

Biomédica 2017;37(Supl.1):112-20
doi: <http://dx.doi.org/10.7705/biomedica.v37i0.3221>

ARTÍCULO ORIGINAL

Confiabilidad y dimensión del cuestionario de salud del paciente (PHQ-9) para la detección de síntomas de depresión en estudiantes de ciencias de la salud en Cartagena, 2014

Carlos Arturo Cassiani-Miranda^{1,2}, María Camila Vargas-Hernández³, Eduard Pérez-Aníbal³, Mariana Isabel Herazo-Bustos³, Mauricio Hernández-Carrillo¹

¹ Grupo de Investigación en Salud Pública, Universidad Santiago de Cali, Cali, Colombia

² Fundación Universitaria San Martín, Cali, Colombia

³ Semillero de Investigación, Grupo de Investigación en Salud Pública, Universidad de Cartagena, Cartagena, Colombia

Introducción. La detección de los síntomas de depresión en estudiantes universitarios mediante instrumentos válidos es necesaria. El cuestionario de salud del paciente (*Patient Health Questionnaire-9*, PHQ-9) es confiable y ha sido validado para dicha detección en el contexto universitario, pero se desconoce su desempeño psicométrico en Colombia.

Objetivo. Estimar la confiabilidad y la dimensión del PHQ-9 como instrumento de detección de los síntomas de depresión en estudiantes del área de ciencias de la salud de una universidad de Cartagena, Colombia.

Materiales y métodos. Se estudió la estructura factorial y la confiabilidad del PHQ-9 en una muestra de 550 estudiantes para una prevalencia esperada de síntomas depresivos con importancia clínica de 25 %, un nivel de confianza de 95 %, y de error de 3 %. Se utilizó la versión en español del PHQ-9 con autorización de los autores. Se hizo un análisis factorial confirmatorio, y se estimó la consistencia interna con el coeficiente alfa de Cronbach y el omega de McDonald.

Resultados. Se analizaron 541 encuestas. La edad media del grupo fue de 20,18 años (desviación estándar, DE=2,59), 354 (63,77 %) eran mujeres y 196 (36,23 %), hombres. La prevalencia de los síntomas de depresión de importancia clínica fue de 27,3 %. El análisis factorial confirmatorio evidenció un modelo de dos factores, los cuales explicaron el 42,80 % de la varianza total. La proporción de la varianza explicada por los factores estuvo entre 0,243 (ítem 5) y 0,587 (ítem 2). El coeficiente alfa de Cronbach fue de 0,830 y, el omega de McDonald, de 0,89.

Conclusiones. El PHQ-9 resultó ser una herramienta válida y confiable para la detección de síntomas de depresión en estudiantes de ciencias de la salud de una universidad de Cartagena.

Palabras clave: depresión; estudiantes; encuestas y cuestionarios; cribado; análisis factorial; validez de las pruebas; reproducibilidad de resultados; Colombia.

doi: <http://dx.doi.org/10.7705/biomedica.v37i0.3221>

Reliability and dimensionality of PHQ-9 in screening depression symptoms among health science students in Cartagena, 2014

Introduction: Depression among university students must be screened with valid instruments. The Patient Health Questionnaire (PHQ-9) is reliable and valid for screening depression in a university context; however, its psychometric performance in Colombia is unknown.

Objective: To estimate the reliability and dimensionality of PHQ-9 as a screening instrument for depressive symptoms among health sciences students of a university in Cartagena, Colombia.

Materials and methods: We analyzed the factorial structure and reliability of the PHQ-9 in a sample of 550 students for expected prevalence of clinically-meaningful depressive symptoms (CMDS) of 25% with a confidence level of 95% and an error level of 3%. We used the free Spanish version of PHQ-9 for Colombia, authorized by its creators. A confirmatory factorial analysis, and an estimate of internal consistency using Cronbach's alfa and McDonald's omega were obtained.

Results: We analyzed 541 surveys. The average age of the group was 20.18 (SD=2.59) years old, 354 (63.77%) participants were women and 196 (36.23%) men. CMDS prevalence was 27.3%. The

Contribución de los autores:

Carlos Arturo Cassiani-Miranda, María Camila Vargas-Hernández, Eduard Pérez-Aníbal, Mariana Isabel Herazo-Bustos: diseño del proyecto de investigación; obtención, análisis e interpretación de los datos

Mauricio Hernández-Carrillo: procesamiento estadístico de los datos

Todos los autores participaron en la escritura del manuscrito.

confirmatory factorial analysis exhibited a two-factor model, which explained the total variance of 42.80%. The proportion of variance explained by the factors was 0.243 (item 5) and 0.587 (item 2). Cronbach's alpha was 0.830 and McDonald's omega, 0.89.

Conclusions: PHQ-9 was a valid and reliable tool for screening depressive symptoms among health sciences students in a university in Cartagena, Colombia.

Key words: Depression; students; surveys and questionnaires; screening; factor analysis; statistical, validity of tests; reproducibility of results; Colombia.

doi: <http://dx.doi.org/10.7705/biomedica.v37i0.3221>

La depresión es un grave problema de salud pública debido a su impacto social y económico (1). A nivel mundial, la prevalencia del trastorno depresivo grave a lo largo de la vida y en el último año en la población general se sitúa entre 1,5 y 19,0 % y entre 0,8 y 5,8 %, respectivamente (2,3). Sin embargo, en los estudiantes universitarios alrededor del mundo, estas cifras son sustancialmente mayores, con prevalencias que oscilan entre 10 y 85 %, con una media de 30,6 % (4). En Colombia, la prevalencia de síntomas de depresión en estudiantes universitarios también es mayor que en la población general, con un rango entre 25 y 52,3 % (5,6).

