

# Estudios psicométricos de los instrumentos en español para la detección de trastornos mentales comunes: una revisión.

**Psychometric studies of instruments in Spanish for the detection of common mental disorders: A review.**

Johann M. Vega-Dienstmaier <sup>1,a</sup>

## RESUMEN

El objetivo de esta revisión es describir las características psicométricas de las versiones en español de instrumentos utilizados en la detección de trastornos mentales comunes, sobre la base de una búsqueda sistemática en PubMed y Google Académico. Se identificaron los siguientes instrumentos con información pertinente: General Health Questionnaire (GHQ-12 y GHQ-28), Hospital Anxiety and Depression Scale (HADS), Escala de Salud Personal (ESP), Self-Reporting Questionnaire (SRQ-20), Escala de Malestar Psicológico de Kessler (K-6 y K-10), Patient Health Questionnaire (PHQ), Escala de Edimburgo (EPDS), Dep5, Anx5 y Symptom Checklist (SCL-90 y SCL-90-R). Los instrumentos más estudiados fueron el GHQ y la HADS. La mayoría de estudios fueron realizados en España. En Perú se evaluaron el SRQ-20, la ESP, la K-6, la K-10 y el GHQ-28. Las propiedades psicométricas descritas incluyen: consistencia interna, estructura factorial, área bajo la curva Receiver Operating Characteristic (abcROC), sensibilidad, especificidad e índice de Youden (J). El J de los instrumentos estudiados estuvo entre 0,2389 y 0,767 y el abcROC, entre 0,758 y 0,933. Esta revisión puede ser útil en la elección de instrumentos psicométricos en español para la realización de estudios y la práctica clínica, así como en la identificación de necesidades de investigación sistemática en torno al tema.

**PALABRAS CLAVE:** Psicometría, reproducibilidad de los resultados, tamizaje, escalas de valoración psiquiátrica, depresión, ansiedad.

## SUMMARY

The objective of this review is to describe the psychometric characteristics of the Spanish versions of instruments for the detection of common mental disorders, on the basis of a systematic search carried out in PubMed and Google Scholar. The following instruments with pertinent psychometric information were identified: General Health Questionnaire (GHQ-12 and GHQ-28), Hospital Anxiety and Depression Scale (HADS), Personal Health Scale (PHS), Self-Reporting Questionnaire (SRQ -20), Kessler Psychological Distress Scale (K-6 and K-10), Patient Health Questionnaire (PHQ), Edinburgh Scale (EPDS), Dep5, Anx5 and Symptom Checklist (SCL-90 and SCL-90-R). The most studied instruments were GHQ and HADS. Most of the studies were carried out in Spain. In Peru, SRQ-20, PHS, K-6, K-10 and GHQ-28 were evaluated. The psychometric properties described include internal consistency, factorial structure, area under the Receiver Operating Characteristic curve (AUC), sensitivity, specificity and Youden index (J). J of the instruments were between 0.2389 and 0.767, and the AUC, between 0.758 and 0.933. This review could help to choose psychometric instruments in Spanish for studies and clinical practice, as well as to systematically identify research needs about this topic.

<sup>1</sup> Facultad de Medicina Alberto Hurtado, Universidad Peruana Cayetano Heredia. Lima, Perú.

<sup>a</sup> Médico-psiquiatra. Maestro en epidemiología clínica. ORCID ID: 0000-0002-5686-4014

**KEYWORDS:** Psychometrics, reproducibility of results, screening, psychiatric status rating scales, depression, anxiety.

## INTRODUCCIÓN

Los trastornos mentales comunes (TMC) se refieren a los desórdenes depresivos (depresión mayor y distimia) y de ansiedad (trastorno de ansiedad generalizada, trastorno de angustia, fobias, trastorno de ansiedad social, trastorno obsesivo-compulsivo y trastorno de estrés postraumático), los cuales son problemas de salud mental muy prevalentes en la población (1). Tomando en cuenta la CIE-10, los TMC están incluidos dentro de los “trastornos del ánimo” y de los “trastornos neuróticos, relacionados con el estrés y somatomorfos” (2).

Los diversos TMC tienen una estrecha relación entre ellos, evidenciada en una falta de límites definidos entre los síndromes que los caracterizan, una superposición de sus síntomas y una alta comorbilidad (3). Todo ello sustenta la idea de usar un mismo instrumento para la detección de los TMC en su conjunto, en vez de usar diferentes herramientas psicométricas para cada trastorno por separado.

Se han desarrollado diversos instrumentos en el mundo para la identificación de TMC tales como el SRQ-20, el GHQ-12 y la HADS; sin embargo, para poderlos utilizar adecuadamente es importante saber cuáles están validados en el idioma y la población en los que se pretenden aplicar (4).

El objetivo de esta revisión es identificar los instrumentos para detección de TMC en su conjunto que cuentan con versiones en español y describir sus propiedades psicométricas; esta información es importante para poder escoger un instrumento adecuado para realizar investigación o para utilizarse en atención primaria.

## METODOLOGÍA

Se realizaron búsquedas bibliográficas en PubMed y Google Académico en 2 fases:

Primero, para identificar los instrumentos que se utilizan para la detección de trastornos mentales comunes sin restricción de idioma o país, se consideraron las siguientes palabras claves: a) relacionadas con los estudios psicométricos: dentro de ellas tenemos: “valid\*”, “psychometr\*”, “psicometr\*”, “sensitivity”,

“sensibilidad”, “specificity”, “especificidad”, “questionnaire”, “cuestionario”, “scale”, “escala” y “screen\*”; y b) relacionadas con TMC: “common mental disorders”, “depres\*”, “anxi\*”, “ansi\*”.

Para que un instrumento sea considerado en esta revisión, debía estar dirigido a evaluar al conjunto de TMC en general (trastornos ansiosos o depresivos), no solamente a uno de ellos en particular. Así, por ejemplo, una escala dirigida sólo a identificar trastornos depresivos pero no de ansiedad quedaría excluida.

