Comparación de la sensibilidad y la especificidad entre diferentes versiones de la Escala de Depresión Geriátrica

María Izal[a], Ignacio Montorio[a], Roberto Nuevo[b] y Gema Pérez-Rojo[a]

[a]Facultad de Psicología. Universidad Autónoma de Madrid. Madrid. España.
[b]Departamento de Psicología Básica, Clínica y Psicobiología. Universidad Jaume I. Castellón. España.

RESUMEN

Introducción: la Escala de Depresión Geriátrica (GDS) ha sido ampliamente utilizada como instrumento de cribado para detectar depresión en personas mayores. Debido a su longitud (30 ítems), se han propuesto versiones reducidas. El objetivo de este trabajo es comparar la capacidad diagnóstica de 3 versiones de la GDS (GDS-30, GDS-15 y GDS-5) para identificar la presencia de depresión en personas mayores.
Material y métodos: la muestra estuvo compuesta por 246 personas reclutadas en centros de día y hogares para personas mayores. La edad media ± desviación estándar fue de 74,1 ± 9,5 años; el 68,7% eran mujeres. La presencia de depresión (20,2%) se estableció utilizando una entrevista diagnóstica semiestructurada basada en los criterios del DSM-IV.
Resultados: la GDS-30 demostró altas sensibilidad (95,7%) y especificidad (92,4%) para un punto de corte de 13. La capacidad diagnóstica de la GDS-15 disminuyó; presentó una sensibilidad del 95,7% y una especificidad del 84,3% para un punto de corte del 5. Finalmente, para la GDS-5 la sensibilidad obtenida fue del 78,7% y la especificidad del 77,3% para un punto de corte de 2.
Conclusiones: la versión original (GDS-30) parece ser una buena medida de cribado para detectar depresión en personas mayores, aunque la capacidad de diagnóstico disminuye cuando se reduce la longitud de la escala. Se sugiere que la GDS debería revisarse con el objetivo de desarrollar nuevas versiones reducidas que no pierdan capacidad diagnóstica.

Palabras clave
Depresión. Personas mayores. Curva característica operativa del receptor. Cribado. Diagnóstico.

Comparison of the sensitivity and specificity of distinct versions of the Geriatric Depression Scale

ABSTRACT

Introduction: the Geriatric Depression Scale (GDS) has been widely used as a screening instrument for detecting depression in older adults. Given its length (30 items), several abbreviated versions have been proposed. This study aimed to compare the efficacy of three versions of the GDS (GDS-30, GDS-15, and GDS-5) in screening for depression in older adults.
Material and methods: the sample was composed of 246 persons recruited in day centers and homes for older adults. The mean age was 74.1 years (SD = 9.5) and 68.7% were women. Depression (20.2%) was diagnosed with a semi-structured interview based on DSM-IV criteria.
Results: the GDS-30 demonstrated high sensitivity (95.7%) and specificity (92.4%) for a cut-score of 13. The performance of the GDS-15 was somewhat worse with a sensitivity of 95.7% and a specificity of 84.3% for a cut-score of 5. For the GDS-5, the sensitivity was 78.7% and the specificity was 77.3% for a cut-score of 2.
Conclusions: the original version (GDS-30) seems to be a good screening measure for depression in older adults, but the diagnostic performance of the scale suffers when the length of the scale is reduced. We suggest that the GDS should be revised with the goal of developing new abbreviated versions that do not lose diagnostic performance.

Key words
Depression. Older adults. Receiver operating characteristic curve. Screening. Diagnosis.

INTRODUCCIÓN

La detección de la depresión en el grupo de población de más edad es un tema de interés clínico, no sólo por el sufrimiento que provoca en la persona que la padece, sino por su repercusión en el deterioro funcional, la pérdida de la autonomía, la disminución del bienestar y su contribución al incremento de la mortalidad[1]. La epidemiología pone de relieve que la depresión es un problema que afecta a una proporción significativa de la población mayor y será, a juicio de la Organización Mundial de la Salud,

Correspondencia: Dr. I. Montorio.
Departamento de Psicología Biológica y de la Salud.
Facultad de Psicología. Universidad Autónoma de Madrid.
Campus de Cantoblanco. 28049 Madrid. España.
Correo electrónico: ignacio-montorio@uam.es

