

# Propiedades Psicométricas de las Preguntas de Whooley y Arroll para el Cribado de la Depresión Prenatal en Población Española

## Validation of Whooley and Arroll Questions for Screening Prenatal Depression in Spanish Population

María F. Rodríguez-Muñoz<sup>1</sup>, Katina Kovacheva<sup>2</sup>, María Provencio<sup>3</sup>,  
María Dolores López-Salmerón<sup>4</sup>, Cristina Soto Balbuena<sup>5</sup> y Ana Fonseca<sup>6</sup>

### Resumen

La depresión prenatal (DP) es un problema de importancia mundial. Los profesionales sanitarios deberían disponer de herramientas de cribado sencillas, breves y basadas en la evidencia para garantizar las mejores prácticas en el manejo de la salud perinatal. El propósito de este estudio fue evaluar las propiedades psicométricas de las tres preguntas en el cribado de la depresión prenatal mediante el uso del cuestionario Patient Health Questionnaire-9 (PHQ-9) como *gold standard*. Métodos: Se trata de un estudio transversal, con una muestra compuesta por 405 mujeres embarazadas. Se analizaron la validez, fiabilidad y estructura interna de las preguntas de Whooley y Arroll. Se utilizó el PHQ-9 como patrón oro, y se realizaron análisis factoriales, de consistencia interna y curva ROC. Resultados: El análisis factorial indicó un solo factor con cargas entre .73 y .83. La escala mostró una fiabilidad adecuada ( $\omega=.79$ ) y buena validez discriminante (área bajo la curva=.80,  $p<.001$ ). La sensibilidad fue .48 y la especificidad .91, con un punto de corte de 1. Discusión: Las preguntas de Whooley y Arroll presentan propiedades psicométricas adecuadas para su uso como herramienta de cribado de la depresión prenatal en población española.

**Palabras clave:** depresión prenatal, cribado, validez, fiabilidad, estructura factorial

### Abstract

Introduction: Prenatal depression (PD) is a problem of worldwide importance. Health professionals should be provided with simple, brief, and evidence-based screening tools to ensure best practices in the management of perinatal health. The purpose of this study was to assess the validity of the three questions in screening for PD in a sample of Spanish pregnant women, using the Patient Health Questionnaire-9 as the gold standard. Methods: This cross-sectional study had a sample consisting of 405 women. Results: The results showed a single-factor scale with an adequate reliability index ( $\omega=.79$ ), whose area under the curve is .80,  $p<.001$ . Sensitivity and specificity were .48 and .91, respectively, with 1 as the cut-off point. Discussion: The psychometric properties of the Whooley and Arroll questions suggest that they are suitable for use as a screening tool for prenatal depression.

**Keywords:** prenatal depression, screening, sensitivity, and specificity

<sup>1</sup>Doctora en Psicología de la Salud. Catedrática, Departamento de Personalidad, Evaluación y Tratamiento Psicológico. Facultad de Psicología, Universidad Nacional de Educación a Distancia. C/ Juan del Rosal, 10 (28040) Madrid, España. Correo: mfrdruiguez@psi.uned.es. <https://orcid.org/0000-0002-1908-0405>

<sup>2</sup>Doctora en Psicología de la Salud. Departamento de Ciencias del Comportamiento y Salud, Universidad Miguel Hernández. Correo: kilieva@umh.es. <https://orcid.org/0000-0002-3508-8407>

<sup>3</sup>Doctora en Psicología. Profesora Contratada Doctora. Facultad de Psicología y Ciencias de la Salud, Universidad a Distancia de Madrid (UDIMA). Correo: maria.provencio@udima.es. <https://orcid.org/0000-0002-4918-9567>

<sup>4</sup>Doctora en Psicología de la Salud. Facultad de Psicología y Ciencias de la Salud, Universidad a Distancia de Madrid (UDIMA). Correo: mariadolores.lopez.s@udima.es. <https://orcid.org/0000-0003-4633-464X>

<sup>5</sup>Doctora en Psicología de la Salud. Hospital Universitario Central de Asturias. Correo: cristsobal@yahoo.es. <https://orcid.org/0009-0005-5421-0497>

<sup>6</sup>Doctora en Psicología de la Salud. Profesora. Facultad de Psicología y Pedagogía, Universidad de Coimbra. Correo: anadfonseca@fpce.uc.pt. <https://orcid.org/0000-0003-1395-1406>

## Introducción

La depresión perinatal (DPN) es un trastorno del estado de ánimo que afecta a las mujeres durante el embarazo y a las madres hasta 12 meses después del parto (Yu et al., 2021). Es el trastorno más frecuente en la etapa prenatal (Dagher et al., 2021), con una prevalencia que oscila entre el 14% y el 23% (Liu et al., 2022), y ha habido una tendencia creciente en la última década, principalmente como consecuencia de los desafíos impuestos por la pandemia del COVID-19 (Motrico et al., 2022). Existen cada vez más investigaciones sobre la depresión durante el embarazo, ya que desempeña un papel crucial en el establecimiento del estado postnatal (Shin et al., 2020). La DPN se caracteriza por la presencia de síntomas de tristeza, ansiedad, irritabilidad, cambios de humor, dificultades para dormir y falta de interés por actividades que antes resultaban placenteras. Estos síntomas se asocian a la pérdida de identidad y confianza en el rol materno, y las mujeres experimentan muy a menudo sentimientos de culpa y vergüenza. Este trastorno no solo afecta al bienestar y al funcionamiento social de la mujer, reduciendo su calidad de vida, sino que también compromete el bienestar y el desarrollo socioemocional del bebé y afecta a toda la dinámica familiar (Slomian et al., 2019). La literatura ha enfatizado la importancia del apoyo como factor protector contra la aparición de la depresión (Forner-Puntonet et al., 2021).

