

# Validez diagnóstica de la Self-Rating Depression Scale de Zung en pacientes de atención primaria

ARAGONÈS BENAIGES, E.\*; MASDÉU MONTALÀ, R. M.\*; CANDO GUASCH, G.\*\* y COLL BORRÀS, G.\*\*\*

\* Centro de Atención Primaria de Constantí. Instituto Catalán de la Salud. \*\* Centro de Atención Primaria de Bonavista. Tarragona. Instituto Catalán de la Salud. \*\*\* Hospital Psiquiátrico Instituto Pere Mata. Reus (Tarragona).

## Diagnostic validity of Zung's Self-Rating Depression Scale on primary care patients

### Resumen

**Introducción:** Aunque la versión española del cuestionario Self-Rating Depression Scale de Zung (SDS) goza de gran difusión, no existen estudios sobre su validez como prueba diagnóstica en pacientes de atención primaria.

**Metodología:** En una primera fase evaluamos una muestra de 350 pacientes consecutivos de atención primaria con el SDS y, en una segunda fase, aplicamos la Structured Clinical Interview for DSM-IV, en sus apartados de Episodio Depresivo Mayor y Distimia actuales, a una submuestra compuesta por todos los pacientes con resultado positivo en el test y 1/3 aleatorio de los negativos. Utilizamos metodología específica para evitar el sesgo de verificación y calculamos prevalencia, sensibilidad y especificidad, valores predictivos, curva Receiver Operating Characteristic (ROC) y Razones de Verosimilitud Específicas por Estratos (RVEE).

**Resultados:** La prevalencia de depresión mayor fue de 14,7% (IC95%: 10,7%-18,7%) y la de distimia 4,6% (IC95%: 2,4%-6,8%). Los resultados de sensibilidad y especificidad para detectar ambos diagnósticos fueron 0,95 (IC95%: 0,87-1) y 0,74 (IC95%: 0,68-0,79). El área bajo la curva ROC fue de 0,93. La RVEE para las puntuaciones < 50 conduce a una probabilidad post-test de 0,01. En el estrato de puntuaciones > 69 la RVEE lleva a una probabilidad post-test de 0,96. Los estratos intermedios ofrecieron resultados menos concluyentes.

**Conclusiones:** El SDS es eficaz en pacientes de atención primaria y muestra un comportamiento equiparable a otros cuestionarios de detección de la depresión. La aplicación de las RVEE puede ser muy práctica e informativa para estimar la probabilidad de padecer un trastorno depresivo en los pacientes individuales.

**Palabras clave:** Zung's Self-Rating Depression Scale (SDS).

Escalas de valoración psiquiátrica. Trastorno depresivo. Sensibilidad y especificidad. Atención primaria.

### Summary

**Introduction:** Although the Spanish version of the Zung's Self-Rating Depression Scale (SDS) is widely used, there are no studies about its validity as a diagnostic test in primary health care patients.

**Methods:** In a first phase, a sample of 350 consecutive primary care patients was assessed with the SDS. In a second phase, a subsample composed by all the positive test results and 1/3 of the negatives selected at random, was assessed with the modules of current Major Depressive Episode and Dysthymia of the Structured Clinical Interview for DSM-IV. Specific methods to avoid verification bias were used. Prevalence, sensitivity and specificity, predictive values, Receiver Operating Characteristic (ROC) curve, and Stratum Specific Likelihood Ratios (SSLR) were calculated.

**Results:** Prevalence estimations of major depression and dysthymia were 14,7% (IC95%: 10,7-18,7%) and 4,6% (IC95%: 2,4%-6,8%) respectively. Sensitivity and specificity to detect both diagnoses were 0,95 (IC95%: 0,87-1) and 0,74 (IC95%: 0,68-0,79). Area under ROC curve was 0,93. SSLR for scoring < 50 led to a post-test probability of 0,01. In the stratum with scoring > 69 the SSLR generated a post-test probability of 0,96. Less conclusive results were obtained by intermediate strata.

**Conclusions:** The SDS is effective in primary care patients and shows operating characteristics comparable to other depression assessment scales. SSLR provides practical information to estimate the probability of suffer a depressive disorder in individual patients.

**Key words:** Zung's Self-Rating Depression Scale (SDS).

Psychiatric status rating scales. Depressive disorder. Sensitivity and specificity. Primary health care.