Los síntomas de depresión se asocian con un deterioro significativo en el desempeño cognoscitivo, lo cual perjudica el rendimiento académico y representa un factor de estrés adicional para la mayoría de estudiantes (7). A pesar de la significativa carga global de los trastornos depresivos, cerca del 50 % de los pacientes con este problema no se detectan en situaciones no clínicas (8). En este contexto, la salud mental de los estudiantes universitarios tiene serias implicaciones para la salud pública (9), y se ha comprobado la importancia de detectar tempranamente la depresión entre ellos, pues, cuando este tipo de trastorno aparece durante esta etapa de la vida, tiende a hacerse crónico (10). En este sentido, el uso de cuestionarios cortos de detección puede mejorar el reconocimiento de la depresión en situaciones no clínicas (11).

Existen diversos instrumentos o escalas para la detección de los síntomas de depresión en la población general (12-14); sin embargo, la mayoría son en inglés o en otros idiomas diferentes del español, lo cual exige validarlos para ajustar las diferencias idiomáticas y culturales que pueden alterar el proceso de medición implícito en la escala original (15).

Correspondencia:
Mauricio Hernández-Carrillo, Carrera 79B N° 9-18, apartamento C303, Cali, Colombia
Teléfono: (300) 273 3699
mauriciohc@gmail.com

Recibido: 09/02/16; aceptado: 14/01/17

Durante los últimos 20 años en Colombia, se han validado pocos cuestionarios para la detección de los síntomas de depresión. Se siguen empleando aquellos diseñados hace más de tres décadas, los cuales no reflejan los criterios diagnósticos más recientes del Manual Diagnóstico y Estadístico de Enfermedades Mentales (*Diagnostic and Statistical Manual of Mental Disorders*, DSM).

Una alternativa para llenar ese vacío es el cuestionario PHQ-9 (*Patient Health Questionnaire-9*), una versión de nueve ítems del cuestionario de salud del paciente derivado de la *Primary Care Evaluation of Mental Disorders*, PRIME-MD, como herramienta de detección de los síntomas de depresión en la atención primaria basada en los criterios diagnósticos del DSM-IV. Este cuestionario ha demostrado tener adecuadas sensibilidad y especificidad (16,17), y es sustancialmente más corto que la mayoría de los cuestionarios usados para la detección de los síntomas de depresión (18,19). El PHQ-9 se ha traducido a varios idiomas y hay estudios que respaldan su confiabilidad y validez (20,21).

La validación de la versión en castellano del PHQ-9, realizada en Chile en centros de atención primaria, ha demostrado adecuadas consistencia interna (alfa de Cronbach=0,835), sensibilidad (88 %) y especificidad (92 %), comparado con la escala de Hamilton-D, aplicada por los psiquiatras, así como una adecuada validez concurrente y predictiva con respecto a los criterios de la CIE-10 para la depresión (20).

En la población colombiana se ha utilizado la sección de nueve ítems para la detección de síntomas depresivos del PHQ en su versión castellana (22, 23), pero, aunque Londoño, *et al.*, evaluaron la validez aparente del instrumento (22), no informaron sobre su consistencia interna o su estructura factorial. Por otro lado, la versión para Colombia del PHQ-9 (disponible en <http://www.phqscreeners.com/>) (24), también se ha aplicado en el país, pero sin haberla adaptado y sin haber analizado su consistencia interna o su estructura factorial.

En el proceso de validación de constructos psicológicos, es fundamental establecer la confiabilidad y la validez de las escalas (25,26). La validez es la capacidad del instrumento para medir el constructo (26). Una de las formas de evaluar la validez del constructo de una escala es la verificación de su dimensión, es decir, si el instrumento recopila los factores (o dimensiones) que teóricamente hacen parte de la escala (27,28).

En varios estudios en los cuales se han hecho análisis factoriales exploratorios y confirmatorios, se ha encontrado que el PHQ-9 tiene una estructura unidimensional (29,30), lo cual coincide con la teoría unidimensional del constructo "depresión". Sin embargo, en otros análisis factoriales exploratorios (31,32) y confirmatorios (33), se ha evidenciado un modelo de dos factores para el PHQ-9, uno de ellos correspondiente a los ítems somáticos (alteraciones del sueño, cambios en el apetito y fatiga) y, el otro, a los ítems no somáticos o afectivos (humor deprimido, sentimientos de desesperanza y pensamientos suicidas). Estos datos indican que la dimensión de la escala puede verse influenciada por las diferencias en las poblaciones o por el modelo del constructo "depresión" empleado como referencia (34).

En general, en los estudios de validación del PHQ-9 en estudiantes universitarios, se han registrado valores satisfactorios de consistencia interna, con un coeficiente alfa de Cronbach que oscila entre 0,80 y 0,90 (35,36). Por otra parte, la mayoría de los estudios de validación del PHQ-9 han estimado la confiabilidad de la escala con el alfa de Cronbach, a pesar de que los expertos en psicometría recomiendan utilizar, además, estimaciones más precisas de la confiabilidad, como el omega de McDonald (37).

En el país no hay informes sobre la confiabilidad del PHQ-9 estimada con el coeficiente omega de McDonald, ni se conoce la estructura factorial y la consistencia interna del cuestionario. En este contexto, es conducente el uso de coeficientes que indiquen el desempeño psicométrico del PHQ-9 en estudiantes de ciencias de la salud, tales como su dimensión y confiabilidad (alfa de Cronbach, omega de McDonald), y su validez (37). Un apropiado desempeño psicométrico del PHQ-9 permitirá su uso en la detección rápida de posibles episodios depresivos graves en el contexto universitario (4).

