

Validez de apariencia y adaptación de la escala PHQ-9 para la detección de sintomatología depresiva en universitarios de ciencias de la salud de Cartagena (Colombia)

Face validity and adaptation of the PHQ-9 for detection of depressive symptoms in university health sciences students Cartagena (Colombia)

Carlos Arturo Cassiani-Miranda¹, Eduard Pérez-Aníbal², María Camila Vargas-Hernández³, Mariana Herazo-Bustos⁴, Adrián Cabarcas-Tovar⁵

Resumen

Objetivos: Determinar la validez de apariencia del PHQ-9 para la detección de sintomatología depresiva en estudiantes de ciencias de la salud de la Universidad de Cartagena (Colombia).

Materiales y métodos: Estudio transversal que evalúa el comportamiento psicométrico de la escala PHQ-9. Se consultó a 10 expertos en el área de investigación y diagnóstico clínico. Para determinar la validez de apariencia se utilizó la pertinencia y adecuación. La pertinencia se realizó a través de la tasa de validez de contenido. Para evaluar la adecuación, se realizó un grupo de discusión.

Resultados: La pertinencia arrojó que todos los ítems de la escala PHQ-9 eran esenciales para evaluar la sintomatología depresiva en la población estudiada con una CVR igual a 1 en los nueve ítems evaluados (Esencial CVR ≥ 0.62 , $p = 0,05$). Para la adecuación se realizaron los ajustes idiomáticos de todos los ítems teniendo en cuenta las observaciones relevantes de los expertos y las conclusiones del grupo de discusión.

Conclusiones: El PHQ-9 mostró adecuada validez de apariencia, es necesario evaluar el desempeño psicométrico de esta versión de la escala en esta población.

Palabras clave: PHQ-9, validez de apariencia, depresión, estudiantes universitarios.

Fecha de recepción: 4 de septiembre de 2017
Fecha de aceptación: 25 de octubre de 2017

¹ Médico psiquiatra de la Universidad de Cartagena-Grupo de Investigación en Neurociencias-Universidad de Santander (UDES)-Bucaramanga, Colombia. <https://orcid.org/0000-0002-2288-1027>

² Médico residente de primer año de medicina interna de la Universidad de Cartagena-Grupo de Investigación en Psiquiatría y Salud Mental- Universidad de Cartagena. <https://orcid.org/0000-0002-9234-7217>

³ Médico residente de segundo año de medicina interna de la Universidad de Cartagena. Grupo de Investigación en Psiquiatría y Salud Mental- Universidad de Cartagena. <https://orcid.org/0000-0001-5420-2191>

⁴ Médico, residente de segundo año de Psiquiatría de la Pontificia Universidad Javeriana de Bogotá, Colombia. Grupo de Investigación en Psiquiatría y Salud Mental. <https://orcid.org/0000-0003-1383-2220>

⁵ Estudiante de cuarto año de medicina de la Universidad de Cartagena. <https://orcid.org/0000-0002-7273-4113>

Correspondencia: Carlos Arturo Cassiani Miranda. Hospital Psiquiátrico San Camilo. Calle 44 #7-04. Bucaramanga-Santander.

Abstract

Objective To determine the appearance validity of the PHQ-9 for the detection of depressive symptoms among health sciences students at the University of Cartagena, Colombia

Materials and Methods: A cross sectional study that assesses the psychometric properties of the PHQ-9 scale. It was consulted to 10 experts in the areas of research and clinical diagnosis. To determine the appearance validity, they were used the relevance and adequacy. Relevance determination was performed through rate content. To assess the adequacy, a panel discussion was conducted.

Results: The relevance found that all the items of the PHQ-9 scale were essential for depressive symptoms screening in this population. With a CVR equal to 1 in the nine items evaluated (CVR Essential ≥ 0.62 , $p = 0.05$). For adequacy, idiomatic settings of all items were made taking into account the relevant comments from experts and the panel discussion conclusions.

Conclusions: the Phq-9 scale has an adequate appearance validity, it is necessary to evaluate the psychometric properties of this version of the scale in our population.

Keywords: PHQ-9, face validity, depression, university students.

INTRODUCCIÓN

La depresión es un problema mayor de salud pública debido a su impacto social y económico (1). La prevalencia a lo largo de la vida y en el último año de Trastorno Depresivo Mayor (TDM) en la población general se sitúa entre 1.5 a 19.0 % y 0.8 % a 5.8 %, respectivamente, (2, 3). Sin embargo, en los estudiantes universitarios, estas cifras son sustancialmente más altas, con proporciones de prevalencia que oscilan entre 10 % y 85 %, con una media de 30.6 % (4) tanto a nivel mundial como en Colombia, en donde la de sintomatología depresiva en estudiantes universitarios se ubica en el rango entre 25 y 52.3 % (5–10).

