



## Propiedades psicométricas y valores normativos del General Health Questionnaire (GHQ-12) en población general española<sup>1</sup>

Kátia B. Rocha <sup>2</sup> (*Agència de Salut Pública de Barcelona (ASPB) y Universitat Autònoma de Barcelona, España*), Katherine Pérez (*Agència de Salut Pública de Barcelona (ASPB) y CIBER de Epidemiología y Salud Pública (CIBERESP), España*), Maica Rodríguez-Sanz (*Agència de Salut Pública de Barcelona (ASPB) y CIBER de Epidemiología y Salud Pública (CIBERESP), España*), Carme Borrell (*Agència de Salut Pública de Barcelona (ASPB) y CIBER de Epidemiología y Salud Pública (CIBERESP), España*) y Jordi E. Obiols (*Universitat Autònoma de Barcelona, España*)

**RESUMEN.** El objetivo principal del presente estudio es analizar las propiedades psicométricas del *General Health Questionnaire* (GHQ-12) como un instrumento de cribado en una muestra representativa de la población española, y en población mayor de 65 años. Se trata de un estudio instrumental utilizando los datos de la Encuesta Nacional de Salud de España (2006). Fueron entrevistadas 29.476 personas mayores de 16 años, no institucionalizadas. Se realizaron análisis de correlación entre ítems, análisis de consistencia interna, un análisis factorial exploratorio (AFE) y un análisis de validez externa del GHQ-12. Todos los análisis fueron realizados utilizando dos métodos de puntuación (Likert y GHQ). Los resultados muestran que cada uno de los

<sup>1</sup> Agradecimientos: Financiamiento parcial del Ministerio de Sanidad y Consumo -Observatorio de Salud de la Mujer Dirección General de la Agencia de Calidad y Ministerio de Ciencia e Innovación - Instituto de Salud Carlos III (España). Al profesor José Blas Navarro de la Universitat Autònoma de Barcelona (UAB) por la revisión del manuscrito. Al Programa Alàan que ha concedido una beca a la primera autora del artículo. Este artículo es parte de la tesis doctoral de Kátia B. Rocha realizada en el Departament de Psicologia Clínica i de la Salut de la Univesitat Autònoma de Barcelona.

<sup>2</sup> Correspondencia: Plaza Lesseps, 1. 08023 Barcelona (España). E-mail: krocha@aspb.cat

ítems del GHQ se correlaciona con los demás, presentando valores de correlación superiores a 0,51. El análisis factorial exploratorio confirma que gran parte de la variabilidad del instrumento (67% Likert y 73% GHQ) puede ser explicada al reunir todos los ítems en un único factor. El GHQ-12 presenta una elevada consistencia interna en población general y también en población mayor de 65 años. Los análisis factoriales confirman que el GHQ-12 puede ser utilizado como un instrumento unidimensional de cribado y la puntuación más adecuada en este caso es la puntuación GHQ.

**PALABRAS CLAVE.** General Health Questionnaire. Cribado de salud mental. Encuestas de salud. Estudio instrumental.

**ABSTRACT.** The main objective of this study is to analyze the psychometric properties of the General Health Questionnaire (GHQ-12) as a screening instrument in a representative sample of the Spanish population, and among people older than 65 years. This is a instrumental study using data from the National Health Survey of Spain (2006). The sample include 29,476 non-institutionalized interviewed persons over 16 years. Correlation analysis was carried out between items, internal consistency analysis, exploratory factor analysis (EFA) and a discussion of external validity of GHQ-12. All analyzes were performed using two scoring methods (Likert and GHQ). The results showed that the GHQ-12 has a high internal consistency in the general population and in people older than 65 years. Each of the items in the GHQ correlated with others, showing correlation values greater than .51. Exploratory factor analysis confirmed that much of the variability of the instrument (67% Likert and 73% GHQ) may be explained by gathering all the items into a single factor. The factor analysis confirmed that the GHQ-12 can be used as a unidimensional screening and scoring most appropriate in this case.