En consecuencia, el objetivo de este trabajo fue adaptar el PHQ-9 y analizar su consistencia interna (confiabilidad) y su estructura factorial (dimensión)

como instrumento de detección de síntomas de depresión en la población estudiantil colombiana.

Materiales y métodos

Diseño

Se hizo un estudio observacional en el cual se determinó la validez del constructo mediante la estimación de la consistencia interna y el análisis factorial confirmatorio de una adaptación del cuestionario PHQ-9 a partir de la versión traducida para Colombia por Kroenke, *et al.*

Participantes

La muestra se calculó para una población finita de 1.745 sujetos, correspondiente al total de estudiantes activos en el primer periodo académico del 2014 en los programas de Medicina, Enfermería, Odontología y Química Farmacéutica de una universidad de Cartagena. Para una prevalencia esperada de 25 % de síntomas de depresión de importancia clínica (38), con nivel de confianza de 95% y de error de 3 % (39,40), el tamaño muestral fue de 550 sujetos.

Se empleó un muestreo aleatorio estratificado por programa académico, semestre y sexo, en el cual cada estrato se componía de un número de estudiantes proporcional a su tamaño. Este número de participantes fue suficiente para el análisis factorial y la valoración de la consistencia interna, para lo cual se requieren, al menos, diez participantes por cada ítem de la escala (15).

Instrumento

El cuestionario PHQ-9 es un módulo de la PRIME-MD sobre la depresión con nueve ítems ajustados a los criterios del DSM-IV para la depresión grave (16,24). Puede emplearse de manera algorítmica para el diagnóstico probable de un trastorno depresivo grave, o como una medida continua de puntajes que van de 0 a 27, con puntos de corte de 5, 10, 15 y 20, los cuales establecen los niveles de síntomas de depresión como leves, moderados, moderadamente graves o graves (16).

Según Kroenke, *et al.*, el PHQ-9 tiene una sensibilidad y una especificidad de 88 %, una consistencia interna según el alfa de Cronbach de 0,86 a 0,89, un puntaje de prueba-reprueba de 0,84, una concordancia de 84 % entre el cuestionario respondido por los propios participantes y el realizado por el evaluador, y un área bajo la curva de 0,95 (16), lo cual se considera como un desempeño psicométrico apropiado.

En un análisis de más de 38 estudios con más de 32.000 pacientes de atención primaria en su conjunto, Williams, *et al.*, concluyeron que el PHQ-9 era igual o superior a otros instrumentos de medición de la depresión (13). Asimismo, se han hecho validaciones con base en el consenso de expertos y la determinación de la confiabilidad del PHQ-9 en Perú (41) y en Chile (20). Aunque se han desarrollado varias escalas para evaluar los síntomas de la depresión, el grupo de trabajo para trastornos depresivos del DSM-5 ha considerado el módulo de depresión de nueve ítems del PHQ como la medición de preferencia para cuantificar la gravedad de la depresión (18); además, los síntomas nucleares del episodio depresivo en el DSM-5 son básicamente los mismos que en el DSM-IV.

En este estudio, se utilizó la versión del PHQ-9 adaptada por sus creadores en español para Colombia, la cual es de libre acceso y está disponible en: http://www.phqscreeners.com/sites/g/files/g10016261/f/201412/PHQ9_Spanish%20for%20Colombia.pdf.

Procedimiento

El estudio contó con dos fases. En la primera, se verificó la validez aparente del cuestionario mediante la opinión de diez expertos en el área de investigación y diagnóstico clínico, quienes indicaban si los ítems representaban, o eran relevantes en la medición de los criterios para el diagnóstico de un episodio de depresión según el DSM-IV. Cada experto identificaba las palabras que hicieran incomprensible el ítem o indujeran a error a quienes contestaran el cuestionario, o planteaba correcciones gramaticales en las preguntas. A continuación, se hicieron los ajustes idiomáticos de todos los ítems con base en las observaciones relevantes de los expertos y las conclusiones del grupo de discusión a cargo del equipo investigador. Este procedimiento permitió obtener una versión del PHQ-9 para Colombia ajustada a las características de la población de estudio (42).

En la segunda fase en el aula de clases, después de diligenciar el consentimiento informado, los participantes suministraron información demográfica básica sin dar sus nombres; después diligenciaron el PHQ-9 en presencia de asistentes de investigación entrenados para este propósito, quienes dieron las instrucciones para diligenciarlo correctamente y estuvieron atentos a aclarar dudas teniendo cuidado de no sugerir respuestas.

Análisis estadístico

La confiabilidad de la consistencia interna del cuestionario se estimó mediante el coeficiente alfa de Cronbach (43) y el omega de McDonald (44). El alfa de Cronbach es la medida de consistencia interna más popular en los estudios de validación en publicaciones biomédicas (15,45). Sin embargo, el alfa de Cronbach puede subestimar la consistencia interna cuando no se cumple el principio de equivalencia tau, es decir, cuando los incisos muestran coeficientes muy disímiles en la matriz de una solución factorial, caso en el que el omega de McDonald es una mejor medida de la consistencia interna. La recomendación actual es emplear ambos coeficientes (15,45). Además, se estimó la correlación de Pearson en cada ítem y la puntuación total corregida del PHQ-9, es decir, sin incluir el ítem evaluado. Para el cálculo del omega, se utilizó el programa estadístico R, versión 3.2.1. Se calcularon la media, la desviación estándar (DE), el coeficiente de correlación de Pearson y el alfa de Cronbach con la omisión del ítem.

Para evaluar la dimensión de la escala, se hizo un análisis factorial confirmatorio (46). Se utilizaron las pruebas de esfericidad de Bartlett (47) y la de adecuación de la muestra de Kaiser-Meyer-Olkin (48), con el fin de confirmar que la matriz de correlaciones no fuera una matriz de identidad y que se podía continuar con el proceso de estimación de los factores latentes en la escala.

