

Margalida Gili^{1,2}
 Emilio Lopez-Navarro^{1,2}
 Clara Homar^{1,2}
 Adoración Castro^{1,2}
 Mauro García-Toro^{1,2}
 Joan Llobera^{2,3}
 Miquel Roca^{1,2}

Propiedades psicométricas de la versión española de la escala QIDS-SR₁₆ en pacientes con trastorno depresivo

¹Institut Universitari d'Investigació en Ciències de la Salut. Universitat de les Illes Balears. Palma de Mallorca. Illes Balears

²Red de Actividades Preventivas y Promoción de la Salud en Atención Primaria (RediAPP), Spain

³Gerencia de Atención Primaria de Mallorca. Servicio Balear de Salud (Ib-Salut). Illes Balears

Introducción. La depresión es un problema de salud mental a nivel mundial, por lo que mejorar el diagnóstico con pruebas fiables y breves es crucial. En este sentido, la escala *Quick Inventory of Depressive Symptomatology-Self Report* (QIDS-SR₁₆) ha sido validada en distintos países encontrando que es un instrumento con un correcto equilibrio entre fiabilidad y tiempo. El objetivo del estudio es evaluar las propiedades psicométricas de la versión española del QIDS-SR₁₆ y proporcionar puntos de corte para valorar la gravedad del trastorno depresivo.

Metodología. Basado en los datos del estudio RESIST que reclutó 1595 pacientes depresivos en 17 comunidades autónomas. Los instrumentos utilizados fueron la Escala Hamilton para la Evaluación de la Depresión de 17 ítems (HDRS₁₇) y la versión española del QIDS-SR₁₆. El análisis estadístico incluyó procedimientos para determinar la fiabilidad test-retest, la consistencia interna, y explorar la naturaleza dimensional del cuestionario, así como el cálculo de la curva ROC para determinar diferentes puntos de corte.

Resultados. El QIDS-SR₁₆ muestra una buena fiabilidad test-retest y una alta consistencia interna ($\alpha=0.871$), así como una curva ROC cuyo valor es 0.946. El análisis factorial exploratorio indica la existencia de un factor que explica el 46.80% de la varianza. La validez convergente y la sensibilidad al cambio han sido adecuadas.

Conclusiones. Los resultados sugieren que el QIDS-SR₁₆ es un instrumento fiable evaluar la gravedad de la sintomatología depresiva en población española. El punto de corte que ofrece un mejor balance entre sensibilidad y especificidad se sitúa en una puntuación total de 7.

Palabras clave: Depresión, Evaluación, QIDS-SR, Validación, Puntos de corte, Propiedades psicométricas

Actas Esp Psiquiatr 2014;42(6):292-9

Psychometric properties of Spanish version of QIDS-SR₁₆ in depressive patients

Introduction. Depression is a worldwide health problem. Thus, making the diagnosis with reliable and short tests is crucial. In this regard, the Quick Inventory of Depressive Symptomatology-Self Report (QIDS-SR₁₆) has been validated in several countries. It was found that this instrument has a correct balance between time and reliability. This study has aimed to assess psychometric properties of QIDS-SR₁₆ Spanish version, and to calculate several cutoffs to evaluate the depressive disorder severity.

Method. The study was based on the data from the RESIST study that recruited 1595 depressive patients from 17 regional communities. Instruments used were Hamilton Depression Rating Scale (HDRS₁₇) and Spanish version of QIDS-SR₁₆. Statistical analyses included test-retest reliability and internal consistency calculation, and exploratory factor analysis. In addition, ROC curve was calculated in order to determine different cutoff values.

Results. QIDS-SR₁₆ shows adequate test-retest reliability and high internal consistency ($\alpha=0.871$), as well as ROC value of 0.946. Exploratory factor analysis showed a one factor model, which accounted for 46.80% of variance. Convergent validity and sensitivity to change were adequate.

Discussion. The results suggest that the QIDS-SR₁₆ is a reliable test to assess depressive symptom severity in the Spanish population. The cutoff that shows the best sensitivity/specificity rate was a total score of 7.

Keywords: Depression, Assessment, QIDS-SR, Validation, Cut-off points, Psychometric properties

Correspondencia:

Margalida Gili Planas

Institut Universitari d'Investigació en Ciències de la Salut (IUNICS)

Universitat de les Illes Balears

Carretera de Valldemossa km 7.5

07071 Palma de Mallorca, España

Tel.: 34 971173081

Fax: 34 971 259935

Correo electrónico: mgili@uib.es

INTRODUCCIÓN

La depresión constituye un importante problema de salud pública. Los datos epidemiológicos publicados en relación a su prevalencia indican que cerca del 17% de personas adultas experimenta un trastorno depresivo mayor a lo largo de su vida y un 7% durante un periodo de 12 meses¹. Según los datos del estudio ESEMeD-España, el trastorno depresivo mayor es el trastorno mental más frecuente en población general, con una prevalencia del 10,6% a lo largo de toda la vida y del 4,0% en el último año². También en España, el Estudio SCREEN ha hallado que el 29,0% de los pacientes que acuden a atención primaria presenta un trastorno depresivo mayor³.