Recibido el 4-12-2006; aceptado el 21-2-2007.

uno de los principales problemas de salud de los próximos 20 años. Al igual que en otros grupos de edad, la depresión es difícil de identificar, si bien en las personas mayores existen elementos específicos de dificultad ya que tienden a quejarse inicialmente de problemas físicos, dificultad para concentrarse o problemas de memoria más que de depresión en sí misma. Además, existen diferentes circunstancias que pueden estar enmascarando el verdadero alcance de la depresión en este grupo de población, tales como la amplia comorbilidad con otros trastornos, una elevada sintomatología subclínica, estereotipos debidos a la edad en el diagnóstico por parte de profesionales que pueden asumir como normal un grado moderado de depresión asociado al envejecimiento o una confusión entre las propias personas mayores que asocian los síntomas depresivos al proceso del envejecimiento normal, lo que evita la posibilidad de acceso al tratamiento[2].

Desde la psicología clínica aplicada al envejecimiento, así como desde el ámbito de la medicina clínica geriátrica, existe un amplio interés en contar con instrumentos de fácil aplicación y corrección que permitan establecer de manera válida y fiable sospechas de la existencia de una depresión. La evaluación de la depresión es la segunda área en importancia en psicología clínica; entre los distintos métodos empleados, el autoinforme supone más del 90% del total de procedimientos empleados para la evaluación de este trastorno mental. La técnica de autoinforme, derivada de la más amplia y general técnica de autoobservación, se refiere al informe verbal que la persona proporciona sobre sí misma o su comportamiento y permite obtener información sobre las respuestas de las personas acerca de su experiencia subjetiva en relación con respuestas manifiestas, psicofisiológicas u cognitivas, en este caso, habitualmente asociadas a la depresión. La información sobre las respuestas motrices (p. ej., fatiga) y psicofisiológicas (p. ej., pérdida de peso) pueden ser ulteriormente contrastadas por otros métodos y, por tanto, el autoinforme es un método indirecto de evaluar estas manifestaciones clínicas de la depresión. No obstante, las respuestas cognitivas y emocionales (p. ej., tristeza, pensamientos negativos) de la depresión no pueden ser contrastadas mediante otros métodos, por lo que es el único método de recogida de la información de estas respuestas vinculadas a la depresión.

La Escala de Depresión Geriátrica (GDS)[3] fue el primer instrumento diseñado específicamente para población mayor y ha sido extensamente usada como instrumento de cribado para la detección de depresión en personas mayores, tanto en el ámbito comunitario como en el clínico. La GDS se desarrolló inicialmente como una medida de cribado de la depresión en la edad avanzada, evitando la mayoría de los problemas asociados con la medida de la depresión en personas mayores[4]. Para ello se seleccionaron los ítems más adecuados para esta población y se suprimió los ítems referidos a síntomas somáticos que pudieran aumentar la confusión entre el diagnóstico de depresión y enfermedades físicas comunes en las personas mayores; también se adoptó un formato de respuesta dicotómico (sí/no) (rango, 0-30)[5]. La GDS se ha mostrado como un instrumento útil de evaluación de la depresión en personas mayores cuya capacidad cognitiva esté intacta, pero no lo es en personas con deterioro cognitivo[6]. Esta versión original, que constaba de 30 ítems, se ha empleado desde entonces en cientos de estudios empíricos con diferentes muestras de personas mayores (personas hospitalizadas, institucionalizadas, personas que viven en la comunidad, etc.) y ha demostrado buenas propiedades psicométricas en versiones en diferentes idiomas y culturas, y unos valores adecuados en la detección de la depresión en la edad avanzada El estudio original de validación encontró que la escala presentaba un 84% de sensibilidad y un 95% de especificidad para un punto de corte de 11, y un 80% de sensibilidad y un 100% de especificidad para un punto de corte de 14[7]. La utilización de la GDS en el ámbito clínico puede responder a 2 propósitos diferentes. Por una parte, puede utilizarse como una escala de gravedad de la depresión, esto es, el número de diferentes manifestaciones clínicas de la depresión presentes en una persona. Por otra parte, puede utilizarse como una prueba de cribado de la posible presencia de un trastorno de depresión en función del número de ítems positivos con los que la persona se identifica, lo que hace necesario contar con puntos de corte óptimos en función de un patrón diagnóstico, normalmente basados en los criterios de depresión mayores contemplados en los sistemas de clasificación de trastornos mentales (CIE, DSM).

