

ARTÍCULO ORIGINAL

## Validez y confiabilidad de la escala del *Center for Epidemiologic Studies-Depression* en estudiantes adolescentes de Colombia

Paul Anthony Camacho<sup>1</sup>, German Eduardo Rueda-Jaimes<sup>2</sup>, José Fidel Latorre<sup>1</sup>,  
Álvaro Andrés Navarro-Mancilla<sup>2</sup>, Mauricio Escobar<sup>2</sup>, Jorge Augusto Franco<sup>2</sup>

<sup>1</sup> Maestría de Epidemiología, Facultad de Medicina, Universidad Industrial de Santander, Bucaramanga, Colombia

<sup>2</sup> Grupo de Neuropsiquiatría, Centro de Investigaciones Biomédicas, Facultad de Medicina, Universidad Autónoma de Bucaramanga, Bucaramanga, Colombia

**Introducción.** El trastorno depresivo mayor es la segunda causa productora de discapacidad en Latinoamérica. No hay escalas validadas en Colombia para adolescentes.

**Objetivo.** Evaluar la validez y la confiabilidad de la escala del Centro de Estudios Epidemiológicos para la Depresión en adolescentes colombianos en edad escolar.

**Materiales y métodos.** Se llevó a cabo un estudio de validación de la escala con una muestra de adolescentes en edad escolar de Bucaramanga. Se evaluaron con la escala para depresión y la entrevista clínica estructurada. La escala fue aplicada nuevamente 3 a 28 días después. Se determinaron la validez de criterio, la consistencia interna y la reproducibilidad prueba re prueba.

**Resultados.** Se evaluaron 390 adolescentes con edad promedio de  $14,8 \pm 1,22$  años. La prevalencia de trastorno depresivo mayor fue de 11,5%. El índice de consistencia, o alfa de Cronbach, fue de 0,85. El área bajo la curva según las características receptor operador fue de 0,82 y el punto de corte mayor o igual a 23 mostró una sensibilidad de 73,3%. Se encontró: especificidad de 73,6%, valor diagnóstico positivo de 26,6% y valor diagnóstico negativo de 95,5%. El coeficiente de concordancia de Lin fue de 0,75.

**Conclusión.** La validez y la confiabilidad de la versión en español de la escala en adolescentes en edad escolar del Centro de Estudios Epidemiológicos para la Depresión, fueron similares a las informadas anteriormente, aunque cambió el punto de corte.

**Palabras clave:** reproducibilidad de resultados, validez de las pruebas, trastorno depresivo mayor, psiquiatría del adolescente, escalas de valoración psiquiátrica, cribado.

### Validity and reliability of the Center for Epidemiologic Studies-Depression scale on Colombians adolescent students

**Introduction.** Major depressive disorder is the second major cause of adolescent psychological incapacitation in Latin-America. However, scales for detecting these disorders have not been validated for screening adolescents in Colombia.

**Objective.** The validity and reliability of a Spanish translation of the Center for Epidemiologic Studies (CES-D)-Depression scale was assessed in adolescent students.

**Materials and methods.** A validation study for a diagnostic scale was performed with a sample of 390 adolescent students from Bucaramanga, Santander Province, in northwestern Colombia. The students were evaluated by two methods: (a) the CPS-depression scale and (b) a semi-structured clinical interview. Three to 28 days after the interview, the scale was re-applied. Criterion validity, internal consistency and test-retest reliability was analyzed.

**Results.** The mean age was  $14.8 \pm 1.2$  years old. The prevalence of major depressive disorder was 11.5%. Cronbach's alpha was 0.85. The area under the curve produced by the receiver operating characteristic curve was 0.82, and the cut point of  $\geq 23$  showed a sensitivity of 73.3%; specificity, 73.6%; positive predictive value, 26.6%, and negative predictive value, 95.5%. Lin's coefficient of concordance was 0.75.

**Conclusions.** The validity and reliability of the Spanish translation of the CES-D scale were similar to those reported in the international literature although with a higher cut point.

**Key words:** Reproducibility of results, validity of tests; depressive disorder, major; adolescent psychiatric, psychiatric status rating scales.

La Organización Mundial de la Salud, en el informe del 2001 sobre la carga global de las enfermedades, determinó que el trastorno depresivo mayor era la segunda causa más importante de América Latina y el Caribe en años de vida ajustados en función de la discapacidad (*Disability Adjusted Life Years, DALY*) (1).

El trastorno depresivo mayor en los adolescentes causa gran malestar clínico y problemas interpersonales, ausentismo escolar y bajo rendimiento (2); también, aumenta el riesgo de abuso de sustancias, las conductas antisociales, las conductas riesgosas tanto sexuales como en la alimentación, los intentos suicidas y los suicidios (3-5); además, se presenta concomitantemente con otros trastornos mentales (6,7). Todo lo anterior demanda una gran utilización de recursos familiares, médicos y sociales (8).

La prevalencia del trastorno depresivo mayor en adolescentes puede variar entre 0,7% y 6,1% (9-11). Su presentación es cada vez más temprana y se evidencia un aumento en su frecuencia (9).