**KEY WORDS.** General Health Questionnaire. Mental health screening. Health surveys. Instrumental study.

El *General Health Questionnaire* (GHQ-12) es un instrumento auto-administrado de cribado que tiene por objetivo detectar morbilidad psicológica y posibles casos de trastornos psiquiátricos en contextos como la atención primaria o en población general (Goldberg y Williams, 1988). Además del GHQ-12 existen otros instrumentos de cribado en salud mental como el SF-12 (*Mental Health Component Scale*), MCS-121 (*2-ítem Short-Form Survey*), K10 y K6 (*Kessler Psychological Distress Scale*) (Gill, Butterworth, Rodgers y Mackinnon, 2007; Patel *et al.*, 2008). Sin embargo, el GHQ-12 ha sido traducido a más de 11 idiomas (Werneke, Goldberg, Yalcin y Ustun, 2000), siendo uno de los instrumentos de cribado validado más utilizados en todo el mundo (Hewitt, Perry, Adams y Gilbody, 2010) y hay que añadir que es el instrumento recomendado para ser utilizado en encuestas de salud (McDoweel, 2006).

Hay estudios que exploran la existencia de diferentes factores o dimensiones teóricas del GHQ-12. Por ejemplo, hay estudios que describen la existencia de dos

factores: *Depresión/Ansiedad y Disfunción social* (Werneke *et al.*, 2000) y otros que describen la existencia de tres: *Estrategias de afrontamiento, Autoestima y Estrés* (Graetz, 1991; Martín, 1999; Sánchez-López y Dresh, 2008). Sin embargo, hay estudios que afirman que el GHQ-12 debe ser utilizado prioritariamente como instrumento unidimensional de cribado (Hankins, 2008; Hu, Stewart-Brown, Twigg y Weich, 2007; Werneke *et al.*, 2000; Ye, 2009). A partir de los datos de la Encuesta de Salud de Inglaterra, un estudio concluye que los factores encontrados se deben en realidad a los sesgos de los seis ítems positivos y los seis negativos del GHQ-12, que se agrupan en dos factores aislados (Hankins, 2008).

La validación en España del GHQ (60, 30 y 12 ítems) fue realizada en la ciudad de Tres Cantos, Madrid. Una de las principales limitaciones de esta validación es que no ha sido publicada en ninguna revista científica, lo que dificulta el acceso a este material. Las propiedades psicométricas del GHQ-12 en España también fueron analizadas en adolescentes (López-Castedo y Fernández, 2005), puérperas (Navarro *et al.*, 2007) y en adultos (Sánchez-López y Dresh, 2008). Según nuestro conocimiento, el estudio más reciente publicado en España con el GHQ-12 fue desarrollado con una muestra de 1.001 participantes, utilizando la puntuación Likert. Los resultados muestran una consistencia interna del GHQ-12 ligeramente inferior a la encontrada en otros estudios. Además, los autores destacan el bajo poder discriminatorio de uno de los ítems y que el GHQ-12 funciona mejor como una escala multidimensional (Sánchez-López y Dresh, 2008).

Cabe destacar que gran parte de los estudios desarrollados en nuestro contexto utilizan el GHQ-12 como un instrumento de cribado, utilizando la puntuación GHQ, no la puntuación Likert, como es el caso de los estudios desarrollados a partir de los datos de encuestas de salud internacionales (Hamer, Stamatakis y Steptoe, 2009; Laaksonen *et al.*, 2007; Molloy, Stamatakis, Randall y Hamer, 2009) y en el contexto español (Artazcoz, Borrell y Benach, 2001; Artazcoz, Borrell, Benach, Cortès y Rohls, 2004; Borrell, Artazcoz, *et al.*, 2010; Borrell, Muntaner *et al.*, 2010; Muntaner, Borrell, Benach, Pasarín y Fernández, 2003; Rocha, Pérez, Rodríguez-Sanz, Borrell y Obiols, 2010). Así, faltan estudios que analicen las propiedades psicométricas del GHQ como instrumento unidimensional de cribado, comparando los diferentes métodos de puntuación del instrumento. Campbell y Knowles (2007) destacan que muchos investigadores no explicitan el método de puntuación utilizado (Likert, GHQ, C-GHQ) y que las diferentes puntuaciones utilizadas influyen en las características psicométricas y factoriales del GHQ-12.