La principal crítica que ha recaído sobre la escala original es su larga extensión, por lo que al aplicarse a personas mayores conlleva el peligro de que se produzcan sesgos en la evaluación debidos a la fatiga o a una pérdida en el mantenimiento de la atención[4], o las limitaciones de tiempo en la utilización en el ámbito de la atención primaria[8]. En esta línea, se desarrolló una versión abreviada de la GDS, que emplea 15 ítems de la versión original[9]. Esta escala (en adelante, GDS-15) ha demostrado también buenas propiedades psicométricas y una correlación elevada con la versión de 30 ítems (entre 0,87 y 0,89[10,11]). Asimismo, existen datos que avalan la capacidad de la GDS-15 como instrumento de cribado de la depresión. Por ejemplo, D´Ath et al[12] encontraron una sensibilidad del 91% y una especificidad del 72% para un punto de corte de 5. Otros estudios, sin embargo, ponen en entredicho la eficacia de la GDS-15 para detectar la depresión. Así, Van Marwijk et al[13] encontraron que un punto de corte de 3 optimizaba simultáneamente la sensibilidad y la especificidad de la escala, pero con valores moderados (67 y 73%, respectivamente). Otro trabajo encontraba, para un punto de corte de 3, una sensibilidad del 100% y una especificidad del 72%[14]. Sin embargo, los datos de este trabajo indican que el valor predictivo positivo para este punto de corte era de tan sólo el 18,4%. En la misma línea, Brown et al[15] encontraban una sensibilidad del 67% y una especificidad del 81% para un punto de corte de 4, pero el valor predictivo positivo era de 0,31.

A pesar de las dudas que ha generado la GDS-15, la necesidad de buscar medidas de evaluación fiables lo

más cortas posibles ha llevado a la propuesta de versiones aún más reducidas de la GDS. Incluso la escala de 15 ítems lleva unos minutos para ser completada y las personas mayores no siempre la aceptan. Así, se han propuesto medidas de 12, 10, 5, 4 o incluso un solo ítem. La versión de 5 ítems (GDS-5), desarrollada por Hoyl et al[16], es de todas ellas la que parece haber tenido una mayor aceptación y presenta unas mejores propiedades psicométricas. Concretamente, en un trabajo reciente[17] se ha encontrado que la GDS-5 funcionaba de manera equivalente a la GDS-15, con una sensibilidad del 94%, una especificidad del 81%, un valor predictivo positivo del 81% y un valor predictivo negativo del 94%, para un punto de corte de 2. Sin embargo, hay aún pocos trabajos que avalen la eficacia de la GDS-5 y se ha criticado la excesiva reducción en el número de ítems que podría estar simplificando la estructura dimensional de un constructo, la depresión, que es necesariamente multidimensional.

La adaptación española de la GDS fue desarrollada por Izal y Montorio[5,18], y tiene valores adecuados en los índices de consistencia interna, fiabilidad dos mitades, validez convergente y discriminante. Los datos normativos con estas muestras reflejan que la puntuación media es de 21 puntos y en personas mayores sin depresión es de 5 puntos. La puntuación de corte se sitúa en 14 puntos, puntuación que denota la presencia de depresión. Asimismo, existen otros trabajos españoles de adaptación de esta escala[19-21].

El presente trabajo se dirige a comparar la capacidad de 3 versiones de la GDS, la original de 30 ítems y las formas reducidas de 15 y 5, para identificar la presencia de depresión en personas mayores. Los datos disponibles son aún controvertidos respecto a la posibilidad de usar las versiones reducidas sin perder capacidad de cribado y, particularmente, existen muy pocos datos al respecto con la versión española de la escala.