El trastorno depresivo mayor es difícil de detectar, tanto por la diversidad en la presentación de los síntomas (somáticos y psicológicos) como por la dificultad para la expresión de los mismos por razones culturales y emocionales, que incluyen el estigma del diagnóstico. Los adolescentes carecen de introspección de su enfermedad y sus padres no detectan el problema, pues el trastorno depresivo mayor genera síntomas "interiorizados", es decir, con poca manifestación hacia los demás, lo cual lleva a una pobre solicitud de los servicios de salud (12-14). De esta forma, el trastorno depresivo mayor en la

adolescencia se diagnostica y trata menos de lo que se podría, a pesar de la existencia de tratamientos efectivos (15). Las consecuencias de estas falencias no son sólo las ya mencionadas, además, están la morbilidad y la disfunción social en la vida adulta (16).

El diagnóstico del trastorno depresivo mayor es complejo e implica la realización de una entrevista por un profesional entrenado. Sin embargo, existen varias escalas "autoinformadas" para tamización, que aumentan la probabilidad de realizar el diagnóstico (17). En la actualidad, existen varias escalas para la detección de síntomas depresivos en población infantil y adolescentes. En la mayoría de los casos, se diseñaron para adultos y, luego, fueron validadas en población menor de 18 años. Entre las más usadas están el inventario de depresión de Beck, la escala "autoaplicada" de depresión de Zung, la escala del *Center for Epidemiologic Studies-Depression* (CES-D), el inventario de depresión de niños y la escala de depresión de adolescentes de Reynolds (18). La escala CES-D se utiliza para la detección de síntomas depresivos en adolescentes, a pesar de haberse diseñado inicialmente para adultos; posee propiedades psicométricas similares a las escalas utilizadas en menores y tiene la ventaja de que puede usarse sin restricción (18,19). Sin embargo, para su utilización en poblaciones específicas como los adolescentes colombianos, es necesario conocer las propiedades psicométricas y ajustar el punto de corte de acuerdo con las variaciones socioculturales e individuales de la población.

El objetivo de este estudio fue evaluar la validez y la confiabilidad de la escala CES-D en adolescentes, con miras a facilitar la tamización, la detección precoz y el tratamiento temprano del trastorno depresivo mayor.

## Materiales y métodos

Se realizó un estudio de validación de una prueba diagnóstica con muestreo transversal

### Correspondencia:

German Eduardo Rueda, Facultad de Medicina, Universidad Autónoma de Bucaramanga, Calle 157 N° 19-55, Cañaveral Parque, Bucaramanga, Colombia  
Teléfonos: (+57 7) 643 6008  
gredu@unab.edu.co

Recibido: 18/12/07; aceptado:23/12/08

(probabilístico). El estudio fue aprobado por el Comité de Ética en Investigación de la Facultad de Medicina de la Universidad Autónoma de Bucaramanga. Se solicitó el consentimiento informado por escrito de los padres y del estudiante antes de iniciar el estudio.

### **Instrumento**

La escala CES-D se construyó para su uso en estudios epidemiológicos en población general. La selección de los ítems se realizó a partir de diversas fuentes, como la escala de depresión de Zung, el inventario de depresión de Beck y la escala de Raskin, entre otras, y se centró fundamentalmente en componentes cognitivos y conductuales de la sintomatología depresiva, dejando una menor presencia a los síntomas de tipo somático (18).

La escala presenta un rango de calificación de 0 a 60 puntos. Cada pregunta de la CES-D cuantifica la frecuencia e intensidad de los síntomas depresivos con una escala tipo Likert de 0 a 3 puntos. La forma de aplicación es "autoadministrada", seleccionando la respuesta que mejor refleje la situación en la semana previa. Los ítems están formulados en sentido negativo, excepto cuatro de ellos que lo están en sentido positivo, con el fin de controlar el sesgo de respuesta (20). La versión utilizada se tradujo por el método de traducción al original (*back-translation*) y se hizo la adaptación cultural; este proceso fue descrito en otra publicación (21).

### **Población de estudio y muestra**

La población la constituyeron los adolescentes matriculados durante 2005 en instituciones de educación secundaria de Bucaramanga, aproximadamente 68.000 estudiantes. El tamaño de la muestra se estimó bajo el supuesto de una prevalencia de 10% de trastorno depresivo mayor entre estudiantes, una sensibilidad de 80% y especificidad de 80% de la escala CES-D en adolescentes (22). Para ello, fue necesario estudiar 370 estudiantes; no obstante, debido a la experiencia en investigaciones con poblaciones similares que mostraron una pérdida por no aceptación de los adolescentes o sus tutores del 20%, se hizo un ajuste por este porcentaje.

### **Procedimiento**

Se realizó una prueba piloto con 108 estudiantes de una institución educativa seleccionada por conveniencia, para evaluar la estimación de la prevalencia del trastorno depresivo mayor, la claridad del cuestionario en la población, la validez de apariencia por el grupo de psiquiatras del estudio y otros posibles problemas. La prevalencia observada fue del 10%.

