

Validación inicial de *Patient Health Questionnaire 2-Item* para detectar síntomas depresivos en diabéticos y/o hipertensos en la atención primaria de salud

PABLO MARTÍNEZ^{1,2,3,4,5,6,a}, GONZALO SOTO-BRANDT^{7,b},
SEBASTIÁN BRANDT^{8,c}, VIVIANA GUAJARDO^{1,2,9},
GRACIELA ROJAS^{1,2,3,10*}

Validation of patient health Questionnaire-2 to detect depressive symptoms in diabetic or hypertensive patients

Background: *Patient Health Questionnaire (PHQ-9)* has nine questions and is used in diabetic or hypertensive patients to detect depressive symptoms. The PHQ-2 uses the first two questions of the PHQ-9 to rapidly detect those patients that should answer the whole questionnaire. **Aim:** To validate the PHQ-2 to detect depressive symptoms in diabetic or hypertensive patients consulting at Primary Health Care (PHC). **Material and Methods:** Secondary analysis of data obtained during the baseline assessment of a clinical trial. Diabetic and hypertensive patients aged 18 years or more, attending a public health care clinic of Metropolitan Santiago, were invited to participate. Those accepting, answered the Patient Health Questionnaire (PHQ-9), consisting in nine questions about depressive symptoms. Demographics and health data were also collected. The PHQ-2 capacity to discriminate PHQ-9 scores equal or higher than 10 and 15 and the correlation between both versions, were assessed. **Results:** Ninety-four participants aged $64 \pm$ years (73% women) answered the questionnaire. A cut-off score of 3 or more points in the PHQ-2 achieved the best tradeoff between sensitivity and specificity for discriminating PHQ-9 scores equal or higher than 10 (area under the receiver operating characteristic curve (ROC) = 0.92, 95% confidence interval [95% CI] = 0.87 to 0.97) and 15 points (area under the ROC curve = 0.95, 95% CI = 0.92 to 0.99). Both versions had a high positive correlation ($r = 0.87$). **Conclusions:** The PHQ-2 allows a stepped, simple and accurate screening for depressive symptoms. Diabetic or hypertensive patients with 3 or more points should be immediately assessed with the remaining questions of the PHQ-9.

(Rev Med Chile 2020; 148: 1614-1618)

Key words: Diabetes Mellitus; Depression; Hypertension; Mass Screening; Primary Health Care.

¹Departamento de Psiquiatría y Salud Mental, Hospital Clínico Universidad de Chile. Santiago, Chile.

²Instituto Milenio para la Investigación en Depresión y Personalidad (MIDAP). Santiago, Chile.

³Núcleo Milenio para Mejorar la Salud Mental de Adolescentes y Jóvenes (Imhay). Santiago, Chile.

⁴Escuela de Psicología, Facultad de Humanidades, Universidad de Santiago de Chile. Santiago, Chile.

⁵CITIAPS, Universidad de Santiago de Chile. Santiago, Chile.

⁶Psicomédica, Clinical & Research Group. Santiago, Chile.

⁷Programa Salud Mental, Escuela de Salud Pública, Universidad de Chile. Santiago, Chile.

⁸Escuela de Psicología, Facultad de Ciencias Sociales, Universidad de Chile. Santiago, Chile.

⁹Servicio de Psiquiatría, Hospital El Pino. Santiago, Chile.

¹⁰Núcleo Milenio en Desarrollo Social (DESOC). Santiago, Chile.

^aPsicólogo, Doctor en Psicología. ^bPsicólogo, Magíster en Salud Pública.

^cEstudiante de Psicología.

Fuente de apoyo financiero Este estudio recibió el apoyo del Fondo Nacional de Desarrollo Científico y Tecnológico, a través del proyecto FONDECYT Regular N° 1180224, y de ANID – Iniciativa Científica Milenio / Instituto Milenio para la Investigación en Depresión y Personalidad-MIDAP ICS13_005.

Los autores declaran no tener conflictos de interés

Recibido 18 de mayo de 2020, aceptado 15 de diciembre de 2020.

Correspondencia a: Graciela Rojas

Avenida La Paz 1003, Santiago, graciela.rojas.castillo@gmail.com

La depresión es un problema de salud pública global y nacional¹. Es sabido que la detección oportuna puede mejorar su pronóstico². Su comorbilidad habitual con problemas médicos crónicos, como la diabetes y/o la hipertensión³, puede dificultar su detección y tratamiento^{4,5}. Por lo tanto, promover la detección adecuada de los síntomas depresivos en estos pacientes con un instrumento de tamizaje breve, facilitando el acceso oportuno a tratamiento, puede impactar positivamente en el curso del cuadro clínico de estas enfermedades crónicas⁶.