No evaluar la DPN y, en consecuencia, no tratarla supone un elevado coste para el sistema sanitario debido a las consecuencias para la madre, como empeoramiento de los síntomas, riesgo de suicidio (Al-Halabí et al., 2021; Legazpi et al., 2022), ansiedad (Soto-Balbuena et al., 2018; Soto-Balbuena et al., 2021), pérdida de productividad y reducción de la calidad de vida, y para el bebé, como parto prematuro, aumento de visitas al pediatra o urgencias hospitalarias, problemas emocionales, problemas de conducta y bajo rendimiento escolar (Gunlicks y Weissman, 2008). Aunque el impacto económico de la DPN no ha sido estimado en el contexto español, un estudio realizado en el Reino Unido estimó un coste de alrededor de 74.000 libras (84.000 euros) por cada caso de DPN (Bauer et al., 2014).

El DPN es un problema de salud pública, por lo que su identificación y tratamiento adecuado es una prioridad para nuestro sistema sanitario. Sin embargo, existen ciertas barreras para su detección, como el estigma asociado a los problemas de salud mental en este periodo vital, el desconocimiento sobre el DPN y sus consecuencias, la falta de un protocolo de actuación al respecto y el escaso tiempo disponible para las consultas médicas (Bina & Glasser, 2019). Para reducir las consecuencias negativas del DPN, debemos incluir la detección de síntomas depresivos en la atención prenatal para derivar los casos graves a tratamiento especializado, ofrecer recursos en los casos menos graves y reducir las consecuencias negativas de la depresión (López-Janer et al., 2021; Marcos-Nájera et al., 2020).

En un reciente análisis exhaustivo de las Guías Clínicas centradas en el DPN, se encontró que, de catorce guías examinadas, diez proporcionaban recomendaciones explícitas para el cribado dentro de este grupo demográfico (Motrico et al., 2022). De hecho, la literatura sostiene que los entornos obstétricos podrían servir como entornos adecuados para implementar cribados periódicos (Rodríguez-Muñoz et al., 2017). Concretamente en España, la Guía de Práctica Clínica sobre Atención al Embarazo y Puerperio (MSSSI, 2014) aborda de forma indirecta la importancia de implementar protocolos de seguimiento relativos a la salud mental materna. A pesar del reconocimiento por parte de organismos nacionales e internacionales de la importancia de las evaluaciones rutinarias de salud mental, especialmente durante el embarazo (Motrico et al., 2022), en general no se utilizan escalas de cribado. Una de las razones de este fenómeno es la creencia de que estas evaluaciones son excesivamente largas. Aproximadamente el 75%-90% de los casos no se detectan durante las revisiones clínicas de atención primaria (Thielke et al., 2007) durante el embarazo.

Para superar las dificultades descritas, las Guidelines on Clinical Management and Service Guidance de NICE (National Institute for Health and Care Excellence, 2007) recomendaron las preguntas de Whooley y Arroll como herramienta de cribado para detectar síntomas depresivos durante el periodo prenatal. Esta herramienta consta de tres ítems, con opciones de respuesta

dicotómicas (sí/no). La administración de la herramienta dura algo menos de tres minutos.

Dos de las preguntas surgen de los criterios de depresión mayor del DSM-5 (APA, 2014) en relación con el estado de ánimo depresivo y la pérdida de placer (Criterio A1 y A2, respectivamente), manteniendo similitud con los criterios del DSM-5-TR (APA, 2022). El apoyo emocional e instrumental es un aspecto psicosocial de gran relevancia en la DPN (Marcos-Nájera et al., 2020), dada la relevancia de la variable "pedir ayuda", Arroll et al. (2005) sugirieron incluir una pregunta para aumentar la especificidad de la herramienta.

Las propiedades psicométricas de las preguntas de Whooley y Arroll han sido evaluadas en diversas poblaciones. En población general, el estudio clásico de Whooley et al. (1997) encontró una sensibilidad del 96% y una especificidad del 57%, demostrando su utilidad como herramienta breve de cribado en atención primaria. Posteriormente, cuando Arroll et al. (2005) añadieron la tercera pregunta sobre la necesidad de ayuda, mejoró la especificidad hasta el 94%, manteniendo una buena sensibilidad. En el contexto perinatal, Darwin et al. (2015) analizaron estas preguntas en mujeres británicas durante el control prenatal, encontrando una sensibilidad del 46% y una especificidad del 92% frente al EPDS como patrón de referencia. En mujeres embarazadas mexicanas, Navarrete et al. (2019) validaron las preguntas de Whooley y Arroll utilizando la SCID (la Entrevista Clínica Estructurada para los Trastornos del Eje I del DSM) como estándar de oro, hallando una sensibilidad del 94.7% y una especificidad del 90% durante el embarazo, lo que demuestra un excelente rendimiento en este grupo poblacional. Dadas sus buenas propiedades y la brevedad de su aplicación, ya se están utilizando en algunas Comunidades Autónomas españolas (Generalitat de Catalunya, 2018) y su uso también se contempla en la Guía de Práctica Clínica sobre el Manejo de la Depresión en el Adulto del Sistema Nacional de Salud, actualizada en 2023 (Ministerio de Sanidad, 2023). Sin embargo, hasta donde sabemos, su validez no ha sido evaluada hasta la fecha.