Se consideraron como TMC los siguientes: depresión mayor, distimia, cualquier otro trastorno depresivo, trastorno de ansiedad generalizada, trastorno de ansiedad social, trastorno de pánico, trastorno de estrés postraumático y cualquier otro trastorno de ansiedad.

En una fase posterior, se utilizaron como palabras de búsqueda: a) los nombres de los instrumentos psicométricos identificados en la primera fase; b) las palabras relacionadas a estudios psicométricos; y c) los nombres de los países en donde se habla español o alusión al idioma español; la instrucción específica fue:

(“valid\*” OR “psychometri\*” OR “develop\*” OR “translation” OR “cultural adaptation” OR “reliability”) AND (“screen\*”[ti] OR “scale”[ti] OR “questionnaire”[ti] OR “instrument”[ti] OR “tool”[ti] OR “test”[ti] OR “checklist”[ti] OR “assessment”[ti] OR “measur\*”[ti]) AND (GHQ\* OR SRQ\* OR K-6 OR K-10 OR “Kessler Psychological Distress Scale” OR SSQ OR EPDS OR MCS-12 OR “Escala de Salud Personal” OR “Personal Health Scale” OR “Zung’s Self-rated Anxiety Scale”) AND (Argentina OR Bolivia OR Chile OR Colombia OR “Costa Rica” OR Cuba OR Ecuador OR “El Salvador” OR España OR Spain OR Guatemala OR “Guinea Ecuatorial” OR Honduras OR México OR Mexico OR Nicaragua OR Panamá OR Panama OR Uruguay OR Perú OR Peru OR “Puerto Rico” OR “República Dominicana” OR Uruguay OR Venezuela OR Spanish).

Se incluyeron los estudios publicados hasta octubre de 2021 sin tomar en cuenta su antigüedad, o los criterios relacionados con la calidad de la investigación

(tamaño o selección apropiada de la muestra, etc.). Las posibles deficiencias metodológicas de los estudios incluidos fueron comentadas en la discusión.

El autor se encargó de la búsqueda y de la selección de los estudios en base al título y contenido de los resúmenes.

Se prestó atención a las estructuras factoriales y a los valores de alfa de Cronbach, sensibilidad (s), especificidad (e) y área bajo la curva Receiver Operating Characteristic (abcROC) de los instrumentos descritos en los estudios psicométricos encontrados.

Para uniformizar los hallazgos de la eficacia del instrumento en la detección de trastornos mentales, si el estudio no indicaba el índice de Youden (J), éste fue calculado a partir de la s y e reportadas:  $J = s + e - 1$  (5).

### INSTRUMENTOS PARA DETECTAR TRASTORNOS MENTALES COMUNES

Se identificaron los siguientes instrumentos utilizados para la detección de trastornos mentales comunes: 1) el General Health Questionnaire (GHQ), en sus versiones de 28 ítems o GHQ-28 (4), 12 ítems o GHQ-12 (4,6,7) y 5 ítems o GHQ-5 (4); 2) la Hospital Anxiety and Depression Scale o HADS (4); 3) el Self-Reporting Questionnaire de 20 ítems o SRQ-20 (4,8); 4) la Escala de Malestar Psicológico de Kessler o Kessler Psychological Distress Scale en sus versiones de 10 ítems o K-10 (4) y 6 ítems o K-6 (4,6); 5) la Personal

Health Scale, PHS, Escala de Salud Personal o ESP (9,10); 6) la Escala de Edimburgo o EPDS (4,7); 7) la Dep5 y Anx5 (11); 8) el Patient Health Questionnaire (PHQ) (12); 9) la Symptom Checklist-90 (SCL-90 y SCL-90-R) (13,14); 10) la Escala de Ansiedad de Zung o Zung's Self-rated Anxiety Scale (ZSAS) (7); 11) el Shona Symptom Questionnaire o SSQ (4,15); 12) la Mental Health Component Scale del 12-item Short-Form Health Survey (MCS-12) (16); y 13) el New Case-Finding Instrument for Common Mental Disorders (CMDQ) (17).

### INSTRUMENTOS PARA DETECCIÓN DE TRASTORNOS MENTALES COMUNES CON VERSIONES EN ESPAÑOL ESTUDIADOS PSICOMÉTRICAMENTE

La estrategia de búsqueda permitió encontrar 606 artículos, de los cuales la gran mayoría se excluyó principalmente porque estudiaban instrumentos que no estaban dirigidos a la detección de TMC o que se enfocaban en sólo uno de ellos en particular (generalmente depresión) sin involucrar al resto.

En la tabla 1 se describen los estudios psicométricos de las versiones en español de instrumentos para detección de TMC. Como podemos observar, los estudios psicométricos de las versiones en español han sido llevados a cabo principalmente en España y las escalas más investigadas fueron el GHQ y la HADS. A continuación, revisaremos la información psicométrica disponible para cada uno de los instrumentos:

**Tabla 1.** Descripción de los estudios psicométricos de versiones en español de instrumentos para la detección de trastornos mentales comunes.

Estudio	País	Individuos estudiados	Instrumento	Hallazgos
Goldberg et al. 2017 (47)	México	408 pacientes de atención primaria	Dep5 y Anx5	Dep5: Detección de depresión sin ansiedad o depresión ansiosa: abcROC=0,79; PC=3: s=76%, e=68%, J=0,44. Anx5: Detección de ansiedad o depresión ansiosa: abcROC=0,78; PC=3: s=83%, e=66%, J=0,49.
Jatuff et al, 2010 (44)	Argentina y Venezuela	227 individuos entre pacientes psiquiátricos y controles de ambos países	ESP	Alfa=0,77 en ambos países. Estructura: 1 factor. PC=6/7. Argentina: s=62%, e=68%, J=0,30. Venezuela: s=39%, e=92%, J=0,31

(página siguiente)

**Tabla 1 (cont).** Descripción de los estudios psicométricos de versiones en español de instrumentos para la detección de trastornos mentales comunes.