El muestreo que se usó para la selección de los individuos fue de múltiples etapas. La primera etapa correspondió a un muestreo aleatorio estratificado según el nivel económico de los colegios de la ciudad. El nivel bajo correspondió a los estratos 1 y 2, el nivel medio, a estratos 3 y 4, y el nivel alto, a los estratos 5 y 6. De 191 colegios de la ciudad, se escogieron 9 colegios: 3 de nivel bajo, 5 del medio y 1 del alto. Posteriormente, se seleccionó un número de conglomerados proporcional al tamaño del colegio (salones de clase por colegio). Por último, se hizo un muestreo por cuotas seleccionando los adolescentes de 13 a 17 años de cada salón de clase (23).

Los estudiantes contestaron la escala en el salón de clase a manera de "autoinforme" y unas preguntas sobre datos demográficos. A quienes contestaron la escala CES-D se les invitó a asistir a una evaluación clínica por parte de un médico psiquiatra, quien aplicó la entrevista clínica estructurada para los trastornos del eje I del DSM-IV (*Diagnostic and Statistical Manual of Mental Disorders*), versión clínica como guía de referencia para confirmar los criterios diagnósticos del DSM-IV-TR (*Text Revision*) para el trastorno depresivo mayor (24). A esta entrevista se le añadió una pregunta referente a la irritabilidad, pues el DSM-IV tiene este síntoma como adicional cuando se trata de adolescentes.

La identidad del entrevistador y del estudiante se "enmascararon" para el resultado de la escala CES-D. En la sala de espera para la entrevista clínica, se aplicó nuevamente la escala a todos los asistentes para evaluar su reproducibilidad. La entrevista clínica se realizó entre 3 y 28 días después de la aplicación del primer cuestionario.

Si el psiquiatra diagnosticaba algún trastorno mental, se le informaba al tutor y éste refería al adolescente al profesional idóneo.

### **Análisis estadístico**

Para el análisis descriptivo se calcularon las frecuencias para las variables cualitativas y se determinaron las medidas de tendencia central y dispersión en función de la distribución de frecuencias (prueba de Shapiro-Wilk). En el análisis de las diferencias en función del trastorno depresivo mayor, se utilizaron la prueba de ji al cuadrado de Mantel-Haenszel o la exacta de Fischer, la t de Student o la de Mann-Whitney, según el caso.

Para establecer la consistencia interna del cuestionario, se estimó el coeficiente  $\alpha$  de Cronbach (25). Se calculó el área bajo la curva según las características del receptor operador (*Receiver Operating Characteristic*, ROC) y su comparación con el método de Sidak (26). De acuerdo con el análisis de la curva ROC, se sugirió el punto de corte (27).

La validez de criterio de la escala CES-D se determinó comparando el resultado de la primera aplicación, considerado como positivo o negativo, con el diagnóstico de trastorno depresivo mayor hecho en la entrevista. Se calcularon la sensibilidad, la especificidad, los valores diagnósticos, la eficiencia y las razones de máxima verosimilitud (28). Además, se estimaron los índices de calidad  $\kappa$  (1,0),  $\kappa$  (0,0) y  $\kappa$  (0,5), correspondientes a la sensibilidad, la especificidad y la eficiencia, respectivamente. Estos estimadores de la calidad de la prueba se obtuvieron de las probabilidades condicionales de los mismos marcadores en función de la prevalencia y el nivel de la prueba. Además, se evaluó la legitimidad de la prueba por medio de la ji al cuadrado (28).

Se evaluó la reproducibilidad prueba-reprueba de la CES-D con el coeficiente de concordancia kappa, el método de límites de acuerdo de Bland y Altman y el coeficiente de concordancia de Lin (29,30). Se estimaron los límites de acuerdo de Bland y Altman, según el tiempo de la segunda aplicación con el test de Bradley-Blackwood (31).

Los formatos se revisaron en forma manual para verificar que estuvieran completos; luego, se digitaron por duplicado en una base de datos controlada electrónicamente y se procesaron en Stata 9 (32). Todas las razones y proporciones se estimaron con sus intervalos de confianza al 95% (IC95%) y se aceptaron como significativas aquellas diferencias cuyo error tipo I fueran inferiores a 5% ( $p < 0,05$ ).