Por ello, el propósito de este estudio fue evaluar las propiedades psicométricas (validez, fiabilidad y estructura factorial) de las preguntas de

Whooley y Arroll utilizando el Cuestionario de Salud del Paciente-9 (PHQ-9) como patrón oro para diagnosticar depresión durante el embarazo en una muestra de mujeres residentes en España.

Se plantea como hipótesis que dichas preguntas presentarán unas propiedades psicométricas adecuadas en esta población, confirmando su utilidad como herramienta de cribado en el ámbito prenatal.

## Método

### Participantes

La muestra del presente estudio transversal estuvo compuesta por 405 mujeres embarazadas reclutadas durante el primer trimestre de gestación en consultas de atención prenatal del Servicio de Obstetricia del Hospital Universitario Central de Asturias (HUCA). Los criterios de elegibilidad fueron los siguientes (a) estar actualmente embarazada, (b) tener 18 años o más, (c) recibir servicios médicos en el hospital respectivo, (d) poseer una comprensión adecuada del idioma español (tanto lectura como escritura) para dar su consentimiento y completar las encuestas. No se aplicaron criterios de exclusión adicionales.

La Tabla 1 indica las características sociodemográficas de la muestra del estudio. La edad media de las participantes era de 33.1 años (DE=4.8). La mayoría contaba con un buen nivel educativo, siendo el 50.5% mujeres con estudios universitarios y el 26.6% con estudios medios. Además, el 91.9% vivía con su pareja, el 72.3% se encontraba en situación de empleo activo y el 83.9% no declaraba padecer ninguna enfermedad.

Tabla 1. Principales características de las participantes (N=405)

Variables sociodemográficas	n (%) o Media (±DT)
Edad	33.1 (4.8)
Nivel de estudios	
Estudios básicos	22.9%
Estudios medios	26.6%
Estudios superiores	50.5%
Convivencia pareja	
Sí	91.9
No	8.1
Enfermedades	
Sí	16.1
No	83.9
Situación de desempleo	
Sí	27.7
No	72.3

Tabla 2. Ítems de las preguntas Whooley y Arroll

Ítem 1	¿En el último mes se ha sentido desanimada, deprimida o sin esperanza?	Sí / No
Ítem 2	¿En el último mes ha sentido que tiene poco interés o que no disfruta?	Sí / No
Ítem 3	Si es así, ¿Siente que necesita ayuda?	Sí / No

## Instrumentos

En la Tabla 2 se presentan las tres preguntas de cribado propuestas por Whooley y Arroll, formuladas para detectar sintomatología depresiva mediante respuestas dicotómicas (“sí” / “no”). Las dos primeras preguntas están orientadas a evaluar el estado de ánimo deprimido y la pérdida de interés o placer, mientras que la tercera pregunta, propuesta por Arroll et al. (2005), indaga sobre la percepción de necesidad de ayuda. Estas preguntas han sido ampliamente utilizadas en el ámbito de la atención primaria y comienzan a implementarse también en el contexto perinatal, destacando por su brevedad y facilidad de administración. Como instrumento de referencia para este estudio se utilizó el Patient Health Questionnaire-9 (PHQ-9) (Kroenke & Spitzer, 2002), una herramienta autoinformada compuesta por 9 ítems diseñados para evaluar la presencia de los síntomas específicos del Trastorno Depresivo Mayor según los criterios del DSM-IV, y cuya estructura ha demostrado ser también adecuada para el DSM-5 (Mitchell et al., 2016). Cada ítem se puntúa en una escala de Likert de cuatro puntos que va de 0 (“nunca”) a 3 (“casi todos los días”), con una puntuación total que oscila entre 0 y 27, donde puntuaciones más altas indican mayor gravedad de los síntomas.

El punto de corte recomendado para identificar riesgo de depresión mayor es  $\geq 10$ , y ha mostrado una sensibilidad y especificidad del 88% (Kroenke et al., 2001). En cuanto a su fiabilidad, el PHQ-9 ha demostrado una consistencia interna elevada, con coeficientes alfa de .89 en su validación original y .86 en estudios posteriores. El cuestionario ha sido validado en diferentes muestras clínicas en España, incluyendo pacientes hospitalizados (Diez-Quevedo et al., 2001) y en atención primaria (Pinto-Meza et al., 2005), y su uso está ampliamente recomendado para la evaluación del estado de ánimo durante el embarazo, en mujeres de distintas nacionalidades (Lara et al., 2015). En población española embarazada hispanohablante también ha mostrado adecuadas propiedades psicométricas (Marcos-Nájera et al., 2018).

## Procedimiento

La investigación fue aprobada previamente por el comité ético (REF.N. 18/18) y realizada de acuerdo con las normas de Buena Práctica Clínica y los principios éticos internacionales aplicables a la investigación en humanos (Declaración de Helsinki). Antes de su participación, se informó a las mujeres sobre el propósito del estudio, orientado a evaluar el bienestar psicológico durante el embarazo. La participación fue completamente voluntaria y aquellas que aceptaron formar parte del estudio firmaron un consentimiento informado.

La recogida de datos se realizó de forma presencial, aprovechando el tiempo de espera en sus citas prenatales, durante el cual las participantes completaron la batería de cuestionarios autoadministrados.