MATERIAL Y MÉTODOS

Sujetos

La muestra estuvo compuesta por 246 personas mayores, no institucionalizadas, reclutadas en centros de día y clubes y hogares para personas mayores. La media de edad ± desviación estándar fue de 74,1 ± 9,5 años; el 68,7% eran mujeres. La sospecha de deterioro cognitivo fue un criterio de exclusión de los participantes en el estudio. Un total de 13 personas fueron excluidas por obtener una puntuación igual o inferior a 25 en la prueba miniexamen cognoscitivo. Un 20,2% de la muestra (47 personas) cumplió los criterios diagnósticos del DSM-IV para un trastorno de depresión mayor (DM).

Medidas

Además de la GDS, descrita con detalle en la introducción de este trabajo, se evaluaron los siguientes aspectos:

– La presencia de depresión clínicamente relevante fue establecida mediante una entrevista diagnóstica estructurada basada en los criterios del DSM-IV[22] para los trastornos de ansiedad y los del estado de ánimo. Esta entrevista cuestiona, siguiendo el orden establecido en los árboles de decisión del DSM-IV, los diferentes criterios de cada trastorno. En cada trastorno se pregunta inicialmente por los criterios clave y, en caso de respuesta afirmativa, por otras características asociadas al trastorno, y trata de establecer la gravedad de la alteración y el cuadro clínico que caracterizan al sujeto. Se valoró también el consumo de fármacos y el padecimiento de algún trastorno orgánico, para descartar la presencia de trastorno depresivo debido a enfermedad médica o a consumo de medicamentos.

El Mini-Examen Cognoscitivo[23] es un instrumento que evalúa diversos aspectos relevantes del funcionamiento cognitivo. Es una prueba habitualmente empleada en valoraciones neuropsicológicas con personas mayores, como un procedimiento rápido y fácil de detección de posibles casos de deterioro cognitivo[24]. La puntuación final oscila entre 0 y 35 puntos y ofrece un índice del grado de funcionamiento cognitivo del sujeto. La prueba se aplicó a la muestra seleccionada como un modo de filtrar posibles casos de deterioro cognitivo, utilizando un punto de corte de 25[24].

Procedimiento

La evaluación fue realizada por licenciados en psicología con formación y experiencia en evaluación gerontológica. Asimismo, la GDS fue aplicada por el entrevistador, cuyo papel fue estrictamente leer las instrucciones y el contenido de los ítems y consignar las respuestas sin que en ningún caso aportara explicaciones adicionales a lo incluido en el instrumento.

Análisis de datos

En primer lugar, se compararon las medias en la puntuación total de las tres versiones de la GDS entre personas diagnosticadas con DM y personas sin diagnóstico, mediante la prueba de comparación de medias para muestras no relacionadas.

Con el fin de comparar la capacidad de las 3 versiones de la GDS en la detección de la depresión se utilizó el análisis de la curva característica operativa del receptor (COR, traducción de las siglas inglesas ROC; *receiver operating characteristics*). En los últimos años se ha generalizado su uso en el ámbito clínico como medio de analizar la capacidad de una prueba para identificar la presencia o ausencia de un determinado trastorno[25]. El análisis COR permite conocer la capacidad de un cuestionario para identificar casos positivos y negativos de un grupo determinado, es decir, aporta información objetiva sobre la capacidad de una prueba como instrumento de cribado para la presencia de trastornos. En concreto, el

análisis COR ofrece datos sobre la sensibilidad y la especificidad asociadas a cada puntuación en el instrumento de evaluación. Esto permite establecer cuantitativamente un punto de corte para la identificación de casos clínicos. Aunque varios criterios se han empleado para tomar la decisión de cuál es el punto de corte más adecuado, el índice de Youden[26] es el más usualmente informado en la literatura médica. Este índice viene dado por el valor obtenido por la fórmula ([sensibilidad + especificidad] – 1). Los valores cercanos a 1 indican que tanto la sensibilidad como la especificidad asociadas a esa puntuación son elevadas. Otro índice de la bondad de una prueba como instrumento de cribado es el área bajo la curva, cuyos valores oscilan entre 0 y 1. Puntuaciones > 0,0 indican que la prueba clasifica correctamente más casos de los esperables por azar y valores cercanos a 1 indican un buen funcionamiento del instrumento. Una ventaja adicional del análisis COR es que la prueba es bastante robusta incluso cuando no puede asumirse la normalidad de las puntuaciones o las muestras son de tamaños diferentes[27]. Todos los análisis se realizaron con el programa SPSS para Windows, versión 13.0