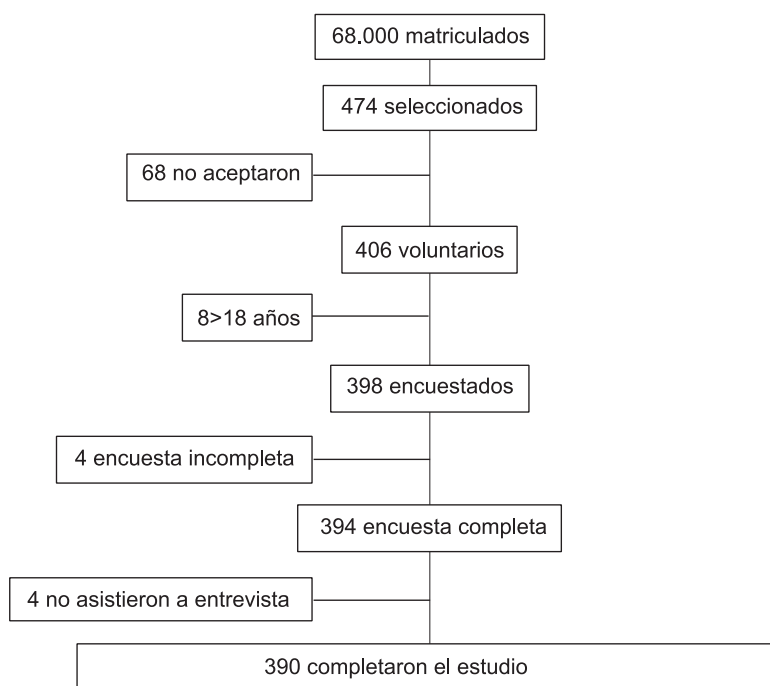
### **Resultados**

En la figura 1 se observa el esquema de captación de adolescentes para el estudio. Completaron la evaluación 217 (55,6%) mujeres y 173 (44,4%) hombres. Las mujeres tenían mayor edad que los hombres (15 años; IC95% 14,9-15,2 Vs. 14,5; IC95% 14,3-14,7;  $p < 0,001$ ). El nivel socioeconómico informado fue bajo en 133 (34,3%) estudiantes; medio, en 227 (58,8%), y alto, en 27 (6,9%). Trescientos cincuenta y un (90,0%) estudiantes estaban matriculados en instituciones públicas. Del total de los estudiantes, 10 (2,56%) estaban en 6º grado, 18 (4,6%) en 7º grado, 92 (23,6%) en 8º grado, 76 (19,5%) en 9º grado, 120 (30,8%) en 10º grado y 74 (18,9%) en 11º grado.

A 45 estudiantes se les diagnosticó trastorno depresivo mayor durante la entrevista clínica (prevalencia de 11,5%; IC95% 8,5-15,1). El trastorno depresivo mayor fue más frecuente en las mujeres de los cursos superiores y con mayor edad, y no hubo diferencias en el estrato socioeconómico o por estudiar en una institución pública (cuadro 1).

Para la primera aplicación del CES-D, se obtuvo un valor medio de  $18,67 \pm 8,74$  (IC95% 17,80-19,54) y un puntaje que osciló entre 0 y 52 puntos. Para la segunda aplicación, los estudiantes respondieron la escala con una media de  $18,62 \pm 8,77$  (IC95% 17,75-19,49) y una amplitud en el puntaje entre 0 y 47 puntos.

En el cuadro 2 se observan los índices de validez de criterio en la población estudiada con el punto de corte sugerido y, en el cuadro 3, con otros puntos de corte; el área bajo la curva ROC de la primera aplicación fue de 0,815 (IC95% 0,748-0,881). En la figura 2 se compara la curva ROC en los dos momentos de aplicación.



**Figura 1.** Esquema del estudio

**Cuadro 1.** Diferencias entre los adolescentes escolarizados con y sin trastorno depresivo mayor.

Variables	Trastorno depresivo mayor		p
	No (n=345)	Sí (n=45)	
Edad (promedio)	14,7 ± 1,2	15,4 ± 1,3	0,001*
Sexo (masculino)	162 (47%)	11 (24,4%)	0,004€
Estrato socio-económico			
Bajo	120 (34,8%)	13 (28,9%)	
Medio	197 (57,1%)	31 (68,8%)	0,228£
Alto	26 (7,6%)	1 (21,2%)	
Colegio público	307 (90%)	44 (97,7%)	0,064£
Nivel educativo			
Básica secundaria	183 (53%)	13 (28,9%)	0,004€
Media vocacional	162 (47%)	32 (71,1%)	

\*Prueba de Mann-Whitney; € ji al cuadrado; £ prueba exacta de Fisher.

Al evaluar la legitimidad de la escala CES-D en cada punto de corte, se visualizó que las discrepancias que se pueden presentar entre la escala CES-D y la SCID-I (Structured Clinical Interview for DSM Disorders) en los verdaderos positivos y el producto de la prevalencia y el

nivel de la prueba, son superiores al nivel de significancia del 5%; por lo tanto, la escala CES-D se puede utilizar para la determinación del trastorno depresivo mayor como una prueba válida (cuadro 4).

Para los índices de calidad de la escala CES-D, la sensibilidad presentó valores de  $0,541 \pm 0,349$  (mínimo: 0,000; máximo: 1,000), con un kappa máximo de la sensibilidad para el punto de corte  $\geq 10$  con concordancia del 100%. El kappa máximo de la especificidad se obtuvo para el punto de corte  $\geq 46$  con acuerdo de 100%, que es similar al patrón de referencia (cuadro 4).

La consistencia interna de la escala CES-D medida por el alfa de Cronbach fue de 0,856. La concordancia diagnóstica de la prueba medida por kappa fue 0,65 (IC95% 0,55-0,75) y el coeficiente de correlación de Lin de 0,75 (IC95% 0,71-0,79).