No hemos hallado ningún estudio publicado en España que explore las características psicométricas del GHQ-12 en población mayor de 65 años. Por otro lado, se observa un interés creciente en utilizar este instrumento de cribado en esta población (Bowling, 2007; Costa, Barreto, Uchoa, Firmo y Lima-Costa, 2006; Netuveli, Wiggins, Montgomery, Hildon y Blane, 2008; Salama-Younes, Montazeri, Ismaíl y Roncin, 2009). La literatura describe que el GHQ-12 tiende a presentar más falsos positivos en la población mayor de 65 años, ya que los síntomas físicos son muchos más comunes con el aumento de la edad, lo que puede llevar a una mayor puntuación en el GHQ-12 (Papassotiropoulos y Hein, 1999).

El presente artículo es un estudio de tipo instrumental (Carretero-Dios y Pérez, 2007; Montero y León, 2007) que tiene como objetivo principal analizar las propiedades

psicométricas del GHQ-12 como un instrumento de cribado en una muestra representativa de la población española, y también en población mayor de 65 años. Se plantean como objetivos específicos analizar el poder discriminatorio de cada uno de los 12 ítems del instrumento, investigar su estructura factorial principalmente como instrumento unidimensional de cribado, analizar la consistencia interna del GHQ-12 y, finalmente, hacer un análisis de validez externa. Además, todos los resultados fueron analizados utilizando dos métodos distintos de puntuación, GHQ y Likert, lo que permitirá compararlos.

## Método

### *Participantes*

Participaron en el presente estudio 29.476 personas mayores de 16 años (14.458 hombres,  $M = 44,8$  años, rango de edad: 16-101; y 15.018 mujeres,  $M = 47,1$  años, rango de edad: 16-104), no institucionalizadas, que contestaron a la Encuesta Nacional de Salud de España (ENS) de 2006, y que son representativas de la población española residente en viviendas familiares principales.

### *Instrumento*

El *General Health Questionnaire* (GHQ-12), elaborado por Goldberg y Williams (1988), está formado por 12 ítems, siendo 6 de ellos sentencias positivas y 6 sentencias negativas. Los ítems son contestados a través de una escala tipo Likert de cuatro puntos (0-1-2-3), que puede ser transformada en una puntuación dicotómica (0-0-1-1), llamada puntuación GHQ. El GHQ-12 presenta una buena fiabilidad en los diferentes estudios realizados con alfas de Cronbach que varían entre 0,82 y 0,86 (Goldberg *et al.*, 1997). Un estudio reciente desarrollado en España describe un alfa de Cronbach de 0,76 (Sánchez-López y Dresh, 2008). Los resultados de validación del GHQ-12 realizado en 15 países del mundo son buenos, con resultados de curva ROC que varían entre 83 y 85 (Goldberg *et al.*, 1997).

### *Otras variables de estudio*

Depresión, ansiedad u otro trastorno mental autodeclarado. En la ENS de 2006 hay un apartado que pregunta si el entrevistado ha sufrido alguna enfermedad o problema de salud en los últimos 12 meses. Entre las opciones de respuestas está haber sufrido depresión, ansiedad u otro trastorno mental.

Estado de salud autopercibido. En la ENS hay un ítem que pregunta cómo el entrevistado percibe su estado de salud en los últimos 12 meses. El ítem es contestado a través de una escala Likert de 1-5 (1 = *muy bueno* y 5 = *muy malo*). Este ítem viene siendo utilizado en diferentes estudios como una medida adecuada de la salud global (Borrell, Artacoz *et al.*, 2010; Borrell, Muntaner *et al.*, 2010; Fogel e Israel, 2009; Ramos, Moreno, Rivera y Pérez, 2010).

Variables sociodemográficas:

- Edad. La muestra será dividida en dos grupos de edad, los que tienen entre 16 y 64 años ( $n = 23.759$ ) y los que tienen 65 años o más ( $n = 5.718$ ).

- Sexo. Para el análisis de la fiabilidad del GHQ-12 la muestra fue analizada en conjunto y también estratificada por sexo.

### *Procedimiento*

La muestra de estudio de la ENS de 2006 fue seleccionada mediante un muestreo polietápico estratificado. La unidad de primera etapa son las secciones censales y la unidad de segunda etapa son las viviendas familiares. Fueron seleccionadas cerca de 31.000 viviendas representativas de las 17 Comunidades Autónomas. El 96% de la muestra teórica fue entrevistada; sin embargo, para completar este porcentaje, algunas viviendas fueron sustituidas por otras (31%). La información fue recogida entre junio de 2006 y junio de 2007. Las entrevistas fueron realizadas en la casa de los entrevistados por entrevistadores entrenados (Ministerio de Sanidad y Consumo, 2006).

