Factor structure of PHQ-9 in perinatal period: A longitudinal study

Estructura factorial del PHQ-9 en el periodo perinatal: Un estudio longitudinal

María F. Rodríguez-Muñoz [1] https://orcid.org/0000-0002-1908-0405
Katina Kovacheva [2]* https://orcid.org/0000-0002-3508-8407
María Dolores López-Salmerón [3] https://orcid.org/0000-0003-4633-464X
María Provencio [4] https://orcid.org/0000-0002-4918-9567
Cristina Soto-Balbuena [1] https://orcid.org/0009-0005-5421-0497
Marta Martínez-Marquínez [5]

[1] Universidad Nacional de Educación a Distancia, Madrid, España. E-mail: mfrodriguez@psi.uned.es; cristsobal@yahoo.es
[2] Universidad Miguel Hernández de Elche, Elche, España. E-mail: kilieva@umh.es
[3] Universidad Camilo José Cela, Villanueva de la Cañada, España. E-mail: mdolores.lopezs@ucjc.edu;
[4] Universidad a Distancia de Madrid, Madrid, España. E-mail: maria.provencio@udima.es
[5] Servicio de Salud del Principado de Asturias. E-mail: marta.martinez@sespa.es

ABSTRACT

Background: Screening for perinatal depression (PD) is challenging due to the overlap of symptoms with pregnancy and postpartum. The PHQ-9 is a commonly used diagnostic tool, and understanding its factor structure is crucial. **Aim:** To assess the three-factor factor structure of the PHQ-9 and its longitudinal stability during the perinatal stage. **Methods:** The study involved 205 women from a Spanish hospital, collecting data during the third trimester, immediate postpartum and six weeks postpartum. Statistical analyses, including confirmatory factor analysis, were performed to assess the stability of the factor structure and reliability of the PHQ-9. **Results:** The three-factor model of the PHQ-9, comprising cognitive-affective, somatic and pregnancy-related symptoms, was stable during pregnancy and postpartum, except in the immediate postpartum period. Analyses supported the validity of the PHQ-9 during the perinatal stage. **Conclusions:** Understanding the factors contributing to depressive symptoms during the perinatal period is crucial. Although the three-factor model is robust, caution is advised in the immediate postpartum period due to symptom overlap. The PHQ-9 is a reliable tool for assessing and intervening in depressive symptoms during the perinatal period.
Keywords: perinatal period, depression, diagnosis, PHQ-9.

RESUMEN

Antecedentes: La detección de la depresión perinatal (DP) es desafiante debido a la superposición de síntomas con el embarazo y el posparto. El PHQ-9 es una herramienta de diagnóstico utilizada, y comprender su estructura factorial es crucial. **Objetivo:** Evaluar la estructura factorial de tres factores del PHQ-9 y su estabilidad longitudinal durante la etapa perinatal. **Método:** El estudio involucró a 205 mujeres de un hospital español, recolectando datos durante el tercer trimestre, el posparto inmediato y seis semanas después del parto. Se realizaron análisis estadísticos, incluyendo el análisis factorial confirmatorio, para evaluar la estabilidad de la estructura factorial y la fiabilidad del PHQ-9. **Resultados:** El modelo de tres factores del PHQ-9, que comprende síntomas cognitivo-afectivos, somáticos y relacionados con el embarazo, se mantuvo estable durante el embarazo y el posparto, excepto en el período inmediato posparto. Los análisis respaldaron la validez del PHQ-9 durante la etapa perinatal. **Conclusiones:** Comprender los factores que contribuyen a los síntomas depresivos durante el período perinatal es crucial. Aunque el modelo de tres factores es robusto, se recomienda precaución en el período inmediato posparto debido a la superposición de síntomas. El PHQ-9 es una herramienta confiable para evaluar e intervenir en los síntomas depresivos durante la etapa perinatal.
Palabras clave: período perinatal, depresión, diagnóstico, PHQ-9.

M. F. Rodríguez-Muñoz, K. Kovacheva, M. D. López-Salmerón, M. Provencio, C. Soto-Balbuena, M. Martínez-Marquínez

Introduction

Depression is a mental disorder characterized by low mood, a loss of interest or pleasure in daily activities for several weeks, and a variety of additional symptoms (American Psychiatric Association Publishing, 2022), including changes in appetite and sleep or even suicidal ideation or behavior. Depression is a common disorder, although often misdiagnosed, that can affect women during the perinatal period (Marcos-Nájera et al., 2021). Approximately one in five women develop perinatal depression ([PD]; National Childbirth Trust, 2017). Particularly, in Spain, the prevalence of depression in women during the antenatal period is approximately 15%, being higher in the first and third trimesters (Míguez and Vázquez, 2021). Meanwhile, in the postnatal period, the prevalence amounts to approximately 27%, with many cases of depression in untreated pregnancy (Besteiro et al., 2001). Approximately two months after delivery, the maximum peak of prevalence was observed, decreasing over time, except for cases of major depression, which follow an upward trend (Vázquez and Míguez, 2022). According to Motrico et al. (2022), these figures increased after the onset of the COVID-19 pandemic, reaching a prevalence of 47%. PD is associated with negative repercussions on the woman herself, the baby, family dynamics, and society (Slomian et al., 2019). Hence, PD identification and treatment is a public health priority (Meaney, 2018).