Los síntomas depresivos se asocian con deterioro significativo en el desempeño cognoscitivo, el cual afecta en forma perjudicial el rendimiento académico y representa un estresor negativo adicional para la mayoría de estudiantes (11). A pesar de la importante carga global de los trastornos depresivos, cerca del 50 % de los pacientes con este problema no son identificados en contextos no clínicos (12). Teniendo en cuenta lo anterior y que la salud mental de los estudiantes universitarios tiene importantes

implicaciones para la salud pública (13), se ha establecido la importancia de identificar tempranamente la depresión en los estudiantes universitarios, ya que además estos trastornos debutan durante estas etapas de la vida y tienden a cronificarse (14). En consecuencia, el uso de cuestionarios de tamizaje cortos puede mejorar el reconocimiento de la depresión en diferentes contextos no clínicos (15).

Existen diversos instrumentos de tamizaje o escalas para la detección de sintomatología depresiva en población general (16,17), sin embargo, la mayoría se han desarrollado en inglés o en otros idiomas diferentes del español, lo que hace necesario realizar el proceso de validación, ya que existen diferencias a nivel idiomático y cultural que pueden alterar el ajuste del proceso de medición que traía la escala original (18).

Durante los últimos 20 años en Colombia se han desarrollado o validado pocos cuestionarios para el tamizaje de depresión. A pesar de su adecuada sensibilidad y especificidad, algunos de estos cuestionarios son muy extensos y dispendiosos para ser usados en

contextos no clínicos (19). Además todos han sido diseñados hace más de 3 décadas y no recogen los criterios diagnósticos más recientes del DSM (Manual Diagnóstico y Estadístico de Enfermedades Mentales, por sus siglas en inglés). Para resolver este problema se dispone recientemente del PHQ-9, que es la versión de 9 ítems del Cuestionario de Salud del Paciente, derivado de la Evaluación de Trastornos Mentales en Cuidado Primario (PRIME-MD), como herramienta de tamizaje de depresión en cuidado primario según los criterios diagnóstico del DSM-IV, el cual demuestra una adecuada sensibilidad y especificidad (20,21), y es sustancialmente más corto que la mayoría de escalas usadas para el tamizaje de depresión (22,23). El PHQ-9 ha sido traducido a varios idiomas, con varios estudios que respaldan su confiabilidad y validez con herramienta de tamizaje (24–27), sin embargo, sus propiedades psicométricas no ha sido evaluadas formalmente en la población colombiana.

La validación de escalas es un proceso continuo, de adaptaciones y ajustes para grupos o poblaciones específicas, sin dejar a un lado la equivalencia lingüística o cultural de la escala del idioma original a la nueva versión (28). Existen diferentes formas de validez, las cuales pueden ser de constructo, convergente, divergente, nomológica, de criterio (29). La validez de apariencia tiene como objetivo determinar si la escala parece medir lo que debe medir (29,30), por lo tanto no representa en sí un concepto estadístico, sino que depende del concepto que los expertos emitan sobre la pertinencia de los ítems de la escala. Para evaluar la validez de apariencia se conforma un grupo de jueces, por lo general expertos que determinan si en su concepto el instrumento en apariencia mide las cualidades deseadas, o por el contrario, otro de personas que van a ser evaluadas por

la escala (31). La importancia de esta forma de validez radica en la aplicabilidad y en la aceptabilidad de la escala en un contexto determinado (32).

Durante muchos años la adaptación se limitaba a la traducción casi literal de las escalas de medición para su posterior aplicación, sin prestar mayor atención a sus contenidos, lo que sin duda tiene importantes implicaciones en los hallazgos y su interpretación (33). Hoy en día, en el campo de la investigación transcultural la adaptación de instrumentos implica un proceso más riguroso que trasciende la traducción del idioma original de un instrumento, particularmente cuando es empleado en un contexto sociocultural y lingüísticamente distinto de aquel para el que fue creado, a fin de alcanzar la equivalencia de significado en ambos contextos (34,35). Sin embargo, los informes sobre adaptación lingüística y cultural del PHQ-9 son escasos.

Para evaluar determinar la validez de apariencia se utilizan algunas medidas, tales como la pertinencia y adecuación (36). La pertinencia hace referencia a la relevancia de cada ítem y su contribución a identificar el constructo evaluado. La adecuación valora la claridad, comprensión, redacción y lenguaje con que se expresa cada ítem, y su adaptación a la población donde se va aplicar la escala. Por lo tanto, la validez de apariencia constituye un paso metodológico importante antes de evaluar las propiedades psicométricas de una escala.

El objetivo de este informe es determinar la validez de apariencia del PHQ-9 para la detección de sintomatología depresiva en estudiantes de ciencias de la salud de la universidad de Cartagena (Colombia).

MATERIALES Y MÉTODOS

Este informe corresponde a la primera fase de un estudio transversal que evalúa el desempeño psicométrico de la escala PHQ-9 en la detección de sintomatología depresiva en estudiantes de ciencias de la salud en una universidad en Cartagena.

Para la realización del estudio se contó con la autorización por escrito para utilizar la versión para Colombia del PHQ-9 (disponible en <http://www.phqscreeners.com/>) por parte de uno de sus autores, Kurt Kroenke. Además, fue aprobado por el Comité de Ética de la Facultad de Medicina de la Universidad de Cartagena.