Estudio	País	Individuos estudiados	Instrumento	Hallazgos
Zapata-Vega et al, 2002 (9)	Perú	50 pacientes psiquiátricos y 50 trabajadores de un hospital	ESP	PC=8/9: s=74%, e=82%, J=0,56. Confiabilidad test-retest: r=0,86.
Molina et al, 2006 (21)	España	Atención primaria	GHQ	Construcción de versiones españolas a partir del GHQ-60. Se proponen: 2 versiones españolas: a) GHQ-30, 6 subescalas: 1) enfermedad general, 2) pérdida de la autoestima, 3) disfunción social, 4) depresión severa, 5) disturbios del sueño, 6) pérdida del interés; y b) GHQ-28, 4 subescalas: 1) depresión severa, 2) disturbios del sueño, 3) síntomas somáticos, 4) disfunción social.
Burrone et al, 2015 (48)	Argentina	854 pacientes de atención primaria	GHQ-12	Alfa=0,80. Estructura: 2 factores.
Villa G. et al, 2013 (29)	Colombia	85 pacientes internados	GHQ-12	Estructura: 2 factores: “bienestar psicológico” y “malestar psicológico”. Alfa=0,84. Se descartó el ítem 11 por no presentar adecuada carga factorial.
Ruiz et al, 2017 (26)	Colombia	1641 participantes de población general y clínica	GHQ-12	Identificación de trastornos emocionales: PC=11/12: s= 82%,e= 63%, J=0,45. Alfa=0,90. Estructura: 1 factor.
Romppel et al, 2017 (28)	Colombia y Alemania	Población general	GHQ-12	Estructura: 1 factor. No equivalencia entre la versión de Alemania y la de Colombia
Sánchez-López & Dresch, 2008 (25)	España	Población general, n=1001 sujetos de 25 a 65 años	GHQ-12	Alfa=0,76. Estructura: 3 factores
Padrón et al, 2012 (31)	España	Adolescentes, n=4146	GHQ-12	Alfa=0,82. Estructura: 3 factores: 1) ansiedad / depresión, 2) anhedonia / disfunción social, 3) pérdida de confianza y autoestima.
López-Castedo & Fernández, 2005 (30)	España	Adolescentes, n=1930	GHQ-12	Alfa=0,85. Confiabilidad test-retest satisfactoria (r=0,79). Estructura: 2 factores: ansiedad y disfunción social.
Aguado et al, 2012 (27)	España	1453 mujeres postparto (6 semanas)	GHQ-12	Estructura: 1 factor.
Navarro et al, 2007 (32)	España	Mujeres postparto	GHQ-12 y EPDS	EPDS: abcROC=0,933; PC=9/10: s=85,5%, e=85,3%, J=0,708. GHQ-12: abcROC=0,904; PC=4/5: s=80,6%, e=80,4%, J=0,61. EPDS mejores resultados que GHQ-12

(página siguiente)

**Tabla 1 (cont).** Descripción de los estudios psicométricos de versiones en español de instrumentos para la detección de trastornos mentales comunes.

Estudio	País	Individuos estudiados	Instrumento	Hallazgos
Delgado-Gomez et al, 2013 (19)	España	Población afectada por terrorismo, n=891	GHQ-28	Alfa=0,955. Estructura: 3 factores (ansioso-somático, disfunción social, depresión).
Lobo et al, 1988 (24)	España	Pacientes con problemas endocrinos	GHQ-28	PC=6/7: s=87,9%, e=88,8%, J=0,767
Retolaza Balsategui & Ballesteros Rodríguez, 2011 (22)	España	202 pacientes de atención primaria	GHQ-28	Versión nueva y previa del GHQ-28 tienen igual capacidad para discriminar casos, abcROC: 0,758 vs 0,762, p=0,63. Estructura: 4 componentes principales: 1) sueño, vigor,energía; 2) competencia social, 3) ideas suicidas,desesperanza; 4) ansiedad y síntomas somáticos.
Lobo et al, 1986 (23)	España	Pacientes ambulatorios de medicina interna (n=100)	GHQ-28	PC= 5,5: s=84,6%, e=82,0%, J=0,666
López-Castedo & Domínguez, 2010 (20)	España	2010 adolescentes	GHQ-28	Alfa=0,93. Confiabilidad test-retest=0,84. Estructura: 4 factores: depresión severa, ansiedad/insomnio, síntomas somáticos, disfunción social.
Ames-Guerrero et al, 2020 (18)	Perú	434 individuos	GHQ-28	Alfa=0,829. Estructura: 4 factores correlacionados: síntomas somáticos, ansiedad/insomnio, disfunción social y depresión severa.
Failde et al, 2000 (49)	España	185 pacientes con probable enfermedad cardiaca isquémica	GHQ-28 y "Mental health" (MH) del SF-36	GHQ-28: Alfa=0,95. MH: Alfa=0,80. GHQ-28 correlación negativa con subescalas del SF-36: MH (-0,61), "vitalidad" (-0,61).
Villoria & Lara, 2018 (38)	Chile	215 pacientes con cáncer	HADS	Estructura: 2 factores: ansiedad y depresión. Consistencia interna: Ansiedad: 0,76; Depresión: 0,84; Total: 0,87. Correlación significativa con DASS-21
Hinz et al, 2014 (39)	Colombia	Población general, n=1500	HADS	Puntajes promedio de HADS similares a los de estudios europeos: ansiedad=4,61±3,64 y depresión=4,30±3,91. Las mujeres tuvieron más ansiedad y depresión que los hombres. Los puntajes de depresión aumentaban con la edad.
Herrero et al, 2003 (33)	España	385 pacientes ambulatorios de medicina general	HADS	Estructura: 2 factores (ansiedad y depresión). Gold Standard: SCID-DSM-IV: PC=13, s=79,8%, e=83,3%, J=0,631

(página siguiente)

**Tabla 1 (cont).** Descripción de los estudios psicométricos de versiones en español de instrumentos para la detección de trastornos mentales comunes.

