Validación de la escala de Zung para depresión en universitarias de Bucaramanga, Colombia

Adalberto Campo-Arias¹
Luis Alfonso Díaz-Martínez²
Germán Eduardo Rueda-Jaimes³
Jaider Alfonso Barros-Bermúdez⁴

Resumen

Antecedentes: la escala de Zung para depresión (EZ-D) se utiliza con mucha frecuencia en investigaciones en Colombia; sin embargo, no contamos con una validación formal en población alguna. *Objetivo*: validar la EZ-D en mujeres universitarias de Bucaramanga, Colombia.

Método: a un grupo de estudiantes universitarias se le solicitó la participación en este estudio. Las estudiantes diligenciaron la EZ-D y posteriormente se realizó una entrevista estructurada (SCID-I) para episodio depresivo mayor (EDM) durante el 2003. *Resultados*: se entrevistaron 110 de universitarias con edad promedio de 22,1 años (DE 3,4). En la EZ-D puntuaron entre 29 y 78 con una media de 49 puntos (DE 11,2), 48 estudiantes (43,6%) presentaban síntomas depresivos con importancia clínica. En la entrevista se confirmó que 19 universitarias (17,3%) reunían criterios para un EDM. La prueba de alfa de Cronbach fue 0,85; la sensibilidad de EZ-D para EDM, 94,7% (IC 95%: 90,5-99,7); la especificidad, 67,0% (IC 95%: 56,3-76,3); el valor predictivo positivo, 37,5% (IC 95%: 24,3-52,7); el valor predictivo negativo, 98,4% (IC 95%: 90,2-99,9); la prueba kappa de Cohen, 0,39 (IC 95%: 0,23-0,54); el área bajo la curva, 0,856 (IC 95%: 0,780-0,933), y la correlación test retest fue 0,848 (p <0,01).

Conclusiones: la EZ-D es útil como instrumento de tamizaje, con alta sensibilidad y aceptable especificidad en la identificación de posibles caso de EDM en mujeres universitarias.

Palabras clave: escala de Zung para depresión, validación, estudiantes, universitarios.

Title: Validation of Zung's Self-Rating Depression Scale Among Universitary Student Women from Bucaramanga, Colombia.

Médico, profesor asociado del Grupo de Neuropsiquiatría, Facultad de Medicina, Universidad Autónoma de Bucaramanga, Bucaramanga, Colombia.

Médico, profesor asociado del Grupo de Neuropsiquiatría, Facultad de Medicina, Universidad Autónoma de Bucaramanga, Bucaramanga, Colombia.

Médico, profesor asociado del Grupo de Neuropsiquiatría, Facultad de Medicina, Universidad Autónoma de Bucaramanga, Bucaramanga, Colombia.

Médico, profesor asociado del Grupo de Neuropsiquiatría, Facultad de Medicina, Universidad Autónoma de Bucaramanga, Bucaramanga, Colombia.

Abstract

Background: Zung's self-rating depression scale (SDS) is used frequently in Colombian investigation. However, we don't count on formal validation of it. *Objective*: To validate SDS among university female students from Bucaramanga, Colombia.

Method: Female students were ask for participating in this study, then fill out SDS and interviewed by psychiatrists using Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I) in to order to identify who meet criteria for major depressive episode (MDE) last month.

Results: We interviewed one hundred ten students; mean age was 22.1 years (SD 3.4). SDS score range 29 to 78 (average 49, SD 11.2). Forty-eight (43.6%) students reported clinically important depressive symptoms, while with SCID-I interview we identified nineteen (17.3%) MDE. Cronbach's alpha test was 0.850, sensibility was 94.7% (CI95% 90.5-99.7), specificity 67.0% (CI95% 56.3-76.3), positive predict value 37.5% (CI95% 24.3-52.7), negative predict value 98.4% (CI95% 90.2-99.9), Cohen's kappa test 0.390 (CI95% 0.23-0.54), area under curve was 0.856 (CI95% 0.780-0.933), and test-retest correlation was 0.848 (p<0.01).

Conclusions: SDS is useful tool for screening MDD among university female students, with high sensibility and good specificity.

Key words: Zung's self-rating depression scale, validation studies, students, university.