La versión de 9 ítems del Cuestionario de Salud del Paciente, o escala PHQ-9 por su nombre en inglés, *Patient Health Questionnaire 9*, es una versión autoadministrada del instrumento de diagnóstico, *PRIME-MD*, para los trastornos mentales comunes en atención primaria (20–22). Corresponde al módulo para tamizaje de depresión de acuerdo con los criterios diagnósticos del DSM-IV. El PHQ-9 puede ser usado de manera algorítmica para realizar el diagnóstico probable de un Trastorno Depresivo Mayor; o como una medida continua de puntajes que van de 0 a 27, con puntos de corte de 5, 10, 15 y 20, representando los niveles de síntomas depresivos como leve, moderado, moderadamente severo y severo, respectivamente (20).

El objetivo de la primera fase de este estudio fue determinar la validez de apariencia de la escala; para ello se consultó a 10 expertos en el área de investigación y diagnóstico clínico de depresión de cuatro regiones diferentes de Colombia. El grupo estuvo conformado por 7 psiquiatras, 2 enfermeras y 1 psicóloga,

con un promedio de 21,4 años de experiencia y suficiencia investigativa en el área y que además representarían las diferentes regiones del país. Se seleccionó el número de 10 expertos, puesto que supera el número mínimo requerido según la fórmula para que los resultados sean confiables. A cada experto se le envió por correo electrónico un formato en el que se les solicitaba la siguiente información: a) Que respondiera, marcando con una X, "SI" o "NO" si consideraban que cada uno de los ítems representaba o es relevante para medir los criterios diagnósticos de episodio depresivo según el DSM-IV. b) Que identificará palabras que hicieran el ítem incomprensible o indujeran errores en quienes contestan el test; o si consideraban hacer alguna corrección en la forma gramatical de plantear lo que el ítem pregunta, para que la medición del síntoma al que hace alusión fuera medido con exactitud según los criterios diagnósticos del DSM. c) Que respondieran si consideraban que en términos generales la escala permitía identificar adecuadamente, como herramienta de tamización, un episodio depresivo mayor en la población de este estudio. Con esta información se construyeron las medidas para evaluar la validez de apariencia: pertinencia y adecuación. La pertinencia valoró la relevancia de cada ítem y su contribución a detectar sintomatología depresiva. La adecuación evaluó la claridad, comprensión, redacción y lenguaje con que se expresa cada ítem y su adaptación a los estudiantes universitarios.

La determinación de la pertinencia se realizó a través del modelo estadístico propuesto por Lawshe y modificado en 2008 por Tristán (37) mediante la aplicación de la siguiente fórmula: $CVR = (E - (N/2)) / (N/2)$, donde CVR es la tasa de validez de contenido, E es el número de expertos que consideran esencial el ítem y N es el número total de expertos consultados

(38). El resultado se comparó de acuerdo con el número de expertos consultados con la el valor mínimo de CVR según el método de Lawshe. Para este estudio se determinó que un ítem era esencial si al aplicar la fórmula de CVR el valor obtenido era mayor que 0.62.

Para evaluar la adecuación se realizó un grupo de discusión conformado por los autores del proyecto (2 psiquiatras, 1 residente de Psiquiatría, 1 magíster en salud pública, 1 epidemiólogo y 2 estudiantes de último año de medicina), teniendo igual de valor por cada votante, donde se tuvo en cuenta las observaciones expresadas por los expertos para cada ítem. Para llegar a un consenso de la redacción de cada uno de los ítems de la escala se sometió a votación la inclusión o exclusión de algunos términos.

RESULTADOS

Pertinencia

La pertinencia de cada ítem, evaluada con la Razón de Validez de Contenido (0), arrojó que todos los ítems de la escala PHQ-9 eran esenciales para el tamizaje de sintomatología depresiva en la población estudiada, de acuerdo con la opinión de los expertos consultados, con una CVR igual a 1 en los nueve ítems evaluados (Esencial $CVR \geq 0.62$, $p = 0,05$) (tabla 1)

$CRV = \frac{e}{N}$; donde e = número de panelistas que tienen acuerdo en la categoría “esencial”, y N = número total de panelistas.

Tabla 1. Pertinencia de los ítems del PHQ-9 para tamizaje de sintomatología depresiva en estudiantes de ciencias de la salud de la Universidad de Cartagena CVR: Content Value Ratio – Razón de Validez de Contenido

Item	Esencial	No esencial	CVR
1	10	0	1
2	10	0	1
3	10	0	1
4	10	0	1
5	10	0	1
6	10	0	1
7	10	0	1
8	10	0	1
9	10	0	1

Fuente: datos tabulados por los autores.

Adecuación

En el análisis por ítem se resumen a continuación las observaciones relevantes de los expertos, que después del debate interno en el grupo investigador tuvieron el mayor peso a la hora de ajustar la estructura gramatical de las preguntas de la escala (tabla 2). En la columna de decisión de la tabla se indica con un “SI” si el consenso del grupo de discusión llevó a cabo el cambio propuesto por los expertos en la estructura de la escala.