Además, los trastornos depresivos generan una alta discapacidad. La Organización Mundial de la Salud señala que la depresión constituye el cuarto estado de salud que contribuye a la carga global de enfermedad, y se espera que en el año 2030 sea la primera en los países más industrializados⁴. También se asocia con una mayor tasa de mortalidad, altos niveles de utilización de servicios sanitarios, enormes costes económicos y constituye el principal factor de riesgo para el suicidio. Todas estas razones han llevado a considerar la depresión como un importante problema de salud pública a nivel mundial que requiere un mejor diagnóstico y tratamiento⁵.

En este contexto resulta de gran relevancia clínica mejorar la disponibilidad de instrumentos que faciliten la identificación temprana. Se requieren instrumentos válidos, fiables y breves tanto para el screening y diagnóstico de este trastorno, como para la valoración de su gravedad y cambios en su evolución. La determinación de la gravedad de la depresión tiene importantes implicaciones. Una de ellas es la elección del tipo de tratamiento (terapia psicológica, farmacológica o ambas) y así se refleja en numerosas guías⁶⁻¹¹. Por otro lado, la evaluación periódica de la gravedad permite monitorizar la evolución y los cambios en la sintomatología, por lo que constituye una buena medida para valorar la eficacia o efectividad del tratamiento, tanto en la práctica clínica habitual como en el campo de la investigación.

Existen numerosos instrumentos para evaluar la gravedad de los síntomas depresivos. Los más utilizados son la Escala Hamilton para la Depresión (HDRS)¹², la Escala de Depresión Montgomery-Asberg¹³ o el Inventario de Depresión de Beck (Beck et al., 1961)¹⁴. Las versiones breves de estos instrumentos ayudan a disminuir los recursos necesarios para su aplicación y contribuyen a una identificación más rápida de la depresión en diferentes niveles asistenciales. Igualmente, resultan de utilidad para su uso en investigación, puesto que en la mayoría de estudios el tiempo y el cansancio de los participantes resultan de gran importancia. Entre los instrumentos breves más utilizados que evalúan la gravedad de la depresión encontramos la versión de 17 ítems de la Escala Hamilton¹⁵. El *Patient Health Questionnaire* (PHQ-9)¹⁶ ha sido otro instrumento ampliamente utilizado,

pues permite valorar la gravedad de la depresión en solo 9 ítems, que hacen referencia a los 9 síntomas centrales de la depresión propuestos por la clasificación DSM¹⁷.

Un instrumento recientemente utilizado en el estudio americano Star*D es el *Inventory for Depressive Symptomatology* (IDS-30)^{18,19}. Existen dos versiones reducidas de la IDS, tanto en formato heteroaplicado como autoaplicado. Ambas versiones contienen solamente aquellos ítems de la versión completa necesarios para evaluar los 9 criterios nucleares de la depresión propuestos por el DSM-5¹⁷. Tanto la versión aplicada por el clínico, la *Quick Inventory of Depressive Symptomatology-clinician-rated* (QIDS-C₁₆)²⁰ como la versión autoaplicada, la *Quick Inventory of Depressive Symptomatology-Self Report* (QIDS-SR₁₆)²⁰ pueden ser administradas en pocos minutos (5-7) por lo que resultan coste-eficaces²¹. La QIDS-SR₁₆, cuyas propiedades psicométricas se evalúan en este trabajo, incluye 16 ítems en los que el sujeto selecciona la afirmación que mejor describe cómo se ha encontrado en los últimos 7 días, puntuando de 0 a 3 en función de la gravedad. Los ítems incluyen alteraciones del sueño (4 ítems), estado de ánimo triste (1 ítem), cambios en el apetito y peso (4 ítems), concentración y/o toma de decisiones (1 ítem), opinión de uno mismo (1 ítem), ideación suicida (1 ítem), interés (1 ítem), nivel de energía (1 ítem) e inquietud (2 ítems). La puntuación total de la QIDS-SR₁₆ oscila entre 0 y 27 y puntuaciones más altas indican mayor gravedad de la sintomatología depresiva.

Existe considerable evidencia de la validez y fiabilidad del QIDS en pacientes con depresión en diferentes países y lenguas^{20,22-24}. También se han propuesto diferentes puntos de corte respecto a la gravedad del trastorno basados en el instrumento original: sin depresión (0-5), depresión leve (6-10), depresión moderada (11-15), depresión grave (16-20), depresión muy grave (≥ 21)²⁰. No conocemos datos en España ni tampoco antes se habían proporcionado puntos de corte en población española para determinar la gravedad de la sintomatología depresiva. Sería de gran utilidad ofrecer una validación en español de este instrumento y proporcionar puntos de corte de la medida en población española.