The detection of depression during the perinatal stage presents some challenges due to the similarity of the symptomatology of pregnancy and postpartum (Marcos-Nájera et al., 2018). Determining whether sleeping difficulties, changes in appetite or weight, and fatigue are specific somatic symptoms of this stage or are due to depression is important (Yonkers et al., 2009).

The number of validated tools for PD diagnosis is increasing, among which the Patient Health Questionnaire-9 (PHQ-9) stands out (Kroenke et al., 2001; Kroenke and Spitzer, 2002). The PHQ-9 is a simple and quick self-assessment instrument validated in different contexts and languages, including Spanish (Diez-Quevedo et al., 2001; Zhong et al., 2014). Several studies have already studied its psychometric properties (Lara et al., 2015) and factorial structure (Familiar et al., 2015; Marcos-Nájera et al., 2018; Zhong et al., 2014). However, the results showed differences regarding the number of factors observed, with some indicating a unifactorial model (Familiar et al., 2015), others bifactorial (somatic and non-somatic; Zhong et al., 2014), and even trifactorial (cognitive-affective, somatic, and pregnancy-related symptoms; Marcos-Nájera et al., 2018). The latter structure was chosen for our analysis due to its specific validation in Spanish samples of pregnant women. Although one of the factors consists of only two items, we believe these items capture specific and clinically relevant symptoms associated with the perinatal period. Thus, we consider this factor significant despite its limited number of items.

In the trifactorial model proposed by Marcos-Nájera et al. (2018), the items loading on their cognitive-affective factor include four symptoms: anhedonia, depression/hopelessness, feelings of failure, and thoughts of suicide or self-harm. Meanwhile, those on the somatic factor 2 include concentration difficulties and psychomotor retardation/agitation, and pregnancy-related symptoms 3 include sleep problems, fatigue, and lack of appetite. Marcos-Nájera et al. (2018) compared the two-factor and the three-factor models. The researchers observed that the latter is superior, which agrees with another study that also studied the factor structure of pregnant women with the BDI-II questionnaire (Alexander et al., 2014). Both studies concluded that the factor, composed of items related to pregnancy-related problems, could facilitate differentiation with the symptoms present during PD. The global score of the scale allows us to visualize the overall severity of depressive symptoms. However, the identification of the factors present could facilitate clinicians to approach their intervention. Given that the three-factor structure proposed by Marcos-Nájera et al. (2018) has been the only one tested in Spanish samples of pregnant women, we decided to focus our analysis on this model. We acknowledge that one of the factors consists of only two items, which could be seen as a limitation. However, we believe that these items represent clinically relevant symptoms specific to the perinatal period that warrant further exploration. The lack of concordance between studies on the factors present in the PHQ-9, in addition to the evolution of these factors in pregnancy and postpartum, requires a considerable number of studies to clarify the factors of the questionnaire.

The aim of this study is to evaluate the factor structure proposed by Marcos-Nájera et al. (2018), to determine whether these factor structures are maintained longitudinally throughout the perinatal stage in a sample of women receiving healthcare during pregnancy and postpartum.

Method

Participants

A total of 205 women were recruited in the obstetrics service of the Hospital Universitario de Asturias, a public health center in the Principality of Asturias, Spain, through a convenient sampling method. The study received ethics committee approval (REF.N. 18/18). Each participant completed a series of questionnaires privately.

The sample size was estimated based on a prevalence of 13%, as indicated by Rodríguez-Muñoz et al. (2017) during the antenatal period and Vázquez & Miguez (2022) in the postnatal period, with a margin of error of 5%. Based on this approach, approximately 248 participants were needed to ensure adequate representativeness across the three measurement points.

As shown in Figure 1, out of the 205 who completed the measures in the first trimester (Time 1), 138 (67.32%) completed the measures in the immediate postpartum period (Time 2), of which 89 (43.41%) completed the measures at six weeks postpartum (Time 3).
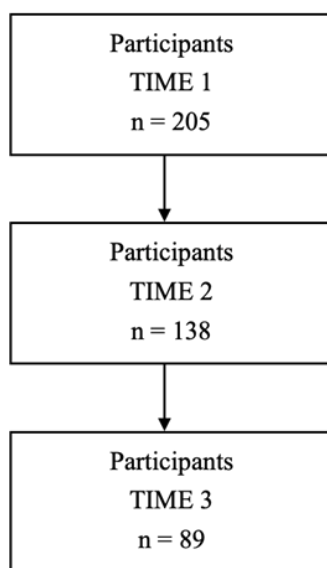

Figure 1. Participants

Despite not reaching the ideal sample size, we believe that **205 participants** is sufficient for meaningful analyses of the factor structure of the PHQ-9 at different time points. This figure reflects a significant commitment from participants throughout the study, even considering the common attrition rates in studies involving this population (Le et al., 2008).