## Análisis de datos

El análisis de los datos se realizó mediante el programa IBM SPSS Statistics, versión 24. En primer lugar, se llevaron a cabo análisis descriptivos para caracterizar el perfil sociodemográfico de la muestra. A continuación, se evaluó la fiabilidad interna de las preguntas de Whooley y Arroll mediante el coeficiente omega de McDonald, considerado una estimación más robusta que el alfa de Cronbach para escalas breves y con ítems no tau-equivalentes (Hayes & Coutts, 2020; Toro et al., 2022). Se consideraron indicadores de fiabilidad aceptables aquellos con valores comprendidos entre .70 y .90 (Campo-Arias et al., 2008).

Posteriormente, se evaluó la adecuación de los datos para el análisis factorial mediante el índice de Kaiser-Meyer-Olkin (KMO), cuyos valores pueden oscilar entre 0 y 1. Según Kaiser (1974), se consideran aceptables valores a partir de .50, y se clasifican en:  $\geq .90$  (excelente),  $\geq .80$  (meritorio),  $\geq .70$  (moderado),  $\geq .60$  (mediocre),  $\geq .50$  (negligente) y  $< .50$  (inaceptable). Se optó por un análisis factorial exploratorio (AFE) en lugar de un análisis factorial confirmatorio debido a la ausencia de estudios previos que analicen la estructura factorial de la

escala completa en población española durante el embarazo. El AFE se realizó utilizando el método de componentes principales con rotación varimax, siguiendo las recomendaciones de Tabachnick y Fidell (1996).

Para estudiar la validez convergente, se calcularon dos tipos de análisis. Por un lado, se calcularon correlaciones biseriales puntuales entre cada uno de los ítems dicotómicos de las preguntas de Whooley y Arroll y la puntuación total del PHQ-9, utilizando el coeficiente de correlación de Pearson adaptado para variables binarias, con el objetivo de examinar la relación entre cada pregunta individual y la gravedad de los síntomas depresivos. Por otro lado, se evaluaron las correlaciones de cada pregunta con la puntuación total del PHQ-9 como evidencia adicional de validez convergente.

Finalmente, para determinar la capacidad discriminativa de las preguntas de cribado, se realizó un análisis de la curva ROC (receiver operating characteristic) utilizando el PHQ-9 como patrón oro. Se calcularon los valores de sensibilidad, especificidad, valor predictivo positivo (VPP) y valor predictivo negativo (VPN) para distintos puntos de corte, con el fin de estimar el rendimiento diagnóstico del instrumento.

## Resultados

En esta muestra de mujeres embarazadas, aproximadamente una cuarta parte respondió afirmativamente a las dos primeras preguntas de cribado. Concretamente, el 22,5% refirió haber experimentado estado de ánimo deprimido y el 23,5% manifestó haber perdido el interés o el disfrute en las actividades habituales. En cambio, solo el 7% respondió afirmativamente a la pregunta relacionada con la percepción de necesidad de ayuda.

### Fiabilidad y evidencias de validez de las preguntas de Whooley y Arroll

La Tabla 3 muestra las características descriptivas de los ítems, la correlación de cada ítem con la puntuación total de la escala y los índices de fiabilidad. La consistencia interna del instrumento fue adecuada, con un coeficiente omega de

McDonald de .79, lo que sugiere que los ítems presentan una coherencia interna suficiente para ser considerados parte de un mismo constructo.

Para examinar la estructura interna del cuestionario, se llevó a cabo un análisis factorial exploratorio (AFE). El índice de adecuación muestral Kaiser-Meyer-Olkin (KMO) fue de .646, lo que indica un nivel moderado de adecuación para el análisis factorial, mientras que el test de esfericidad de Bartlett resultó significativo ( $\chi^2=179,901$ ;  $gl=3$ ;  $p<.0001$ ), evidenciando que las correlaciones entre los ítems eran lo suficientemente fuertes para justificar el análisis. El AFE se realizó mediante el método de componentes principales con rotación varimax, siguiendo las recomendaciones metodológicas de Tabachnick y Fidell (1996).

Siguiendo la regla de Kaiser (autovalores  $>1$ ), se extrajo un único factor que explicó el 53% de la varianza total, lo que sugiere una estructura unidimensional del instrumento en esta muestra. Las cargas factoriales de los ítems oscilaron entre .73 y .83, indicando que todos los ítems contribuyen significativamente al factor común. El ítem relacionado con la percepción de ayuda mostró la menor carga factorial, aunque permaneció dentro de un rango aceptable, lo que podría reflejar que este ítem introduce una dimensión ligeramente distinta, centrada en la demanda subjetiva de apoyo más que en la sintomatología afectiva.

Se analizaron las correlaciones entre cada una de las preguntas de Whooley y Arroll y la puntuación total del PHQ-9 como evidencia de validez convergente (Tabla 4). Los resultados mostraron correlaciones positivas y estadísticamente significativas en todos los casos ( $p<.001$ ). La primera pregunta ("sentirse triste o deprimida") mostró una correlación moderada ( $r=.499$ ) con la puntuación total del PHQ-9. La segunda pregunta ("pérdida de interés o placer") presentó la correlación más alta ( $r=.553$ ). La tercera pregunta ("deseo de recibir ayuda") también correlacionó significativamente, aunque con una magnitud algo menor ( $r=.437$ ). Estos resultados apoyan la validez convergente de las preguntas de Whooley y Arroll en población embarazada española.