El objetivo de este estudio es evaluar las propiedades psicométricas, en términos de validez y fiabilidad, de la versión en español de la escala QIDS-SR₁₆, utilizando la escala HDRS₁₇ como prueba de referencia y, además, proporcionar puntos de corte para determinar la gravedad de la sintomatología depresiva y la evolución de la enfermedad.

METODOLOGÍA

Diseño y muestra

Para este trabajo se han utilizado los datos del estudio RESIST, un estudio epidemiológico nacional, prospectivo

vo con dos evaluaciones, naturalístico y multicéntrico²⁵. Se seleccionó una muestra geográficamente estratificada de 400 psiquiatras proporcionalmente distribuidos según las 17 comunidades autónomas españolas. Cada psiquiatra invitó a participar a 4 o 5 pacientes. Los pacientes debían tener más de 18 años, cumplir criterios DSM-IV para depresión mayor²⁶ y haber firmado el consentimiento informado. Las evaluaciones se llevaron a cabo durante 2 visitas rutinarias. Las evaluaciones se llevaron a cabo después de 6-8 semanas de tratamiento y a las 10±2 semanas después de la primera evaluación. La muestra estuvo compuesta por 1595 pacientes con TDM que acudían a consultas ambulatorias de psiquiatría. De los 1870 pacientes iniciales, se excluyeron 275 por diferentes motivos: cambio de tratamiento (9.1%), pacientes sin segunda evaluación (3.6%) y datos incompletos o perdidos (1.9%).

Este estudio recibió la aprobación del Comité Ético de Investigación Clínica de la Fundación Teknon (y sigue las normas éticas de la Declaración de Helsinki con una revisión del año 2000). Los detalles completos de la metodología utilizada pueden consultarse en Roca et al, 2011.

Instrumentos

Cuaderno de recogida de datos (CRD): cumplimentado por el psiquiatra. Incluía los criterios DSM-IV de trastorno depresivo mayor. También se recogieron los datos sociodemográficos (edad, sexo, estado civil, situación laboral, nivel de estudios y lugar de residencia); características clínicas del TDM (edad de inicio del primer episodio depresivo, duración del episodio, número de episodios previos) comorbilidad psiquiátrica y médica.

Escala Hamilton para la evaluación de la depresión versión de 17 ítems (HDRS-17 Hamilton Depression Rating Scale)¹² es una escala heteroaplicada de 17 ítems, diseñada para ser utilizada en pacientes con diagnóstico de TDM con el objetivo de evaluar la gravedad de la depresión y los cambios del paciente. Se cumplimenta de acuerdo con la información obtenida en la entrevista clínica. Cada ítem tiene entre 3-5 opciones de respuesta en función de la gravedad. Se utilizó esta escala como prueba de referencia. Siguiendo las sugerencias de la literatura previa^{27,28}, el criterio de remisión clínica se situó en una puntuación total igual o inferior a 7.

*Versión española del QIDS-SR16*²⁰: la versión auto-aplicada de 16 ítems que se ha utilizado en este estudio fue extraída de la escala IDS-SR30^{18,19} que puede consultarse en la página web oficial de los cuestionarios IDS/QIDS²⁹. Existen diferentes traducciones al español realizadas en países de América Latina y España. Para este trabajo se seleccionó la versión específica para población española. Este instrumento se utiliza como instrumento de cribaje y para valorar la gravedad de la sintomatología depresiva. Es sensible a cambios debido a la medicación, psicoterapia o tratamientos somáti-

cos, por lo que resulta útil tanto para objetivos clínicos como de investigación.

Análisis estadísticos

En primer lugar se generaron descriptivos de las características sociodemográficas y clínicas. Para el cálculo de la fiabilidad test-retest se procedió a realizar una correlación entre las puntuaciones de cada ítem respecto al inicio y a los 6 meses. Con objeto de estimar la consistencia interna del cuestionario, se utilizó el coeficiente de fiabilidad de Cronbach y adicionalmente se calculó utilizando el método de las dos mitades con la corrección de Spearman-Brown.

Con el fin de estimar los diferentes puntos de corte se procedió al cálculo de la curva ROC. A partir de los datos de ésta, se computaron los valores de sensibilidad, especificidad, valor predictivo positivo y valor predictivo negativo, para cada uno de los puntos de corte propuestos, además del indicado en la versión original inglesa.

Para estimar el número de factores que subyacen al cuestionario, se utilizó un procedimiento de análisis factorial exploratorio, en el que se estableció como criterio de selección un valor propio de cada factor mayor a 1. Por lo que respecta al cálculo de la validez convergente del cuestionario, se optó por realizar una correlación de Pearson entre los ítems y la puntuación total del QIDS-SR con la puntuación total de la prueba de referencia.

Finalmente, para el cálculo de la sensibilidad al cambio se procedió a categorizar a los participantes en dos grupos, remisión clínica y depresión, de acuerdo con las puntuaciones en la prueba de referencia y utilizando el punto de corte propuesto en la literatura²⁸. Posteriormente, se utilizó una prueba t para comparar las medias de las puntuaciones en el QIDS-SR de los dos grupos.