En Colombia, las escalas para cuantificar síntomas depresivos más comunes, la escala de Zung y el inventario de Beck para depresión, se han utilizado para determinar la prevalencia de trastornos depresivos en estudiantes universitarios (1),(2),(3), (4),(5). En otros países estos instru-

mentos han mostrado una buena validez y confiabilidad en diferentes contextos (6),(7); sin embargo, carecen de una validación formal en ésta y en otras poblaciones colombianas, por ello se necesita mostrar su utilidad en este sector de la población, debido a que las propiedades psicométricas de estos instrumentos varían de acuerdo con la muestra investigada (8).

Los síntomas depresivos con importancia clínica y los trastornos depresivos más frecuentes -trastorno depresivo mayor (TDM) y trastorno distímico— tienen altas prevalencias en la población general de Colombia (9),(10). Los datos disponibles muestran que estas prevalencias son significativamente mayores en los estudiantes universitarios (1),(2),(3),(4), (5). Estas manifestaciones se asocian con un deterioro significativo en el funcionamiento cognoscitivo, el cual afecta generalmente, en forma perjudicial, el rendimiento académico (11). Esto a su vez representa un estresor negativo para la mayoría de los estudiantes (12). La identificación temprana de estos trastornos no sólo minimizaría la posibilidad de fracaso académico, sino que también reduciría en forma substancial otras conductas de riesgo para la salud, como el consumo de cigarrillo o de alcohol o las conductas alimentarias desadaptivas asociadas con la presencia de trastornos depresivos, fenómeno más frecuente en las mujeres jóvenes (13),(14),(15).

En consecuencia, es de capital importancia contar con la validación formal de un instrumento de auto-informe de síntomas depresivos en estudiantes universitarios que ayude a una identificación válida y confiable de posibles casos de episodio depresivo mayor (EDM).

El objetivo de este estudio fue validar la escala de Zung para depresión en estudiantes universitarios de la ciudad de Bucaramanga, Colombia. En este informe se presenta la consistencia interna y la validez concurrente encontradas.

Método

El presente es un estudio de validación concurrente de una escala, prueba diagnóstica, en estudiantes universitarias durante el segundo semestre de 2003. Este proyecto fue evaluado y aprobado por el Comité de Ética de la Facultad de Medicina de la Universidad Autónoma de Bucaramanga (Unab) y representaba un riesgo mínimo para quienes aceptaran ser parte de éste (16). Se solicitó la participación enteramente voluntaria de los estudiantes después de explicar los objetivos de la investigación, garantizar la confidencialidad e informar que no recibirían ningún incentivo adicional por la inclusión en el estudio.

Se solicitó la cooperación de estudiantes de todas las facultades,

diurna y nocturnas, de la Unab, con excepción de Medicina y Psicología, que se excluyeron por conveniencia. Para la fecha, la Unab contaba con 5.687 estudiantes matriculados en 18 carreras. Por las características del estudio, se necesitaba una muestra mínima de cien sujetos (17).

La escala de Zung para depresión (EZ-D) es un instrumento autoadministrado tipo Likert, que consta de veinte ítems que se diligencian marcando con una equis la frecuencia de los síntomas durante los últimos quince días y dos últimas semanas. Da puntuaciones de uno a cuatro para cada ítem y globales entre veinte y ochenta. Tradicionalmente, las puntuaciones iguales o superiores a cuarenta se consideran síntomas depresivos con importancia clínica o sugestivos de un EDM. Por lo general, el puntaje se multiplica por 1,25, a fin de obtener puntuaciones entre 25 y 100, donde 50 puntos o más (punto de corte) sugieren EDM (18). Se utilizó la versión usada en el Estudio nacional de salud mental y consumo de sustancias psicoactivas de Colombia de 1993 (9).

Los estudiantes que aceptaron diligenciar la escala lo hicieron en el aula de clase y posteriormente se les citó, entre uno y quince días, con una media de 4,03 días (DE 2,3), para una entrevista con un especialista en psiquiatría, a fin de comparar los resultados de la escala, momento en el que se aplicó el módulo para identificar EDM de la entrevista estruc-

turada para los trastornos del eje I (SCID-I) (19). Se utilizó esta entrevista estructurada para garantizar homogeneidad en el proceso de evaluación, ya que con este tipo de instrumento se alcanza el mayor grado de precisión diagnóstica que con una entrevista clínica habitual o no estructurada (20). Además, este tipo de herramientas diagnósticas no requieren una validación formal (8). Tanto estudiantes como entrevistadores estaban enmascarados respecto al resultado de la prueba EZ-D. Los estudiantes que reunieron criterios para EDM o cualquier otro trastorno fueron remitidos a Bienestar Universitario para seguimiento por psicología o remitidos a psiquiatría para manejo especializado por su empresa de salud.