Estudio	País	Individuos estudiados	Instrumento	Hallazgos
Quintana et al, 2003 (34)	España	Pacientes ambulatorios con diferentes problemas de salud y estudiantes universitarios	HADS	Buena confiabilidad test-retest (>0,85). Alfa=0,86 para ansiedad, Alfa=0,86 para depresión. Estructura: 2 factores (ansiedad y depresión). Alta validez concurrente con BDI, STAI y los dominios mentales del Short-Form Health Survey.
Vallejo et al, 2012 (37)	España	Pacientes con fibromialgia, n=301	HADS	Estructura: 3 factores: ansiedad, agitación y depresión. Alfa, subescala de ansiedad=0,83. Alfa, subescala de depresión=0,87
Forjaz et al, 2009 (50)	España	Pacientes con Parkinson	HADS	Los resultados respaldan el uso de la escala total como medida de distrés y de la subescala de ansiedad pero no de la subescala de depresión.
Rodriguez-Blazquez, 2009 (35)	España	387 pacientes con enfermedad de Parkinson	HADS	Alfa: Subescala de ansiedad=0,81. Subescala de depresión=0,83. Estructura: 2 factores. HADS correlacionó estrechamente con medidas de calidad de vida.
Yamamoto-Furusho et al, 2018 (36)	México	112 pacientes con enfermedad inflamatoria intestinal	HADS	Estructura: 2 factores. Alfa=0,88.
Brenlla & Aranguren, 2010 (40)	Argentina	Población general (n=2517) y pacientes psiquiátricos (n=76)	K-10	abcROC=0,782; PC=24,5, s=0,724, e=0,756, J=0,48. Alfa=0,88. Correlacionada significativamente con BDI-II, la STAI-R y el SCL-90-R. Estructura: 1 dimensión.
Vargas Terrez et al, 2011 (41)	México	280 pacientes de consulta externa del primer nivel de atención	K-10	Para depresión (MINI): abcROC=0,87; PC=21, s=78,7%, e=79%, J=0,577. Para ansiedad (MINI): abcROC=0,824; PC=22: s=72,4%, e=73,8%, J=0,462.
Arias-Gallegos et al, 2019 (42)	Perú	165 estudiantes universitarios	K-10	Estructura: 1 factor. Alfa=0,899; Omega=0,901
Dominguez-Lara S & Alarcón-Parco D, 2020 (43)	Perú	528 estudiantes de psicología	K-6	Estructura: 2 factores. Problemas con el ítem 5.
Diez-Quevedo et al, 2001 (12)	España	Pacientes internados en hospital general (n=1003)	PHQ	s=87%, e=88%, J=0,75
Bonicatto et al, 1997 (13)	Argentina	Personas sanas (n=484) y pacientes con cáncer (n=105)	SCL-90	Alfa total=0,96. Alfa de las dimensiones >0,7 excepto para ideación paranoide y psicoticismo.

(página siguiente)

**Tabla 1 (cont).** Descripción de los estudios psicométricos de versiones en español de instrumentos para la detección de trastornos mentales comunes.

Estudio	País	Individuos estudiados	Instrumento	Hallazgos
Felipe-Castaño et al, 2015. (14)	España	1663 adolescentes de comunidad	SCL-90-R	Se reporta datos normativos según edad y sexo.
Pendergast et al, 2014 (46)	Múltiples países incluyendo Perú	Mujeres postparto (n=299, Perú)	SRQ	4 ítems somáticos fueron excluidos quedando una escala de 16 ítems. Estructura: 1 factor.
Véliz, 1998 (45)	Perú	64 pacientes psiquiátricos y 157 supuestos sanos	SRQ-20	Pacientes psiquiátricos: s=0,8421, e=0,8077, J= 0,6498. Supuestos sanos: s=0,5968, e= 0,6421, J= 0,2389

PC=punto de corte; abcROC= Área bajo la curva ROC; s=sensibilidad; e=especificidad; J= índice de Youden.

### General Health Questionnaire (GHQ)

El GHQ se ha estudiado en España, Colombia, Argentina y Perú; en población general, adolescentes, pacientes de atención primaria y con enfermedades físicas; principalmente las versiones de 28 y 12 ítems. Para el GHQ-28 se ha encontrado un alfa de Cronbach entre 0,829 (18) y 0,955 (19); y las siguientes estructuras de 3 o 4 factores:

- Ansioso-somático, disfunción social y depresión (19).
- Depresión severa, ansiedad/insomnio, síntomas somáticos y disfunción social (18,20).
- Depresión severa, disturbios del sueño, síntomas somáticos y disfunción social (21).
- “Sueño / vigor / energía”, “competencia social”, “ideas suicidas / desesperanza” y “ansiedad / síntomas somáticos” (22).

Para la detección de TMC, el GHQ-28 ha tenido índices J de 0,666 en pacientes ambulatorios de medicina interna (23) a 0,767 en pacientes con problemas endocrinos (24); y un abcROC de 0,76 en pacientes de atención primaria (22).

En el caso del GHQ-12, el alfa ha estado entre 0,76 (25) y 0,9 (26); la confiabilidad test-retest fue 0,79; y se han hallado estructuras de uno (26–28), 2 y 3 factores:

- “bienestar psicológico” y “malestar psicológico” (29).
- “ansiedad” y “disfunción social” (30).
- “ansiedad/depresión”, “anhedonia/disfunción social” y “pérdida de confianza y autoestima” (31).