El *Patient Health Questionnaire 2-item* (PHQ-2)⁷ utiliza las dos primeras preguntas de PHQ-9 para la detección de síntomas cardinales de la depresión, constituyéndose en una alternativa breve para el tamizaje de la depresión⁸. En el estudio de validación original de PHQ-2, efectuado en 580 pacientes que asistieron a establecimientos de APS en Estados Unidos, un puntaje de 3 o más presentó las mejores características de clasificación para discriminar casos con un diagnóstico de depresión mayor⁷.

Una opción a este instrumento son las *Whooley questions*⁹, dos preguntas idénticas a las de PHQ-2, aunque con una mayor ventana de tiempo para la exploración de los síntomas y un formato de respuesta dicotómico⁹. En meta-análisis diagnósticos^{10,11}, PHQ-2 ha resultado ser más específico (i.e., menos falsos positivos), aunque menos sensible, que las *Whooley questions*. De todas formas, las recomendaciones sugieren que la administración de un instrumento de estas características sea seguida de una entrevista clínica que permita la confirmación diagnóstica⁹.

Con el objetivo de validar inicialmente PHQ-2 en pacientes diabéticos y/o hipertensos, se presentan los análisis secundarios de datos de 94 usuarios que asistieron a establecimiento de APS en Santiago, Chile.

Material y Método

Análisis secundarios de los datos basales de un estudio piloto para un ensayo clínico aleatorio. El estudio piloto correspondió a un estudio de factibilidad, de diseño pre y post sin grupo control, realizado en un establecimiento de APS de la comuna de Macul, entre junio y septiembre de

2019. Los sujetos del estudio fueron elegibles si tenían 18 años o más, y recibían atención por el programa de salud cardiovascular del establecimiento APS. Se excluyeron quienes rechazaron contestar/participar del estudio. Reclutadores entrenados tomaron contacto de manera sistemática y en sala de espera con potenciales participantes, para la administración in situ de una breve encuesta para confirmar los criterios de elegibilidad y la aplicación del consentimiento informado. Todos los participantes firmaron el consentimiento informado. El estudio fue aprobado por el Comité Ético Científico del Hospital Clínico de la Universidad de Chile (acta de aprobación N° 36, del 06 de junio de 2018).

La encuesta de reclutamiento registró datos demográficos (sexo y edad) y clínicos (p.ej., recibir atención por diabetes/hipertensión en el establecimiento), y se evaluaron los síntomas depresivos con PHQ-9. Este instrumento ha sido validado para la población consultante de APS en Chile¹², y es uno de los cuestionarios más utilizados en el mundo para el autoinforme de síntomas depresivos. Consta de 9 preguntas, puntuadas en una escala tipo Likert de 0 a 3 puntos, basadas en los criterios para episodio depresivo del Manual Diagnóstico y Estadístico de los Trastornos Mentales, cuarta edición (DSM-IV)¹². Las puntuaciones de 5, 10, 15, y 20 representan los puntos de corte para depresión leve, moderada, moderada a grave, y grave, respectivamente¹². En el estudio de validación original, un puntaje de 10 o más en PHQ-9 logró una sensibilidad y especificidad de 88% para depresión mayor¹².

Se efectuaron análisis exploratorios, de carácter descriptivo, para caracterizar la muestra, y análisis de discriminación para la construcción de curvas de características operativas del receptor (ROC), que grafica la sensibilidad de una prueba diagnóstica versus la tasa de falsos positivos, y es utilizada para determinar la precisión diagnóstica de este tipo de pruebas¹³. En este caso, se evaluó la capacidad de PHQ-2 -las dos primeras preguntas del PHQ-9, con un rango entre 0 a 6 puntos-, para discriminar puntajes de PHQ-9 iguales o superiores a 10 y a 15 puntos. Adicionalmente, se evaluó la correlación entre los puntajes de ambas versiones con el coeficiente de correlación de Pearson (r). Los análisis estadísticos fueron realizados con el software Stata 14.0¹⁴.

Tabla 1. Comparación entre puntajes del PHQ-2 y PHQ-9 (n = 94)

n (%)	Puntaje PHQ-2	Puntaje PHQ-9			
		Media	Mínimo – máximo	≥ 10 puntos	≥ 15 puntos
26 (27,7)	0	1,8	0 – 12	1 (3,8)	0 (0)
18 (19,1)	1	6,2	2 – 13	4 (22,2)	0 (0)
8 (8,5)	2	9,5	5 – 16	3 (37,5)	2 (25)
13 (13,8)	3	13,7	4 – 21	9 (69,2)	7 (54)
5 (5,3)	4	12,8	8 – 17	3 (60)	2 (40)
5 (5,3)	5	15,0	12 – 21	5 (100)	3 (60)
19 (20,2)	6	20,2	15 – 26	19 (100)	19 (100)

Datos en últimas dos columnas son frecuencia (porcentaje).