Los trastornos depresivos son frecuentes en el ámbito de la atención primaria, con una prevalencia para la depresión mayor entre el 6% y el 13% en diferentes estudios (1, 2). A menudo representan condiciones persis-

tentes o recurrentes y tienen repercusión en diversas áreas vitales de quien los padece (3, 4). Los pacientes deprimidos con frecuencia manifiestan su sufrimiento a través de síntomas y molestias somáticas de origen os-

curo (5) y pueden convertirse en grandes utilizadores de los servicios sanitarios, especialmente de la atención primaria (6). A pesar de la disfuncionalidad que origina la depresión y del hecho que existen opciones terapéuticas efectivas (7) en una consulta de atención primaria pueden pasar desapercibidos más del 50% de los pacientes deprimidos (8, 9).

Se han investigado las características y utilidad de algunos cuestionarios autoaplicados para el cribaje de la depresión en este nivel asistencial (10) y se ha constatado que su uso aumenta sensiblemente la tasa de detección (11, 12). Uno de estos cuestionarios es el Zung's Self-Rating Depression Scale (SDS), que data de 1965 (13) y que desde entonces ha tenido una amplia difusión y ha sido usado pródigamente en el ámbito de la atención primaria en Estados Unidos (12, 14). Una primera versión castellana de Conde data de 1969 (15) aunque está más difundida una revisión posterior del mismo autor (16).

Hay datos que señalan que el espectro y la naturaleza de la depresión que se presenta en las consultas de atención primaria tiene características diferentes a la depresión vista en el nivel psiquiátrico especializado: predominan formas de naturaleza clínica más leve, abundan formas con sintomatología depresiva por debajo del umbral diagnóstico, a menudo la presentación sintomática se focaliza en molestias somáticas y existen importantes relaciones de comorbilidad con otros trastornos psiquiátricos, como la ansiedad o el abuso de alcohol, y con diversas enfermedades somáticas (17). Por este motivo los datos obtenidos en estudios de validación de las versiones españolas de este cuestionario, llevados a cabo en el ámbito psiquiátrico, pueden no ser directamente aplicables a la utilización de estas pruebas en atención primaria.

Por otra parte, las clasificaciones y los criterios nosológicos han evolucionado y las categorías diagnósticas de los sistemas normalizados vigentes (p. ej. Diagnostic and Statistic Manual for Mental Disorders, Fourth Edition: DSM-IV) no son coincidentes ni equiparables a las de otros sistemas (18), como el Global Rating o «juicio clínico», Criterios de Kielholz o Criterio Hospitalario tomados como referencia para analizar la validez del cuestionario en los estudios previos (19-21).

Este estudio pretende determinar la fiabilidad y la validez como prueba diagnóstica del cuestionario SDS de Zung para la detección de la depresión en los pacientes de atención primaria, utilizando como referencia las categorías diagnósticas del DSM-IV y empleando los instrumentos estadísticos para el análisis de validez de pruebas diagnósticas más actuales.

## METODOLOGÍA

### Obtención de la muestra y recogida de datos

El estudio se llevó a cabo en ocho consultas médicas de cuatro centros de atención primaria, tres de ellos en la ciudad de Tarragona y el cuarto en el municipio de Constantí.

Durante el período de estudio, en cada una de las consultas estudiadas, el médico de familia que las atendía solicitaba sistemáticamente la participación en el estudio a los pacientes consecutivos que visitaba por cualquier motivo, tras comprobar los criterios de inclusión y exclusión. Se consideraron los siguientes criterios de inclusión: 1) Estar adscrito al centro de atención primaria, 2) Edad comprendida entre 18 y 70 años y 3) No haber acudido a la consulta por motivos exclusivamente administrativos ni por delegación de terceras personas. Se consideró motivo de exclusión la existencia de limitación física, psíquica, cultural o enfermedad concurrente que hiciese imposible la aplicación de las pruebas del estudio. También debían excluirse los pacientes afectos de trastorno psicótico conocido.

A los pacientes seleccionados para la participación y que consentían en ella, en una primera fase, se les entregaba el cuestionario SDS de Zung que debían cumplimentar en el mismo momento bajo supervisión y asistencia, si fuera necesaria, de un miembro del equipo investigador. En esta fase se recogían mediante un cuestionario específico los datos sociodemográficos de los individuos de la muestra.

La puntuación del SDS se realizaba inmediatamente después de ser cumplimentado y a continuación, en una segunda fase, se aplicaba la prueba diagnóstica de referencia a una submuestra compuesta por la totalidad de los pacientes con resultado positivo (índice SDS  $\geq$  50%) y un tercio de los pacientes con resultado negativo seleccionados aleatoriamente.