When analyzing the sociodemographic characteristics of women who did not complete follow-ups at the time points compared with those who completed all three measurements, the results indicated that the differences in terms of age ($F = 0.910$, $p < 0.404$), education level ($\chi2 = 3.575$, $p < 0.893$), employment status ($\chi2 = 4.556$, $p < 0.602$), and cohabitation with partner ($\chi2 = 7.912$, $p < 0.442$) were insignificant.

Procedure

To participate in the study, we established certain eligibility criteria for the participants. These criteria included being in the third trimester of pregnancy, receiving prenatal services at the hospital, and having Spanish language skills.

Prior to participation, participants were informed of the purpose of the study, which was to assess the psychological well-being of pregnant women and mothers. Participants were asked for their written consent to participate in the study.

While waiting for their prenatal appointment, participants completed a series of questionnaires. The PHQ-9, along with other questionnaires, was administered by midwives during routine prenatal check-ups and postpartum follow-ups. Subsequently, participants were followed up for a second measurement just after delivery and a third measurement one year after delivery.

Instruments

PHQ-9 (Kroenke et al., 2001; Kroenke and Spitzer, 2002) is a nine-item assessment tool designed to assess the specific symptoms of Major Depressive Disorder using DSM-IV criteria. The structure of the PHQ-9 is also appropriate for DSM-5, which is the updated version (Mitchell et al., 2016). Each item is scored on a four-point Likert scale, ranging from 0 (never) to 3 (almost every day). Higher scores indicate greater severity of depressive symptoms; the total score ranges from 0 to 27. Researchers recommended a cut-off point of 10 to be at risk for major depressive disorder (Gilbody et al., 2007; Kroenke and Spitzer, 2002; Manea et al., 2012, 2015). The PHQ-9 has been validated in samples of hospitalized patients in medical centers (Diez-Quevedo et al., 2001 and primary care in Spain (Pinto-Meza et al., 2005). The PHQ-9 is also widely recommended for assessing women at any time during pregnancy and of different nationalities (Lara et al., 2015).

Data analysis

For the descriptive analysis of the PHQ-9 items, we provided statistics such as mean, standard deviation, skewness, kurtosis, and the existence of floor effect and ceiling effect. Internal consistency was evaluated using both Cronbach's alpha and McDonald's ordinal omega, with the IBM SPSS statistics program (version 24), setting a confidence level of 95% at three time points.

The factor structure was examined using SPSS Analysis of Moment Structures for confirmatory factor analysis. The factor structure was examined using SPSS Analysis of Moment Structures for confirmatory factor analysis (CFA), following the three-factor model proposed by Marcos-Nájera et al. (2018), which includes: Factor 1: Cognitive-affective (items 1, 2, 6, and 9); Factor 2: Pregnancy symptoms (items 3, 4, and 5); Factor 3: Somatic (items 7 and 8).

To assess the fit of a model properly, we used several indices. One of the relevant indices to determine the fit of a model is the standard chi-square, a parsimonious measure obtained by dividing the chi-square value by the degrees of freedom. A good fit is expected to be <2.00 (Ullman & Bentler, 2003), while an acceptable fit is < 5.00 (Taylor & Todd, 1995). Another important index is the root mean square error of approximation (RMSEA), which is considered acceptable when it is less than or equal to ≤0.08 and a good fit when it is ≤0.05 (Kline, 1998). Two incremental fit indices should also be taken into account: the

comparative fit index (CFI), which is considered a good fit when ≥0.90 (Bentler, 1990), and the normalized fit index (NFI), which is also considered a good fit when ≥0.90 (Kline, 1998).

Factor loadings were obtained through maximum likelihood factor analysis, using Varimax rotation to enhance the interpretability of the factors.

Results

Descriptive statistics

As shown in Table 1, the mean age of the participants was 34 years old at the beginning of the study. Most participants live with their partner (87.5%) and are working (78%). In terms of the level of academic studies, the majority of participants (58.2%) had university studies, 34.6% had secondary studies, and 7.2% had basic studies.

Table 1. Main characteristics of the participants (N=205)

Sociodemographic	% or mean	±SD	Minimum	Maximum
Age	34.92	4.65	20	47
Level of education				
Basic education	7.2%			
Intermediate studies	34.6%			
Higher education	58.2%			
Living with a partner				
Yes	87.5%			
No	12.5%			
Unemployment status				
Yes	22%			
No	78%			

The descriptive statistics of the participants' responses to the PHQ-9 items are shown in Table 2. Considering the three measurement times, the item with the highest score was 4 (time 1: M = 1.30, SD = 0.82; time 2: M = 1.10, SD = 0.82; time 3: M = 1.18, SD = 0.84), while the item with the lowest score was 9 (time 1: M = 0.03, SD = 0.17; time 2: M = 0.01, SD = 0.08; time 3: M = 0.02, SD = 0.15). Regarding the floor and ceiling effect, all items showed a floor effect, that is, lower scores were relatively common. Meanwhile, a ceiling

effect was not observed in any of the items because less than 10% of the participants reached the maximum scores in each item.