Tabla 2. Estadísticos de clasificación para PHQ-2

Puntaje de corte	Sensibilidad	PHQ-9 ≥ 10		Sensibilidad	PHQ-9 ≥ 15	
		Especificidad	Precisión		Especificidad	Precisión
2	88,6	78,0	83,0	100,0	72,1	81,9
3	81,8	88,0	85,1	93,9	82,0	86,2
4	61,4	96,0	79,8	72,7	91,8	85,1

Resultados

La muestra estuvo compuesta por 73,4% de mujeres (n = 69), con una media de edad de 64,1 años (desviación estándar [DE] = 11,0). La media (DE) para los puntajes de PHQ-9 y PHQ-2 en la muestra fueron de 9,9 (7,7) y 2,5 (2,3) puntos, respectivamente. En la Tabla 1 se detalla la comparación entre los puntajes de PHQ-2 y PHQ-9 para el total de la muestra, observándose un incremento prácticamente lineal, continuo y sin disminución, entre los puntajes de ambos instrumentos, una prevalencia de síntomas depresivos moderados (PHQ-9 ≥ 10 puntos) de 46,8% (n = 44) y de síntomas depresivos moderados a graves (PHQ-9 ≥ 15 puntos) de 35,1% (n = 33).

De acuerdo con la Tabla 2, un punto de corte de 3 o más en PHQ-2 logra el mejor balance entre sensibilidad y especificidad para discriminar puntajes de PHQ-9 iguales o superiores a 10 (área bajo la curva ROC = 0,92, intervalo de confianza de 95% [IC 95%] = 0,87 a 0,97) y a 15 puntos (área bajo la curva ROC = 0,95, IC 95% = 0,92 a 0,99). Se obtuvo una correlación positiva alta entre los puntajes de PHQ-2 y PHQ-9 (r = 0,87).

Discusión

La presente comunicación breve ha reportado la validación inicial de PHQ-2⁷, una versión abreviada de PHQ-9⁸, para la evaluación de los síntomas depresivos en pacientes diabéticos y/o hipertensos en la APS en Chile. Los análisis de la capacidad de discriminación de PHQ-2 sugieren que un puntaje de 3 o más permiten clasificar adecuadamente aquellos pacientes del programa de salud cardiovascular que se presentan con síntomas depresivos moderados o más graves en la APS.

En el presente estudio se encontró el mismo punto de corte tomando como criterio de referencia la presencia de síntomas moderados y moderados a graves evaluados con PHQ-9, en una población específica de APS: pacientes del programa de salud cardiovascular. En comparación a los resultados de la presente validación inicial, estudios que han validado el PHQ-2 en otras latitudes han encontrado resultados disímiles debido a la variabilidad clínica y metodológica de los mismos. Por ejemplo, en población general consultante de APS en Auckland, Australia, y en

Letonia, se encontró que un punto de corte de 2 o más permitió clasificar correctamente a una mayor cantidad de pacientes deprimidos^{15,16}, mientras que dos estudios comunitarios en zonas rurales de México y de China se logró el mismo resultado con PHQ-2 ≥ 3 ^{17,18}. Cabe destacar más allá de las diferencias geográficas y culturales, los estudios en Australia, Letonia y China incluyeron entrevistas psiquiátricas como estándares de referencia^{15,16,18}, cuya administración no fue contemporánea a PHQ-2 en uno de estos estudios¹⁶, mientras que en el estudio Mexicano el estándar de referencia fue PHQ-9¹⁷. El estudio chino se realizó exclusivamente en población adulto mayor, la mayoría con enfermedades físicas¹⁸.

El estándar de referencia utilizado en este estudio fue el PHQ-9, instrumento con características de clasificación excelentes para el diagnóstico de depresión mayor, que no reemplaza la entrevista clínica para la confirmación diagnóstica⁸. Futuros estudios debiesen abordar limitaciones de la presente validación inicial, utilizando una entrevista clínica como estándar de referencia para evaluar el desempeño del PHQ-2; involucrar técnicas de muestreo aleatorias para la selección de consultantes del programa de salud cardiovascular que asisten a varios establecimientos de APS, logrando un tamaño muestral mayor; analizar si el comportamiento del instrumento presenta variaciones según sexo, edad, o lugar de residencia (urbano/rural), u otras; así como caracterizar el impacto de ambas comorbilidades sobre el desempeño del instrumento.

Los resultados sugieren que preguntar por los dos síntomas cardinales de depresión (ánimo deprimido y anhedonia) pueden constituirse en una buena herramienta para que los médicos generales en APS efectúen una posterior exploración más detallada de un probable cuadro depresivo en individuos con hipertensión y/o diabetes. Eventualmente, pudiera ser recomendable la utilización inicial del PHQ-2 para facilitar la labor de los médicos generales en la detección de una patología que se asocia a peores resultados clínicos y funcionales entre consultantes del programa de salud cardiovascular⁴⁻⁶.