El criterio de significación estadística se estableció en 0.05.

RESULTADOS

En la Tabla 1 se muestran las características clínicas y sociodemográficas de la muestra. Los análisis se llevaron a cabo en 1595 pacientes, de éstos 553 (34.6%) eran hombres y 1042 (65.3%) mujeres, con una media de edad de 47,7 años (rango 18-88). La mayoría de los pacientes estaban casados (61%), trabajaban en el momento de la evaluación (45%), y vivían en un entorno urbano (72%). Adicionalmente, se calcularon los estadísticos descriptivos para cada uno de los ítems en los dos momentos de medición. En la Tabla 2 se ofrece información detallada sobre los mismos.

| Tabla 1 | Datos sociodemográficos y clínicos de la muestra |
|---------------------------------------|--|
| Variables sociodemográficas | n (%) |
| <i>Sexo</i> | |
| Hombre | 553 (34.6) |
| Mujer | 1042 (65.3) |
| <i>Edad</i> | 47.73 (13.14) |
| ≤30 | 155 (9.71) |
| 31-50 | 780 (48.90) |
| 50≥ | 660 (41.37) |
| <i>Situación Laboral</i> | |
| Empleado | 728 (45.0) |
| Estudiante | 41 (2.4) |
| Desempleado | 189 (11.8) |
| Labores del hogar | 383 (24.0) |
| Jubilado | 254 (15.9) |
| <i>Estado Civil</i> | |
| Soltero | 315 (19.7) |
| Casado | 973 (61) |
| Viudo | 111 (7) |
| Separado | 196 (12.3) |
| <i>Nivel educativo</i> | |
| Primaria incompleta | 296 (18.6) |
| Primaria completa | 529 (33.2) |
| Secundaria | 506 (31.7) |
| Universidad | 264 (16.6) |
| <i>Convivencia</i> | |
| Solo | 280 (17.6) |
| Acompañado | 1315 (82.4) |
| <i>Entorno</i> | |
| Rural | 446 (28) |
| Urbano | 1149 (72) |
| Variables Clínicas | Media (Desviación) |
| Edad primer episodio | 40.31 (13.15) |
| Duración episodio | 14.2 (9.4) |
| Numero de episodios previos | 3.72 (2.9) |
| QIDS-SR Primera evaluación | 17.31 (8.20) |
| HDRS ₁₇ Primera evaluación | 17.3 (8.3) |
| <i>QIDS-SR gravedad</i> | n (%) |
| Ninguna (0 - 6) | 200 (12.5) |
| Baja (8 - 13) | 357 (22.4) |

| Tabla 1 | Continuación |
|-----------------------------------|---------------------------|
| Variables clínicas | Media (desviación) |
| Moderada (14 - 19) | 416 (26.1) |
| Grave (20 - 25) | 347 (21.8) |
| Muy grave (>26) | 275 (17.2) |
| <i>HDRS₁₇ gravedad</i> | |
| Ninguna (0 - 13) | 200 (12.5) |
| Baja (14 - 25) | 286 (17.9) |
| Moderada (26 - 38) | 352 (22.1) |
| Grave (39 - 48) | 626 (39.2) |
| Muy grave (>49) | 131 (8.2) |

Estabilidad temporal

La correlación entre cada ítem en la primera y segunda evaluación se muestra en la Tabla 2. El ítem 4 muestra un alto valor de los estadísticos de asimetría y curtosis al ser mayores a 2.00 y 7.00 respectivamente, lo que es indicador de que el ajuste a la normalidad paramétrica no es adecuado. Es precisamente este ítem, el correspondiente a horas de sueño, el único cuya correlación respecto al retest es inferior a 0.70, sin embargo ésta es significativa a un nivel de confianza de 99%. El resto de ítems muestran unos valores de curtosis y asimetría dentro de los límites de la normalidad, así como una alta fiabilidad test-retest, todas las correlaciones son superiores a 0.70 y significativas a un nivel de confianza de 0.99.

Consistencia interna

La consistencia interna global de la escala, evaluada mediante el estadístico α de Cronbach, es de 0.871. En el análisis ítem por ítem del valor de α la escala se comporta de forma homogénea y no aparecen ítems irrelevantes que perjudiquen al α global de la QIDS. La Tabla 3 muestra el valor del estadístico α para cada ítem en caso de que éste fuera eliminado así como la media y varianza explicada para cada ítem. Adicionalmente, se procedió a comprobar la consistencia interna mediante el método de las dos mitades con la corrección de Spearman-Brown, obteniéndose un valor de 0.850.

En conjunto, tanto el α de Cronbach como la corrección de Spearman-Brown indican una correcta consistencia interna de la escala.