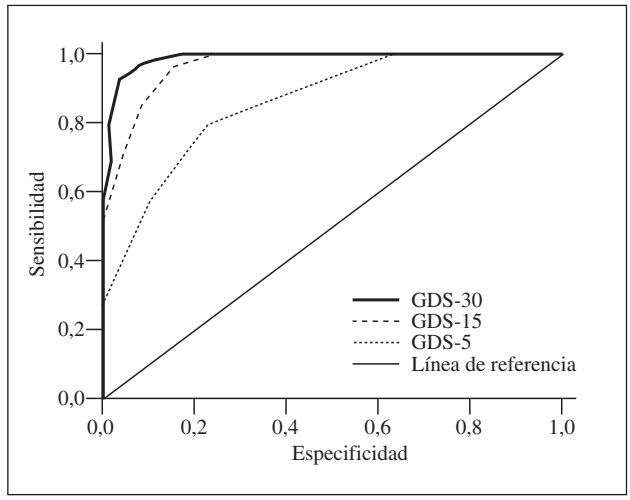

Figura 1. *Curvas COR para GDS-30, GDS-15 y GDS-5 que predicen la presencia de depresión mayor.*

RESULTADOS

En primer lugar, se compararon mediante la prueba de la t para muestras independientes las puntuaciones en las 3 versiones de la GDS entre las 47 personas identificadas con DM y las 186 identificadas como personas sin DM. Los resultados de estos análisis, junto con los datos descriptivos para las diferentes variables, se presentan en la tabla 1, donde puede apreciarse que las personas identificadas con DM presentaron puntuaciones significativamente superiores en las 3 versiones de la GDS ($p < 0,001$).

En un primer análisis COR, con la GDS-30 identificando la presencia de DM, la curva dejó por debajo un área de 0,982 ($p < 0,001$; intervalo de confianza [IC] del 95%, 0,968-0,996), en el supuesto no paramétrico. Un punto de corte de 13 (puntuación igual o superior) maximizó simultáneamente la sensibilidad (95,7%) y la especificidad (92,4%) de la escala para la identificación de la DM (índice de Youden = 0,881). Con este punto de corte, la potencia predictiva positiva de la escala (el porcentaje de personas clasificadas por la GDS-30 con DM que realmente la tienen) fue del 76,3%, mientras la potencia predictiva negativa (el porcentaje de personas clasificadas por la GDS-30 como sin DM que realmente no la tienen) fue del 98,8%. El coeficiente kappa, al evaluar el grado de acuerdo entre la GDS-30 y el diagnóstico mediante la entrevista estructurada, fue de 0,805 ($p < 0,001$). Por otra parte, una puntuación de 11 maximizaba la sensibilidad con una especificidad superior a la esperable por azar (0,50). La sensibilidad para esta puntuación fue del 100% y la especificidad del 81,1% (índice de Youden = 0,811). Con este punto de corte la potencia predictiva positiva fue del 57,3% y la negativa del 100%, con un coeficiente kappa de 0,635 ($p < 0,001$). Finalmente, una puntuación de 17 maximizaba la especificidad (98,9%) reteniendo una sensibilidad superior a la esperable por azar (59,6%). La potencia predictiva positiva con este punto de corte fue del 93,3%, mientras que la negativa fue del 90,6%, con un coeficiente kappa de 0,676 ($p < 0,001$). En la tabla 2 se presentan la sensibilidad, la especificidad y el índice de Youden asociados a todas las puntuaciones en la GDS-30 que se encuentran entre el punto de corte que optimiza la sensibilidad y el que optimiza la especificidad; se destaca el punto de corte que optimiza ambas a la vez (fig. 1).

TABLA 1. Prueba de la t de comparación de medias entre personas con y sin depresión mayor

Instrumentos	*Depresión mayor (n = 47)*		*Sin depresión mayor (n = 186)*		t_{230}	*IC del 95%*
	Media	*DT*	*Media*	*DT*		
GDS-30	17,34	3,21	6,45	3,89	19,86	11,9-9,8
GDS-15	7,64	2,11	2,45	1,90	16,33	5,8-4,6
GDS-5	2,70	1,21	0,98	0,96	9,04	2,1-1,3

DT: desviación típica; IC: intervalo de confianza.
Todas las diferencias fueron estadísticamente significativas ($p < 0,001$).