Table 2. Descriptive statistics of the PHQ-9 items at the three measurement times.

Item	Min-Max	Mean	SD	Skewness	Kurtosis	% of responses with a score of 0	% of responses with a score of 3
Time 1 (N = 205)							
1. Little interest or pleasure in doing things.	0-3	0.47	0.66	1.57	2.95	60.2	2.3
2. You have felt down, depressed, or hopeless.	0-3	0.41	0.55	1.11	1.20	62	0.5
3. Difficulty falling asleep or staying asleep, or you have slept too much.	0-3	1.17	0.92	0.40	-0.68	26.3	9.7
4. You have felt tired or low in energy.	0-3	1.30	0.82	0.74	1.33	12.4	7.8
5. Poor appetite or have overeaten.	0-3	0.76	0.89	1.28	2.00	47.2	4.6
6. Felt bad about yourself - that you are a failure or that you have looked bad to yourself or your family.	0-3	0.40	0.68	1.77	2.84	69.4	1.9
7. Had difficulty concentrating on things 0 1 2 3 such as reading the newspaper or watching television.	0-3	0.39	0.66	1.75	2.75	70	1.4
8. You have been moving or talking so slowly that other people might notice, or conversely - you have been so restless or agitated, that you have been moving around a lot more than usual.	0-3	0.32	0.63	2.10	4.25	75.6	1.4
9. Have you thought you would be better off dead, or have you thought of hurting yourself in any way?	0-3	0.03	0.17	5.31	26.52	96.8	0
Time 2 (N = 138)							
1. Little interest or pleasure in doing things.	0-3	0.43	0.83	2.47	7.46	72.1	2.9
2. You have felt down, depressed, or hopeless.	0-3	0.42	0.62	1.40	1.70	64.7	0.7
3. Difficulty falling asleep or staying asleep, or you have slept too much.	0-3	0.62	0.86	1.29	0.84	58	5.1
4. You have felt tired or low in energy.	0-3	1.10	0.82	0.74	0.33	20.9	8.6

5. Poor appetite or have overeaten.	0-3	0.65	0.85	1.16	0.48	56.1	4.3
6. Felt bad about yourself - that you are a failure or that you have looked bad to yourself or your family.	0-3	0.34	0.64	2.20	5.25	73.4	2.2
7. Had difficulty concentrating on things 0 1 2 3 such as reading the newspaper or watching television.	0-3	0.47	0.81	1.75	2.21	69.8	4.3
8. You have been moving or talking so slowly that other people might notice, or conversely - you have been so restless or agitated, that you have been moving around a lot more than usual.	0-3	0.22	0.60	3.07	9.68	84.8	2.2
9. Have you thought you would be better off dead, or have you thought of hurting yourself in any way?	0-3	0.01	0.08	11.79	139.00	99.3	0.7
Time 3 (N = 89)							
1. Little interest or pleasure in doing things.	0-3	0.42	0.65	1.28	0.46	66.7	0
2. You have felt down, depressed, or hopeless.	0-3	0.44	0.65	1.43	1.89	63.3	1.1
3. Difficulty falling asleep or staying asleep, or you have slept too much.	0-3	0.61	0.94	1.34	0.59	64.4	6.7
4. You have felt tired or low in energy.	0-3	1.18	0.84	0.68	0.15	17.8	10
5. Poor appetite or have overeaten.	0-3	0.61	0.83	1.20	0.59	57.8	3.3
6. Felt bad about yourself - that you are a failure or that you have looked bad to yourself or your family.	0-3	0.39	1.29	6.59	52.54	80.4	1.1
7. Had difficulty concentrating on things 0 1 2 3 such as reading the newspaper or watching television.	0-3	0.50	0.76	1.59	2.16	63.7	3.3
8. You have been moving or talking so slowly that other people might notice, or conversely - you have been so restless or agitated, that you have been moving around a lot more than usual.	0-3	0.22	0.59	3.15	10.55	84.4	2.2
9. Have you thought you would be better off dead, or have you thought of hurting yourself in any way?	0-3	0.02	0.15	6.47	40.90	97.7	0

Internal validity tests

Table 3 presents the item sum means, standard deviations, Cronbach's α, and McDonald's omega coefficient for the PHQ-9 across the three temporal measures. The internal consistency of the PHQ-9 ranged from 0.78 to 0.88, thereby suggesting good internal reliability. This result indicated that the variables in the scale are related and measure a common construct. In addition, good fit indices from the factor analysis indicate that the PHQ-9 maintains its structure across measurement points.

Table 3. Analysis of the scale in the three measurement times.