Puntos de corte

Para poder calcular la sensibilidad y especificidad del cuestionario para distintos puntos de corte calculamos la

Tabla 2 Test-retest. Correlación de los ítems al inicio y a los 6 meses

| Ítem QIDS-SR | Primer momento | | | Segundo momento | | | Correlación | Valor p |
|--------------|----------------|-----------|----------|-----------------|-----------|----------|-------------|---------|
| | Media (D) | Asimetría | Curtosis | Media (D) | Asimetría | Curtosis | | |
| Ítem 1 | 1.41 (0.99) | 0.11 | -1.03 | 2.09 (1.28) | -0.32 | -1.02 | 0.78 | ≤0.001 |
| Ítem 2 | 1.44 (1.08) | 0.13 | -1.25 | 1.97 (1.32) | -0.22 | -1.18 | 0.75 | ≤0.001 |
| Ítem 3 | 1.14 (1.15) | 0.45 | -1.28 | 1.78 (1.37) | 0.01 | -1.33 | 0.76 | ≤0.001 |
| Ítem 4 | 0.22 (0.57) | 2.99 | 9.27 | 0.20 (0.46) | 2.56 | 7.55 | 0.34 | ≤0.001 |
| Ítem 5 | 1.92 (0.91) | -0.33 | -0.88 | 2.62 (1.14) | -0.69 | -0.27 | 0.81 | ≤0.001 |
| Ítem 6 y 7 | 1.03 (0.87) | 0.52 | -0.40 | 1.57 (1.27) | 0.19 | -1.16 | 0.75 | ≤0.001 |
| Ítem 8 y 9 | 1.07 (1.07) | 0.57 | -0.99 | 1.26 (1.27) | 0.55 | -1.00 | 0.76 | ≤0.001 |
| Ítem 10 | 1.65 (0.87) | 0.07 | -0.81 | 2.40 (1.13) | -0.44 | -0.57 | 0.74 | ≤0.001 |
| Ítem 11 | 1.33 (0.88) | 0.24 | -0.64 | 2.29 (1.17) | -0.39 | -0.69 | 0.72 | ≤0.001 |
| Ítem 12 | 0.77 (0.77) | 0.77 | 0.15 | 1.53 (1.38) | 0.23 | -1.38 | 0.75 | ≤0.001 |
| Ítem 13 | 1.69 (0.97) | 0.01 | -1.12 | 2.29 (1.17) | -0.39 | -0.69 | 0.73 | ≤0.001 |
| Ítem 14 | 1.59 (0.84) | 0.10 | -0.67 | 2.42 (1.15) | -0.44 | -0.63 | 0.71 | ≤0.001 |
| Ítem 15 | 1.13 (0.86) | 0.41 | -0.47 | 1.80 (1.22) | -0.12 | -1.10 | 0.73 | ≤0.001 |
| Ítem 16 | 0.91 (0.78) | 0.69 | 0.24 | 1.56 (1.22) | 0.12 | -1.13 | 0.72 | ≤0.001 |

curva ROC. Tomando como punto de corte el valor original propuesto, esto es, una puntuación total igual o superior a 6, el resultado del área por debajo de la curva ROC es 0.946. Para este mismo punto de corte el valor de sensibilidad diagnóstica se ubica en un 98% mientras que la especificidad alcanza el 63%. En cuanto al valor predictivo de la prueba, tomando una puntuación igual o superior a 6 como criterio de depresión, el valor predictivo positivo es igual a 0.95 mientras que el negativo es 0.86.

Tomando los distintos valores arrojados por la curva ROC, se calcularon los resultados de sensibilidad, especificidad, valor predictivo positivo, y valor predictivo negativo para distintos puntos de corte. En la Tabla 4 se muestra información detallada sobre los resultados de estos cálculos.

Análisis factorial exploratorio

El procedimiento exploratorio de factores indicó un adecuado valor del test de Kaiser, Meyer y Olkin para la muestra, (KMO=0.894). Por su parte, la prueba de esfericidad de Bartlett arrojó un resultado estadísticamente significativo ($\chi^2=11691,39$; $p\leq 0.001$), lo que unido al resultado de la prueba KMO lleva a concluir que la distribución de los datos es

adecuada para un procedimiento de análisis factorial. El análisis factorial indica la existencia de un único factor que explica el 46.80% de la varianza, y cuyo valor propio se sitúa en 6.084.

Validez Convergente

La correlación de los ítems de la escala QIDS-SR muestra que todos ellos se relacionan significativamente con el HDRS₁₇, salvo el ítem número 4 que corresponde a la cantidad de horas de sueño. La relación de este ítem con la prueba de referencia es igual a -0.016 aunque no es una relación estadísticamente significativa ($p=0.535$). En cuanto a la puntuación total del cuestionario QIDS-SR, ésta se relaciona positivamente con la puntuación total del HDRS₁₇ ($r_{xy}=0.845$), alcanzando esta relación la significación estadística ($p\leq 0.001$). Finalmente, la correlación entre la puntuación en HDRS₁₇ y los factores extraídos mediante análisis factorial, evidencia que el factor de mayor valor propio muestra una alta relación con la prueba de referencia de un modo significativo ($r_{xy}=0.852$; $p\leq 0.001$). La Tabla 5 muestra información detallada sobre la relación entre los ítems de la QIDS-SR y el HDRS₁₇, así como de la puntuación total de QIDS y el factor extraído de ésta.