Para obtener la matriz factorial, se utilizó el método de factorización por ejes principales, método que se recomienda cuando los puntajes de la escala se distribuyen de manera no paramétrica (prueba de Kolmogorov-Smirnov, $p < 0,05$) (49). Se usó el método de rotación oblicua de Promax de los factores del PHQ-9 con gran correlación entre ellos (46,49).

Para la estimación de la varianza compartida, se calcularon las comunalidades, los valores propios, la correlación (r) entre los factores observados y el porcentaje de varianza total explicada de los factores relevantes. Por último, se estimó el coeficiente alfa tradicional (43) de los factores detectados y el alfa ajustado para el número de ítems, lo cual permitió una mejor comparación de los coeficientes (27). El análisis multivariado se hizo con el paquete estadístico SPSS, versión 15.

Consideraciones éticas

La participación en esta investigación representó un riesgo mínimo para los estudiantes según

la Resolución 8430 del Ministerio de Salud de Colombia, dado que no se realizaron intervenciones farmacológicas o psicosociales.

El estudio fue aprobado por el Comité de Ética de la Facultad de Medicina de la Universidad de Cartagena y se obtuvieron los permisos académicos respectivos para abordar a los estudiantes. Asimismo, se respetaron los derechos de autor del instrumento, pues el PHQ-9 se puede usar sin restricción en investigaciones y en la práctica clínica, e incluso, existe una versión libre para Colombia autorizada por sus creadores.

Resultados

Características sociodemográficas de la población

De los 550 estudiantes, participaron finalmente 541, pues nueve encuestas se descartaron debido a errores en su diligenciamiento. La edad promedio del grupo fue de 20,18 (DE=2,59) años. En lo concerniente al sexo, 345 (63,77 %) eran mujeres y 196 (36,23%) eran hombres.

Dimensiones

La prueba de Kaiser-Meyer-Olkin arrojó un resultado de 0,889 (media=0,24–1,04, y desviación típica: 0,527 – 0,960). Los ítems 1 y 2 del PHQ-9 explicaron el 42,80 % de la varianza total. Las comunalidades registraron valores entre 0,243 (ítem 5) y 0,587 (ítem

2). Con respecto a la matriz de estructura factorial y la rotación Promax, se obtuvieron dos factores. Los ítems 1, 2, 6, y 9 conformaron el factor 1 y, los ítems 3, 4, 5, 7 y 8, el factor 2. Se obtuvo un alfa de Cronbach de 0,830 (cuadro 1).

Confiabilidad

Se obtuvieron valores elevados de consistencia interna, con un coeficiente alfa de Cronbach de 0,830, en tanto que el omega de McDonald fue de 0,89. Además, se presentan la media, la desviación estándar, la correlación de Pearson corregida y el alfa de Cronbach con la omisión del ítem (cuadro 2).

Discusión

En la presente investigación se observó que el PHQ-9 presentaba una gran consistencia interna y una estructura bidimensional, con lo cual se confirmó su confiabilidad y validez para la detección de los síntomas de depresión en estudiantes de ciencias de la salud de una universidad de Cartagena.

Dimensión

La matriz de la estructura factorial en este estudio evidenció dos factores, lo cual se constató mediante un análisis factorial confirmatorio, ya que el instrumento no es nuevo y se debía verificar si se mantenía la misma estructura reportada en estudios previos con una muestra distinta (33,34,44). La estimación

Cuadro 1. Análisis factorial confirmatorio del cuestionario PHQ-9 en estudiantes de ciencias de la salud de una universidad de Cartagena, 2014

Punto	Comunalidad	Factor	
		1	2
1. Poco interés o placer en hacer las cosas	0,310	0,544	0,460
2. Sentirse desanimado/a, deprimido/a o sin esperanzas	0,587	0,760	0,593
3. Problemas para dormir o mantenerse en el sueño o dormir demasiado	0,564	0,426	0,742
4. Sentirse cansado/a o con poca energía	0,529	0,558	0,722
5. Sentir poco apetito o comer en exceso	0,243	0,404	0,483
6. Sentirse mal acerca de sí mismo o tener un sentimiento de fracaso o de abandono propio o de la familia	0,525	0,724	0,513
7. Dificultad para concentrarse en diferentes actividades tales como leer el periódico o ver televisión	0,382	0,610	0,490
8. Moverse o hablar tan despacio que otras personas lo han notado o bien, por el contrario, estar tan inquieto/a o intranquilo/a que se mueve mucho más de lo normal	0,390	0,619	0,487
9. Pensamientos acerca de que sería mejor estar muerto/a o deseos de lastimarse de alguna forma	0,321	0,548	0,273
KMO	0,889		
Valor propio	3,372		0,481
Varianza (%)	37,464		5,344

KMO: prueba de Kaiser-Meyer-Olkin
Los valores en negrilla indican la pertenencia del ítem al factor.