### El Self-Rating Depression Scale de Zung

El SDS es un cuestionario autoaplicado para la detección y la medida de la gravedad de la depresión que consta de 20 ítems planteados en forma positiva o negativa que exploran síntomas depresivos y evalúan si la característica está presente en el momento actual o recientemente «raramente», «a veces», «frecuentemente» o «siempre». Cada ítem puntúa de 1 a 4 puntos. La puntuación bruta obtenida (p) se convierte a un índice porcentual (índice SDS =  $p \times 100 / 80$ ). Los puntos de corte usuales son:  $\geq$  50% para la depresión leve,  $\geq$  60% para la depresión moderada y  $\geq$  70% para la depresión grave. En este estudio utilizamos la segunda versión castellana de Conde (16).

### Prueba diagnóstica de referencia

Como prueba diagnóstica de referencia se utilizó el módulo de Trastornos Afectivos de la entrevista psiquiátrica Structured Clinical Interview for DSM-IV Axis I Disorders SCID-I (DSM-IV) en versión española (22). Se consideraron los diagnósticos de Episodio Depresivo Mayor actual (F32.x) y Trastorno Distímico actual (F34.1) teniendo en cuenta la norma jerárquica según la cual si existe Episodio Depresivo Mayor actual no se explora el diagnóstico de Trastorno Distímico actual.

La SCID fue aplicada por cuatro investigadores, médicos de familia, que habían recibido formación y entrenamiento específico en el uso de la SCID siguiendo la metodología propuesta en el Manual del Usuario de la entrevista.

Éstos no habían participado en la aplicación del SDS e ignoraban, en el momento de la entrevista, el resultado previo del cuestionario en el paciente entrevistado para evitar así el sesgo de revisión.

### Análisis de la fiabilidad (23)

Obtuvimos el coeficiente de fiabilidad utilizando el método de las dos mitades (Split-Half). Éste consiste en partir el test en dos mitades que representan dos formas supuestamente paralelas de un test con la mitad de la longitud original, en nuestro caso ítems pares-ítems impares, y comprobar la correlación entre ambas puntuaciones mediante el coeficiente de Pearson. Dado que un test aumenta su fiabilidad con el número de ítems que lo integran, la fiabilidad obtenida será inferior a la del test completo. Por este motivo se usa la corrección de Spearman-Brown del coeficiente estimando su valor con relación a la longitud original del test.

Un aspecto particular de la fiabilidad es la consistencia interna: hasta que punto los diferentes ítems del test tienden a medir un mismo contenido. Este grado de acuerdo lo medimos con el coeficiente  $\alpha$  de Cronbach.

### Análisis de la validez diagnóstica

Cuando no todos los pacientes a los que se ha aplicado el test diagnóstico objeto del estudio son seleccionados para la aplicación de la prueba de referencia y se calculan los índices de validez diagnóstica sólo con la submuestra de pacientes verificados, se produce un sesgo de verificación que, en nuestro caso, con la totalidad de los resultados positivos y sólo un tercio de los negativos verificados, elevaría erróneamente la prevalencia y la sensibilidad y disminuiría la especificidad.

Bajo la asunción de independencia condicional: la selección de pacientes para la verificación sólo depende del resultado del test y no de otras variables, Zhou (24) demuestra que los valores predictivos calculados directamente de la submuestra verificada no están sesgados. Para la sensibilidad, especificidad y prevalencia este autor propone unos estimadores no sesgados de «máxima verosimilitud» (maximum likelihood estimator), así como métodos para calcular sus varianzas (25). Las razones de verosimilitud se han calculado a partir de la prevalencia corregida y de los estimadores insesgados de los valores predictivos (26).

En un instrumento que no se limita a ofrecer un resultado positivo o negativo sino cuyo resultado es una variable cuantitativa, al dicotomizar el resultado en positivo o negativo, despreciamos información útil. Para superar esta limitación se han propuesto nuevas técnicas en la

evaluación de pruebas diagnósticas, como son la curva ROC y las razones de verosimilitud específicas por estratos (RVEE).

La curva ROC (27) se construye mediante un gráfico de dos ejes donde, para cada punto de corte, en el eje de ordenadas se expresa la sensibilidad y en el eje de abscisas el complementario de la especificidad. Los puntos resultantes unidos por una línea dibujan una curva el área por debajo de la cual refleja la valoración global de la validez prueba diagnóstica.