| Tabla 3 | Consistencia interna de la escala QIDS-SR | | |
|--|---|------------|----------------|
| Ítem | Media† | Varianza † | α de Cronbach† |
| Ítem 1: Conciliación de sueño | 15.89 | 58.61 | 0.864 |
| Ítem 2: Sueño durante la noche | 15.87 | 56.81 | 0.861 |
| Ítem 3: Despertar temprano | 16.16 | 56.02 | 0.861 |
| Ítem 4: Horas de sueño | 17.08 | 66.83 | 0.879 |
| Ítem 5: Tristeza | 15.38 | 56.51 | 0.854 |
| Ítem 6 y 7: Alteración del apetito | 16.27 | 58.62 | 0.881 |
| Ítem 8 y 9: Alteración del peso | 16.23 | 57.56 | 0.885 |
| Ítem 10: Concentración/Toma de decisiones | 15.65 | 57.34 | 0.855 |
| Ítem 11: Visión de sí mismo | 15.98 | 57.73 | 0.857 |
| Ítem 12: Pensamientos de muerte o suicidio | 16.53 | 59.58 | 0.860 |
| Ítem 13: Interés general | 15.61 | 55.69 | 0.852 |
| Ítem 14: Nivel de energía | 15.71 | 57.36 | 0.855 |
| Ítem 15: Sensación de lentitud | 16.18 | 57.77 | 0.857 |
| Ítem 16: Sentimientos de inquietud | 16.40 | 61.92 | 0.869 |
| Total QIDS-SR | 17.31 | 67.22 | 0.871 |

†: Valor de tendencia central, dispersión y consistencia interna si se elimina el ítem

| Tabla 4 | Propiedades diagnósticas para distintos puntos de corte | | | |
|----------------|---|-------------------|---------|---------|
| Punto de Corte | Sensibilidad (%) | Especificidad (%) | VPP (%) | VPN (%) |
| 6 | 98 | 63 | 95 | 86 |
| 7 | 97 | 75 | 96 | 77 |
| 8 | 94 | 79 | 97 | 65 |
| 9 | 91 | 84 | 97 | 56 |

| Tabla 5 | Validez convergente - Correlación QIDS-SR con HDRS ₁₇ | |
|--|--|---------|
| | Correlación | Valor p |
| Ítem 1: Conciliación de sueño | 0.523 | ≤0.001 |
| Ítem 2: Sueño durante la noche | 0.570 | ≤0.001 |
| Ítem 3: Despertar temprano | 0.586 | ≤0.001 |
| Ítem 4: Horas de sueño | -0.016 | 0.535 |
| Ítem 5: Tristeza | 0.662 | ≤0.001 |
| Ítem 6 y 7: Alteración del apetito | 0.528 | ≤0.001 |
| Ítem 8 y 9: Alteración del peso | 0.472 | ≤0.001 |
| Ítem 10: Concentración/Toma de decisiones | 0.614 | ≤0.001 |
| Ítem 11: Visión de sí mismo | 0.588 | ≤0.001 |
| Ítem 12: Pensamientos de muerte o suicidio | 0.630 | ≤0.001 |
| Ítem 13: Interés general | 0.635 | ≤0.001 |
| Ítem 14: Nivel de energía | 0.631 | ≤0.001 |
| Ítem 15: Sensación de lentitud | 0.620 | ≤0.001 |
| Ítem 16: Sentimientos de inquietud | 0.431 | ≤0.001 |
| Total QIDS-SR | 0.845 | ≤0.001 |
| Factor 1 | 0.852 | ≤0.001 |

Sensibilidad al cambio

Para establecer la capacidad de la escala para detectar cambios en la sintomatología depresiva, se realizó una comparación de medias para muestras relacionadas sobre la puntuación total de la QIDS-SR. La muestra se dividió en pacientes deprimidos (n=703) y pacientes remitidos (n=892) en función de la puntuación en HDRS₁₇. Tal y como se ha establecido para los criterios de remisión, si la puntuación total en la prueba de referencia era inferior a 7 se consideró que el paciente se encontraba en remisión clínica. El resultado de la comparación de medias indica que existen diferencias significativas entre ambos grupos (t=33.29; p≤0.001), siendo

el grupo de pacientes remitidos el que obtenía de media una puntuación inferior (\bar{X} =5.78; σ =3.88) respecto al grupo de pacientes deprimidos (\bar{X} =13.22; σ =5.04).

CONCLUSIONES

La conclusión principal de nuestro estudio es que la versión en español de la escala QIDS-SR presenta adecuadas propiedades psicométricas y, por tanto, es un instrumento útil para evaluar la gravedad de la sintomatología depresiva en población española.