Tabla 3. Análisis de fiabilidad de la escala

Ítem	M	DT	Asimetría	Curtosis	r1	ω	Carga factorial	Comunalidades
Escala total Whooley y Arroll	.59	.91	1.30	.48		.70		
Tristeza y desesperanza	.26	.44	1.32	-.26	.85**	.79	.83	.70
Falta de interés y placer	.27	.47	1.25	-.44	.84**	.66	.80	.63
Ayuda	.07	.26	3.37	9.44	.65**	.53	.73	.53

Nota. r1 Correlaciones entre cada ítem y la escala total. ω: El coeficiente Omega. \*\* $p < .001$

Tabla 4. Correlaciones entre las preguntas de Whooley y Arroll ya la puntuación total del PHQ-9

Ítem	r	p
Tristeza y desesperanza	.499	< .001
Falta de interés y placer	.553	< .001
Ayuda	.437	< .001

Nota. r=coeficiente de correlación de Pearson.

Tabla 5. Análisis de la curva ROC

Ítem	Área bajo la curva	p	95% IC
Escala total Whooley y Arroll	.795	.000	.73 - .86
Tristeza y desesperanza	.728	.000	.66 - .80
Falta de interés y placer	.745	.000	.67 - .82
Ayuda	.622	.001	.54 - .70

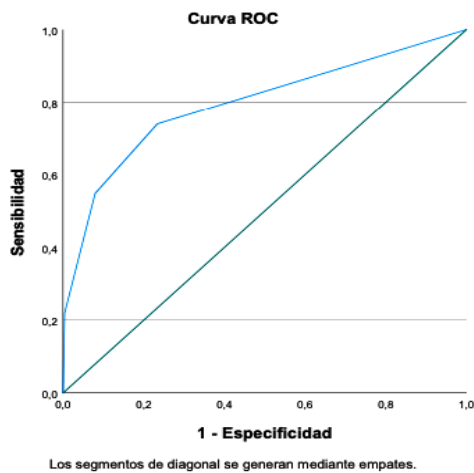


Figura 1. Curva ROC de las preguntas Whooley y Arroll

### Sensibilidad, especificidad, valor predictivo y cociente de probabilidad de las preguntas de Whooley y Arroll

Se realizó el análisis de la curva ROC para cada uno de los ítems y la escala en su conjunto, obteniéndose buenos valores, entre .70 y .90 (Hosmer & Lemeshow, 2000), para la escala total y los dos primeros ítems (Tabla 5). Este resultado sugirió que la escala proporciona una buena discriminación entre pacientes deprimidos y no deprimidos (Figura 1). El valor 1 se definió como el punto de diagnóstico óptimo con el mejor par (sensibilidad-especificidad). Obteniendo una sensibilidad y especificidad de .48 y .91, respectivamente además de un valor predictivo positivo del 74% y un valor predictivo negativo del 77% (Tabla 6).

Tabla 6. Validez de las preguntas Whooley y Arroll como primer screening de DP. Validez %

Sensibilidad	Especificidad	PPV	NPV
48	91	74	77

Nota. PPV: valor predictivo positivo; NPV: valor predictivo negativo

## Discusión

El objetivo de este estudio fue evaluar las propiedades psicométricas de las preguntas de Whooley y Arroll en mujeres españolas durante la etapa prenatal, utilizando el PHQ-9 como patrón oro. Nuestros hallazgos son consistentes con investigaciones previas en otros contextos internacionales. En población general, Whooley et al. (1997) y Arroll et al. (2005) hallaron niveles elevados de sensibilidad y especificidad moderada, mientras que, en mujeres embarazadas británicas, Darwin et al. (2015) encontraron una sensibilidad más baja y una especificidad similar a la de nuestro estudio. En el caso de la validación en población mexicana (Navarrete et al., 2019), se reportaron valores muy elevados tanto de sensibilidad como de especificidad. En nuestra muestra, las preguntas de Whooley y Arroll mostraron una especificidad elevada (.91), lo cual coincide con estos antecedentes y refuerza su utilidad como herramienta para descartar falsos positivos. Asimismo, el coeficiente omega obtenido (0,79) indica una buena consistencia interna, en línea con otras adaptaciones internacionales. Estos datos

respaldan la estructura unidimensional del instrumento y su aplicabilidad en el contexto clínico perinatal español.

Los resultados mostraron que las tres preguntas son un instrumento adecuado para el cribado de la depresión durante el embarazo, ya que el valor de especificidad obtenido fue de .91, indicando que el 91% de las encuestadas sin depresión fueron correctamente clasificadas. Los resultados también indicaron que las tres preguntas corresponden a la medición de un único factor, aunque la tercera pregunta se añade posteriormente y no sigue el diagnóstico según el DSM-5-TR, sino que refleja un aspecto más específico de la depresión en el periodo prenatal, como es la relevancia del sistema de apoyo.