Las RVEE son un instrumento promovido por diversos autores desde el paradigma de la medicina basada en la evidencia como más informativo y práctico que los tradicionales índices de sensibilidad y especificidad y que permiten relacionar la probabilidad pre-test de padecer el trastorno en cuestión (prevalencia), el resultado del test y la probabilidad post-test (28). La razón de verosimilitud (RV) es el cociente entre la probabilidad de obtener determinado resultado en el test cuando el trastorno está presente y la probabilidad de obtener el mismo resultado cuando el trastorno no existe. La RV se integra en el Teorema de Bayes de la siguiente forma:

Odds previa  $\times$  Razón de verosimilitud = Odds posterior.

Y la conversión de odds en probabilidad y viceversa se consigue:

$$\text{Probabilidad} = \text{Odds} / \text{Odds} + 1$$

$$\text{Odds} = \text{Probabilidad} / 1 - \text{Probabilidad}$$

## RESULTADOS

Se excluyeron del estudio a 14 pacientes: cinco de ellos por enfermedad que impedía la aplicación de las pruebas en el momento del estudio (un cuadro febril agudo, un enfermo neoplásico, un vértigo agudo, dos crisis psicógenas agudas), seis por limitación cultural (tres extranjeros y tres analfabetos), dos por limitación física (un larigectomizado y un ciego), uno por limitación psíquica (trastorno cognitivo) y uno por trastorno psicótico activo (esquizofrenia).

Se propuso la participación en el estudio a 380 pacientes que cumplieran las condiciones de inclusión y no tenían motivo de exclusión de los que 30 rehusaron (7,9%). Así pues, 350 pacientes formaron la muestra inicial en la que cabe destacar una relación mujer varón de 2:1.

De los 350 pacientes que cumplimentaron el SDS, 141 (40,3%; IC 95%: 35,1%-45,4%) dieron un resultado positivo. De estos, tres pacientes declinaron participar en la segunda fase: un varón de 56 años y dos mujeres de 61 y 53 años con puntuaciones de 59, 59 y 50% respectivamente. Se aplicó la entrevista SCID a 138 pacientes con SDS positivo y a 67 con SDS negativo. Las características sociodemográficas de la muestra inicial y de la submuestra que pasó a la segunda fase se muestran en la tabla I.

El coeficiente de fiabilidad hallado fue  $r_{x, x'(\text{corregido})} = 0,80$  y la consistencia interna se manifestó en un coeficiente  $\alpha$  de Cronbach = 0,78.

TABLA I Características demográficas de los pacientes en la muestra inicial y en la submuestra de la segunda fase

Variable	Categoría	Muestra primera fase (n= 350)		Muestra segunda fase (n= 205)	
		Número	Porcentaje	Número	Porcentaje
Sexo	Varón	123	35,1	62	30,1
	Mujer	227	64,9	143	69,4
Edad	Media (DE)	45,9 (14,5)		46,2 (14,2)	
	18-29	61	17,4	34	16,6
	30-39	57	16,3	32	15,6
	40-49	80	22,9	53	25,9
	50-59	77	22	41	20
	60-70	75	21,4	45	22
Estado civil	Soltero	53	15,1	30	14,6
	Casado/Vive en pareja	272	77,7	160	78,
	Divorciado/Separado	11	3,1	5	2,4
	Viudo	14	4	10	4,9
Nivel educativo	Analfabeto	25	7,1	12	5,9
	Sabe leer/escribir	64	18,3	38	18,5
	Estudios primarios	169	48,3	105	51,2
	Estudios secundarios	70	20	38	18,5
	Estudios universitarios	22	6,3	12	5,9

Corrigiendo los datos para evitar el sesgo de verificación, según se ha descrito en la metodología, la prevalencia hallada para Episodio Depresivo Mayor actual fue de 14,7% (IC 95%: 10,7%-18,7%) y para Trastorno Distímico actual de 4,6% (IC 95%: 2,4%-6,8%). La prevalencia global de ambos trastornos fue de 19,3% (IC 95%: 14,9%-23,7%). Los resultados de los índices de validez obtenidos para el punto de corte convencional (índice SDS ≥ 50) se expresan en la tabla II.

En la figura 1 se muestra la curva ROC, obtenida con los valores de sensibilidad y especificidad corregidos para evitar el sesgo de verificación.