Los niveles de acuerdo en la primera y la segunda aplicación de la escala CES-D se evaluaron con los límites de acuerdo de Blant y Altman. En la figura 3 se puede observar que la línea

**Cuadro 2.** Índices de validez de criterio de la escala CES-D en adolescentes.

Escala CES-D	Trastorno depresivo mayor por entrevista psiquiátrica		Total
	Positivo	Negativo	
Negativo	33	91	124
Positivo	12	254	266
Total	45	345	390
Punto de corte 23			IC95%
		Límite inferior	Límite superior
Prevalencia de la enfermedad	11,54	8,6	15,2
Eficiencia	73,59	68,9	77,8
Sensibilidad	73,33	57,8	84,9
Especificidad	73,62	68,6	78,1
Valor diagnóstico positivo	26,61	19,27	35,44
Valor diagnóstico negativo	95,49	92,05	97,54
Cociente de probabilidades positivo	2,78	2,17	3,57
Cociente de probabilidades negativo	0,36	0,22	0,59

**Cuadro 3.** Razones de probabilidad del CES-D para diferentes puntos de corte.

Punto de corte	Sen	Esp	VPP	VPN	Eficiencia	LR+	LR-
≥16	0,91	0,44	0,18	0,97	0,49	1,63	0,20
≥17	0,91	0,50	0,19	0,98	0,54	1,81	0,18
≥18	0,89	0,54	0,20	0,97	0,58	1,95	0,20
≥19	0,89	0,59	0,22	0,98	0,62	2,14	0,19
≥20	0,84	0,63	0,23	0,97	0,66	2,29	0,25
≥21	0,78	0,67	0,24	0,96	0,68	2,37	0,33
≥22	0,73	0,70	0,24	0,95	0,71	2,48	0,38
≥23*	0,73	0,74	0,27	0,95	0,74	2,78	0,36
≥24	0,69	0,76	0,27	0,95	0,75	2,90	0,41
≥25	0,62	0,81	0,30	0,94	0,79	3,35	0,46
≥26	0,58	0,87	0,37	0,94	0,84	4,43	0,49
≥27	0,51	0,89	0,37	0,93	0,84	4,52	0,55
≥28	0,49	0,91	0,41	0,93	0,86	5,27	0,56

Sen: sensibilidad; Esp: especificidad; VPP: valor diagnóstico positivo; VPN: valor diagnóstico negativo; LR+: cociente de probabilidad positivo; LR-: cociente de probabilidad negativo; \*punto de corte.

correspondiente al promedio de las diferencias entre las pruebas fue de 0,097, con un test de Bradley Blackwood de  $p=0,946$ . Al evaluar por intervalos de tiempo, tampoco se observó una tendencia que afectara el acuerdo de la segunda y la primera aplicación (<7 días,  $p=0,644$ ; 7-21 días,  $p=0,873$ , y >21 días,  $p=0,360$ ).

### Discusión

En este estudio se demuestra que la versión en español de la escala CES-D tiene una adecuada validez de criterio, una excelente consistencia

interna y una muy buena reproducibilidad en la detección del trastorno depresivo mayor en una muestra representativa de adolescentes colombianos en edad escolar.

El punto de corte de una escala de tamización debe ser ajustado para una población particular (33). El punto de corte tradicional de la escala CES-D es 16, derivado del estudio original *Community Mental Health Assessment* en adultos estadounidenses (34). Algunos estudios han demostrado que las puntuaciones en

adolescentes y adultos jóvenes son mayores debido a factores normativos propios de la edad, como el exceso de síntomas transitorios, interpersonales y anímicos (35,36). A pesar de esto, el punto de corte en adolescentes tardó en establecerse y la CES-D se utilizó para hallar la prevalencia en adolescentes, estimando valores

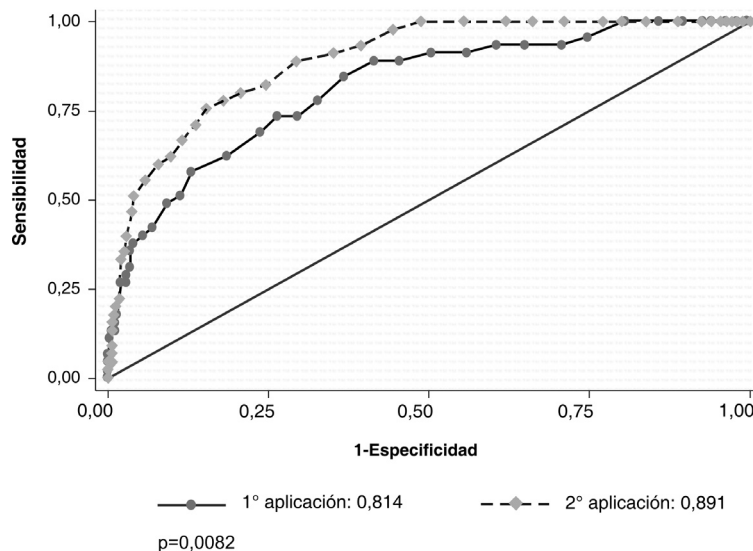
entre 24% y 50% (37-40), lo cual sugiere un gran número de falsos positivos. Sólo una investigación buscó el punto de corte en adolescentes estadounidenses y fue  $\geq 24$  (41).

En el caso de los adolescentes colombianos en edad escolar, el punto de corte ( $\geq 23$ ) fue mayor que el observado en adultos estadounidenses

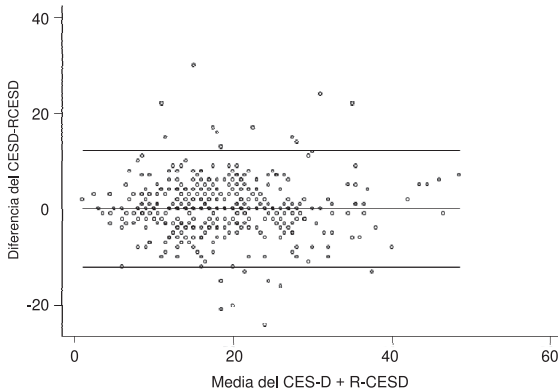
**Cuadro 4.** Índices de calidad de la CES-D para los diferentes puntos de corte.