La consistencia interna global de la versión en español de QIDS-SR fue adecuada ($\alpha=0.871$), y el análisis ítem por ítem indica que no existen elementos que alteren la consistencia general de la escala, a pesar de la asimetría y curtosis del ítem 4. Adicionalmente, el procedimiento de test-retest corroboró estos datos acerca de la consistencia interna de la escala. Los resultados encontrados son convergentes con los hallazgos de estudios previos ($\alpha=0.86$)¹¹; ($\alpha=0.769$)³⁰, ($\alpha=0.86$)²²; ($\alpha=0.81$)²⁰. En cuanto a la estabilidad temporal evaluada mediante el procedimiento de test-retest, los ítems mantienen una alta asociación entre ambos momentos temporales lo que sugiere que la estabilidad temporal es buena. El único ítem que no ha mostrado una alta asociación, a pesar de ser ésta significativa, es el que corresponde a las horas de sueño, que se caracterizó también por una asimetría elevada. La estabilidad temporal coincide con los datos publicados previamente¹¹.

Nuestros resultados sugieren la existencia de un único factor que explica casi el 46.8% de la varianza observada. Sin embargo este valor está por debajo del obtenido por otros autores^{11,24,28}. La validez convergente de la escala QIDS-SR respecto a la prueba de referencia es elevada tanto para cada ítem por separado como para la puntuación global de la escala. El único elemento que no se relaciona en la misma dirección es de nuevo el correspondiente a las horas de sueño. En cuanto al factor extraído, también mantiene una relación significativa y en la misma dirección que las puntuaciones del HDRS₁₇. Este resultado coincide con el estudio realizado por Rush et al., 2003, ($c=0.81$)²⁰ en el que también se encontró una estrecha correlación entre ambos instrumentos.

Respecto a las propiedades diagnósticas el área bajo la curva ofrece un elevado valor que permite calificar a la prueba con una muy buena precisión diagnóstica. En cuanto al punto de corte original propuesto de 6²⁰ éste obtiene un buen valor de sensibilidad pero la especificidad de la prueba se resiente resultando en una proporción de sujetos sin depresión identificados correctamente igual al 63%. Teniendo en cuenta los datos de nuestro estudio, el punto de corte óptimo se sitúa en 7, pues mantiene una sensibilidad solamente un 1% inferior al punto de corte original y eleva la especificidad hasta el 75%. Los valores predictivos asociados a este punto de corte indican una correcta probabilidad de asociación entre la presencia o ausencia de diagnóstico y la existencia de cuadro depresivo. Adicionalmente, hemos considerado los datos de otros puntos de corte, sin embargo, ninguno de ellos ofrece un mejor equilibrio entre índices diagnósticos que el punto de corte igual a 7. Este dato no coincide con otras validaciones realizadas previamente. Es el caso de la validación turca del instrumento, que sitúa el punto de corte en 9³⁰ o la validación china, que lo sitúa en 5²⁴. Finalmente, los resultados del análisis de sensibilidad al cambio sugieren que el QIDS-SR puede ser un instrumento adecuado para detectar si el trastorno depresivo remite, coincidiendo con Rush et al., 2003²⁰.

En conclusión, los hallazgos obtenidos en este estudio indican que la versión en español del instrumento QIDS-SR presenta adecuadas propiedades psicométricas, y puntos de corte válidos para determinar la gravedad de la sintomatología depresiva y la evolución de la enfermedad, comparado con la versión original del instrumento²⁰ y otras validaciones en otros países. El QIDS-SR es un instrumento idóneo para el uso clínico y de investigación con satisfactorias propiedades en población española.

BIBLIOGRAFÍA

1. Kessler D, Sharp D, Lewis G. Screening for depression in primary care. *Br J Gen Pract.* 2005;55(518):659-60.
2. Haro JM, Palacín C, Vilagut G, Martínez M, Bernal M, Luque I, et al. Prevalence of mental disorders and associated factors: Results from the ESEMeD-Spain study. *Med Clin.* 2006;126(12):445-51.
3. Roca M, Gili M, García-García M, Salva J, Vives M, García-Campayo J, et al. Prevalence and comorbidity of common mental disorders in primary care. *J Affect Disord.* 2009;119:52-8.
4. Mathers CD, Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. *PLoS Med.* 2006;3(11):e442.
5. Lopez AD, Mathers CD, Ezzati M, Jamison DT, Murray CJL. Global and regional burden of disease and risk factors, 2001: systematic analysis of population health data. *Lancet.* 2006;367:1747-57.
6. NHS National Institute for Clinical Excellence. Depression: the treatment and management in adults; 2009 [update].
7. Anderson IM, Baldwin RC, Cowen PJ, Howard L, Lewis G, Matthews K, et al. Evidence-based guidelines for treating depressive disorders with antidepressants: a revision of the 2000 British Association for Psychopharmacology guidelines. *J Psychopharmacol.* 2008;22(4):343-96.
8. Bauer M, Bschor T, Pfenning A, Whybrow PC, Angst J, Versiani M, et al. World Federation of Societies of Biological Psychiatry (WFSBP) guidelines for biological treatment of unipolar depressive disorders in primary care. *World J Biol Psychiatry.* 2007;8(2):67-104.
9. New Zealand Guidelines Group. Identification of common mental disorders and management of depression in primary care. Ministry of Health, 2008.
10. Patten SB, Kennedy SH, Lam RW, O'Donovan C, Filteau MJ, Parikh SV, et al. Canadian Network for Mood and Anxiety Treatments (CANMAT) clinical guidelines for the management of major depressive disorder in adults. I. Classification, burden and principles of management. *J Affect Disord.* 2009;117(Suppl.1):S5-14.
11. Cameron IM, Crawford JR, Cardy AH, du Toit SW, Lawton K, Hay S, et al. Psychometric properties of the Quick Inventory of Depressive Symptomatology (QIDS-SR) in UK primary care. *J Psychiatr Res.* 2013;47:592-8.
12. Hamilton MA. Hamilton rating scale for depression. *J Neurol Neurosurg Psychiatry.* 1960;23:56-62.
13. Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. *Br J Psychiatry.* 1979;134:382-9.
14. Beck AT, Ward C, Mendelson M, Mock J, Erbaugh J. An inventory for measuring depression. *Arch Gen Psychiatry.* 1961;4:561-71.
15. Hamilton M. Development of a rating scale for primary depressive illness. *Br J Soc Clin Psychol.* 1967;6(4):278-96.
16. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med.* 2001;16(9):606-13.
17. American Psychiatric Association. Diagnostic and statistical