La zona de la curva ROC próxima al vértice superior izquierdo de la gráfica es donde los valores de sensibilidad y especificidad alcanzan sus valores más altos. En la

tabla III se muestran estos valores para diferentes puntos de corte en esta zona, de manera que el lector interesado pueda seleccionar un punto de corte óptimo teniendo en cuenta la importancia relativa que pueda atribuir a los falsos positivos y a los falsos negativos en la aplicación de la prueba.

En la tabla IV se expresan para diferentes niveles de puntuación las RVEE y las probabilidades post-test to-

TABLA II Índices de validez del Zung's Self-Rating Depression Scale (punto de corte ≥ 50)

	Episodio Depresivo Mayor (DSM-IV)	Episodio Depresivo Mayor o Trastorno Distímico (DSM-IV)
Prevalencia	0,15 (0,11-0,19)	0,19 (0,15-0,24)
Sensibilidad	0,94 (0,83-1)	0,95 (0,87-1)
Especificidad	0,70 (0,64-0,75)	0,74 (0,68-0,79)
Valor predictivo positivo	0,35 (0,27-0,43)	0,46 (0,38-0,55)
Valor predictivo negativo	0,99 (0,96-1)	0,99 (0,96-1)
Razón de verosimilitud del resultado +	3,09 (2,56-3,72)	3,61 (2,95-4,42)
Razón de verosimilitud del resultado -	0,09 (0,01-0,59)	0,06 (0,01-0,43)

Nota: Entre paréntesis se expresan los intervalos de confianza al 95%.

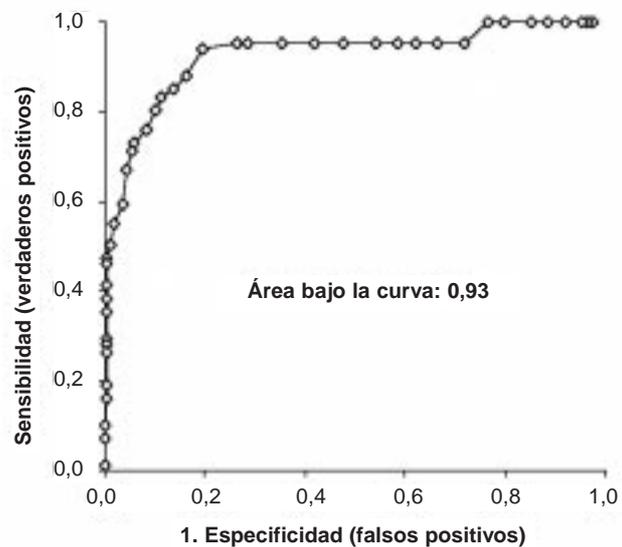


FIG. 1.—Curva ROC del Zung's Self-Rating Depression Scale para detectar Episodio Depresivo Mayor o Trastorno Distímico (DSM-IV).

**TABLA III** Sensibilidad y especificidad del Zung's Self-Rating Depression Scale en diferentes puntos de corte para detectar Episodio Depresivo Mayor o Trastorno Distímico (DSM-IV)

Punto de corte	Sensibilidad	Especificidad
≥ 50	0,95	0,74
≥ 51	0,94	0,80
≥ 52	0,88	0,84
≥ 54	0,85	0,86
≥ 55	0,83	0,89
≥ 56	0,80	0,90
≥ 58	0,76	0,92
≥ 59	0,73	0,94
≥ 60	0,71	0,95
≥ 61	0,67	0,96

mando como probabilidad pre-test para nuestra muestra la prevalencia hallada para los trastornos depresivos del 19,3%, que corresponde a una odds pre-test de 0,24.

### CONCLUSIONES

Las pérdidas de pacientes pueden introducir un sesgo en los resultados, presumiblemente infravalorando la prevalencia real, pues los pacientes afectos de trastornos mentales pueden ser más reacios en participar (29). Sin embargo, la participación conseguida es relativamente alta (92,1% en la primera fase y 97,9% en la segunda), lo que nos hace confiar en la validez de nuestros datos. Esto lo atribuimos a que la colaboración de los pacientes la solicitaba su propio médico de familia y a que todo el proceso de recogida de datos se realizaba en un mismo momento a continuación de la visita.