Punto de corte	Q (%)	$\kappa$ (1,0)%	$\kappa$ (0,0)%	$\kappa$ (0,5,0)%	$\chi^2(3,84)$
$\geq 10^*$	87,44	100*	2,71	5,27	10,552
$\geq 12$	77,18	75,00	3,56	6,79	10,405
$\geq 14$	68,21	81,29	5,87	10,95	18,618
$\geq 16$	60,00	80,19	8,51	15,39	26,626
$\geq 18$	50,51	79,07	11,67	20,33	35,971
$\geq 20$	42,31	64,19	13,69	22,57	34,270
$\geq 22$	34,62	60,90	17,04	26,63	40,475
$\geq 24$	28,97	50,56	21,36	30,03	42,120
$\geq 26$	18,21	41,87	28,89	34,19	47,179
$\geq 28$	13,85	35,06	36,91	35,96	50,467
$\geq 30\&$	9,49	32,40	48,95	38,99 <sup>a</sup>	61,860
$\geq 32$	7,18	26,19	47,83	33,85	48,851
$\geq 34$	5,90	22,28	48,62	30,56	42,249
$\geq 36$	4,87	14,94	56,52	23,64	32,939
$\geq 38$	2,82	11,05	54,78	18,39	23,614
$\geq 40$	2,31	11,52	71,74	19,85	32,226
$\geq 43$	2,05	9,72	81,16	17,36	30,773
$\geq 46^{**}$	1,54	5,94	100**	11,22	23,178

Q: nivel de la prueba;  $\kappa$  (1,0): kappa máxima de la sensibilidad;  $\kappa$  (0,50): kappa al azar comparada con el SCID;  $\kappa$  (0,0): kappa máxima de la especificidad;  $\chi^2(3,84)$ : prueba de legitimidad de la escala; \*punto de corte óptimo para la sensibilidad; \*\*punto de corte óptimo para la especificidad; & punto de corte óptimo para diagnóstico.



**Figura 2.** Comparación de las curvas de receptor operador (ROC) de la primera y segunda aplicación de la CES-D.



**Figura 3.** Límites de Bland y Altman para la reproducibilidad prueba-reprueba del CES-D en adolescentes colombianos.

( $\geq 16$ ) y colombianos ( $\geq 20$ ), tal como se esperaba. (21,36) Además, el puntaje promedio de la CES-D en los adolescentes estudiados fue similar al de los adolescentes estadounidenses y superior al hallado en adultos. (21,36,40).

Una prueba de tamización ideal debe tener una sensibilidad por encima del 90%, lo cual no se evidenció en este estudio. Esto podría explicarse por la alta puntuación de síntomas transitorios, anímicos e interpersonales en adolescentes sin trastorno depresivo mayor, posiblemente similar a la puntuación de adolescentes con esta alteración de seriedad leve, que son la mayoría en estudios de población (35,36). Este hecho podría afectar, además, la especificidad. Sin embargo, el cuestionario es útil en la medida que el análisis de la curva receptor operador es bueno, ya que las áreas bajo esta curva superiores a 0,75 indican que la capacidad diagnóstica de una prueba es apropiada (27).

La sensibilidad de la CES-D en otros estudios de población fue superior o similar (60% al 90%) a la presentada en este trabajo, al igual que la especificidad (73% al 94%). El valor diagnóstico positivo fue superior en esta muestra, lo cual se explica por la mayor prevalencia del trastorno depresivo mayor en los adolescentes estudiados (21,41-44).

Dependiendo del punto de corte, la CES-D podría utilizarse como prueba de tamización, con un punto de corte de 23, o de ayuda diagnóstica, con un punto de corte de 30.

Se estima que el coeficiente alfa de Cronbach es bueno cuando se encuentra entre 0,70 y 0,90 (42). En esta muestra estuvo por encima de 0,80, lo cual indica una consistencia interna excelente, como la informada en otras poblaciones tanto de adultos como adolescentes (21,36,39-44).

La reproducibilidad de la prueba fue muy buena. Sin embargo, las condiciones de aplicación de las pruebas no fueron similares y algunos estudiantes durante la primera aplicación, a pesar de las recomendaciones, respondieron la CES-D con interferencia pues se hallaban en el aula de clase; no obstante, en la segunda aplicación en la sala de espera, pusieron mayor dedicación y empeño en sus respuestas. Probablemente, esto también explique la discrepancia en las curvas ROC y en los límites de acuerdo Bland y Altman. Esta discrepancia conduce a recomendar aplicar la CES-D en un ambiente privado que facilite la concentración y el interés del adolescente. Por tanto, se debe preferir la utilización de esta prueba de tamización en la consulta (búsqueda de casos), más que como una herramienta de tamización de población.