Según lo señalan estudios de validación del instrumento en otros países¹⁵⁻¹⁸, falta explorar si el PHQ-2 presenta o no las mismas características de clasificación (y punto de corte) en la población general que asiste a la APS, así como otros grupos

específicos (p.ej., embarazadas o mujeres en el posparto).

El PHQ-2 permitiría un tamizaje escalonado, simple y preciso, de síntomas depresivos, en que los pacientes diabéticos y/o hipertensos con 3 o más puntos en este instrumento deben ser luego sometidos a una entrevista clínica para confirmar el diagnóstico de depresión⁹. Debido a la estrecha relación entre enfermedades crónicas y depresión³⁻⁶, este instrumento abreviado puede promover la detección oportuna de la depresión en enfermos crónicos en la APS, facilitando la labor de los médicos generales en un contexto de alta demanda clínica.

Referencias

1. GBD 2017 DALYs and HALE Collaborators. Global, regional, and national disability-adjusted life-years (DALYs) for 359 diseases and injuries and healthy life expectancy (HALE) for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* 2018; 392 (10159): 1859-922.
2. Ghio L, Gotelli S, Marcenaro M, Amore M, Natta W. Duration of untreated illness and outcomes in unipolar depression: a systematic review and meta-analysis. *J Affect Disord* 2014; 152-154: 45-51.
3. Read JR, Sharpe L, Modini M, Dear BF. Multimorbidity and depression: a systematic review and meta-analysis. *J. Affect. Disord* 2017; 221: 36-46.
4. Stein MB, Cox BJ, Afifi TO, Belik SL, Sareen J. Does co-morbid depressive illness magnify the impact of chronic physical illness? A population-based perspective. *Psychol Med* 2006; 36: 587-96.
5. Katon WJ. Epidemiology and treatment of depression in patients with chronic medical illness. *Dialogues. Clin Neurosci* 2011; 13: 7-23.
6. Sartorius N. Depression and diabetes. *Dialogues. Clin Neurosci* 2018; 20: 47-52.
7. Kroenke K, Spitzer RL, Williams JB. The Patient Health Questionnaire-2: validity of a two-item depression screener. *Med Care* 2003; 41: 1284-92.
8. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med* 2001; 16: 606-13.
9. Whooley MA. Screening for depression – a tale of two questions. *JAMA. Intern Med* 2016; 176: 436-8.
10. Manea L, Gilbody S, Hewitt C, North A, Pummer F, Richardson R, et al. Identifying depression with the

- PHQ-2: a diagnostic meta-analysis. *J. Affect. Disord.* 2016; 203: 382-95.
11. Bosanquet K, Bailey D, Gilbody S, Harden M, Manea L, Nutbrown S, et al. Diagnostic accuracy of the Whooley questions for the identification of depression: a diagnostic meta-analysis. *BMJ Open* 2015; 5: e008913.
 12. Baader T, Molina JL, Venezian S, Rojas C, Farías R, Fierro-Freixenet, et al. Validación y utilidad de la encuesta PHQ-9 (Patient Health Questionnaire) en el diagnóstico de depresión en pacientes usuarios de atención primaria en Chile. *Rev Chil Neuro-Psiquiatr* 2012; 50: 10-22.
 13. Altman DG, Bland JM. Statistics notes: diagnostic tests 3: receiver operating characteristic plots. *BMJ* 1994; 309: 188.
 14. Stata Corp. 2015. *Stata Statistical Software: release 14*. College Station, TX: StataCorp LP.
 15. Arroll B, Goodyear-Smith F, Crengle S, Gunn J, Kerse N, Fishman T, Falloon K, Hatcher S. Validation of PHQ-2 and PHQ-9 to screen for major depression in the primary care population. *Ann Fam Med* 2010; 8: 348-53.
 16. Rancans E, Trapencieris M, Ivanovs R, Vrublevska J. Validity of the PHQ-9 and PHQ-2 to screen for depression in nationwide primary care population in Latvia. *Ann Gen Psychiatry* 2018; 17: 33.
 17. Arrieta J, Aguerrebere M, Raviola G, Flores H, Elliott P, Espinosa A, et al. Validity and utility of the Patient Health Questionnaire (PHQ)-2 and PHQ-9 for screening and diagnosis of depression in rural Chiapas, Mexico: a cross-sectional study. *J Clin Psychol* 2017; 73: 1076-90.
 18. Liu ZW, Yu Y, Hu M, Liu HM, Zhou L, Xiao SY. PHQ-9 and PHQ-2 for screening depression in Chinese rural elderly. *PLoS One* 2016; 11: e0151042.