Existe una dificultad metodológica para establecer un diagnóstico de referencia de incontrovertible validez. Hemos usado los criterios diagnósticos del DSM-IV evaluados mediante una entrevista estructurada y con esto, al menos, garantizamos la fiabilidad y la comparabilidad de los diagnósticos (30). Sin embargo, se ha criticado la validez de estos criterios, derivados principalmente de los cuadros clínicos de pacientes atendidos en el ámbito especializado, cuando son aplicados a individuos fuera de este nivel asistencial (31). El principal problema al usar estos criterios en la comunidad o en atención primaria es definir el umbral de la significación clínica, que en un paciente que llega a una consulta psiquiátrica está implícito, pero que no puede suponerse del mismo modo en un paciente de atención primaria, teniendo en cuenta los diferentes mecanismos que dirigen a un paciente a uno u otro ámbito asistencial. La exigencia, en la entrevista normalizada, de señalar un criterio de significación clínica o repercusión funcional para llegar al diagnóstico, apelando a la opinión del paciente y al juicio clínico del entrevistador, puede no ser suficiente para algunos autores (32).

**TABLA IV** Razones de verosimilitud específicas por estratos (RVEE) y probabilidades post-test en el Zung's Self-Rating Depression Scale (SDS) para detectar Episodio Depresivo Mayor o Trastorno Distímico (DSM-IV)

Índice SDS	RVEE	Odds post-test*	Probabilidad post-test
< 50	0,06	0,02	0,01
50-59	1,13	0,27	0,21
60-69	6,55	1,57	0,61
> 69	108,43	26	0,96

\* Obtenidos a partir de una probabilidad pre-test (prevalencia) de 19,3% que corresponde a una odds pre-test de 0,24.

Al obtener la muestra a partir de las visitas consecutivas que se realizaban en las consultas de atención primaria, ésta aparece como desequilibrada en algunos aspectos como en la mayor proporción de mujeres o por la mayor probabilidad para reclutar a los pacientes más frecuentadores (33). Sin embargo, esto no lo consideramos como una limitación sino como una representación de la realidad de la consulta de atención primaria. Así pues, al utilizar una muestra representativa obtenida directamente de las visitas consecutivas podemos asegurar que el espectro clínico de la enfermedad, tal y como se presenta en el ámbito que pretendemos estudiar, está correctamente representado; cosa que no sucedería, y podría haber sido causa de un sesgo, si hubiésemos utilizado una muestra estratificada o un diseño con un grupo de casos deprimidos y otro de controles sanos.

Los coeficientes de fiabilidad y de consistencia interna hallados,  $r_x, x'$  (corregido) = 0,80 y  $\alpha = 0,78$ , se encuentran en el rango considerado aceptable (de 0,7 a 0,9) (23) y con valores equiparables a los hallados en otros estudios sobre el SDS de Zung llevados a cabo en diferentes ámbitos y que en la revisión de Naughton et al (34) se sitúan en rangos de 0,73-0,85 para el coeficiente de fiabilidad y de 0,72-0,84 para la consistencia interna.

En cuanto a la validez del SDS en la detección de los trastornos depresivos, usando el punto de corte habitual (SDS ≥ 50%) se obtiene una elevada sensibilidad (0,95), que puede ser prioritaria en un trabajo de cribaje, aunque la especificidad está limitada (0,74) y por tanto nos encontraremos una alta proporción de falsos positivos. Si en la utilización del test consideramos más importante la especificidad podemos elegir otro punto de corte, por ejemplo SDS ≥ 60%, con una elevada especificidad (0,95) pero con una sensibilidad pobre (0,71). Un punto de corte equilibrado puede ser SDS ≥ 54% con una sensibilidad de 0,85 y una especificidad de 0,86. Estos resultados se aproximan a los obtenidos en otros estudios sobre la validez del SDS en pacientes de atención primaria: Zung et al (35) hallan, para detectar depresión mayor, una sensibilidad del 0,97 y una especificidad de 0,63 para el punto de corte SDS ≥ 50% y Spitzer et al (36), también para depresión mayor y con el mismo punto de

corte, señalan una sensibilidad de 0,86 y una especificidad de 0,74. En la revisión de Murlow et al (10) sobre 18 estudios acerca de diversos instrumentos para la detección de la depresión en atención primaria, entre ellos el SDS, no hallan diferencias significativas entre las características operacionales de los diversos instrumentos analizados y se apuntan unos índices globales de sensibilidad de 0,84 y de especificidad de 0,72 para detectar depresión mayor. Los instrumentos en los que se analiza su validez diagnóstica para detectar depresión mayor o distimia poseen unos índices globales de sensibilidad 0,78 y especificidad 0,77.

