118*, 101–112. https://doi.org/10.1016/j.jad.2009.02.004

R Core Team. (2018). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rojas, G., Fritsch, R., Solís, J., González, M., Guajardo, V., & Araya, R. (2006). Calidad de vida de mujeres deprimidas en el posparto. *Revista Médica de Chile*, *134*, 713–720. https://doi.org/10.5067/S0034-98872006000600006

Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, … Dunbar, G. C. (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I.): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *The Journal of Clinical Psychiatry*, *59*, 22–33.

StataCorp. (2015). *Stata Statistical Software: Release 14*. College Station, TX: StataCorp.

Stein, A., Pearson, R., Goodman, S. H., Rapa, E., Rahman, A., McCallum, M., … Pariante, C. M. (2014). Effects of perinatal mental disorders on the fetus and child. *The Lancet*, *384*, 1800–1819. https://doi.org/10.1016/S0140-6736(14)61277-0

Steyerberg, E. W. (2009). *Clinical prediction models: A practical approach to development, validation, and updating*. New York, NY: Springer.

Thompson, C. G., Kim, R. S., Aloe, A. M., & Becker, B. J. (2017). Extracting the variance inflation factor and other multicollinearity diagnostics from typical regression results. *Basic and Applied Social Psychology*, *39*, 81–90. https://doi.org/10.1080/01973533.2016.1277529

Vilagut, G., Forero, C. G., Pinto-Meza, A., Haro, J. M., Graaf, R., de, Bruffaerts, R., … ESEMeD Investigators (2013). The mental component of the Short-Form 12 Health Survey (SF-12) as a measure of depressive disorders in the general population: Results with three alternative scoring methods. *Value in Health*, *16*, 564–573. https://doi.org/10.1016/j.jval.2013.01.006

Vrieze, S. I. (2012). Model selection and psychological theory: A discussion of the differences between the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC). *Psychological Methods*, *17*, 228–243. https://doi.org/10.1037/a0027127

Ware, J. E., Jr., & Sherbourne, C. D. (1992). The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Medical Care*, *30*, 473–483.

Yawn, B. P., LaRusso, E. M., Bertram, S. L., & Bobo, W. V. (2015). When screening is policy, how do we make it work? In J. Milgrom, & A. W. Gemmill (Eds.), *Identifying perinatal depression and anxiety: Evidence-based practice in screening, psychosocial assessment, and management* (pp. 32–50). West Sussex, UK: Wiley-Blackwell.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Martínez P, Magaña I, Vöhringer PA, Guajardo V, Rojas G. Development and validation of a three-item version of the Edinburgh Postnatal Depression Scale. *J Clin Psychol*. 2020;1–14. https://doi.org/10.1002/jclp.23041