TABLA 2. Especificidad, sensibilidad e índice de Youden para puntuaciones en la GDS-30

Puntuación	*Sensibilidad*	*Especificidad*	*Índice de Youden*
11	1	0,811	0,811
12	0,979	0,876	0,855
13	0,957*	0,924*	0,881*
14	0,915	0,957	0,872
15	0,787	0,978	0,765
16	0,681	0,978	0,659
17	0,596	0,989	0,585

*Puntuación que maximiza a la vez la sensibilidad y la especificidad; la primera puntuación es la que maximiza la sensibilidad manteniendo una especificidad superior a lo esperable por azar, la última es la puntuación que maximiza la especificidad manteniendo una sensibilidad superior al azar.

TABLA 3. Especificidad, sensibilidad e índice de Youden para puntuaciones en la GDS-15

Puntuación	*Sensibilidad*	*Especificidad*	*Índice de Youden*
4	1	0,724	0,724
5	0,957	0,843	0,800
6	0,851	0,914	0,765
7	0,660	0,968	0,628

En el caso de la GDS-15 la curva COR dejó por debajo un área de 0,962 ($p < 0,001$; IC del 95%, 0,940-0,984), en el supuesto no paramétrico. Una puntuación de 5 maximiza simultáneamente la sensibilidad (95,7%) y la especificidad (84,3%) de la escala para la identificación de la DM (índice de Youden = 0,800). Con este punto de corte, la potencia predictiva positiva de la escala fue del 60,8%, mientras la potencia predictiva negativa fue del 98,7% y el coeficiente kappa de 0,659 ($p < 0,001$). Por otra parte, una puntuación de 4 optimizaba la sensibilidad de la GDS-15 (100%) con una especificidad (72,4%) superior a lo esperable por azar (índice de Youden = 0,724). Con este punto de corte, la potencia predictiva positiva fue del 48% y la negativa del 100%, con un coeficiente kappa de 0,516 ($p < 0,001$). Finalmente, una puntuación de 7 maximizaba la especificidad (96,8%) reteniendo una sensibilidad (66%) superior a la esperable por azar (índice de Youden = 0,628). La potencia predictiva positiva con este punto de corte fue del 83,8%, mientras que la negativa fue del 91,8%, con un coeficiente kappa de 0,681 ($p < 0,001$). En la tabla 3 se presentan la sensibilidad, la especificidad y el índice de Youden asociados a todas las puntuaciones en la GDS-15 que se encuentran entre el punto de corte que optimiza la sensibilidad y el que optimiza la especificidad, con el punto de corte que optimiza ambas a la vez.

En el caso de la GDS-5, la curva COR dejó por debajo un área de 0,856 ($p < 0,001$; IC del 95%, 0,799-0,912), en el supuesto no paramétrico. Una puntuación de 2 maximiza la sensibilidad (78,7%) y especificidad (77,3%) de la escala para la identificación del TAG (índice de Youden = 0,560). Con este punto de corte, la potencia predictiva positiva de la escala fue del 46,8%, mientras la potencia predictiva negativa fue del 93,5% y el coeficiente kappa de 0,447 ($p < 0,001$). Por otra parte, una puntuación de 3 optimizaba la especificidad (89,7%) reteniendo una sensibilidad (57,4%) superior a la esperable por azar (índice de Youden = 0,471). La potencia predictiva positiva con este punto de corte fue del 58,7%, mientras que la negativa fue del 89,2%, con un coeficiente kappa de 0,476 ($p < 0,001$).

DISCUSIÓN

El objeto de este estudio era comprobar la validez psicométrica de tres de las versiones más utilizadas de la escala de depresión geriátrica que varían en función de su longitud, compuestas por 30, 15 y 5 ítems. Para ello se utilizaron diferentes puntos de corte para cada una de las 3 versiones y se comparó la eficacia diagnóstica con un patrón diagnóstico habitual, el criterio diagnóstico de DM según DSM IV-TR.