Time	M	SD	Asymmetry	Kurtosis	α	ω	χ 2/df	RMSEA	CFI	NFI
PHQ1	5.32	3.81	1.04	0.76	.78	.79	1.86	0.05	0.95	0.90
PHQ2	4.30	4.09	1.69	3.12	.80	.82	2.79	0.67	0.87	0.82
PHQ3	4.37	4.68	1.75	2.59	.88	.89	1.11	0.02	0.99	0.93

α: Cronbach's alpha; ω: The Omega coefficient, χ 2/df= chi-square/degrees of freedom; RMSEA = root mean square error of approximation; CFI = comparative fit index; NFI = normalized fit index;

Meanwhile, in terms of the factor analysis indices, the measures taken during the pregnancy (RMSEA = 0.05 good fit, CFI = 0.95 excellent fit, NFI = 0.90 good fit) and postpartum (RMSEA = 0.02 good fit, CFI = 0.99 excellent fit, NFI = 0.93 good fit) had a good fit to the three-factor model. For the measure taken in the immediate postpartum period (RMSEA = 0.67 moderate fit, CFI = 0.87 moderate fit, NFI = 0.82 moderate fit), the fit indices provided indicate a moderate fit of the model, making it appropriate to consider rejecting the model in this period.

Reliability

The reliability of the scores, estimated through Cronbach's α and McDonald's omega coefficients (Table 4), showed a value of >0.78 for the scale at the three measurement times (α = 0.78 and ω = 0.79 at time 1; α = 0.81 and ω = 0.82 at time 2; and α = 0.88 and ω = 0.89 at time 3). These values indicate that the PHQ-9 maintains adequate internal consistency across all measurement points. The item-test correlation ranged from 0.32 (item 9) to 0.70 (item 6) at time 1, and from 0.23 (item 9) to 0.73 (item 4) and 0.38 (item 9) to 0.81 (item 6) at times 2 and 3, respectively. The removal of one item did not show an increase in reliability coefficients, except with the removal of item 9 in times 2 and 3, where it would show an increase of 0.01.

M. F. Rodríguez-Muñoz, K. Kovacheva, M. D. López-Salmerón, M. Provencio, C. Soto-Balbuena, M. Martínez-Marquínez

Table 4. Item-Test correlations and Cronbach´s Alpha and McDonald´s Omega coefficients

Variables	Time 1 (*n* = 205)	Time 2 (*n* = 138)	Time 3 (*n* = 89)
Cronbach´s alpha/ McDonald´s omega			
PHQ-9 Total score	.78 / .79	.81 / .82	.88 / .89
Cronbach´s alpha item – test correlation			
Item1	0.56	0.54	0.72
Item2	0.65	0.70	0.77
Item3	0.64	0.72	0.76
Item4	0.68	0.73	0.78
Item5	0.70	0.67	0.74
Item6	0.67	0.68	0.81
Item7	0.60	0.72	0.78
Item8	0.55	0.58	0.77
Item9	0.32	0.23	0.38
Cronbach´s alpha if item is deleted			
Item1	0.76	0.80	0.86
Item2	0.75	0.78	0.86
Item3	0.76	0.78	0.86
Item4	0.74	0.77	0.86
Item5	0.75	0.79	0.85
Item6	0.75	0.78	0.86
Item7	0.76	0.78	0.86
Item8	0.76	0.79	0.86
Item9	0.78	0.82	0.89

In addition, we assessed the internal consistency of the PHQ-9 scale across the three measurement points, incorporating the three dimensions: cognitive-affective, pregnancy symptoms, and somatic. Cronbach's alpha (α) and McDonald's omega (ω) coefficients were calculated for each time point.

As shown in Table 5, at the first measurement point, the scale demonstrated adequate reliability with an α of 0.694 and a ω of 0.711. At the second time point, reliability improved, yielding an α of 0.776 and a ω of 0.794. By the third measurement, internal consistency remained strong, with an α of 0.822 and a ω of 0.837.

Overall, these findings indicate that the PHQ-9 scale exhibits satisfactory to good internal consistency across all measurement points, confirming its reliability for assessing depression in the studied population. The gradual increase in reliability suggests that the scale items effectively capture the construct of depression over time.

Factor Loadings

Table 6 presents the factor loadings of the PHQ-9 items at the three measurement points (first trimester, immediate postpartum, and six weeks postpartum).

The PHQ-9 scale is designed to assess depressive symptoms in the perinatal population, capturing a range of emotional states and behaviors associated with depression. Factor loadings indicate the strength and direction of the relationship between each item and the underlying depression factor, with values above 0.30 considered significant. In the first trimester, all items demonstrated acceptable loadings, suggesting that they adequately represent the depression construct. Notably, items such as item 6 and item 4 displayed the highest loadings, indicating their strong association with the overall depression factor.

Across the three measurement points, the factor loadings remained relatively stable, reflecting the consistency of the instrument in capturing the essence of depressive symptoms over time. However, item 9, which asks, *"Have you thought you would be better off dead, or have you thought of hurting yourself in any way?"*, consistently exhibited lower loadings across all measurement points. This suggests that it may contribute less to the overall construct of depression in this context, highlighting its more specific and sensitive nature regarding suicidal ideation.