Una limitación de este estudio podría ser el intervalo entre las dos aplicaciones, pues en algunos de los estudiantes apenas había transcurrido tres días, lo que podría generar sesgo de memoria y sobrevaloración de la confiabilidad de la prueba. Sin embargo, no hubo diferencias en la tendencia del acuerdo de Bland y Altman en los diferentes intervalos de tiempo de la segunda aplicación con respecto a la primera.

Otra limitación estuvo asociada con el muestreo; la selección en múltiples etapas y el muestreo por cuotas, garantizaron que se conservaran las proporciones de la población objeto, pero no que la muestra fuera probabilística, y el estrato socioeconómico alto no tuvo el poder para reflejar la realidad del comportamiento del trastorno depresivo mayor en este subgrupo de población. Dentro del contexto de una evaluación de pruebas diagnósticas, esto no afectaría la validez externa de los resultados observados, pero la prevalencia no puede generalizarse.

En conclusión, la CES-D es una prueba adecuada para la tamización de adolescentes en



edad escolar de Bucaramanga con un punto de corte mayor que el utilizado comúnmente. Esta prueba sería útil en adolescentes en la sala de espera de la consulta pediátrica o en los planteles educativos, aunque sería deseable su aplicación en un ambiente privado para mejorar su confiabilidad.

### Conflicto de intereses

Los autores declaran no tener conflictos de intereses con respecto al estudio.

### Financiación

Centro de Investigaciones Biomédicas, Universidad Autónoma de Bucaramanga.

### Referencias

- López AD, Mathers CD, Ezzati M, Jamison DT, Murray CJ.** Global burden of disease and risk factors. Washington, D.C.: Oxford University Press; 2006.
- Fergusson DM, Woodward LJ.** Mental health, educational, and social role outcomes of adolescents with depression. *Arch Gen Psychiatry.* 2002;59:225-31.
- Richardson JL, Radziszewska B, Dent CW, Flay BR.** Relationship between after-school care of adolescents and substance use, risk taking, depressed mood, and academic achievement. *Pediatrics.* 1993;92:32-8.
- Shrier LA, Harris SK, Beardslee WR.** Temporal associations between depressive symptoms and self-reported sexually transmitted disease among adolescents. *Arch Pediatr Adolesc Med.* 2002;156:599-606.
- Evans E, Hawton K, Rodham K, Deeks J.** The prevalence of suicidal phenomena in adolescents: a systematic review of population-based studies. *Suicide Life Threat Behav.* 2005;35:239-50.
- Newman DL, Moffitt TE, Caspi A, Magdol L, Silva PA, Stanton WR.** Psychiatric disorder in a birth cohort of young adults: prevalence, comorbidity, clinical significance, and new case incidence from ages 11 to 21. *J Consult Clin Psychol.* 1996;64:552-62.
- Lewinsohn PM, Rohde P, Seeley JR.** Adolescent psychopathology: III. The clinical consequences of comorbidity. *J Am Acad Child Adolesc Psychiatry.* 1995;34:510-9.
- Stover E, Fenton W, Rosenfeld A, Insel TR.** Depression and comorbid medical illness: the National Institute of Mental Health perspective. *Biol Psychiatry.* 2003;54:184-6.
- Fergusson DM, Horwood LJ, Lynskey MT.** Prevalence and comorbidity of DSM-III-R diagnoses in a birth cohort of 15-year-olds. *J Am Acad Child Adolesc Psychiatry.* 1993;32:1127-34.
- Regier DA, Farmer ME, Rae DS, Myers JK, Kramer M, Robins LN, et al.** One-month prevalence of mental disorders in the United States and sociodemographic characteristics: the Epidemiologic Catchment Area study. *Acta Psychiatr Scand.* 1993;88:35-47.
- Blazer DG, Kessler RC, McGonagle KA, Swartz MS.** The prevalence and distribution of major depression in a national community sample: The National Comorbidity Survey. *Am J Psychiatry.* 1994;151:979-86.
- Kessler D, Lloyd K, Lewis G, Gray DP.** Cross sectional study of symptom attribution and recognition of depression and anxiety in primary care. *Br Med J.* 1999;318:436-9.
- Froom J, Schlager DS, Steneker S, Jaffe A.** Detection of major depressive disorder in primary care patients. *J Am Board Fam Pract.* 1993;6:5-11.
- Wu P, Hoven CW, Bird HR, Moore RE, Cohen P, Alegria M, et al.** Depressive and disruptive disorders and mental health service utilization in children and adolescents. *J Am Acad Child Adolesc Psychiatry.* 1999;38:1081-90.
- Bauer M, Whybrow PC, Angst J, Versiani M, Moller HJ, World Federation of Societies of Biological Psychiatry (WFSBP) Task Force on Treatment Guidelines for Unipolar Depressive Disorders.** World Federation of Societies of Biological Psychiatry (WFSBP) guidelines for biological treatment of unipolar depressive disorders. Part 2: Maintenance treatment of major depressive disorder and treatment of chronic depressive disorders and subthreshold depressions. *World J Biol Psychiatry.* 2002;3:69-86.
- Rao U, Ryan ND, Birmaher B, Dahl RE, Williamson DE, Kaufman J, et al.** Unipolar depression in adolescents: Clinical outcome in adulthood. *J Am Acad Child Adolesc Psychiatry.* 1995;34:566-78.
- Rey JM, Grayson D, Mojarrad T, Walter G.** Changes in the rate of diagnosis of major depression in adolescents following routine use of a depression rating scale. *Aus N Z Psychiatry.* 2002;36:229-33.
- McDowell I.** Measuring Health. Depression. In: McDowell I, Newell C. A guide to rating scales and questionnaires. Second edition. Washington, D.C.: Oxford University Press; 1996.
- Wilcox H, Field T, Prodromidis M, Scafidi F.** Correlations between the BDI and CES-D in a sample of adolescent mothers. *Adolescence.* 1998;33:565-74.
- Radloff L.** A self-report depression scale for research in the general population. *Applied Psychological Measure.* 1977;1:385-401.
- Campo-Arias A, Díaz-Martínez LA, Rueda-Jaimes GE, Cadena-Afanador LP, Hernández NL.** Psychometric properties of CES-D scale among Colombian adults from the general population. *Rev Col Psiquiatr.* 2007;36:664-74.
- Obuchowski NA.** Sample size calculations in studies of test accuracy. *Stat Methods Med Res.* 1998;7:371-92.