Para la identificación de TMC, el GHQ-12 ha mostrado un J=0,45 en población general y clínica (26); y específicamente para pacientes postparto, se ha encontrado un abcROC de 0,904 y un J=0,61 (32).

Finalmente, a partir de la versión de 60 ítems del GHQ se ha propuesto una versión española de 30 ítems compuesta de 6 subescalas: “enfermedad general”, “pérdida de la autoestima”, “disfunción social”, “depresión severa”, “disturbios del sueño” y “pérdida del interés” (21).

### La Hospital Anxiety and Depression Scale (HADS)

La HADS ha sido evaluada en España, Colombia, Chile y México. Su estructura de 2 factores (ansiedad y depresión) ha sido respaldada por la mayoría de estudios (33–36); sin embargo, también se ha planteado un modelo de 3 factores (“ansiedad”, “agitación” y “depresión”) para los pacientes con fibromialgia (37). La consistencia interna para la subescala de ansiedad ha estado entre 0,76 (38) y 0,86 (34); y para la de depresión, entre 0,83 (35) y 0,87 (37).

En cuanto a la capacidad de identificar TMC, se ha encontrado un J=0,631 (33) entre pacientes ambulatorios de medicina general. Los puntajes de la HADS han correlacionado significativamente con las Depression, Anxiety and Stress Scales (DASS-21) (38), el Inventario de Depresión de Beck (BDI) (34), el Cuestionario de Ansiedad Estado-Rasgo (STAI) (34), los dominios mentales del Short-Form Health Survey (34) y las medidas de calidad de vida (35).

En la población general de Colombia, se encontraron puntajes de la HADS similares a los

reportados en estudios europeos (39).

### **Escala de Malestar Psicológico de Kessler**

La Escala de Malestar Psicológico de Kessler se ha estudiado en Argentina, México y Perú.

En Argentina, la versión de 10 ítems (K-10) mostró una estructura unidimensional y un alfa de 0,88. Para la discriminación entre pacientes psiquiátricos y población general se encontró un abcROC de 0,782 y un J de 0,48. La K-10 correlacionó significativamente con el BDI-II, la STAI-R y la SCL-90-R (40).

En México, en pacientes de ambulatorios e internados, se evaluó la capacidad de la K-10 para la detección de depresión y ansiedad utilizando como patrón de referencia la Mini International Neuropsychiatric Interview (MINI). Para la depresión, el abcROC fue 0,87 y el J=0,577; y para ansiedad, el abcROC=0,824 y el J=0,462 (41).

En Perú, se ha estudiado la K-10 y la versión de 6 ítems (K-6) en estudiantes universitarios. La K-10 ha mostrado una estructura de 1 solo factor, un alfa de 0,899 y un omega de 0,901 (42). Por otro lado, para la K-6, la estructura de 2 factores (ansiedad y depresión) tiene un mejor respaldo y el ítem 5 (“todo era un esfuerzo”) no mostró buen desempeño (43).

### **Escala de Salud Personal (ESP) o Personal Health Scale (PHS)**

La capacidad de la ESP para discriminar entre pacientes psiquiátricos y controles ha sido estudiada en Perú, Argentina y Venezuela, hallándose Js de 0,56 en Perú (9), 0,30 en Argentina (44) y 0,31 en Venezuela (44). La ESP tuvo un alfa de 0,77, una estructura de un solo factor (44) y una confiabilidad test-retest de 0,86 (correlación de Pearson) (9).

### **El Self-Reporting Questionnaire (SRQ)**

En el Perú, la capacidad de identificar trastornos mentales de la versión de 20 ítems (SRQ-20) ha sido estudiada en pacientes psiquiátricos encontrándose un J=0,6498 y en supuestos sanos con un J=0,2389 (45). Asimismo, un estudio internacional realizado en mujeres postparto que incluía 299 participantes de Perú encontró una estructura unifactorial para una versión de 16 ítems que excluía los síntomas somáticos (46).

### **La Escala de Edimburgo o Edinburgh Postnatal Depression Scale (EPDS)**

La EPDS, una escala diseñada con la finalidad de detectar casos sospechosos de depresión en puérperas, también ha sido útil para identificar otros trastornos mentales en esta población. En un estudio español con mujeres postparto, la EPDS ha mostrado una eficacia para la detección de TMC (trastornos del ánimo, de ansiedad y de ajuste) algo superior al GHQ-12, con un abcROC=0,933 y un J=0,708 (32).

### **Patient Health Questionnaire (PHQ)**

El PHQ, a diferencia del PHQ-9, evalúa la presencia no sólo de depresión mayor, sino también de trastorno de pánico, otros trastornos de ansiedad, bulimia nerviosa, otros trastornos depresivos, probable abuso o dependencia de alcohol y trastorno por atracones. La utilidad del PHQ para la detección de TMC, ha sido estudiada en España en pacientes internados de un hospital general encontrándose un J=0,75 (12).

### **Dep5 y Anx5**

En un estudio internacional se han investigado escalas cortas para ansiedad (Anx5) y depresión (Dep5). En la muestra mexicana de este estudio, la Dep5 tiene un abcROC de 0,79 y J=0,44 para la detección de depresión con o sin ansiedad; y la Anx5 tiene un abcROC de 0,78 y J=0,49 para la identificación de ansiedad con o sin depresión (11).

### **La Symptom Checklist-90 (SCL-90)**

La SCL-90 y su versión revisada (SCL-90-R) ha sido evaluada en Argentina (13) y España (14), respectivamente. La SCL-90 tuvo un alfa total de 0,96, y valores de alfa mayores de 0,7 para sus dimensiones con la excepción de las subescalas de ideación paranoide y psicoticismo (13). El estudio español reportó los datos normativos de la SCL-90-R según edad y sexo en adolescentes de la comunidad (14).

## **DISCUSIÓN**

La mayoría de los estudios encontrados se realizaron en España, existiendo investigaciones en Perú, Argentina, Colombia, México, Chile y Venezuela.