De este modo, se estableció la consistencia interna de las escalas por medio del valor de alfa de Cronbach (21); se determinaron sensibilidad, especificidad, valor predictivo positivo y valor predictivo negativo (22); se halló el área bajo la curva receptor operador (ROC) (23); se estableció la concordancia producto del azar con la prueba kappa de Cohen (24), y para averiguar la correlación test-retest se usó el coeficiente de correlación del momento del producto de Pearson (r) (25).

Todas las pruebas estadísticas fueron realizadas con el paquete estadístico Stata 8.0, junto con sus intervalos de confianza del 95% (IC 95%) o desviación estándar (DE), según el caso (26).

Resultados

El estudio se cerró con el inicio de las vacaciones académicas para los estudiantes de la institución. Para la fecha se habían entrevistado 110 de las universitarias con edad entre 16 y 33 años, promedio de 22,1 años (DE 3,4). En la EZ-D puntuaron entre 29 y 78, con una media de 49 puntos (DE 11,2). Se determinó que 48 estudiantes (43,6%) presentaban síntomas depresivos con importancia clínica (puntaje de 50 o más). En la entrevista se confirmó que 19 universitarias (17,3%) reunían criterios para un EDM. La prueba de alfa de Cronbach fue 0,85. La sensibilidad de la EZ-D para EDM fue 94,7% (IC 95%: 90,5-99,7); la especificidad, 67,0% (IC 95%: 56,3-76,3); el valor predictivo positivo, 37,5% (IC 95%: 24,3-52,7), y el valor predictivo negativo, 98,4% (IC 95%: 90,2-99,9), con una kappa de Cohen de 0,39 (IC 95%: 0,23-0,54). La fuente de estos resultados aparece en la Tabla 1.

La correlación test-retest presentó una r de 0,848 (p <0,01). Las razones de probabilidad positiva, así como la sensibilidad y especificidad para algunos puntos de corte se muestran en la Tabla 2.

El área bajo la curva ROC fue 0,856 (IC 95%: 0,780-0,933), tal como se aprecia en la Figura 1.

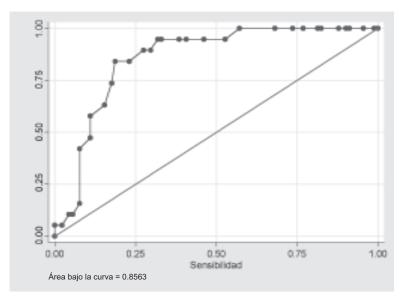
Tabla 1. Relación entre los hallazgos de la escala de Zung para depresión y la entrevista psiquiátrica

Escala de Zung	Trastorno depresivo mayor por entrevista psiquiátrica		Total
	Positivo	Negativo	Total
Positivo	18	30	18
Negativo	1	61	48
Total	19	91	110

Tabla 2. Razones de probabilidad para diferentes puntos de corte

Punto de corte	Sensibilidad (%)	Especificidad (%)	Razón de probabilidad positiva
44	100,0	42,9	1,75
45	94,7	47,3	1,79
46	94,7	53,9	2,05
48	94,7	59,3	2,33
49	94,7	61,5	2,46
51	94,7	67,0	2,87
53	94,7	68,1	2,97
54	89,5	70,3	3,01
55	89,5	72,5	3,25
56	84,2	76,9	3,65
58	84,2	81,3	4,51
59	73,7	82,4	4,19
60	63,2	84,6	4,10

Figura 1. Curva receptor operador de la EZ-D entre estudiantes universitarias



Discusión

Los resultados de este estudio muestran que la EZ-D es un instrumento con una alta sensibilidad, una aceptable especificidad y una modesta concordancia producto del azar en estudiantes universitarias. La EZ-D ha sido validada en otras poblaciones y ha mostrado una sensibilidad que oscila entre el 92% y 95%, y una especificidad entre 74% y 87,5%, (6),(7), mientras que su consistencia interna se ha informado entre 0,58 y 0,84 (27),(28). Indudablemente, estas variaciones se relacionan con las características socioculturales de las poblaciones objeto de estudio e indican que deben hacerse ajustes en los puntos de corte, dependiendo los objetivos que se hayan propuesto y de la comunidad que se vaya a analizar (29).