Cuadro 2. Media, desviación estándar (DE), correlación de Pearson y alfa de Cronbach con la omisión del ítem para el cuestionario PHQ-9

Ítem	Media	DE	R	A
1. Poco interés o placer en hacer las cosas	0,78	0,83	0,63	0,87
2. Sentirse desanimado/a, deprimido/a o sin esperanzas	0,79	0,76	0,76	0,86
3. Problemas para dormir o mantenerse el sueño o dormir demasiado	1,03	0,96	0,66	0,87
4. Sentirse cansado/a o con poca energía	1,06	0,89	0,72	0,87
5. Sentir poco apetito o comer en exceso	1,04	0,94	0,59	0,88
6. Sentirse mal acerca de sí mismo o tener un sentimiento de fracaso	0,72	0,81	0,71	0,86
7. Dificultad para concentrarse en diferentes actividades tales como leer el periódico o ver televisión	0,80	0,80	0,66	0,87
8. Moverse o hablar tan despacio que otras personas lo han notado	0,60	0,71	0,66	0,87
9. Pensamientos acerca de que sería mejor estar muerto/a	0,24	0,53	0,50	0,88
Escala del PHQ	7,07	4,76	1,00	0,88

R: correlación de Pearson corregida del inciso con la puntuación total

A: alfa de Cronbach para la escala con la omisión del inciso

del grado de correlación entre los ítems se hizo mediante la prueba de adecuación de la muestra de Kaiser-Meyer-Olkin (0,889), así como la distribución no paramétrica de Kolmogorov-Smirnov ($p < 0,05$) de las puntuaciones de los ítems.

Los valores de las comunalidades fluctuaron entre 0,243 y 0,587, y expresaron la proporción de la variable que podía explicarse con el modelo factorial obtenido (7,49). Con estos cálculos se pretendía estimar más detalladamente la varianza total y la compartida por todos los ítems, para así evaluar el comportamiento de los ítems entre sí y el peso que cada uno aportaba por sí mismo a la matriz de correlaciones (49).

Los análisis factoriales confirmatorios del PHQ-9 han registrado modelos con uno (28) y dos factores (32-34), sin que exista consenso con respecto a uno u otro. Sin embargo, en los trabajos de Elhai, *et al.* (34), Krause, *et al.* (33), y Richardson, *et al.* (32), se evidenció que los modelos factoriales que consideraban el trastorno depresivo grave como un constructo conformado por dos factores (somático y no somático) presentaban un mayor ajuste estadístico que los de un solo factor. Además, en estos estudios también se observó que los síntomas de trastorno depresivo grave detectados mediante el PHQ-9 se representaban mejor con un modelo que incluyera los factores somáticos y no somáticos que el simple modelo unidimensional. No obstante, la presencia de uno o dos factores se explicaría por el tipo de población, ya que los estudios llevados a cabo en contextos clínicos diferentes, por ejemplo, en pacientes con enfermedades somáticas (31), se adaptaron mejor a una estructura factorial bidimensional, pues, al parecer, la población estaba

mejor representada por diferentes ítems somáticos y no somáticos, así que la forma de entender la depresión debía ajustarse al contexto clínico en el cual se hacía la evaluación.

La validez de la estructura teórica propuesta por Spitzer, *et al.*, la cual suministra un algoritmo para el diagnóstico del trastorno depresivo grave, así como un punto de corte en la escala total para detectarlo, difiere del análisis factorial del presente estudio, cuyo resultado fue una estructura bidimensional, lo cual implicaría que era razonable evaluar los puntos de corte para cada factor por separado, procedimiento imposible en esta investigación, ya que no se validaron los criterios con respecto a un estándar de referencia.

Aunque el PHQ-9 se ha considerado como una herramienta unidimensional, los resultados del presente estudio son coherentes con la teoría, las investigaciones y la práctica clínica sobre la depresión considerada como un constructo de múltiples dimensiones en la población universitaria. Esto no significa, sin embargo, que el modelo propuesto por Spitzer, *et al.*, no sea válido, ya que se usa ampliamente y se ha validado en muchos estudios como herramienta de detección de los síntomas de la depresión.

Con base en los resultados presentados, se puede considerar un proceso de dos pasos en la aplicación del PHQ-9. El primer paso sería la interpretación de la escala completa y, el segundo, una puntuación independiente para los factores somáticos y los no somáticos, enfoque que permitiría prestar la debida atención a los diferentes grupos de síntomas que caracterizan la depresión. La posibilidad de recomendar dicho enfoque requeriría hacer un estudio

que compare la detección de los síntomas de depresión con las escalas de un factor y de dos frente a un estándar de referencia.

Confiabilidad

La consistencia interna es una medida directa de la confiabilidad e indirecta de la validez de un constructo, dado que es una estimación de la correlación que existe entre los ítems (15). En el presente estudio, la consistencia interna del PHQ-9 estuvo dentro del rango deseado (alfa de Cronbach de 0,83), en congruencia con lo informado en estudios precedentes en el ámbito universitario (0,83, 0,85) (35,36) y en la población general, en los cuales el valor del alfa de Cronbach osciló entre 0,79 a 0,89 (20,50-52). La consistencia interna es aceptable si el alfa de Cronbach o el omega de McDonald se encuentran entre 0,70 y 0,79; bajo condiciones más exigentes se prefieren valores entre 0,80 y 0,90 (15).

Actualmente, las publicaciones especializadas recomiendan hacer una segunda medición de la consistencia interna mediante el omega de McDonald para garantizar la confiabilidad de la medición (44). Los estudios de validación del PHQ-9 disponibles en población universitaria (35,36) no informan sobre el valor del omega de McDonald y era necesario calcularlo, puesto que en ciencias sociales y de la salud las puntuaciones pocas veces muestran una distribución normal y, en consecuencia, el alfa de Cronbach tiende a subestimar la consistencia interna porque las puntuaciones presentan distribuciones sesgadas, particularmente si el patrón de respuesta es ordinal (45). En este análisis, el elevado valor del omega de McDonald (0,89) demostró una excelente consistencia interna. Con base en los valores de los coeficientes omega y alfa obtenidos en esta investigación, puede afirmarse que el PHQ-9 tuvo muy buena consistencia interna para la población de estudio.

Esta investigación es la primera en mostrar la dimensión y la estructura interna del PHQ-9 en la población colombiana. Además, se empleó por primera vez el coeficiente omega para el análisis de su consistencia, el cual arroja una mejor estimación de la confiabilidad cuando no se cumple el principio de tau equivalencia (35,37,45).