Estos datos acompañan el objetivo perseguido por los profesionales de la salud, que es lograr una detección precoz, dado que las preguntas permiten saber que sólo el 9% de las mujeres con depresión no son detectadas. El valor de sensibilidad no es tan elevado, sin embargo, al tratarse de un instrumento de cribado una vez detectados los casos, éstos deben ser evaluados con una entrevista clínica por un profesional especializado para establecer un diagnóstico preciso. Otros instrumentos de medida ampliamente utilizados en depresión prenatal, la Escala de Depresión Postnatal de Edimburgo (Cox et al., 1987) o el Inventario de Depresión de Beck versión II (Sanz et al., 2003), tampoco son extremadamente sensibles. Sin embargo, estas herramientas son muy específicas y similares a las preguntas de Whooley y Arroll, ya que se consideran herramientas útiles para descartar la falsedad. El cribado es más rentable que permitir que los problemas de DPN prosperen (Rodríguez-Muñoz et al., 2019). Por tanto, las preguntas de Whooley y Arroll, en mayor medida, serán útiles para su aplicación en contextos de salud obstétrica porque pueden ser administradas por profesionales sin especialización en salud mental, donde es de gran importancia descartar los casos no positivos.

Dada la falta de recursos formativos en salud mental prenatal y la falta de tiempo en las consultas, debemos proporcionar a los profesionales sanitarios herramientas de cribado sencillas, breves y basadas en la evidencia para garantizar las mejores prácticas en el manejo de la salud prenatal. En este estudio, podemos

confirmar que las preguntas de Whooley y Arroll cumplen los criterios de sensibilidad y especificidad en la población española. Además, las preguntas de Whooley y Arroll tienen la ventaja de que se administran en muy poco tiempo y constan de preguntas fáciles de entender y responder.

La limitación de este estudio es el hecho de que la muestra no es probabilística, lo que significa que hay que tener cuidado a la hora de generalizar los resultados. En futuros estudios, podemos llevar a cabo investigaciones longitudinales para evaluar la fiabilidad de las preguntas test-retest.

La depresión durante el periodo prenatal no debe continuar siendo infradiagnosticada. Por ello, deben establecerse líneas de actuación claras para su evaluación, prevención y tratamiento. Los ámbitos obstétrico (Kovacheva et al., 2022; Rodríguez-Muñoz et al., 2017) y pediátrico (Earls et al., 2019) son contextos adecuados para la implementación de cribado sistemático. En este sentido, resulta especialmente relevante destacar la facilidad con la que las tres preguntas pueden incluirse de forma rutinaria en las revisiones prenatales, dada su brevedad y facilidad de aplicación por parte de profesionales sin formación especializada en psicología clínica.

Los resultados de este estudio respaldan el uso de las preguntas de Whooley y Arroll como herramienta breve y eficaz para el cribado de la depresión durante el embarazo en población española. Su alta especificidad, buena consistencia interna y estructura unidimensional las convierten en una opción especialmente útil en contextos de atención obstétrica con limitaciones de tiempo o recursos. Aunque su sensibilidad es moderada, su utilidad clínica radica en la posibilidad de descartar con fiabilidad los casos no positivos y detectar aquellas mujeres que podrían beneficiarse de una evaluación clínica más profunda. Estas evidencias apoyan la incorporación sistemática del instrumento en los protocolos de salud prenatal, siempre complementado con una valoración profesional. Futuras investigaciones deberían continuar evaluando su estabilidad temporal y su aplicabilidad en otros contextos perinatales.

## Referencias

- American Psychiatric Association (2014). *Diagnostic and statistical manual of mental disorders, 5th Edition*. American Psychiatric Association Publishing.
- American Psychiatric Association (2022). *Diagnostic and statistical manual of mental disorders-Text Revision (DSM-5-TR)*. American Psychiatric Association Publishing.
- Al-Halabí, S., García-Haro, J., Rodríguez-Muñoz, M. de la F., Fonseca-Pedrero, E., Universidad de Oviedo, Servicios de Salud del Principado de Asturias., Universidad Nacional de Educación a Distancia, & Universidad de La Rioja. (2021). Conducta Suicida y Periodo Perinatal: entre el Tabú y la Incomprensión. *Papeles del Psicólogo*, 42(2).  
<https://doi.org/10.23923/pap.psicol.2963>
- Arroll, B., Smith, F. G., Kerse, N., Fishman, T., & Gunn, J. (2005). Effect of the addition of a “help” question to two screening questions on specificity for diagnosis of depression in general practice: diagnostic validity study. *BMJ*, 331(7521), 884.  
<https://doi.org/10.1136/bmj.38607.464537.7c>
- Bauer, A., Parsonage, M., Knapp, M., Iemmi, V., & Adelaja, B. (2014). The costs of perinatal mental health problems. Unpublished.  
<https://doi.org/10.13140/2.1.4731.6169>
- Bina, R., & Glasser, S. (2019). Factors associated with attitudes toward seeking mental health treatment postpartum. *Women & Health*, 59(1), 1–12.  
<https://doi.org/10.1080/03630242.2017.1421286>
- Campo-Arias, A., & Oviedo, H. C. (2008). Propiedades Psicométricas de una Escala: la Consistencia Interna. *Revista de Salud Pública*, 10(5).  
<https://doi.org/10.1590/s0124-00642008000500015>
- Dagher, R. K., Bruckheim, H. E., Colpe, L. J., Edwards, E., & White, D. B. (2021). Perinatal depression: Challenges and opportunities. *Journal of Women's Health (2002)*, 30(2), 154–159.  
<https://doi.org/10.1089/jwh.2020.8862>
- Darwin, Z., McGowan, L., & Edozien, L. C. (2016). Identification of women at risk of depression in pregnancy: Using women’s accounts to understand the poor specificity of the Whooley and Arroll case finding questions in clinical practice. *Archives of Women's Mental Health*, 19(1), 41–49.  
<https://doi.org/10.1007/s00737-015-0508-1>
- Diez-Quevedo, C., Rangil, T., Sánchez-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. *Psychosomatic Medicine*, 63(4), 679–686.
- Earls, M. F., Yogman, M. W., Mattson, G., Rafferty, J., Baum, R., Gambon, T., Lavin, A., Wissow, L., & Committee on Psychosocial Aspects of Child and Family Health. (2019). Incorporating recognition and management of perinatal depression into pediatric practice. *Pediatrics*, 143(1).  
<https://doi.org/10.1542/peds.2018-3259>
- Forner-Puntonet, M., Fidel-Kinori, S. G., Beneria, A., Delgado-Arroyo, M., Perea-Ortueta, M., Closa-Castells, M. H., Estelrich-Costa, M. de L. N., Daigre, C., Valverde-Collazo, M. F., Bassas-Bolibar, N., Bosch, R., Corrales, M., Dip-Pérez, M. E., Fernández-Quirós, J., Jacas, C., Lara-Castillo, B., Lugo-Marín, J., Nieva, G., Sorribes-Puertas, M., ... Ramos-Quiroga, J. A. (2021). La Atención a las Necesidades en Salud Mental de los Profesionales Sanitarios durante la COVID-19. *Clinica y Salud*, 32(3), 119–128.  
<https://doi.org/10.5093/clysa2021a13>
- Generalitat de Catalunya (2018). *Protocol de Segment de l'embaràs a Catalunya*. Departament de Salut. [File PDF].  
[https://salutpublica.gencat.cat/web/.content/mi\\_nisite/aspcat/promocio\\_salut/embaras\\_part\\_pu\\_erperi/protocol\\_seguint\\_embaras/protocol-seguiment-embaras-2018.pdf](https://salutpublica.gencat.cat/web/.content/mi_nisite/aspcat/promocio_salut/embaras_part_pu_erperi/protocol_seguint_embaras/protocol-seguiment-embaras-2018.pdf)
- Cox, J. L., Holden, J. M., & Sagovsky, R. (1987). Detection of Postnatal Depression: Development of the 10-item Edinburgh Postnatal Depression Scale. *The British Journal of Psychiatry: The Journal of Mental Science*, 150(6), 782–786.  
<https://doi.org/10.1192/bjp.150.6.782>
- Gunlicks, M. L., & Weissman, M. M. (2008). Change in child psychopathology with