Respecto a la eficacia diagnóstica para TMC de los instrumentos psicométricos evaluados, tenemos Js que

van desde 0,2389 para el SRQ en Perú (45) hasta 0,767 para el GHQ-28 en pacientes españoles con problemas endocrinos (24); y abcROC entre 0,758 para el GHQ-28 en pacientes de atención primaria en España (22) y 0,933 para la EPDS en mujeres postparto de España (32).

Los instrumentos cuyas versiones en español han sido más estudiadas son el GHQ y la HADS, y la información sobre su eficacia diagnóstica puede resumirse de la siguiente manera: el J es 0,666 - 0,767 para el GHQ-28 (23,24), 0,45-0,61 para el GHQ-12 (26,32) y 0,631 para la HADS (33); asimismo, el abcROC es 0,904 para el GHQ-12 (32) y 0,758 para el GHQ-28 (22).

En el Perú, los instrumentos cuyas propiedades diagnósticas han sido evaluadas son sólo el SRQ-20 (J=0,2389-0,6498) (45) y la ESP (J=0,56) (9), en ambos casos con muestras pequeñas. Otras escalas estudiadas en nuestro país con relación a otras propiedades psicométricas fueron la K-6, la K-10 y el GHQ-28.

La revisión se limita a instrumentos para la detección TMC en general, no de algún trastorno en particular, de este modo no se incluyeron, por ejemplo, estudios centrados en la detección de depresión mayor con escalas tales como la CES-D, la Escala de Zung o el Inventario de Depresión de Beck; sin embargo, sí se incluyó un estudio que evaluaba la EPDS porque examinaba su capacidad de identificar trastornos del ánimo, ansiedad y ajuste (32).

En los estudios sobre eficacia diagnóstica debemos considerar la forma en que se seleccionaron los individuos “sanos” y los “enfermos”. Es diferente partir de una muestra y separar a sus integrantes en “sanos” y “enfermos” que tomar 2 muestras independientes: una de “sanos” y otra de “enfermos”. En el último caso, podría excluirse a una buena cantidad de individuos con severidad intermedia del problema, generando una diferencia más clara entre los grupos y resultando mejores valores de eficacia diagnóstica respecto al primer caso. Teniendo en cuenta lo anterior es difícil hacer comparaciones entre instrumentos si sus valores de eficacia diagnóstica provienen de diferentes estudios que usan metodología distinta. En los estudios sobre la ESP (9, 44), uno de ellos realizado en Perú (9), los “enfermos” y los “sanos” tienen diferente origen, los primeros son pacientes con claros diagnósticos psiquiátricos y los “sanos” son trabajadores de la salud o estudiantes; lo cual

probablemente excluya a individuos con trastornos mentales de severidad intermedia que no han tenido contacto con los servicios de salud.

Otro punto para tomar en cuenta es el tamaño de los grupos, si éste es muy pequeño el intervalo de confianza (IC) de la *s* y la *e* resultaría muy amplio y esto pondría en duda los hallazgos. Por ejemplo, en el estudio del SRQ en el Perú (45), si consideramos los resultados para los pacientes psiquiátricos (*s*=84,21% y *e*=80,77%), vemos que fueron calculados a partir de una muestra de 38 pacientes sintomáticos y 26 asintomáticos, lo cual implica para *s* un IC al 95% de 68,75-93,98% y para *e* un IC al 95% de 60,65% a 93,45%. Si tomamos en cuenta al grupo de “supuestos sanos”, el IC al 95% incluye valores para *s* por debajo del 50% (*s*=59,68%, IC al 95%: 46,45-71,95%) y valores para *e* que casi llegan al 50% (*e*=64,21, IC al 95%: 53,72-73,79%). Un instrumento cuya *s* y *e* están alrededor de 50% no tendría utilidad diagnóstica.

La Escala de Distrés de Kessler es un instrumento que se usa con frecuencia en el mundo para la detección de TMC (51), sin embargo, la eficacia diagnóstica de sus versiones en español ha sido muy poco estudiada, encontrándose sólo dos estudios para la K10 -en Argentina (40) y México (41)- y ninguno para la K-6.

## CONCLUSIONES

Esta revisión puede ser de utilidad para seleccionar instrumentos para la identificación de TMC en la atención clínica o en el ámbito de la investigación. Asimismo, se evidencia la necesidad de realizar más estudios para evaluar las propiedades psicométricas de instrumentos para detección de TMC en el Perú.

*Conflictos de interés:* El autor declara no tenerlos.

*Financiamiento:* por el propio autor.

## Correspondencia:

Johann M. Vega-Dienstmaier

Correo electrónico: johann.vega.d@upch.pe

## REFERENCIAS BIBLIOGRÁFICAS

1. Depresión y otros trastornos mentales comunes. Estimaciones sanitarias mundiales. Washington D.C: Organización Panamericana de la Salud; 2017.
2. World Health Organization. Clasificación de los trastornos mentales y del comportamiento. CIE-10: CDI-10. Madrid: Editorial Médica Panamericana SA;