Conclusions

Given the high PD prevalence and severity, conducting research that provides a detailed and comprehensive understanding of this disorder is crucial. In this way, we can improve the prediction accuracy and enable timely intervention in PD treatment (Žutić, 2023). The present study evaluated the reliability and structural model of three factors of the PHQ-9 in a sample of Spanish-speaking women throughout the perinatal stage in an urban hospital in Spain. Identifying and distinguishing these factors, such as cognitive-affective, somatic, and postpartum factors, are fundamental because they allow an adequate derivation and a complete understanding of depressive symptoms in this specific context. We can obtain information about the general severity of depressive symptoms by reviewing the overall PHQ-9 score. However, by identifying the composite dimensions and latent factors of the PHQ-9, we can expand its utility and allow researchers and clinicians to approach depression as a multidimensional construct. This detailed and disaggregated perspective of the PHQ-9

M. F. Rodríguez-Muñoz, K. Kovacheva, M. D. López-Salmerón, M. Provencio, C. Soto-Balbuena, M. Martínez-Marquínez

may provide an accurate understanding of specific aspects of depression and target interventions and treatments effectively.

Consistent with the study by Marcos-Najera et al. (2018), our results supported the appropriateness of the three-factor model during pregnancy. Moreover, this pattern was observed in the postpartum period but not in the immediate postpartum period. Therefore, the second factor, composed of pregnancy-related symptoms, such as sleep problems, feeling tired or lack of energy, and lack of appetite, can also be interpreted as postpartum symptoms due to adaptation to motherhood. However, in the immediate postpartum period, the intense physical and emotional involvement associated with this period may cause these symptoms to overlap to a greater extent. These changes include body alterations, fatigue, and an emotional burden due to the decrease in steroid and peptide hormones (Carrillo-Mora et al., 2021; Carrizo et al., 2020). At that time, low mood and mild depressive symptoms were known as maternity blues. PD assessment instruments are unsuitable for this context, and tools specifically designed for this period are needed as a changeable and transitory period (Tosto et al., 2023).

We have strengthened the discussion on the dimensionality of the PHQ-9 scale. By identifying the cognitive-affective, somatic, and pregnancy/postpartum-related factors, this approach provides a more nuanced understanding of perinatal depression. This multidimensional perspective ensures that depression is not viewed as a singular construct, but rather as a complex disorder with distinct domains. Addressing each factor is crucial for delivering more targeted and effective clinical interventions, tailored to the specific needs of women during the different stages of the perinatal period.

The present study has several limitations. First, we did not employ other measures of depression, such as a self-report and a clinical diagnosis. Consequently, we cannot provide information on the concurrent and criterion validity of the PHQ-9. Both measures have been previously validated in previous studies, such as those conducted by Flynn et al. (2011), Zhong et al. (2014), and Gilbody et al. (2007). The sample in our study was mostly composed of couples with high educational levels, working adults, and residing in urban settings in Spain. Low socioeconomic status constitutes a significant risk factor for PD, as previous studies have indicated (Beck, 2001; Robertson et al., 2004). Therefore, extending the research to diverse populations with different socioeconomic levels would be valuable. This research direction would provide a complete picture of the relationship between the factors analyzed and PD, thereby improving future mental health interventions and support strategies during the perinatal period. Furthermore, the relatively small sample size is a limitation, as it may impact the generalizability of the findings. Future research should aim to include larger and more diverse samples to validate the current findings in broader contexts.

Additionally, another limitation of the study lies in our decision to focus primarily on the three-factor model proposed by Marcos-Nájera et al. (2018), based on its validation in

Spanish samples and model fit indices. Although this model has relevance in the perinatal context, its exclusive use may limit the generalizability of the findings, as alternative structures (such as the one-factor and two-factor models) could provide valuable insights. Future studies should explore multiple factorial models to ensure a comprehensive understanding of the PHQ-9's structure in perinatal populations.

Given the high PD prevalence and severity, research that provides a detailed and comprehensive understanding of this disorder, which is the aim of this study, is encouraged to accelerate scientific advances and promote the provision of high-level perinatal mental healthcare.

On the other hand, to highlight the importance of future research addressing the health care cost of underdiagnosis of PD. Such research will be essential for a deep understanding of maternal health and child development implications and for assessing the economic costs associated with treatment. The findings obtained will be critical for the implementation of effective strategies to prevent, detect, and treat PD, resulting in significant improvement in the health and well-being of mothers and their children.

The clinical implications of PD diagnosis are of paramount importance, and the possibility that we are overestimating cases must be addressed; it does not necessarily represent a serious problem if we are indeed overestimating. Avoiding overestimation is preferable to underdiagnosis because early and accurate diagnosis is essential to ensure adequate and timely care (Rodríguez-Muñoz et al., 2023).

The factors identified in this study are maintained in pregnancy and the postpartum period, suggesting that the underlying structure of depressive symptoms in PD is robust. However, caution is needed to consider the specific context of the immediate postpartum period, providing the necessary support and understanding for accurate assessment and effective intervention at this sensitive and physiologically charged stage.

In conclusion, using the PHQ-9, a psychometrically sound tool, and its appropriate interpretation enables healthcare professionals to assess depressive symptoms and provide timely interventions systematically. We must understand the specific needs of women, design tailored policies, and promote perinatal mental health by strengthening support systems and ensuring equitable access to medical and psychological care services.