En este trabajo se utilizó el punto de corte habitual empleado en otros estudios realizados en población universitaria colombiana (1), (2),(3),(4),(5). Los datos sugieren que una de cada tres mujeres que puntúan cincuenta o más en la EZ-D presenta un EDM si se confirma el hallazgo con una entrevista estructurada. A la vez, muestran que las universitarias en quienes se encuentran puntajes menores de cincuenta existe una posibilidad muy baja de encontrar un EDM (22). Es necesario tener en cuenta que los síntomas depresivos con importancia clínica están presentes, obviamente, en otros

trastornos depresivos distintos al TDM, al igual que en algunos trastornos de adaptación, en varios trastornos de personalidad e, incluso, en trastornos psicóticos como la esquizofrenia y el trastorno esquizoafectivo (30),(31).

El diagnóstico de un EDM es netamente clínico; no obstante, las escalas autoadministradas como la EZ-D avudan a identificar casos de EDM y mejoran el pronóstico global cuando se acompañan de las medidas necesarias para asegurar un adecuado tratamiento y seguimiento (32). El aporte más importante de esta investigación fue evaluar la validez de criterio o concurrente de la EZ-D mediante una entrevista estructurada en estudiantes universitarias y su potencial utilidad como instrumento de tamizaje en grandes poblaciones de características similares (33). Sin duda, una entrevista estructurada tiene mayor validez diagnóstica que una entrevista clínica habitual, al investigar sistemáticamente los síntomas (20). Sin embargo, quedan por establecer las propiedades psicométricas de la EZ-D, haciendo ajustes en el punto de corte, lo mismo que conocer la utilidad de esta escala en varones universitarios, cuando se sabe que el autoinforme de síntomas depresivos presenta particularidades por sexo (34).

Es muy importante tener en cuenta que la validez de una prueba diagnóstica incluye varios aspectos

(8): la llamada validez de apariencia es decir, si los ítems que componen la escala evalúan la concepción teórica que se explora, que generalmente es un consenso de expertos (35) —la EZ-D muestra una buena validez de apariencia—; la validez de contenido, que alude a si la escala cubre el o los diferentes dominios que forman el concepto (36), y la validez concurrente -presentada en este artículo—, conocida también como validez de criterio o predictiva, que no es más que la comparación de la escala con el patrón de referencia, en este caso la entrevista clínica estructurada realizada por psiguiatra (8),(34).

Finalmente, el diseño del estudio impide la validación de la concepción, lo que significa contrastar el instrumento evaluado con validadores externos. En el caso de los trastornos depresivos, puede ser la evolución en el tiempo o la respuesta al tratamiento. Este tipo de validación abarca la validez de convergente (cuando se compara con otro instrumento que mide la misma concepción), la validez divergente (cuando se contrasta con una escala que mide otra idea, en este caso, por ejemplo, ansiedad), la validez discriminante (la propiedad de la escala de diferenciar cuadros clínicos con algunos síntomas comunes) y la sensibilidad al cambio que, de la misma manera, no se podía explorar (8),(34).

Por su parte, la confiabilidad de una escala se refiere a la consisten-

cia o a la repetitividad con la que los sujetos son diferenciados en una u otra categoría propuesta. Para una escala la confiabilidad incluye consistencia interna y la confiabilidad test-retest (8),(36). La primera cuantifica la correlación entre los ítems que componen la escala, y se espera que el valor oscile entre 0,80 y 0,90 (37), como se encontró en el presente trabajo. Asimismo, la consistencia interna medida con alfa de Cronbach es una característica de la población estudiada y no una propiedad inherente del instrumento evaluado, de tal suerte que el resultado del alfa de Cronbach de una prueba debe ser informado siempre que ésta sea empleada (38),(39). En cuanto a la confiabilidad test-retest, ésta compara los resultados o puntaies de un instrumento autoadministrado en dos momentos en el tiempo (8); las correlaciones mayores de 0,5 se consideran sustanciales o más fuertes que lo usual (40).