- 1994.
3. Kessler RC, Chiu WT, Demler O, Merikangas KR, Walters EE. Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Arch Gen Psychiatry*. 2005;62: 617–627.
  4. Ali G-C, Ryan G, De Silva MJ. Validated Screening Tools for Common Mental Disorders in Low and Middle Income Countries: A Systematic Review. *PLoS One*. 2016;11: e0156939.
  5. Youden WJ. Index for rating diagnostic tests. *Cancer*. 1950;3: 32–35.
  6. Makanjuola VA, Onyeama M, Nuhu FT, Kola L, Gureje O. Validation of short screening tools for common mental disorders in Nigerian general practices. *Gen Hosp Psychiatry*. 2014;36: 325–329.
  7. Tran TD, Tran T, Fisher J. Validation of three psychometric instruments for screening for perinatal common mental disorders in men in the north of Vietnam. *J Affect Disord*. 2012;136: 104–109.
  8. Barreto do Carmo MB, Santos LM dos, Feitosa CA, Fiaccone RL, Silva NB da, Santos DN dos, et al. Screening for common mental disorders using the SRQ-20 in Brazil: what are the alternative strategies for analysis? *Braz J Psychiatry*. 2017;40: 115–122.
  9. Zapata-Vega MI, Mezzich JE, Mazzotti G, De La Grecca R, Llanos JP, Hernández J. Estudio de validación de la Escala de Salud Personal (ESP) en Lima, Perú. *Psiquiatría y Salud Integral*. 2001;1: 54–58.
  10. Mezzich JE, Caldera JT, Berganza CE. Psychiatric Diagnosis in Primary Care and the Personal Health Scale. *Past, Present and Future of Psychiatry*. World Scientific; 1994. pp. 212–218.
  11. Goldberg DP, Prisciandaro JJ, Williams P. The primary health care version of ICD-11: the detection of common mental disorders in general medical settings. *Gen Hosp Psychiatry*. 2012;34: 665–670.
  12. Diez-Quevedo C, Rangil T, Sanchez-Planell L, Kroenke K, Spitzer RL. Validation and utility of the patient health questionnaire in diagnosing mental disorders in 1003 general hospital Spanish inpatients. *Psychosom Med*. 2001;63: 679–686.
  13. Bonicatto S, Dew MA, Soria JJ, Seghezze ME. Validity and reliability of Symptom Checklist '90 (SCL90) in an Argentine population sample. *Soc Psychiatry Psychiatr Epidemiol*. 1997;32: 332–338.
  14. Felipe-Castaño E, León-del-Barco B, López-Pina JA. Symptom CheckList-90-R: proposed scales in a probability sample of adolescents from the general population. *Span J Psychol*. 2015;17: E107.
  15. Patel V, Simunyu E, Gwanzura F, Lewis G, Mann A. The Shona Symptom Questionnaire: the development of an indigenous measure of common mental disorders in Harare. *Acta Psychiatr Scand*. 1997;95: 469–475.
  16. Gill SC, Butterworth P, Rodgers B, Mackinnon A. Validity of the mental health component scale of the 12-item Short-Form Health Survey (MCS-12) as measure of common mental disorders in the general population. *Psychiatry Res*. 2007;152: 63–71.
  17. Christensen KS, Fink P, Toft T, Frostholt L, Ornbøl E, Olesen F. A brief case-finding questionnaire for common mental disorders: the CMDQ. *Fam Pract*. 2005;22: 448–457.
  18. Ames-Guerrero RJ, Barreda-Parra VA, Huamani-Cahua JC. Psychometric properties and factor invariance for the General Health Questionnaire (GHQ-28): study in Peruvian population exposed to the COVID-19 pandemic. *medRxiv*. 2020. Available: <https://www.medrxiv.org/content/10.1101/2020.11.10.20229435v1.abstract>
  19. Delgado-Gomez D, Lopez-Castroman J, de Leon-Martinez V, Baca-Garcia E, Cabanas-Arrate ML, Sanchez-Gonzalez A, et al. Psychometrical assessment and item analysis of the General Health Questionnaire in victims of terrorism. *Psychol Assess*. 2013;25: 279–287.
  20. López-Castedo A, Domínguez J. Exploratory factor analysis and psychometric properties of the General Health Questionnaire in Spanish adolescents. *Psychol Rep*. 2010;107: 120–126.
  21. Molina JD, Andrade-Rosa C, González-Parra S, Blasco-Fontecilla H, Real MA, Pintor C. The factor structure of the General Health Questionnaire (GHQ): a scaled version for general practice in Spain. *Eur Psychiatry*. 2006;21: 478–486.
  22. Retolaza Balsategui A, Ballesteros Rodriguez J. (Is there room for improvement in the Spanish version of the scaled General Health Questionnaire (GHQ-28)?). *Aten Primaria*. 2011;43: 662–667.
  23. Lobo A, Pérez-Echeverría MJ, Artal J. Validity of the scaled version of the General Health Questionnaire (GHQ-28) in a Spanish population. *Psychol Med*. 1986;16: 135–140.
  24. Lobo A, Pérez-Echeverría MJ, Jiménez-Aznárez A, Sancho MA. Emotional disturbances in endocrine patients. Validity of the scaled version of the General Health Questionnaire (GHQ-28). *Br J Psychiatry*. 1988;152: 807–812.
  25. Sánchez-López M del P, Dresch V. The 12-Item General Health Questionnaire (GHQ-12): reliability, external validity and factor structure in the Spanish population. *Psicothema*. 2008;20: 839–843.
  26. Ruiz FJ, García-Beltrán DM, Suárez-Falcón JC. General Health Questionnaire-12 validity in Colombia and factorial equivalence between clinical and nonclinical participants. *Psychiatry Res*. 2017;256: 53–58.
  27. Aguado J, Campbell A, Ascaso C, Navarro P, Garcia-Esteve L, Luciano JV. Examining the factor structure and discriminant validity of the 12-item General Health Questionnaire (GHQ-12) among Spanish