La versión original (GDS-30) muestra ser una buena medida de cribado para detectar la posible presencia de depresión en personas mayores, con buenas sensibilidad (95,7%) y especificidad (92,4%) para un punto de corte de 13, aunque la escala presenta una potencia predictiva positiva relativamente baja (76,3%). La capacidad diagnóstica de la escala empeora notablemente cuando se reduce su longitud, especialmente en la versión más corta. La GDS-15 presenta valores de sensibilidad y especificidad peores que la GDS completa y, sobre todo, para el punto de corte de 5, que maximiza simultáneamente ambos índices, la potencia predictiva positiva es del 60,8%, por lo que el uso de esta versión de la escala implica asumir una elevada probabilidad de obtener falsos positivos. Esta relativamente elevada tasa de inespecificidad de la GDS-15 podría representar una limitación en su uso, dependiendo del contexto clínico en el que se enmarca. Si únicamente se utiliza como un primer filtro en un estudio epidemiológico o en un trabajo para identificar depresión tratable, podría no representar un problema, ya que en una fase posterior sería descartado como un verdadero caso de depresión. Si, por el contrario, la utilización de la GDS-15 supone un fin en sí mismo con objeto de identificar casos de depresión para un posible tratamiento, un 16% de la muestra podría recibir una etiqueta diagnóstica o tratamiento no empíricamente justificados. Los resultados de la GDS-5 son ya claramente peores. Para un punto de corte de 2 la potencia predictiva positiva fue de sólo el 46,8%. Al margen de que pueden hacerse comentarios similares respecto a la versión de 5 elementos, con un 23% de casos diagnosticables

de depresión que, en realidad, no lo son, hay que añadir que más de una de cada 5 personas mayores con depresión no serían identificadas como tales. Esta evidencia invalida el uso clínico de la GDS-5, ya que no responde mínimamente a cualquiera de los dos objetivos que puede tener un autoinforme de depresión. Por una parte, con un número tan reducido de ítems no podría cuantificarse la gravedad de un trastorno y, por otra parte, respecto al objetivo que ha guiado la realización de este trabajo, la posibilidad de usar las versiones GDS como prueba de cribado, tampoco sería adecuado ya que no responde a la principal característica que deber tener una prueba de esta naturaleza: ser altamente sensible respecto al fenómeno que pretende identificar. Más de un 20% de los casos no detectados superaría el umbral de una buena práctica clínica.

En resumen, la evidencia obtenida en este trabajo nos lleva a sugerir que la GDS debería revisarse con el fin de desarrollar una nueva escala que combine utilidad diagnóstica con una muy elevada sensibilidad, una adecuada especificidad y economía en su aplicación, mediante la selección de los ítems de mayor capacidad discriminativa.