References

Alexander, L.-A., Rodríguez, M. M., Perry, D. F., & Le, H. N. (2014). The latent symptom structure of the Beck depression inventory: second edition in Latina pregnant women. *Maternal and Child Health Journal, 18*, 1132–1141. https://doi.org/10.1007/s10995-013-1343-5

American Psychiatric Association Publishing. (2022). *Diagnostic and statistical manual of mental disorders-Text Revision* (DSM-5-TR). American Psychiatric Association Publishing. https://doi.org/10.1176/appi.books.9780890425787

Beck, C. T. (2001). Predictors of postpartum depression: an update. *Nursing Research*, *50*, 275–285. https://doi.org/10.1097/00006199-200109000-00004

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*(2), 238-246. https://doi.org/10.1037/0033-2909.107.2.238

Besteiro, E. M., Aliaga, M. T., Pineda, M., Alegre, J. M., Berlanga, J. F., & Martín, P. B. (2001). La depresión postparto: un estudio exploratorio con una muestra de mujeres valencianas. *Revista de Psicopatología y Psicología Clínica*, *6*(1), 37-48. https://doi.org/10.5944/rppc.vol.6.num.1.2001.3903

Carrillo-Mora, P., García-Franco, A., Soto-Lara, M., Rodríguez-Vásquez, G., Pérez-Villalobos, J., & Martínez-Torres, D. (2021). Cambios fisiológicos durante el embarazo normal. *Revista de la Facultad de Medicina* (México), *64*(1), 39-48. https://doi.org/10.22201/fm.24484865e.2021.64.1.07

Carrizo, E., Domini, J., Quezada, R. Y. J., Serra, S. V., Soria, E. A., & Miranda, A. R. (2020). Variaciones del estado cognitivo en el puerperio y sus determinantes: una revisión narrativa. *Ciência & Saúde Coletiva*, *25*(8), 3321–3334. https://doi.org/10.1590/1413-81232020258.26232018

Diez-Quevedo, C., Rangil, T., Sanchez-Planell, L., Kroenke, K., & Spitzer, R. L. (2001). Validation and utility of the Patient Health Questionnaire in diagnosing mental disorders in 1003 general hospital Spanish inpatients. *Psychosomatics Medicine, 63*, 679–686. https://doi.org/10.1097/00006842-200107000-00021

Familiar, I., Ortiz-Panozo, E., Hall, B., Vieitez, I., Romieu, I., Lopez-Ridaura, R., & Lajous, M. (2015). Factor structure of the Spanish version of the Patient Health Questionnaire-9 in Mexican women. *International Journal of Methods in Psychiatric Research*, *24*, 74–82. https://doi.org/10.1002/mpr.1461

Flynn, H. A., Sexton, M., Ratliff, S., Porter, K., & Zivin, K. (2011). Comparative performance of the Edinburgth Postnatal Depression Scaleand the Patient Health Questionnaire-9 in pregnant and pospartum women seeking psychiatric services. *Psychiatry Research*, *187*(1-2), 130–134. https://doi.org/10.1016/j.psychres.2010.10.022

Gilbody, S., Richards, D., & Barkham, M. (2007). Diagnosing depression in primary care using self-completed instruments: UK validation of PHQ-9 and CORE-OM. *The British Journal of General Practice: the Journal of the Royal College of General Practitioners*, *57*(541), 650–652.

Kline, R. B. (1998). *Principles and practice of structural equation modeling.* (3rd ed.). Guilford Press.

Kroenke, K., & Spitzer, R. (2002). The PHQ -9: a new depression diagnostic and severity measure. *Psychiatric Annals 32,* 509–515. https://doi.org/10.3928/0048-5713-20020901-06

Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). The PHQ-9: validity of a brief depression severity measure. *Journal of General Internal Medicine 16,* 606–613. https://doi.org/10.1046/j.1525-1497.2001.016009606.x

Lara, M. A., Navarrete, L., Nieto, L., Martín, J. P., Navarro, J. L., & Lara-Tapia, H. (2015). Prevalence and incidence of perinatal depression and depressive symptoms among Mexican women. *Journal of affective disorders, 175,* 18–24. https://doi.org/10.1016/j.jad.2014.12.035

Le, H. N., Lara, M. A., & Perry, D. F. (2008). Recruiting Latino women in the U.S. and women in Mexico in postpartum depression prevention research. *Archives of women's mental health*, *11*(2), 159–169. https://doi.org/10.1007/s00737-008-0009-6

Marcos-Nájera, R., Le, H.-N., Rodríguez-Muñoz, M. F., Olivares Crespo, M. E., & Izquierdo Mendez, N. (2018). The structure of the Patient Health Questionnaire-9 in pregnant women in Spain. *Midwifery*, *62*, 36–41. https://doi.org/10.1016/j.midw.2018.03.011

Meaney, M. J. (2018). Perinatal Maternal Depressive Symptoms as an Issue for Population Health. *The American Journal of Psychiatry*, *175*(11), 1084–1093. https://doi.org/10.1176/appi.ajp.2018.17091031