Tabla 2. Re sumen de las observaciones de los expertos y el resultado final del debate interno del grupo investigador sobre la estructura gramatical del ítem

Ítem	Corrección	Experto	Votos a favor	Votos en contra	Decisión
1	Precisar pasatiempos, deseo o actividad sexual en lugar de cosas	ACA	2	5	No
	Reemplazar placer por "Poco placer"	CAA	0	7	No
	Agregar "Ha sentido" al inicio del enunciado	CLJ	7	0	Sí
	Iniciar el enunciado con las palabras "Tiene" o "Tienes"	ZC	0	7	No
	Incluir "Habitualmente le gustan" para finalizar el enunciado	ZC	7	0	Sí
2	Incluir "irritabilidad" o "mal humor" en el ítem	ACA	7	0	Sí
	Agregar "Ha sentido" al inicio del enunciado	CLJ	7	0	Sí
	Anteceder al enunciado las sílabas Se o Te	ZC	0	7	No
	Eliminar "o/a" de la redacción del ítem	RC	2	5	No
3	Reemplazar la palabra "deprimido" por "triste"	CAG	7	0	Sí
	Incluir "más de lo habitual" en la redacción del ítem	ACA	7	0	Sí
	Ha tenido dificultad para dormir o permanecer dormido	CLJ	7	0	Sí
	Dificultad para dormir o dormir más de lo habitual	CAA	7	0	Sí
	"Mantenerse en el sueño" puede inducir a error	EH	7	0	Sí
4	Mantenerse lo hace complejo	RC	7	0	Sí
	Precisar "más de lo usual" o "sin razón aparente" al final del ítem	ACA	7	0	Sí
	Agregar "se ha sentido" al inicio del ítem	CLJ	7	0	Sí
	Incluir retraimiento y aislamiento	MV	1	6	No
5	Los o/a son innecesarios. Se hacen documentos por género	RC	2	5	No
	Agregar "más de lo habitual" al final del ítem	ACA	7	0	Sí
6	Se ha sentido sin apetito o ha comido demasiado	CLJ	7	0	Sí
	Cambiar el enunciado por "Se ha sentido mal con usted mismo, o ha sentido que es un fracaso, o que ha quedado mal con usted mismo o con su familia"	CLJ	7	0	Sí
	Remover "sentirse mal acerca de sí mismo"	EH	2	5	No
	Cambiar "mal acerca de sí mismo" por "mal consigo mismo"	GR	5	2	Sí
	Cambiar el enunciado por: "Me siento mal conmigo mismo, incluso, he pesado que soy un fracaso y he percibido abandono propio o de la familia", además de dividir el ítem en dos.	ZC	0	7	No
7	En la redacción del ítem colocar sentimientos en vez de sentimiento	RC	0	7	No
	Dificultad para concentrarse en diferentes actividades, tales como leer el periódico, ver televisión, entre otros.	NH	2	5	No
	Se puede eliminar la palabra "periódico" dado lo infrecuente en la actualidad de esa actividad en el contexto mundial.	ACA	3	4	No
	Agregar "ha tenido dificultad" al inicio del ítem	CLJ	7	0	Sí
8	Mencionar la dificultad para concentrarse mientras estudia y/o realiza actividades académicas	MV	1	6	No
	Agregar al ítem "se ha movido o ha hablado"	CLJ	7	0	Sí
	Remover del ítem "bien, por el contrario"	CAA	1	6	No
	Los o/a son innecesarios. Se hacen documentos por género	RC	2	5	No

Continúa...

Ítem	Corrección	Experto	Votos a favor	Votos en contra	Decisión
9	Ha tenido pensamientos de que estaría mejor muerto o ha tenido deseos de lastimarse de alguna manera	CLJ	7	0	No
	Considera que "Deseos de lastimarse" puede inducir a error	EH	7	0	Sí
	Cambiar : "Estar muerto" por "no estar vivo"	CAG	2	5	No

Fuente: datos tabulados por los autores.

Ítem 1

Los expertos recomiendan agregar la expresión "Ha sentido" al inicio del enunciado con el objetivo de personalizar la pregunta y que el entrevistado perciba que le están preguntando directamente a él o ella. Se adopta esta recomendación debido que por estar la escala original en inglés, al traducirla al español no se tuvo en cuenta las diferencias lingüísticas de ambos idiomas, por lo cual la traducción original muestra impersonalidad y se hace necesario personalizarla para obtener un mejor resultado al ser aplicado en un país hispano. Esta misma observación aplica para los ítems 1

al 6. Además sugieren agregar la expresión "que habitualmente le gustan" al finalizar el enunciado. Por consenso se acepta esta recomendación, porque con ello se le da mayor especificidad a la pregunta, ya que lo que se desea es determinar una alteración cuyo patrón se viene manifestando hace dos semanas y rompe con el patrón habitual de funcionamiento del individuo. Así, el enunciado es modificado *¿Ha sentido poco interés o placer en hacer las cosas que habitualmente le gustan? A pesar de que pudiera ser redundante debido a que al iniciar la escala hay una instrucción general, esta habitualmente es olvidada o no tomada en cuenta en los individuos cuando contestan las escalas de este tipo.*