- manual of mental disorders (5th ed.). Arlington, VA: American Psychiatric Publishing, 2013.
18. Rush AJ, Gilles DE, Schlessler MA, Fulton CL, Weissenburger J, Burns C. The Inventory for Depressive Symptomatology (IDS): preliminary findings. *Psychiatry Res.* 1986;18:65-87.
 19. Rush AJ, Guillion CM, Basco MR, Jarret RB, Trivedi MH. The Inventory of Depressive Symptomatology (IDS): psychometric properties. *Psychol Med.* 1996;26:477-86.
 20. Rush AJ, Trivedi MH, Ibrahim HI, Carmody TJ, Arnow B, Klein DN, et al. The 16-item Quick Inventory of Depressive Symptomatology (QIDS) Clinician Rating (QIDS-C) and Self-Report (QIDS-SR): psychometric evaluation in patients with chronic major depression. *Biol Psychiatry.* 2003;54:573-83.
 21. Bernstein IH, Rush AJ, Suppes T, Trivedi MH, Woo A, Kyutoku Y, et al. A psychometric evaluation of the clinician-rated Quick Inventory of Depressive Symptomatology (QIDS-C16) in patients with bipolar disorder. *Int J Methods Psychiatr Res.* 2009;18(2):138-46.
 22. Trivedi MH, Rush AJ, Ibrahim HM, Carmody TJ, Biggs MM, Suppes T, et al. The Inventory of Depressive Symptomatology, Clinician Rating (IDS-C) and Self-Report (IDS-SR), and the Quick Inventory of Depressive Symptomatology, Clinician Rating (QIDS-C) and Self-Report (QIDS-SR) in public sector patients with mood disorders: a psychometric evaluation. *Psychol Med.* 2004;34(1):73-82.
 23. Carmody TJ, Rush AJ, Bernstein IH, Brannan S, Husain MM, Trivedi MH. Making clinicians lives easier: guidance on use of the QIDS self-report in place of the MADRS. *J Affect Disord.* 2006;95(1-3):115-8.
 24. Liu J, Xiang YT, Lei H, Wang Q, Wang G, Ungvari GS, et al. Guidance on the conversion of the Chinese versions of the Quick Inventory of Depressive Symptomatology-Self-Report (C-QIDS-SR) and the Montgomery-Asberg Scale (C-MADRS) in Chinese patients with major depression. *J Affect Disord.* 2014;152-154:530-3.
 25. Roca M, García-Toro M, García-Campayo J, Vives M, Armengol S, García-García M, et al. Clinical differences between early and late remission in depressive patients. *J Affect Disord.* 2011;134:235-41.
 26. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders, Text Revised (4th ed.)*. Washington, DC: American Psychiatric Publishing, 2000.
 27. Frank E, Prien RF, Jarrett JB, Keller MB, Kupfer DJ, Lavori P, et al. Conceptualization and rationale for consensus definitions of terms in major depressive disorder: response, remission, recovery, relapse, and recurrence. *Arch Gen Psychiatry.* 1991;48:851-5.
 28. Rush AJ, Kraemer HC, Sackeim HA, Fava M, Trivedi MH, Frank E, et al. Report by the ACNP Task Force on response and remission in major depressive disorder. *Neuropsychopharmacology.* 2006;31(9):1841-53.
 29. IDS/QIDS [update 2014; citation July 9 2014]. Available at: <http://www.ids-qids.org/>
 30. Mergen H, Tavli T, Öngel K. Validity and reliability of the QIDS-SR₁₆-Turkish in comparison with BDI-II-Turkish among young outpatient sample of Turkish university students. *Archives of Neuropsychiatry.* 2012;49:1-5.