Míguez, M. C., y Vázquez, M. B. (2021). Prevalence of depression during pregnancy in Spanish women: Trajectory and risk factors in each trimester. *International Journal of Environmental Research and Public Health*, *18*(13), 6789. https://doi.org/10.3390/ijerph18136789

Mitchell, J., Trangle, M., Degman, B., Gabert, T., Haight, B., Kessler, D., Mack, N., Mallen, E., Novak, H., Rossmiller, D., Setterlund, L., Somers, K., Valentino, N., & Vicent, S. (2016). Adult Depression in Primary Care. *Institute for Clinical Systems Improvement.*

Motrico, E., Domínguez-Salas, S., Rodríguez Domínguez, C., Gómez Gómez, I., Rodríguez Muñoz, M. D. L. F., y Gómez Baya, D. (2022). The impact of the COVID-19 pandemic on perinatal depression and anxiety: a large cross-sectional study in Spain. *Psicothema 34*(2), 200–208. https://doi.org/10.7334/psicothema2021.380

National Childbirth Trust. (2017). *The hidden half: Bringing postnatal mental illness out of hiding.* National Childbirth Trust. https://allcatsrgrey.org.uk/wp/wpfb-file/739-nct-thehiddenhalf-report-150dpi_0-pdf/

Pinto-Meza, A., Serrano-Blanco, A., Peñarrubia, M. T., Blanco, E., & Haro, J. M. (2005). Assessing depression in primary care with the PHQ-9: can it be carried out over the telephone?. *Journal of General Internal Medicine*, *20*(8), 738–742. https://doi.org/10.1111/j.1525-1497.2005.0144.x

Robertson, E., Grace, S., Wallington, T., Stewart, D. E. (2004). Antenatal risk factors for postpartum depression: a synthesis of recent literature. *General Hospital Psychiatry*, (26), 289–295. https://doi.org/10.1016/j.genhosppsych.2004.02.006

Rodríguez-Muñoz, M. F., Castelao-Legazpi, P. C., Olivares-Crespo, M. E., Soto-Balbuena, C., Izquierdo-Méndez, N., Ferrer-Barrientos, F. J., & Le, H. N. (2017). PHQ-2 como primer instrumento de cribado de la depresión prenatal. *Revista Española de Salud Públi*ca, *91*, 30.

Rodríguez-Muñoz, M. F., Motrico, E., Miguez, C., Chaves, C., Suso-Ribera, C., Duque, A., Salinas, M. G., Caparrós-González, R. A., Martin-Agudiez, N., Kovacheva, K., García-López, H. S., Vazquez-Batan, P., Peñacoba, C., & Osma, J. (2023). Perinatal Depression in the Spanish Context: Consensus Report from the General Council of Psychology of Spain. *Clínica y Salud*, *34*(2), 51 - 63. https://doi.org/10.5093/clysa2023a15

Slomian, J., Honvo, G., Emonts, P., Reginster, J., & Bruyère, O. (2019). Consequences of maternal postpartum depression: A systematic review of maternal and infant outcomes. *Women's Health (London, England)*, 15, 1745506519844044. https://doi.org/10.1177/1745506519844044

Taylor, S., & Todd, P.A. (1995). Understanding information technology usage: A test of competing models. *Information Systems Research, 6,* 144-176. https://doi.org/10.1287/isre.6.2.144

Tosto, V., Ceccobelli, M., Lucarini, E., Tortorella, A., Gerli, S., Parazzini, F., & Favilli, A. (2023). Maternity blues: A narrative review. *Journal of Personalized Medicine*, *13*(1). https://doi.org/10.3390/jpm13010154

Ullman, J. B., & Bentler, P. M. (2003). Structural equation modeling. In J. A. Schinka & W. F. Velicer (Eds.), *Handbook of psychology: Research methods in psychology* (Vol. 2, pp. 607–634). John Wiley & Sons.

Vázquez, M. B., & Míguez, M. C. (2022). *La depresión en el primer año posparto: periodos críticos* [Free paper presentation]. X Jornada de Salud Mental Perinatal, *Psicosomàtica y Psiquiatría*, (21).

Yonkers, K. A., Wisner, K. L., Stewart, D. E., Oberlander, T. F., Dell, D. L., Stotland, N., Ramin, S., Chaudron, L., & Lockwood, C. (2009). The management of depression during pregnancy: a report from the American Psychiatric Association and the American College of Obstetricians and Gynecologists. *General Hospital Psychiatry, 31*(5), 403–413. https://doi.org/10.1016/j.genhosppsych.2009.04.003

Zhong, Q., Gelaye, B., Fann, J. R., Sanchez, S. E., & Williams, M. A. (2014). Cross-cultural validity of the Spanish version of PHQ-9 among pregnant Peruvian women: a Rasch item response theory analysis. *Journal of Affective Disorders*, *158*, 148–153. https://doi.org/10.1016/j.jad.2014.02.012

Žutić, M. (2023). Biopsychosocial Models of Peripartum Depression: A Narrative Review. *Clínica y Salud*, *34*(2), 91 - 99. https://doi.org/10.5093/clysa2023a16