Ítem	Escala original	Escala propuesta
Encabezado	Durante las últimas 2 semanas, ¿con qué frecuencia ha sentido molestias por los siguientes problemas?	Durante las últimas 2 semanas, ¿con qué frecuencia ha sentido molestias por los siguientes problemas?
1	Poco interés o placer en hacer las cosas	Ha sentido poco interés o placer en hacer las cosas que habitualmente le gustan
2	Sentirse desanimado/a, deprimido/a o sin esperanzas	Se ha sentido desanimado/a, triste, irritable o sin esperanzas
3	Problemas para dormir o mantenerse el sueño o dormir demasiado	Ha tenido dificultad para dormir o permanecer dormido(a) o ha dormido más de lo habitual
4	Sentirse cansado/a o con poca energía	Se ha sentido cansado/a o con poca energía sin razón aparente
5	Sentir poco apetito o comer en exceso	Se ha sentido con poco apetito o ha comido demasiado o más de lo habitual
6	Sentirse mal acerca de sí mismo o tener un sentimiento de fracaso o de abandono propio o de la familia	Se ha sentido mal con usted mismo, ha sentido que es un fracaso, o que ha quedado mal con usted mismo o con su familia
7	Dificultad para concentrarse en diferentes actividades tales como leer el periódico o ver televisión	Ha tenido dificultad para concentrarse en diferentes actividades, tales como leer el periódico o ver televisión
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	Se ha movido o ha hablado tan despacio que otras personas lo han notado o, por el contrario, ha estado tan inquieto/a o intranquilo/a que se ha movido mucho más de lo habitual
9	Pensamientos acerca de que sería mejor estar muerto/a o deseos de lastimarse de alguna forma	Ha pensado que estaría mejor muerto o ha pensado en lastimarse o hacerse daño de alguna manera

Ítem 2

Se decidió cambiar el término “Deprimido” por “Triste”, ya que el primero denota un estado patológico, que constituye la esencia del constructo de episodio depresivo mayor según el DSM y es un término que engloba muchos aspectos del fenómeno “estar deprimido”; sin embargo, el término triste, es más específico porque hace referencia al estado de ánimo triste que es lo que trata de indagar este ítem de la escala. Por consiguiente, se considera más pertinente incluir el último término. Adicionalmente, los expertos sugieren agregar el término “Irritabilidad” a la construcción del enunciado, ya que este hace parte de los criterios diagnósticos de depresión incluidos en el DSM-IV y que cobra importancia en población joven y adolescente, pudiendo llegar a ser la única manifestación en este grupo específico. Además, el ánimo triste y la irritabilidad, a pesar de ser dos síntomas diferentes, hacen parte del mismo ítem dentro de criterio A de trastorno depresivo mayor del DSM. Así, el enunciado es modificado por *Se ha sentido desanimado/a, triste, irritable o sin esperanzas.*

Ítem 3

Los expertos recomendaron cambiar la expresión “Problemas para dormir” por “Dificultad para dormir” por considerar que el último ofrece un mejor estilo de redacción para el ítem y lo hace más comprensible y específico según su experiencia. Con respecto a la expresión “Mantenerse en el sueño”, utilizada en la versión colombiana de la escala PHQ-9, los expertos consideran que dicha terminología podría inducir a error por dificultad en la comprensión del ítem, lo cual fue aceptado por consenso, y el grupo de trabajo propone como alternativa “Permanecer dormido”. Finalmente, los expertos consideran que es pertinente cambiar la expresión “Dormir

demasiado” por “Dormido(a) más de lo habitual”, ya que la redacción de la escala debe ser realizada de la forma más clara posible para los entrevistados, y la expresión “Más de lo habitual” refleja, de acuerdo con los expertos, de una forma más fidedigna un cambio en el patrón de sueño, puesto que dormir demasiado puede ser un hábito frecuente en la población de estudio. De esta forma, el enunciado es modificado por *Ha tenido dificultad para dormir o permanecer dormido(a) o ha dormido más de lo habitual.*

Ítem 4

Los expertos recomiendan finalizar el enunciado con la frase “Sin razón aparente”, puesto que el sentirse cansado o sin energía puede tener una explicación clara para el entrevistado, pero al no enfatizar en si conoce la razón de su estado, puede inducir a puntuar de forma positiva en el ítem y sobreestimar el número de personas que puntúan con SDIC. Estas recomendaciones son aceptadas por consenso y el ítem se modifica por *Se ha sentido cansado/a o con poca energía sin razón aparente.*

Ítem 5

Los expertos recomiendan cambiar la expresión “Comer demasiado” por “Ha comido demasiado o más de lo habitual”; estas recomendaciones se realizan con el fin de detectar cambios en el patrón de alimentación y fueron aceptadas en consenso; por lo que el ítem se modifica a: *Se ha sentido con poco apetito, ha comido demasiado o más de lo habitual.*