- improvement in parental depression: A systematic review. *Journal of the American Academy of Child and Adolescent Psychiatry*, 47(4), 379–389.  
<https://doi.org/10.1097/chi.0b013e3181640805>
- Hayes, A. F., & Coutts, J. J. (2020). Use omega rather than cronbach's alpha for estimating reliability. *But.... Communication Methods and Measures*, 14(1), 1–24.  
<https://doi.org/10.1080/19312458.2020.1718629>
- Hosmer, D. W. & Lemeshow, S. (2000). *Applied logistic regression*. Wiley-Interscience.
- Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, 39, 31–36.
- Kovacheva, K., Rodríguez-Muñoz, M. F., Baya, D. G., Salas, S. D., & Motrico, E. (2022). The relationship between healthcare experience and perinatal depression during COVID-19 pandemic. *Clinica y Salud*, 33(3), 117–125.  
<https://doi.org/10.5093/clysa2022a13>
- Kroenke, K., & Spitzer, R. L. (2002). The PHQ-9: A new depression diagnostic and severity measure. *Psychiatric Annals*, 32(9), 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(9), 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>
- Legazpi, P. C. C., Rodríguez-Muñoz, M. F., Le, H.-N., Balbuena, C. S., Olivares, M. E., & Méndez, N. I. (2022). Suicidal ideation: Prevalence and risk factors during pregnancy. *Midwifery*, 106(103226), 103226.  
<https://doi.org/10.1016/j.midw.2021.103226>
- Liu, X., Wang, S., & Wang, G. (2022). Prevalence and risk factors of postpartum depression in women: A systematic review and meta-analysis. *Journal of Clinical Nursing*, 31(19–20), 2665–2677.  
<https://doi.org/10.1111/jocn.16121>
- López-Janer, C., Quirós, P., & Cabestrero, R. (2021). Propiedades Psicométricas de la Escala de Depresión de Edimburgo en una Muestra Argentina. *Revista Iberoamericana de Diagnóstico y Evaluación – e Avaliação Psicológica*, 61(4), 33–47.  
<https://doi.org/10.21865/ridep61.4.03>
- 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>
- Marcos-Nájera, R., Rodríguez-Muñoz, M. de la F., Soto Balbuena, C., Olivares Crespo, M. E., Izquierdo Méndez, N., Le, H.-N., & Escudero Gomis, A. (2020). The prevalence and risk factors for antenatal depression among pregnant immigrant and native women in Spain. *Journal of Transcultural Nursing*, 31(6), 564–575.  
<https://doi.org/10.1177/1043659619891234>
- Ministerio de Sanidad. (2023). *Guía de práctica clínica sobre el manejo de la depresión en el adulto*. Sistema Nacional de Salud.  
[https://portal.guiasalud.es/wp-content/uploads/2023/05/gpc\\_534\\_depresion\\_adulto\\_avaliat\\_compl-1.pdf](https://portal.guiasalud.es/wp-content/uploads/2023/05/gpc_534_depresion_adulto_avaliat_compl-1.pdf)
- Ministerio de Sanidad, Servicios Sociales e Igualdad (MSSSI, 2014). *Guía de práctica clínica de atención en el embarazo y puerperio*. [File PDF].  
[https://portal.guiasalud.es/wp-content/uploads/2018/12/GPC\\_533\\_Embarazo\\_AETSA\\_compl.pdf](https://portal.guiasalud.es/wp-content/uploads/2018/12/GPC_533_Embarazo_AETSA_compl.pdf)
- 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. [https://www.isci.org/\\_asset/fnhdm3/Depr.pdf](https://www.isci.org/_asset/fnhdm3/Depr.pdf)
- Motrico E, Domínguez-Salas S, Rodríguez-Domínguez C, Gómez-Gómez I, Rodríguez-Muñoz MF, & 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.  
<http://dx.doi.org/10.7334/psicothema2021.380>