23. **Silva L.** Diseño razonado de muestras y captación de datos para la investigación sanitaria. Primera edición. Madrid: Díaz de Santos; 2000.
24. **First MB, Spitzer RL, Gibbon M, Williams JB.** Entrevista clínica estructurada para los trastornos de eje I del DSM-IV (versión clínica) SCID-I. Barcelona: Masson; 1999.
25. **Cronbach LJ.** Coefficient alpha and the internal structure of test. *Psychometrika*. 1951;16:297-334.
26. **DeLong ER, DeLong DM, Clarke-Pearson DL.** Comparing the areas under two or more correlated receiver operating curves: a nonparametric approach. *Biometrics*. 1988;44:837-45.
27. **Beck JR, Shultz EK.** The use of relative operating characteristics (ROC) curves in test performance evaluation. *Arch Lab Pathol Med*. 1986;110:13-20.
28. **Kraemer H.** Sensitivity and specificity: The signal detection approach. Evaluating medical tests: Objective and quantitative guidelines. First edition. California: Sage Publications; 1992.
29. **Lin LI-K.** A concordance correlation coefficient to evaluate reproducibility. *Biometrics*. 1989;45:255-68.
30. **Bland JM, Altman DG.** Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1:307-10.
31. **Bradley EL, Blackwood LG.** Comparing paired data: a simultaneous test for means and variances. *American Statistician*. 1989;43:234-5.
32. **StataCorp.** Stata statistic software: release 9.0. College Station, TX: StataCorp LP; 2005.
33. **Irwing L, Bossuyt P, Glasziou P, Lijmer J.** Designing studies to ensure that estimates of test accuracy are transferable. *BMJ*. 2002;324:669-71.
34. **Radloff LS, Locke BZ.** The Community Mental Health Assessment Survey and the CES-D scale. En: Weissman M, Myers J, Ross C, editors. *Community surveys*. New Brunswick, NJ: Rutgers University Press; 1986.
35. **Schoenbach VJ, Kaplan BH, Grimson RC, Wagner EH.** Use of a symptom scale to study the prevalence of a depressive syndrome in young adolescents. *Am J Epidemiol*. 1982;116:791-800.
36. **Radloff LS.** The use of the Center for Epidemiologic Studies Depression scale in adolescents and young adults. *J Youth Adolesc*. 1991;20:149-66.
37. **Roberts RE, Andrews JA, Lewinsohn PM, Hops H.** Assessment of depression in adolescents using the Center for Epidemiologic Studies Depression Scale. *Psychol Assessment*. 1990;2:122-8.
38. **Rushton JL, Forcier M, Schectman RM.** Epidemiology of depressive symptoms in the National Longitudinal Study of Adolescent Health. *J Am Acad Child Adolesc Psychiatry*. 2002;41:199-205.
39. **Schoenbach VJ, Kaplan BH, Grimson RC, Wagner EH.** Use of a symptom scale to study the prevalence of a depressive syndrome in young adolescents. *Am J Epidemiol*. 1982; 116:791-800.
40. **Fava GA.** Assessing depressive symptoms across cultures: Italian validation of the CES-D self-rating scale. *J Clin Psychol*. 1983;39:249-51.
41. **Roberts RE, Lewinsohn PM, Seeley JR.** Screening for adolescent depression: A comparison of depression scales. *J Am Acad Child Adolesc Psychiatry*. 1991;30:58-66.
42. **Beekman AT, Deeg DJ, van Limbeek J, Braam AW, De Vries MZ, van Tilburg W.** Criterion validity of the Center for Epidemiologic Studies Depression scale (CES-D): results from a community-based sample of older subjects in The Netherlands. *Psychol Med*. 1997;27:231-5.
43. **Boyd JH, Weissman MM, Thompson D, Myers JK.** Screening for depression in a community sample. *Arch Gen Psychiatry*. 1982;39:195-9.
44. **Breslau N.** Depressive symptoms, major depression, and generalized anxiety: a comparison of self-reports on CES-D and results from diagnostic interviews. *Psychiatry Res*. 1985;15:219-29.