Ítem 6

Los expertos sugieren continuar con el estilo personalizado de redacción en el resto del ítem al cambiarlo por “Se ha sentido mal con usted mismo, o ha sentido que es un fracaso,

o que ha quedado mal con usted mismo o con su familia” y agregar algunas palabras que hicieran la redacción del ítem más entendible para el contexto. De esta forma, la redacción del ítem es modificada por *Se ha sentido mal con usted mismo, ha sentido que es un fracaso, o que ha quedado mal con usted mismo o con su familia.*

Ítem 7

Los expertos recomiendan agregar “Ha tenido” en el encabezado del ítem para darle tono personalizado a la pregunta, por lo tanto el ítem es modificado a: *Ha tenido dificultad para concentrarse en diferentes actividades, tales como leer el periódico o ver televisión.* Uno de los expertos sugirió eliminar la palabra “periódico” debido a que esta es una actividad infrecuente en el ámbito mundial, sin embargo, no hubo consenso sobre esta recomendación y el grupo de discusión no consideró pertinente el cambio con base en el argumento de no producir cambios idiomáticos en la escala que no partieran de una recomendación relevante y fruto de consenso a la hora de la identificación del constructo.

Ítem 8

Al igual que en ítems anteriores, los expertos recomiendan personalizar el estilo de redacción del ítem cambiando “Moverse o hablar tan despacio” por “Se ha movido o ha hablado tan despacio”. Además, el grupo de trabajo decide cambiar “O bien por el contrario” por “Por el contrario” por considerar que genera mayor claridad para el entrevistado; así el ítem es modificado por *¿Se ha movido o ha hablado tan despacio que otras personas lo han notado o, por el contrario, ha estado tan inquieto/a o intranquilo/a que se ha movido mucho más de lo habitual?*

Ítem 9

Los expertos recomiendan personalizar el estilo de redacción del ítem, separándolo por ideas; asimismo, consideran incluir al enunciado la frase “Hacerse daño”, además de la idea de “lastimarse”, ya que según ellos se puede entender que el término “lastimarse” hace alusión directa a formas de daño físico, mientras que la expresión “Hacerse daño” busca indagar por otras formas de autoagresión que pueden tener connotación clínica significativa en pacientes deprimidos con ideas suicidas; de forma que el enunciado queda de la siguiente manera: *Ha pensado que estaría mejor muerto o ha pensado en lastimarse o hacerse daño de alguna manera,* y es aceptado por consenso. En este ítem en particular existió la discusión con respecto a cuál es la mejor forma de indagar por las ideas de muerte, si preguntando “Ha pensado que estaría mejor muerto” o “Ha pensado que es mejor no estar vivo”, lo cual permitió ratificar que de hecho indagar las ideaciones de muerte preguntando si “estaría mejor muerto” es la forma más acertada, ya que evoca más fácilmente las ideas de muerte en el entrevistado según la experiencia de los expertos y los participantes en el grupo de discusión.

DISCUSIÓN

Según la opinión de los expertos consultados, los ítems de la escala PHQ-9 son esenciales para evaluar la de sintomatología depresiva en la estudiantes universitarios de ciencias de salud de Cartagena.

El procedimiento generalmente utilizado para evaluar la validez de apariencia de las escalas en salud mental es el juicio de expertos y el índice cuantitativo disponible más usado hasta el momento es el CVR (Content Validi-

ty Ratio) de Lawshe (37,38). En este trabajo hubo un 100 % de acuerdo entre los expertos en cuanto a la pertinencia de los ítems, y se obtuvo un CVR total de 1 ($P < 0.05$), lo cual indica una adecuada validez de apariencia.

Resultados similares se han obtenido en Latinoamérica; en el contexto de sistema de atención primaria en Perú, Calderón et al., mediante una reunión técnica con la presencia de especialistas en salud mental conformado por de 23 psiquiatras, 3 psicólogos y una enfermera, informaron que el PHQ-9 es un instrumento con adecuada validez de apariencia, sin embargo, los autores no informan el método utilizado para determinar la pertinencia de los ítems, a diferencia de este trabajo, en el que todos los ítems son esenciales aplicando el CVR (39).

El proceso de adecuación permitió realizar algunos ajustes gramaticales en el instrumento a fin de hacerlo más comprensible en el contexto peruano; tales como: en el ítem 2 se consideró incluir la palabra “triste” para reforzar el concepto de deprimido; en nuestro trabajo se decidió reemplazar deprimido por estar triste, puesto que el estado de ánimo triste es el indagado por la escala de tamizaje; de la misma manera, en ambas escalas, peruana y colombiana, en el ítem 9 se añadió la frase “hacerse daño”, ya que la palabra “lastimarse” no es considerada de amplio uso en su contexto para indagar sobre formas de autoagresión. En la versión peruana se recomendó no especificar la palabra “periódico” en el ítem 7; en nuestro trabajo, un experto sugirió eliminarla, sin embargo, es una actividad de uso rutinario en nuestro país.