Una de las principales limitaciones de este estudio consistió en que no se evaluó el comportamiento psicométrico del PHQ-9 frente a un estándar de referencia, es decir, la entrevista psiquiátrica estructurada, y que solo se evaluó en una población

específica. En futuros estudios debería considerarse el análisis de la validez del criterio con respecto a uno de referencia, y hallar los puntos de corte para ambos factores.

Los hallazgos del presente estudio pueden ser útiles para investigadores y profesionales de atención primaria interesados en la detección temprana de trastornos depresivos. Además, confirman la necesidad de revisar permanentemente el desempeño psicométrico de los instrumentos de medición en salud, dadas las importantes variaciones que se pueden observar en el patrón de respuesta a los cuestionarios entre diferentes poblaciones (53).

Es útil contar con instrumentos de medición en salud válidos y confiables que garanticen, al menos, la validez interna de los hallazgos (54) y, en este sentido, se puede concluir que el PHQ-9 fue una herramienta válida, confiable y fácil de usar para la detección de los síntomas de depresión en estudiantes de ciencias de la salud de una universidad de Cartagena.

Agradecimientos

Al Semillero de Investigación en Psiquiatría y Salud Mental (Psycomed) de la Universidad de Cartagena, por el apoyo en la recolección de la información. A los estudiantes de los programas de Medicina, Enfermería, Odontología y Química Farmacéutica de dicha universidad que participaron en el estudio.

Conflicto de intereses

Los autores declaramos que no existe conflicto de intereses en este manuscrito. Este informe hace parte del trabajo de grado para optar al título de psiquiatra del primer autor, Carlos Arturo Cassiani-Miranda.

Financiación

El estudio se financió con recursos propios de los investigadores.

Referencias

1. **Thota AB, Sipe TA, Byard GJ, Zometa CS, HahnRA, McKnight-Eily LR, *et al.*** Collaborative care to improve the management of depressive disorders: A community guide systematic review and meta-analysis. *Am J Prev Med.* 2012;42:525-38. <http://dx.doi.org/10.1016/j.amepre.2012.01.019>
2. **Bromet E, Andrade LH, Hwang I, Sampson NA, Alonso J, De Girolamo G, *et al.*** Cross-national epidemiology of DSM-IV major depressive episode. *BMC Med.* 2011;29:90. <http://dx.doi.org/10.1186/1741-7015-9-90>