De la misma forma, es necesario considerar que la sensibilidad, la especificidad y la concordancia esperada por azar (kappa) de una prueba de tamizaje como la EZ-D es notablemente afectada por la prevalencia del trastorno en cuestión. Por ejemplo, cuando la prevalencia del problema estudiado es relativamente baja la sensibilidad de la escala es menoscabada (38). Se concluye, así, que la EZ-D es útil como instrumento de tamizaje estudiantes universitarias con alta sensibi-

lidad, aceptable especificidad y excelente valor predictivo negativo.

Agradecimientos

A los estudiantes de medicina de la Universidad Autónoma de Bucaramanga, Alexander Ballesteros, Mayerli Barcasnegras, Diego Laverde, Jorge Ortiz, Andrea Pertuz y Jaider Salcedo, por su valiosa colaboración en el proceso logístico de esta investigación como parte de las actividades de investigación formativa que se desarrollan en la Facultad de Medicina de la Unab.

Bibliografía

- González A, Campo A, Haydar R. Síntomas depresivos en estudiantes de psicología. Acta Psiquiat Psicol Am Lat. 1999:45:234-9.
- Arboleda A, Gutiérrez JC, Miranda CA. Prevalencia de síntomas depresivos en médicos internos y residentes de la Universidad del Valle, 1999. Médicas UIS. 2001:15:4-7.
- Miranda CA, Gutiérrez JC, Bernal F, Escobar CA. Prevalencia de depresión en estudiantes de medicina de la U. del Valle. Rev Col Psiquiatría. 2000;29: 251-60.
- Campo-Cabal G, Gutiérrez JC. Psicopatología en estudiantes universitarios de la Facultad de Salud, Univalle. Rev Col Psiquiatría. 2001;30:351-8.
- Amézquita ME, González RE, Zuluaga D. Prevalencia de la depresión, ansiedad y comportamiento suicida en la población estudiantil de pregrado de la Universidad de Caldas, año 2000. Rev Col Psiquiatría 2003;32:341-56.
- Lee HC, Chiue K, Wing YK, Leung CM, Kwong PK, Chung DWS. The Zung selfrating depression scale: screening for

- depression among the Hong Kong chinese elderly. J Geriatr Psychiatry Neurol. 1994;7:216-20.
- Aragonés E, Masdeu RM, Cando G, Coll G. Validez diagnóstica de la self-rating depression scale de Zung en pacientes de atención primaria. Actas Esp Psiquiatr. 2001;29:310-6.
- 8. Blacker D, Endicott J. Psychometric properties: concepts of reliability and validity. In: Rush AJ, Pincus HA, First MB, et al, editors. Handbook of psychiatric measures. Washington: American Psychiatric Association; 2002 (CD-ROM).
- Posada JA, Torres Y. Estudio nacional de salud mental y consumo de sustancias psicoactivas, Colombia, 1993. Bogotá: Ministerio de Salud; 1995.
- Posada-Villa JA, Trevisi 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.
- American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4th ed. Washington: American Psychiatric Association; 2000.
- Fergusson DM, Woodward LJ. Mental health, educational, and social role outcomes of adolescents with depression. Arch Gen Psychiatry. 2002;59:225-31.
- Patton GC, Carlin JB, Coffey C, Wolfe R, Hibbert M, Bowes G. Depression, anxiety, and smoking initiation: a prospective study over 3 years. J Am Public Health. 1998;88:1518-22.
- Torikka A, Kaltiala-Heino R, Rimpelä A, Rimpeläm, Rantanen P. Depression, drinking, and substance use among 14to 16-year-old Finish adolescents. Nord J Psychiatry. 2001;55:351-7.
- Kirkcaldy BD, Sherpard RJ, Siefen RG. The relationship between physical activity and self-image and problem behavoiur among adolescents. Soc Psychiatry Psychiatr Epidemiol. 2002;37:544-50.
- Resolución 008430 de 1993, por la cual se establecen las normas científicas, técnicas y administrativas para la investigación en salud. Bogotá: Ministerio de Salud, 1993. Rev Col Psiquiatría. 1996:25:38-59.