En un estudio de validación hondureña (40) del PHQ-9 en atención primaria, Wulsin et al. describen que el único ítem que precisó

modificación fue el 7, que indaga concentración, en donde la expresión “leer el periódico o ver televisión” fue sustituida por actividades alternativas como “tareas de la casa” o “conversar” teniendo en cuenta los hábitos de cada paciente, debido a que la mayoría de la población evaluada no leía periódico o no veía televisión, lo cual hace interpretar que variaba mucho este ítem; a pesar de ello, la población evaluada no informó mayores dificultades en el diligenciamiento y ningún ítem requirió modificaciones lingüísticas mayores para su comprensión o ajuste cultural, con lo que se concluyó una adecuada validez de apariencia sin descripción de la pertinencia de los ítems (40). Situación distinta es lo observado en la población universitaria de esta ciudad, en la que estas actividades sí son frecuentes; por lo que a la sugerencia de un experto sobre eliminar la palabra “periódico” en el grupo de discusión no hubo consenso sobre esta recomendación; además de considerar que no tenía mayor relevancia a la hora de identificar el síntoma que se pretende evocar.

Por otra parte, Sherina et al. en Malasia informaron que mediante un panel de expertos conformado por psiquiatras, especialistas en salud pública, clínicos e investigadores en una prueba piloto conformada por 80 sujetos en la que se determinaron los “valores de apariencia” del instrumento en una localidad diferente de la del estudio pero con características socioeconómicas similares, se encontró una adecuada validez de apariencia y aceptabilidad del instrumento sin cambios significativos en la estructura gramatical de los ítems (41).

Sin embargo existen otras metodologías para evaluar la validez de apariencia que se caracterizan por que no consultan con expertos sino que ponen el instrumento a

consideración de las población a la que será aplicada para que emitan su juicio sobre la aceptabilidad, comprensión, etc. En este sentido existen varias publicaciones que abordan el problema de la validez de apariencia en poblaciones con particularidades lingüísticas y culturales como los aborígenes. Así, un estudio realizado en una comunidad aborígen de Australia con metodología de grupos focales valoró la aceptabilidad y la validez de apariencia del PHQ-9 y se encontró que en general esta comunidad consideró adecuada esta herramienta de tamizaje para sintomatología depresiva (42).

Si bien los creadores de la escala PHQ-9 presentaron una versión oficial en castellano de uso libre para Colombia, no existen datos publicados que evalúen sus proceso de validez de apariencia o adaptación cultural; por lo que la primera fase de este estudio consistió en evaluar las características gramaticales según las observaciones de los expertos y el panel de discusión del grupo investigador. Este proceso condujo a la obtención de una versión adaptada para la población de estudiantes ciencias de salud de una universidad en Cartagena cuyas propiedades psicométricas serán informadas posteriormente.

CONCLUSIONES

Los resultados obtenidos en este estudio permiten afirmar que el instrumento cuenta con una adecuada validez de apariencia, y a pesar de que se modificaron algunos de los ítems del instrumento, se logró obtener consistencia en los mismos en lo que respecta al constructo evaluado. Esta validez le confiere a la escala aplicabilidad en los estudiantes universitarios del área de las ciencias de la salud en esta universidad de Cartagena, sin embargo, es necesario continuar con la validación de las

propiedades psicométricas de esta versión de la escala PHQ-9 adaptada para Colombia, que permitirá a clínicos e investigadores disponer de una herramienta de tamizaje breve, fácil de usar y confiable para la identificación de síntomas depresivos en poblaciones de características similares dentro del país.

Financiación: Recursos propios de los investigadores

Conflicto de intereses: Ninguno que declarar

Agradecimientos

ACA – Adalberto Campo Arias

CAA – César Augusto Arango Dávila

CLJ - Carlos López Jaramillo

ZC – Zuleima Cogollo

RC – Roberto Chaskel

CAG – Christian Ayola Gómez

EH – Edwin Herazo

MV – María Paula Villarreal

GR – Germán Rueda

NH – Nataly Herazo

REFERENCIA

1. Thota AB, Sipe TA, Byard GJ, Zometa CS HR, 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.
2. Evelyn Bromet LHA, Hwang I, Sampson NA, Alonso J, Girolamo G de. Cross-national epidemiology of DSM-IV major depressive episode. 2011.
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(2):476–93.