3. **Steel Z, Marnane C, Iranpour C, Chey T, Jackson JW, Patel V, et al.** The global prevalence of common mental disorders: A systematic review and meta-analysis 1980-2013. *Int J Epidemiol.* 2014;43:476-93. <http://dx.doi.org/10.1093/ije/dyu038>
4. **Ibrahim AK, Kelly SJ, Adams CE.** A systematic review of studies of depression prevalence in university students. *J Psychiatr Res.* 2013;47:391-400. <http://dx.doi.org/10.1016/j.jpsychires.2012.11.015>
5. **Amézquita-Medina ME, González-Pérez RE, Zuluaga-Mejía D.** Prevalencia de la depresión, ansiedad y comportamiento suicida en la población estudiantil de pregrado de la Universidad de Caldas, año 2000. *Rev Colomb Psiquiatr.* 2003;32:341-56.
6. **Falla G, Marcela P, Sanabria-Ferrand PA.** Prevalencia de la sintomatología depresiva en una población estudiantil de la Facultad de Medicina de la Universidad Militar Nueva Granada, Bogotá - Colombia. *Rev Fac Med.* 2006;54:76-87.
7. **Fergusson DM, Woodward LJ.** Mental health, educational, and social role outcomes of adolescents with depression. *Arch Gen Psychiatry.* 2002;59:225-31. <http://dx.doi.org/10.1001/archpsyc.59.3.225>
8. **Lepine J, Gastpar M, Mendlewicz J, Tylee A.** Depression in the community: The first pan-European study DEPRES (Depression Research in European Society). *Int Clin Psychopharmacol.* 1997;12:19-29.
9. **Vinas-Poch F, Villar E, Caparros B, Juan J, Cornella M, Pérez I.** Feelings of hopelessness in a Spanish university population - Descriptive analysis and its relationship to adapting to university, depressive symptomatology and suicidal ideation. *Soc Psychiatry Psychiatr Epidemiol.* 2004;39:326-34. <http://dx.doi.org/10.1007/s00127-004-0756-2>
10. **Kessler RC, Berglund P, Demler O, Jin R, Walters EE.** Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the national comorbidity survey replication. *Arch Gen Psychiatry.* 2005;62:593-602. <http://dx.doi.org/10.1001/archpsyc.62.6.593>
11. **Thombs BD, Ziegelstein RC.** Depression screening in primary care: Why the Canadian task force on preventive health care did the right thing. *Can J Psychiatry.* 2013;58:692-6.
12. **Lowe B, Spitzer RL, Grafe K, Kroenke K, Quenter A, Zipfel S, et al.** Comparative validity of three screening questionnaires for DSM-IV depressive disorders and physicians' diagnoses. *J Affect Disord.* 2004;78:131-40. [http://dx.doi.org/10.1016/S0165-0327\(02\)00237-9](http://dx.doi.org/10.1016/S0165-0327(02)00237-9)
13. **Williams JW, Pignone M, Ramírez G, Pérez-Stellato C.** Identifying depression in primary care: A literature synthesis of case-finding instruments. *Gen Hosp Psychiatry.* 2002;24:225-37. [http://dx.doi.org/10.1016/S0163-8343\(02\)00195-0](http://dx.doi.org/10.1016/S0163-8343(02)00195-0)
14. **Maurer DM.** Screening for depression. *Am Fam Physician.* 2012;85:139-44.
15. **Campo-Arias A, Oviedo HC.** Propiedades psicométricas de una escala: la consistencia interna. *Rev Salud Pública (Bogotá).* 2008;10:831-9. <http://dx.doi.org/10.1590/S0124-00642008000500015>
16. **Kroenke K, Spitzer RL, Williams JB.** The PHQ-9: Validity of a brief depression severity measure. *Gen Intern Med.* 2001;16:606-13. <http://dx.doi.org/10.1046/j.1525-1497.2001.016009606.x>
17. **Williams N.** PHQ-9. *Occup Med (Lond).* 2014;64:139-40. <http://dx.doi.org/10.1093/occmed/kqt154>
18. **Kocalevent RD, Hinz A, Brähler E.** Standardization of the depression screener Patient Health Questionnaire (PHQ-9) in the general population. *Gen Hosp Psychiatry.* 2013;35:551-5. <http://dx.doi.org/10.1016/j.genhosppsych.2013.04.006>
19. **Sung SC, Low CC, Fung DS.** Screening for major and minor depression in a multiethnic sample of Asian primary care patients: A comparison of the nine-item Patient Health Questionnaire (PHQ-9) and the 16-item Quick Inventory of Depressive Symptomatology - Self-Report (QIDS-SR16). *Asia Pac Psychiatry.* 2013;5:249-58. <http://dx.doi.org/10.1111/appy.12101>
20. **Baader T, Molina J, Venezian S, Rojas C, Farías R, Fierro-Freixenet C, 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-Psiquiat.* 2012;9:10-22. <http://dx.doi.org/10.4067/S0717-92272012000100002>
21. **Londoño A, Romero P, Casas G.** The association between armed conflict, violence and mental health: A cross sectional study comparing two populations in Cundinamarca department, Colombia. *Confl Health.* 2012;6:2-6. <http://dx.doi.org/10.1186/1752-1505-6-12>
22. **Castro-Camacho L, Escobar JM, Sáenz-Moncaleano C, Delgado-Barrera L, Aparicio-Turbay S, Molano JC, et al.** Salud mental en el hospital general: resultados del Cuestionario de Salud del Paciente (PHQ) en cuatro servicios de atención. *Rev Colomb Psiquiatr.* 2012;41:61-85. [http://dx.doi.org/10.1016/S0034-7450\(14\)60069-X](http://dx.doi.org/10.1016/S0034-7450(14)60069-X)
23. **Arango-Lasprilla JC, Plaza SL, Drew A, Romero JL, Pizarro JA, Francis A, et al.** Family needs and psychosocial functioning of caregivers of individuals with spinal cord injury from Colombia, South America. *Neuro Rehabilitation.* 2010;27:83-93. <http://dx.doi.org/10.3233/NRE-2010-0583>
24. **Wittkamp KA, Naeije L, Schene AH, Huyser J, van Weert HC.** Diagnostic accuracy of the mood module of the Patient Health Questionnaire: A systematic review. *Gen Hosp Psychiatry.* 2007;29:388-95. <http://dx.doi.org/10.1016/j.genhosppsych.2007.06.004>
25. **Devon HA, Block ME, Moyle-Wright P, Ernst DM, Hayden SJ, Lazzara DJ, et al.** A psychometric toolbox for testing validity and reliability. *J Nurs Sch.* 2007;39:155-64. <http://dx.doi.org/10.1111/j.1547-5069.2007.00161.x>
26. **Sánchez R, Echeverry J.** Validación de escalas de medición en salud. *Rev Salud Pública.* 2004;6:302-18. <http://dx.doi.org/10.1590/S0124-00642004000300006>
27. **Campo-Arias A, Herazo E, Oviedo H.** Análisis de factores: fundamentos para la evaluación de instrumentos de medición en salud mental. *Rev Colomb Psiquiatr.* 2012;41:657-71. [http://dx.doi.org/10.1016/S0034-7450\(14\)60036-6](http://dx.doi.org/10.1016/S0034-7450(14)60036-6)
28. **Batista-Foguet JM, Coenders G, Alonso J.** Confirmatory factor analysis. Its role on the validation of health related questionnaires. *Med Clin (Barc).* 2004;122(Suppl.1):21-7.
29. **Dum M, Pickren J, Carter-Sobell L, Sobell MB.** Comparing the BDI-II and the PHQ-9 with outpatient substance abusers. *Addict Behav.* 2008;33:381-7. <http://dx.doi.org/10.1016/j.addbeh.2007.09.017>