Sin embargo, agrupar el abanico de diferentes puntuaciones en el test en dos únicas categorías: test positivo y test negativo, equivale a ignorar información útil sobre el paciente que ha cumplimentado el test. Aquí es donde radica la utilidad de las RVEE. En la tabla IV podemos ver el beneficio de dividir los posibles resultados en varios estratos para interpretar el resultado individual de un test: si el cuestionario puntúa en el estrato superior (SDS  $\geq$  69%), en virtud de la altísima RVEE hallada, nuestra probabilidad pre-test (la prevalencia) de 19,3%, pasa a una probabilidad post-test del 96%, que nos permite confirmar el diagnóstico con un alto nivel de certeza. Si el resultado hallado es SDS < 50%, la probabilidad pre-test deviene en una probabilidad post-test del 1%, lo que prácticamente nos permite descartar el diagnóstico. En el estrato SDS= 60-69%, la RVEE nos lleva a una probabilidad post-test de 61%, probablemente útil pero no diagnóstica. En el estrato SDS= 50-59% una RVEE próxima a la unidad nos conduce a una probabilidad post-test del 21%, similar a la probabilidad pre-test; por tanto, un resultado en este estrato ofrece poca información y será poco útil para confirmar o para descartar el diagnóstico, a pesar de que en una dicotomización del resultado siguiendo los tradicionales conceptos de sensibilidad y especificidad podía ser considerado como un resultado positivo o negativo, sin más matices, según el punto de corte escogido.

## AGRADECIMIENTOS

Al Dr. Antonio Labad Alquézar, profesor titular de Psiquiatría de la Facultat de Medicina i Ciències de la Salut de la Universitat Rovira i Virgili, su dirección en el adiestramiento de los investigadores en los criterios del DSM-IV y la entrevista SCID. A la dirección y los médicos de familia de los Centros de Atención Primaria de La Granja, Bonavista, Sant Pere i Sant Pau y Constantí que han permitido nuestra intromisión en sus consultas y a quienes debemos la alta tasa de participación de sus pacientes en nuestro estudio.

## BIBLIOGRAFÍA

1. Katon W, Schulberg H. Epidemiology of depression in primary care. *Gen Hosp Psychiatry* 1992;14:237-42.

2. Coyne JC, Fechner-Bates, Schwenk TL. Prevalence, nature, and comorbidity of depressive disorders in primary care. *Gen Hosp Psychiatry* 1994;16:267-76.

3. Ormel J, Von Korff M, Üstün TB, Pini S, Korten A, Oldehinkel T. Common mental disorders and disability across cultures. Results from the WHO Collaborative Study on Psychological Problems in General Health Care. *JAMA* 1994;272:1741-8.

4. Spitzer RL, Kroenke K, Linzer M, Hahn SR, Williams JBW, Verloin deGruy III E, et al. Health related quality of life in primary care patients with mental disorders. Results from the PRIME-MD 1000 Study. *JAMA* 1995;274:1511-7.

5. Katon WJ, Walker EA. Medically unexplained symptoms in primary care. *J Clin Psychiatry* 1998;59 Supl: 15-21.

6. Karlsson H, Lehtinen V, Joukamaa M. Psychiatric morbidity among frequent attender patients in primary care. *Gen Hosp Psychiatry* 1995;17:19-25.

7. Coulehan JL, Schulberg HC, Block MR, Madonia MJ, Rodríguez E. Treating depressed primary care patients improves their physical, mental, and social functioning. *Arch Intern Med* 1997;157:1113-20.

8. Chocrón Bentata L, Vilalta Franch J, Legazpi Rodríguez I, Auquer K, Franch L, Ramírez R. Diagnóstico de trastornos psiquiátricos por el médico de atención primaria. *Aten Primaria* 1996;18:22-6.

9. Herrán A, Vázquez-Barquero JL, Artal J, García J, Iglesias C, Montejo J, Díez-Manrique JE. El reconocimiento de la enfermedad mental en atención primaria y sus factores determinantes. *Actas Esp Psiquiatr* 1999; 27:87-95.

10. Mulrow CD, Williams JW Jr, Gerety MB, Ramírez G, Montiel O, Kerber C. Case-finding instruments for depression in primary care settings. *Ann Intern Med* 1995;122:913-21.

11. Magruder-Habib K, Zung WW, Feussner JR. Improving physicians' recognition and treatment of depression in general medical care. Results from a randomized clinical trial. *Med Care* 1990;28:239-50.

12. Zung WWK. The role of rating scales in the identification and management of the depressed patient in the primary care setting. *J Clin Psychiatry* 1990;51 Supl:72-6.