4. Ibrahim AK, Kelly SJ, Adams CE, Glazebrook C. A systematic review of studies of depression prevalence in university students. *J Psychiatr Res.* 2013;47(3):391–400.
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(4):341–56.
6. Campo-Cabal, G., & Gutierrez JC. Psicopatología en estudiantes universitarios de la Facultad de Salud-Univalle. *Rev colomb Psiquiatr.* 2001;30(4):351–8.
7. Arboleda MA, Gutiérrez JC, Miranda CA. Prevalencia de síntomas depresivos en médicos internos y residentes de la Universidad del Valle. *Médicas UIS.* 2000;15(1):4–7.
8. Miranda C, Gutierrez J E AL. Prevalencia de depresión en estudiantes de Medicina de la U. del Valle. *Rev colomb Psiquiatr.* 1998;29(3):251–60.
9. Jimenez TG, Ortiz L. Caracterización de la depresión en jóvenes universitarios. *Univ Psychol Bogotá (Colombia).* 2004;3(1):17–26.
10. Falla G, Marcela P SFP. 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(2):76–87.
11. Fergusson DM WL. Mental health, educational, and social role outcomes of adolescents with depression. *Arch Gen Psychiatry.* 2002;59(3):225–31.
12. 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.
13. Vinas Poch F, Villar E, Caparros B, Juan J, Cornella M PI. 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(4):326–34.
14. 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(6):593–602.
15. 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(12):692–6.
16. Lowe B, Spitzer RL, Gr??fe 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(2):131–40.
17. Williams JW, Pignone M, Ramirez G, Perez Stellato C. Identifying depression in primary care: A literature synthesis of case-finding instruments. *Gen Hosp Psychiatry.* 2002;24(4):225–37.
- 18.
19. 18. Campo-Arias A, Oviedo HC. Propiedades psicométricas de una escala: la consistencia interna. *Rev Salud Pública.* 2008;10(5):831–9.
20. Maurer DM. Screening for depression. *Am Fam Physician.* 2012;85(2):139–44.
21. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9: Validity of a brief depression severity measure. *Gen Intern Med.* 2001;16(9):606–13.
22. Kroenke K, Spitzer RL, Williams JBW, L??we B. The Patient Health Questionnaire Somatic, Anxiety, and Depressive Symptom Scales: A systematic review. *Gen Hosp Psychiatry.* 2010;32(4):345–59.
23. 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(5):551–5.
24. Sung SC, Low CC, Fung DS CY. 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(5):249–58.
25. Huang FY, Chung H, Kroenke K, Delucchi KL, Spitzer RL. Using the Patient Health Questionnaire-9 to measure depression among racially and ethnically diverse primary care patients. *J Gen Intern Med*. 2006;21(6):547–52.
 26. Omoro S, Fann J, Weymuller E, Macharia I, Yueh B. Swahili Translation and Validation of Thepatient Health Questionnaire-9 Depressionscale in the Kenyan Head and Neck Cancerpatient Population. *Int J Psychiatry Med*. 2006;36(3):367–81.
 27. 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(5):388–95.
 28. Baader T, Molina J, Venezian S, Rojas C, Farías R, Fierro-freixenet C. BM& MC. 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(1):10–22.
 29. 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(2):155–64.
 30. Sánchez R, Echeverry J. Validación de escalas de medición en salud. *Rev Salud Publica*. 2004;6(3):302–18.
 31. Sanchez-Pedraza R G-RC. Conceptos básicos sobre validación de escalas. *Rev Col Psiquiatr*. 1998;27(2):121–30.
 32. Lamprea J, Gómez-Restrepo C. Validez en la evaluación de escalas. *Rev colomb Psiquiatr [Internet]*. 2007;36(2):340–8. Available from: <http://bases.bireme.br/cgi-bin/wxislind.exe/iah/online/?IsisScript=iah/iah.xis&src=google&base=LILACS&lang=p&nextAction=lnk&exprSearch=456969&indexSearch=ID>
 33. Streiner DL. A checklist for evaluating the usefulness of rating scales. [Internet]. *Can J Psychiatry*. 1993; 38:| 140–8. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/8467441>
 34. Escobar Bravo MÁ. Adaptación transcultural de instrumentos de medida relacionados con la salud. *Rev Asoc Med Bahía Blanca*. 2004;14(2):102–6.
 35. Mammadova F, Sultanov M, Hajiyeva A, Aichberger M, Heinz A. Translation and adaptation of the Zung Self-Rating Depression Scale for application in the bilingual Azerbaijani population. *Eur Psychiatry*. 2012;27(Suppl.2):S27–31.
 36. Stansfield CW. Test translation and adaptation in public education in the USA. *Lang Test*. 2003;20(2):189–207.
 37. Álvarez R. Estadística aplicada a las ciencias de la salud. 2007.
 38. Tristán-López A. Modificación al modelo de Lawshe para el dictamen cuantitativo de la validez de contenido de un instrumento objetivo. *Av en medición*. 2008;6:37–48.
 39. Lawshe C. A Quantitative Approach To Content Validity. *Pers Psychol [Internet]*. 1975;(1):563–75. Available from: <http://www.blackwell-synergy.com/doi/abs/10.1111/j.1744-6570.1975.tb01393.x>
 40. Calderón M, Gálvez-Buccollini JA, Cueva G, Ordoñez C, Bromley C FF. Validación de la versión peruana del PHQ-9 para el diagnóstico de depresión. *Rev Peru Med Exp Salud Publica*. 2012;29(4):578–9.
 41. Wulsin L, Somoza E HJ. The Feasibility of Using the Spanish PHQ-9 to Screen for Depression in Primary Care in Honduras. *Prim Care Companion J Clin Psychiatry*. 2002;4(5):191–5.
 42. Sherina MS, Arroll B, Goodyear-Smith F. Criterion validity of the PHQ-9 (Malay version) in a primary care clinic in Malaysia. *Med J Malaysia*. 2012;67(3):309–15.
 43. Esler DM, Johnston F, Thomas D. The acceptability of a depression screening tool in an urban, aboriginal community-controlled health service. *Aust N Z J Public Heal*. 2007;31(3):259–63.