- Norman GR, Streiner DL. Bioestadística. Madrid: Mosby-Doyma Libros; 1996. p. 129-48.
- Zung WWK. Self-report depression scale. Arch Gen Psychiatry. 1965;12:63-70.
- First MB, Spitzer RL, Gibbon M, Williams JBW. Entrevista clínica estructurada para los trastornos de eje I del DSM-IV (versión clínica) SCID-I. Barcelona: Masson; 1999.
- Rogers R. Standardizing DSM-IV diagnoses: the clinical application of structured interviews. J Person Asses. 2003;81:220-5.
- Cronbach LJ. Coefficient alpha and the internal structure of test. Psychometrika. 1951;16:297-334.
- Sánchez R, Gómez C. Conceptos básicos sobre la validación de escalas. Rev Col Psiquiatría. 1998;27:121-30.
- Rey JM, Morris-Yates A, Stanislaw H. Measuring the accuracy of diagnostic test using receiver operating characteristics (ROC) analysis. Int J Methods Psychiatr Res. 1992;2:39-50.
- Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Meas. 1960;20: 27-46.
- 25. Pearson ES. Studies in the story of probability and statistics. XIV. Some incidents in the early history of biometry and statistics, 1980-94. Biométrica; 1965;52:3-18.
- 26. STATA for windows 8.0. Stata Corporation, College Station. Texas; 2003.
- Dugan W, McDonald MV, Passik SD, Rosenfield BD, Theobald D, Edgerton S. Use of the Zung self-rating depression scale in cancer patients: feasibility as a screening tool. Psycho-Oncology. 1998; 7:483-93.
- Powell R. Psychometric properties of the Beck Depression Inventory and the Zung self rating depression scale in adults with mental retardation. Mental Retardation. 2002;41:88-95.
- 29. Rubio-Stipec M, Hicks MH-R; Tsuang MT. Cultural factors influencing the selection,

- use, and interpretation of psychiatric measures. Handbook of psychiatric measures. Washington: American Psychiatric Association; 2002 (CD-ROM).
- Wassink TH, Flaum M, Nopoulos P, Andreasen NC. Prevalence of depressive symptoms early in the course of schizophrenia. Am J Psychiatry. 1999;156: 315-6.
- 31. Siris SM. Depression in schizophrenia: perspective in the era of "atypical" antipsychotic agents. Am J Psychiatry. 2000;157:1379-89.
- 32. Pignone MP, Gaynes BN, Rushton JL, Burchell CM, Orleans CT, Mulrow CD, et al. Screening for depression in adults: a summary of the evidence for the U. S. Preventive Services Task Force. Ann Intern Med. 2002;136:765-76.
- Greenfield SF, Reizes JM, Magruder KM, Muenz LR, Kopans B, Jacobs DG. Effectiveness of community-based screening for depression. Am J Psychiatry. 1997;154:1391-7.
- Jackson JL, Chamberlin J, Kroenke K. Gender and symptoms in primary cares practices. Psychosomatics. 2003;44:359-66.
- 35. Bland JM, Altman DG. Validating scales and indexes. Br Med J. 2002;324: 606-7.
- Morgan GA, Gliner JA, Harmon RJ. Measurement validity. J Am Acad Child Adolesc Psychiatry. 2001;40:729-31.
- 37. Gliner JA, Morgan GA, Harmon RJ. Measurement reliability. J Am Acad Child Adolesc Psychiatry. 2001;40:486-8.
- 38. Streiner DL. Starting at the beginning: an introduction to coefficient alpha and internal consistency. J Person Asses. 2003;81:99-103.
- Streiner DL. Diagnosing tests: using and misusing diagnostic and screening tests. J Person Asses. 2003;81:209-19.
- 40. Gliner JA, Morgan GA, Harmon RJ. Basic associational designs: analysis and interpretation. J Am Acad Child Adolesc Psychiatry. 2002;41:1256-8.

Recibido para publicación: 1 de septiembre de 2004 Aceptado para publicación: 18 de marzo de 2005

Correspondencia Adalbero Campo-Arias Faculad de Medicina, Universidad Autónoma de Bucaramanga Calle 157 Nº 19-55, Cañaveral Parque, Bucaramanga, Colombia Correo electrónico: acampoar@unab.edu.co