13. Zung WWK. A self-rating depression scale. *Arch Gen Psychiatry* 1965;12:63-70.

14. Zung WWK, Broadhead WE, Roth ME. Prevalence of depressive symptoms in primary care. *J Fam Pract* 1993;37:337-44.

15. Conde V, Sánchez de Vega J. La escala autoaplicada para la depresión de Zung. *Arch Neurobiol (Madrid)* 1969;33:335-58.

16. Conde V, Esteban Chamorro T. Revisión crítica de dos adaptaciones castellanas de la «Self-Rating Depression Scale» (SDS) de Zung. *Arch Neurobiol (Madrid)* 1973;36:375-92.

17. Schulberg HC, Magruder KM, deGruy E. Major depression in primary medical care practice. *Gen Hosp Psychiatry* 1996;18:395-406.

18. Glass RM, Vergare MJ. Diagnosis and treatment of psychiatric disorders. A history of progress. *JAMA* 1994;272:1792.
19. Conde López V, Esteban Chamorro T. Validez de la SDS (Self-Rating Depression Scale) de Zung. *Arch Neurobiol (Madrid)* 1975;38:225-46.
20. Ramos Brieva JA. La escala de Zung-Conde para la depresión: su validez predictiva. *Actas Luso Esp Neurol Psiquiatr Cienc Afines* 1986;14:123-7.
21. Ramos Brieva JA, Lafuente López R, Montejo Iglesias ML, Moreno Sarmiento A, Ponce de León Hernández C, Méndez Barroso R, et al. Validez predictiva de la escala de Zung en deprimidos ancianos. *Actas Luso Esp Neurol Psiquiatr Cienc Afines* 1991;19:122-6.
22. First MB, Spitzer RL, Gibbon M, Williams JBW. Entrevista Clínica Estructurada para los Trastornos del Eje I del DSM-IV (versión clínica) SCID-I. Barcelona: Masson; 1999.
23. Renom Pinsach J. Diseño de tests. Barcelona: Idea; 1992.
24. Zhou XH. Effect of verification bias on positive and negative predictive values. *Stat Med* 1994;13:1737-45.
25. Zhou XH. Correcting for verification bias in studies of a diagnostic test's accuracy. *Stat Methods Med Res* 1998;7:337-43.
26. Choi BCK. Sensitivity and specificity of a single diagnostic test in the presence of work-up bias. *J Clin Epidemiol* 1992;45:581-6.
27. Burgueño MJ, García-Bastos JL, González-Buitrago JM. Las curvas ROC en la evaluación de pruebas diagnósticas. *Med Clin (Barc)* 1995;104:661-70.
28. Sackett DL, Straus S. On some clinically useful measures of the accuracy of diagnostic tests. *ACP J Club* 1998;129:A-17.
29. Kessler RC, Little RJA, Groves RM. Advances in strategies for minimizing and adjusting for survey non-response. *Epidemiol Rev* 1995;17:192-204.
30. Rush AJ. Problems associated with the diagnosis of depression. *J Clin Psychiatry* 1990;51 Supl:15-22.
31. Regier DA, Kaelber CT, Rae DS, Farmer ME, Knauper B, Kessler RC, et al. Limitations of diagnostic criteria and assessment instruments for mental disorders. *Arch Gen Psychiatry* 1998;55:109-15.
32. Frances A. Problems in defining clinical significance in epidemiological studies. *Arch Gen Psychiatry* 1998;55:119.
33. Gervás J, Pérez M, García P, Abaira V. Utilización de servicios sanitarios: pacientes y visitas. *Aten Primaria* 1990;7:346-8.
34. Naughton MJ, Wiklund I. A critical review of dimension-specific measures of health-related quality of life in cross-cultural research. *Qual Life Res* 1993;2:397-432.
35. Zung WWK, Magruder-Habib K, Velez R, Alling W. The comorbidity of anxiety and depression in general medical patients: a longitudinal study. *J Clin Psychiatry* 1990;51 Supl:77-80.
36. Spitzer RL, Williams JB, Kroenke K, Linzer M, Verloin deGruy III E, Hahn SR, et al. Utility of a new procedure for diagnosing mental disorders in primary care. The PRIME-1000 Study. *JAMA* 1994;272:1749-56.

Correspondencia:  
 Enric Aragonès Benaiges  
 CAP de Constantí  
 C/ dels Horts, 6  
 43120 Constantí (Tarragona)  
 E-mail: earagones@ptarragona.scs.es