30. **Baas K, Cramer A, Koeter M, van De Lisdonk E, van Weert H, et al.** Measurement invariance with respect to ethnicity of the Patient Health Questionnaire-9 (PHQ-9). *J Affect Disord.* 2011;129:229-35. <http://dx.doi.org/10.1016/j.jad.2010.08.026>
31. **Krause J, Reed K, McArdlee JJ.** Factor structure and predictive validity of somatic and nonsomatic symptoms from the Patient Health Questionnaire-9: A longitudinal study after spinal cord injury. *Arch Phys Med Rehabil.* 2010;91:1218-24. <http://dx.doi.org/10.1016/j.apmr.2010.04.015>
32. **Richardson EJ, Richards JS.** Factor structure of the PHQ-9 screen for depression across time since injury among persons with spinal cord injury. *Rehabil Psychol.* 2008;53:243-9.
33. **Krause JS, Bombardier C, Carter R.** Assessment of depressive symptoms during inpatient rehabilitation for spinal cord injury: Is there an underlying somatic factor when using the PHQ? *Rehabil Psychol.* 2008;53:513-20. <http://dx.doi.org/10.1037/a0013354>
34. **Elhai JD, Contractor AA, Tamburrino M, Fine TH, Prescott MR, Shirley E, et al.** The factor structure of major depression symptoms: A test of four competing models using the Patient Health Questionnaire-9. *Psychiatry Res.* 2012;199:169-73. <http://dx.doi.org/10.1016/j.psychres.2012.05.018>
35. **Yoon S, Lee Y, Han C, Pae CU, Yoon HK, Patkar A, et al.** Usefulness of the patient health questionnaire-9 for Korean medical students. *Acad Psychiatry.* 2014;38:661-7. <http://dx.doi.org/10.1007/s40596-014-0140-9>
36. **Adewuya AO, Ola BA, Afolabi O.** Validity of the Patient Health Questionnaire (PHQ-9) as a screening tool for depression amongst Nigerian university students. *J Affect Disord.* 2006;93:89-93. <http://dx.doi.org/10.1016/j.jad.2006.05.021>
37. **Zinbarg RE.** Coefficients alpha, beta, omega, and the glb: Comments on Sijtsma. *Psychometrika.* 2009;74:145-54. <http://dx.doi.org/10.1007/s11336-008-9102-z>
38. **Posada-Villa JA, Trevisi-Fuentes C.** Prevalencia, severidad y necesidades no satisfechas del tratamiento de los trastornos de ansiedad, relacionados con sustancias, del estado de ánimo y del control de impulsos en adultos según el Estudio Nacional de Salud Mental, Colombia, 2003. *Medunab.* 2004;7:65-72.
39. **Sánchez-Pedraza R, Gómez-Restrepo C.** Conceptos básicos sobre validación de escalas. *Rev Col Psiquiatr.* 1998;27:121-30.
40. **Aguiar-Barojas S.** Fórmulas para el cálculo de la muestra en investigaciones de salud. *Salud en Tabasco.* 2005;11:333-8.
41. **Calderón M, Gálvez-Buccollini JA, Cueva G, Ordóñez C, Bromley C, Fiestas F.** Validación de la versión peruana del PHQ-9 para el diagnóstico de depresión. *Rev Peru Med Exp Salud Pública.* 2012;29:578-9. <http://dx.doi.org/10.1590/S1726-46342012000400027>
42. **Cassiani-Miranda CA, Pérez-Anibal E, Vargas-Hernández MC, Herazo-Bustos MI, Villarreal-Villanueva MP, Tuirán-Tuirán J, et al.** Desempeño psicométrico de la escala PHQ-9 para el tamizaje de sintomatología depresiva en estudiantes universitarios de Cartagena, año 2014. Tesis. Cartagena: Universidad de Cartagena; 2015.
43. **Cronbach LJ.** Coefficient alpha and the internal structure of tests. *Psychometrika.* 1951;16:297-334. <http://dx.doi.org/10.1007/BF02310555>
44. **McDonald RP.** Theoretical foundations of principal factor analysis, canonical factor analysis, and alpha factor analysis. *Br J Math Stat Psychol.* 1970;23:1-21. <http://dx.doi.org/10.1111/j.2044-8317.1970.tb00432.x>
45. **Zinbarg R, Revelle W, Yovel I, Li W.** Cronbach's alpha, Revelle's beta, and McDonald's omega hierarchical: Their relations with each other and two alternative conceptualizations of reliability. *Psychometrika.* 2005;70:123-33. <http://dx.doi.org/10.1007/s11336-003-0974-7>
46. **Kang H.** A guide on the use of factor analysis in the assessment of construct validity. *J Korean Acad Nurs.* 2013;43:587-94. <http://dx.doi.org/10.4040/jkan.2013.43.5.587>
47. **Bartlett MS.** Test of significance in factor analysis. *Br J Math Stat Psychol.* 1950;3:77. <http://dx.doi.org/10.1111/j.2044-8317.1950.tb00285.x>
48. **Kaiser HF.** An index of factorial simplicity. *Psychometrika.* 1974;34:31-6. <http://dx.doi.org/10.1007/BF02291575>
49. **Kahn JH.** Factor analysis in counseling psychology research, training, and practice: Principles, advances, and applications. *Couns Psychol.* 2006;34:684-718. <http://dx.doi.org/10.1177/0011000006286347>
50. **Yu X, Tam WW, Wong PT, Lam T, Stewart SM.** The Patient Health Questionnaire-9 for measuring depressive symptoms among the general population in Hong Kong. *Compr Psychiatry.* 2012;53:95-102. <http://dx.doi.org/10.1016/j.comppsy.2010.11.002>
51. **Chen S, Fang Y, Chiu H, Fan H, Jin T, Conwell Y.** Validation of the nine-item Patient Health Questionnaire to screen for major depression in a Chinese primary care population. *Asia Pac Psychiatry.* 2013;5:61-8. <http://dx.doi.org/10.1111/appy.12063>
52. **Wang W, Bian Q, Zhao Y, Li X, Wang W, Du J, et al.** Reliability and validity of the Chinese version of the Patient Health Questionnaire (PHQ-9) in the general population. *Gen Hosp Psychiatry.* 2014;36:539-44. <http://dx.doi.org/10.1016/j.genhosppsych.2014.05.021>
53. **Keszei AP, Novak M, Streiner DL.** Introduction to health measurement scales. *J Psychosom Res.* 2010;68:319-23. <http://dx.doi.org/10.1016/j.jpsychores.2010.01.006>
54. **Terwee CB, Bot S, De Boer MR, van der Windt DA, Knol DL, Dekker J, et al.** Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol.* 2007;60:34-42. <http://dx.doi.org/10.1016/j.jclinepi.2006.03.012>