BIBLIOGRAFÍA

1. Jongenelis K, Pot AM, Eisses AMH, Gerritsen DL, Derksen M, Beekman ATF, et al. Diagnostic accuracy of the original 30-Item and shortened versions of the Geriatric Depression Scale in nursing home patients. Int J Geriatr Psychiatry. 2005;20:1067-74.
2. Izal M, Márquez M, Losada A, Montorio I, Nuevo R. Una revisión sobre las intervenciones cognitivo-conductuales para la depresión en la edad avanzada. Rev Esp Geriatr Gerontol. 2003;38:34-45.
3. Brink TL, Yesavage JA, Lum O, Heersema PH, Adey M, Rose M. Screening tests for geriatric depression. Clin Gerontol. 1982;1:37-43.
4. Herrman N, Mittmann N, Silver IL, Shulman KI, Busto UA, Shear NH, et al. A validation study of the Geriatric Depression Scale Short Form. I. Int J Ger Psych. 1996;11:457-60.
5. Izal M, Montorio I. Adaptación en nuestro medio de la Escala de Depresión Geriátrica (GDS) en distintos subgrupos: residentes en la comunidad y asistentes a hospitales de día. Rev Gerontol. 1996;6:329-37.
6. Incalzi RA, Cesari M, Pedone C. Construct validity of the 15-item Geriatric Depression Scale in older medical inpatients. J Geriatr Psychiatry Neurol. 2003;16:23-8.
7. Yesavage JA, Brink TL, Rose TL, Lum O, Huang V, Adey M. Development and validation of Geriatric Depression Screening Scale: A preliminary report. J Psychiatr Res. 1983;17:37-49.
8. Heisel MJ, Gordon D, Flett I, Duberstein PR, Lyness JM. Does the Geriatric Depression Scale (GDS) distinguish between older adults with high versus low levels of suicidal ideation? Am J Geriatr Psychiatry. 2005;13:876-83.
9. Sheikh JI, Yesavage JA. Geriatric Depression Scale (GDS): Recent evidence and development of a shorter version. Clin Gerontol. 1986;5:165-73.
10. Lesher EL, Berryhill JS. Validation of the Geriatric Depression Scale-Short Form among inpatients. J Clin Psychol. 1994;50:256-60.
11. Wall JR, Lichtenberg PA, MacNeill SE. Depression detection in geriatric rehabilitation: Geriatric Depression Scale Short Form versus Long Form. Clin Gerontol. 1999;20:13-21.
12. D'Ath P, Katona P, Mullan E, Evans S, Katona C. Screening detection and management of depression in elderly primary care attendees. I: The acceptability and performance of the 15-item Geriatric Depression Scale (GDS15) and the development of short versions. Family Practice. 1994;11:260-66.
13. Van Marwijk HWJ, Wallace P, De Bock GH, Kaptein AA, Mulder JD. Evaluation of the feasibility, reliability and diagnostic value of shortened versions of the geriatric depression scale. Br J Gen Prac.1995;45:195-9.
14. Arthur A, Jagger C, Lindesay J, Graham C, Clarke M. Using an annual over-75 health check to screen for depression: validation of the short Geriatric Depression Scale (GDS15) within general practice. Int J Geriatr Psychiatry. 1999;14:431-9.
15. Brown LM, Schinka JA. Development and initial validation of a 15-item informant version of the Geriatric Depression Scale. Int J Geriatr Psychiatry. 2005;20:911-8.
16. Hoyl MT, Alessi CA, Harker JO, Josephson KR, Pietruszka FM, Koelfgen M. Development and testing of a five-item version of the Geriatric Depression Scale. J Am Geriatr Soc. 1999;47:873-8.
17. Rinaldi P, Mecocci P, Benedetti C, Ercolani S, Bregnocchi M, Menculini G. Validation of the five-item Geriatric Depression Scale in elderly subjects in three different settings. J Am Geriatr Soc. 2003;51:694-8.
18. Izal M, Montorio I. Adaptation of the Geriatric Depression Scale in Spain: a preliminary study. Clin Gerontologist. 1993;13:83-91.
19. De Dios del Valle R, Hernández Sánchez AM, Rexach Cano LI, Cruz Jentoft AJ. Validación de una versión de cinco ítems de la Escala de Depresión Geriátrica de Yesavage en población española. Rev Esp Geriatr Gerontol. 2001;36:276-80.
20. Fernández-San Martín MI, Andrade C, Molina J, Muñoz PE, Carretero B, Rodríguez M. Validation of the Spanish version of the geriatric depression scale (GDS) in primary care. Int J Geriatr Psychiatry. 2002;17:279-87.
21. Rodríguez-Testal JF, Valdés M. Brief-Time-Series Analysis of Elderly People's Depressive Symptomatology. Span J Psychol. 2003;6:35-50.
22. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4th ed. Washington; 1994.
23. Lobo A, Esquerra J, Gómez Burganda F, Sala JM, Seva Díaz A. El miniexamen cognoscitivo (un «test» sencillo, práctico, para detectar alteraciones intelectuales en pacientes médicos). Actas Luso-Espa Neurol, Psiquiatr. 1979;7:189-202.
24. Del Ser T, Peña-Casanova J. Objetivos y métodos de la evaluación de la demencia. En: Del Ser T, Peña-Casanova J, editores. Evaluación neuropsicológica y funcional de la demencia. Barcelona: J.R. Prous Editores; 1994
25. Shapiro DE. The interpretation of diagnostic tests. Statistical Methods in Medical Research. 1999;8:113-34.
26. Youden WJ. Index for rating diagnostic test. Cancer. 1950;3:32-5.
27. Rice ME, Harris GT. Violent recidivism: Assessing predictive validity. J Consult Clin Psychol. 1995;63:737-48.