- Motrico, E., Moreno-Peral, P., Uriko, K., Hancheva, C., Brekalo, M., Ajaz, E., Apter, G., Bramante, A., Conejo-Cerón, S., Christoforou, A., Dikmen-Yildiz, P., Evagorou, O., Fonseca, A., Lupattelli, A., Radoš, S. N., al Maach, N., Rodríguez-Muñoz, M. F., Žutić, M., & Lambregtse - van den Berg, M. P. (2022). Clinical practice guidelines with recommendations for peripartum depression: A European systematic review. *Acta Psychiatrica Scandinavica*, 146(4), 325–339.  
<https://doi.org/10.1111/acps.13478>
- National Collaborating Centre for Mental Health (2007). *Antenatal and Postnatal Mental Health: the NICE guideline on clinical management and service guidance*. UK: The British Psychological Society. [File PDF]  
<http://www.nice.org.uk/nicemedia/pdf/CG45fullguideline.pdf>
- Navarrete, L., Nieto, L., Lara, M. A., & Lara, M. del C. (2018). Sensitivity and specificity of the three Whooley and Arroll questions for detecting perinatal depression in Mexican women. *Salud Publica de Mexico*, 61(1, ene-feb), 27. <https://doi.org/10.21149/9083>
- 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>
- Rodríguez-Muñoz, M. F., Legazpi, C., Crespo, O., Balbuena, S., Méndez, I., & Barrientos, F. (2017). PHQ-2 como primer instrumento de cribado de la depresión prenatal. *Revista Española de Salud Pública*, 91(33), 1–8.
- Rodríguez, M.F., Le, N.H., Lara, A. & Jaffe, A. (2019). Depresión perinatal. En M.F. Rodríguez-Muñoz (Coord.). *Psicología Perinatal. Teoría y Práctica* (pp. 25-54). Pirámide.  
<https://www.edicionespiramide.es/libro.php?id=5889570>
- Sanz, J., Perdígón, A. L., & Vázquez, C. (2003). Adaptación española del Inventario para la Depresión de Beck-II (BDI-II): 2. Propiedades psicométricas en población general. *Clinica y Salud*, 14(3), 249-280.
- Shin, D., Lee, K. J., Adeluwa, T., & Hur, J. (2020). Machine learning-based predictive modeling of postpartum depression. *Journal of Clinical Medicine*, 9(9), 2899.  
<https://doi.org/10.3390/jcm9092899>
- Slomian, J., Honvo, G., Emonts, P., Reginster, J.-Y., & Bruyère, O. (2019). Consequences of maternal postpartum depression: A systematic review of maternal and infant outcomes. *Women's Health (London, England)*, 15, 174550651984404.  
<https://doi.org/10.1177/1745506519844044>
- Soto-Balbuena, C., Rodríguez-Muñoz, M. F., Escuder-Gomis, A. I., Ferrer-Barrientos, F.J., Le, H. N. & Pmb-Huca, G. (2018). Incidence, prevalence and risk factors related to anxiety symptoms during pregnancy. *Psicothema*, 30(3), 257–263.  
<http://dx.doi.org/10.7334/psicothema2017.379>
- Soto-Balbuena, C., Rodríguez-Muñoz, M. F. & Le, H. N. (2021). Validation of the Generalized Anxiety Disorder screener (GAD-7) in Spanish pregnant women. *Psicothema*, 33(1), 164–170.  
<http://dx.doi.org/10.7334/psicothema2020.167>
- Tabachnick, B. G. & Fidell, L. S. (1996). *Using Multivariate Statistics*. Harper Collins.
- Toro, R., Peña-Sarmiento, M., Avendaño-Prieto, B. L., Mejía-Vélez, S., & Bernal-Torres, A. (2022). Análisis Empírico del Coeficiente Alfa de Cronbach según Opciones de Respuesta, Muestra y Observaciones Atípicas. *Revista Iberoamericana de Diagnóstico y Evaluación – e Avaliação Psicológica*, 63(2), 17.  
<https://doi.org/10.21865/ridep63.2.02>
- Whooley, M. A., Avins, A. L., Miranda, J., & Browner, W. S. (1997). Case-finding instruments for depression: Two questions are as good as many. *Journal of General Internal Medicine*, 12(7), 439–445.  
<https://doi.org/10.1046/j.1525-1497.1997.00076.x>
- Yu, Y., Liang, H.-F., Chen, J., Li, Z.-B., Han, Y.-S., Chen, J.-X., & Li, J.-C. (2021). Postpartum depression: Current status and possible identification using biomarkers. *Frontiers in Psychiatry*, 12.  
<https://doi.org/10.3389/fpsy.2021.620371>