- postpartum women. *Assessment*. 2012;19: 517–525.
28. Romppel M, Hinz A, Finck C, Young J, Brähler E, Glaesmer H. Cross-cultural measurement invariance of the General Health Questionnaire-12 in a German and a Colombian population sample. *Int J Methods Psychiatr Res*. 2017;26. doi:10.1002/mpr.1532
29. Villa G IC, Zuluaga Arboleda C, Restrepo Roldán LF. Propiedades psicométricas del Cuestionario de Salud General de Goldberg GHQ-12 en una institución hospitalaria de la ciudad de Medellín. *Avances en Psicología Latinoamericana*. 2013;31: 532–545.
30. López-Castedo A, Fernández L. Psychometric properties of the Spanish version of the 12-item General Health Questionnaire in adolescents. *Percept Mot Skills*. 2005;100: 676–680.
31. Padrón A, Galán I, Durbán M, Gandarillas A, Rodríguez-Artalejo F. Confirmatory factor analysis of the General Health Questionnaire (GHQ-12) in Spanish adolescents. *Qual Life Res*. 2012;21: 1291–1298.
32. Navarro P, Ascaso C, Garcia-Esteve L, Aguado J, Torres A, Martín-Santos R. Postnatal psychiatric morbidity: a validation study of the GHQ-12 and the EPDS as screening tools. *Gen Hosp Psychiatry*. 2007;29: 1–7.
33. Herrero MJ, Blanch J, Peri JM, De Pablo J, Pintor L, Bulbena A. A validation study of the hospital anxiety and depression scale (HADS) in a Spanish population. *Gen Hosp Psychiatry*. 2003;25: 277–283.
34. Quintana JM, Padierna A, Esteban C, Arostegui I, Bilbao A, Ruiz I. Evaluation of the psychometric characteristics of the Spanish version of the Hospital Anxiety and Depression Scale. *Acta Psychiatr Scand*. 2003;107: 216–221.
35. Rodríguez-Blázquez C, Frades-Payo B, Forjaz MJ, de Pedro-Cuesta J, Martínez-Martin P, Longitudinal Parkinson's Disease Patient Study Group. Psychometric attributes of the Hospital Anxiety and Depression Scale in Parkinson's disease. *Mov Disord*. 2009;24: 519–525.
36. Yamamoto-Furushio JK, Sarmiento-Aguilar A, García-Alanis M, Gómez-García LE, Toledo-Mauriño J, Olivares-Guzmán L, et al. Hospital Anxiety and Depression Scale (HADS): Validation in Mexican Patients with Inflammatory Bowel Disease. *Gastroenterol Hepatol*. 2018;41: 477–482.
37. Vallejo MA, Rivera J, Esteve-Vives J, Rodríguez-Muñoz MF, Grupo ICAF. Use of the Hospital Anxiety and Depression Scale (HADS) to evaluate anxiety and depression in fibromyalgia patients. *Rev Psiquiatr Salud Ment*. 2012;5: 107–114.
38. Villoria E, Lara L. Assessment of the Hospital Anxiety and Depression Scale for cancer patients. *Rev Med Chil*. 2018;146: 300–307.
39. Hinz A, Finck C, Gómez Y, Daig I, Glaesmer H, Singer S. Anxiety and depression in the general population in Colombia: reference values of the Hospital Anxiety and Depression Scale (HADS). *Soc Psychiatry Psychiatr Epidemiol*. 2014;49: 41–49.
40. Brenlla ME, Aranguren M. Adaptación argentina de la Escala de Malestar Psicológico de Kessler (K10). *Revista de Psicología*. 2010;28: 309–340.
41. Vargas BE, Villamil V, Rodríguez C, Pérez J, Cortés J. Validación de la escala Kessler 10 (K-10) en la detección de depresión y ansiedad en el primer nivel de atención. *Propiedades psicométricas. Salud Ment*. 2011;34: 323–331.
42. Arias-Gallegos W, Rivera-Calcina R, Ceballos-Canaza K. Confiabilidad y Estructura Factorial de la Escala de Malestar Psicológico de Kessler en estudiantes universitarios de Arequipa (Perú). *Archivos de Medicina (Manizales)*. 2019;19: 387–395.
43. Dominguez-Lara S, Alarcón-Parco D. Análisis estructural de la escala de malestar psicológico de Kessler (K6) en universitarios peruanos. *Educ Médica*. 2020;21: 155–156.
44. Jatuff D, Montenegro R, Mezzich JE, González-Mora J, Zapata-Vega MI. (A validation study of the Personal Health Scale in Argentina and Venezuela). *Actas Esp Psiquiatr*. 2010;38: 189–195.
45. Véliz J. Validez y confiabilidad del cuestionario autoreporte en personal de una institución militarizada. *An Salud Ment*. 1998;14: 85–102.
46. Pendergast LL, Scharf RJ, Rasmussen ZA, Seidman JC, Schaefer BA, Svensen E, et al. Postpartum depressive symptoms across time and place: structural invariance of the Self-Reporting Questionnaire among women from the international, multi-site MAL-ED study. *J Affect Disord*. 2014;167: 178–186.
47. Goldberg DP, Reed GM, Robles R, Minhas F, Razzaque B, Fortes S, et al. Screening for anxiety, depression, and anxious depression in primary care: A field study for ICD-11 PHC. *J Affect Disord*. 2017;213: 199–206.
48. Burrone MS, Abeldaño A, Lucchese M, Susser L, Enders JE, Alvarado R, et al. (Psychometric properties and reliability of the general health questionnaire(GHQ-12) for adult patients in primary care centers in Cordoba, Argentina). *Rev Fac Cien Med Univ Nac Cordoba*. 2015;72: 236–242.
49. Failde I, Ramos I, Fernández-Palacín F. Comparison between the GHQ-28 and SF-36 (MH 1-5) for the assessment of the mental health in patients with ischaemic heart disease. *Eur J Epidemiol*. 2000;16: 311–316.
50. Forjaz MJ, Rodríguez-Blázquez C, Martínez-Martin P, Longitudinal Parkinson's Disease Patient Study Group. Rasch analysis of the hospital anxiety and depression scale in Parkinson's disease. *Mov Disord*.

2009;24: 526–532.

51. Stolk Y, Kaplan I, Szwarc J. Clinical use of the Kessler psychological distress scales with culturally diverse groups. *Int J Methods Psychiatr Res.* 2014;23: 161–183.

Recibido: 18/11/2021

Aceptado: 17/04/2022