### *Análisis de los datos*

Los datos del presente estudio fueron analizados a partir de la estratificación de la muestra en dos grupos, los menores de 65 años ( $n = 23.759$ ) y los que tenían 65 años o más ( $n = 5.718$ ). Además, todos los análisis se realizaron utilizando dos métodos distintos de puntuación del GHQ-12, Likert y GHQ. Carretero-Dios y Pérez (2007) destacan que un aspecto esencial a considerar al analizar un instrumento son las propiedades métricas de cada uno de los ítems que lo constituyen. En este sentido, fueron realizados análisis del poder discriminatorio de cada uno de los ítems del GHQ-12, mediante el cálculo de correlación ítem-total que mide la relación entre un determinado ítem con los demás ítems de la escala. Son considerados aceptables valores superiores a 0,30. Además, se analizó el alfa de Cronbach del instrumento si fuera excluido cada uno de los ítems. Para obtener evidencias empíricas de la estructura interna del GHQ-12 fue realizado un análisis factorial exploratorio (AFE) para analizar cuanto de la variabilidad del instrumento podría ser explicada al reunir todos los ítems en el primero factor, ya que el GHQ-12 puede ser utilizado como una escala unidimensional. Se considera adecuado si la varianza explicada es igual o superior a 70%. A partir de estos análisis también es posible verificar el peso que cada uno de los ítems tiene en el primer factor (datos sin rotar), así como analizar cuanto aumenta la variabilidad explicada del instrumento al utilizarse la solución de dos o tres factores. Además, como objetivo secundario, se ha forzado la solución factorial con dos y tres factores para, por un lado, verificar si la agrupación de los ítems es semejante a la descrita en la literatura, y por otro, verificar si la agrupación factorial se explicaba en realidad en función de la forma de los ítems (sentencias positivas y negativas). Fueron presentados los resultados rotados por ítems. Cabe destacar que antes de realizar el análisis factorial se calculó la medida de adecuación muestral a partir del *índice de Kaiser-Meyer-Olkin (KMO)*, que permite analizar si la aplicación del análisis factorial resulta pertinente. El análisis de la fiabilidad se realizó a partir del análisis de consistencia interna mediante el cálculo del alfa de Cronbach. El valor mínimo aceptable para el coeficiente de alfa de Cronbach es 0,70 y el valor máximo esperado es 0,90, por encima de ese valor se considera que hay redundancia o duplicación de los ítems (Oviedo y Campo-Arias, 2005). Para finalizar, se ha realizado un análisis de validez externa mediante el análisis *Chi-cuadrado de Pearson*

( $\chi^2$ ) entre la puntuación GHQ-12 y dos variables de la ENS, el estado de salud autopercebido y la auto declaración de sufrir depresión, ansiedad u otro trastorno mental en los últimos 12 meses.

## Resultados

### *Análisis estadístico de los ítems*

Se observa que cuando cada uno de los 12 ítems del GHQ son analizados aisladamente, estos se correlacionan con los demás ítems de la escala, con valores de correlación superiores a 0,30, oscilando en la población menor de 65 años entre 0,43 y 0,69 (Likert) y entre 0,49 y 0,67 (GHQ). Entre la población de 65 años o más, los datos oscilaban entre 0,49 y 0,73 (Likert) y entre 0,54 y 0,69 (GHQ). En la misma dirección, el análisis de consistencia interna muestra que la escala no aumentaría su consistencia interna al excluir algún de los 12 ítems, en ambas puntuaciones (Likert y GHQ) y en los dos grupos de edad, lo que indica que todos los ítems aportan informaciones complementarias y están ligeramente relacionados entre ellos (véase la Tabla 1).

**TABLA 1.** Análisis del alfa de Cronbach si cada ítem de la escala fuera excluido y análisis de la correlación ítem-total entre los 12 ítems del GHQ.

	< 65 años (n = 23.760)				=>65 años (n = 5.718)			
	Likert		GHQ		Likert		GHQ	
	$\alpha$	$r_{i-t}$	$\alpha$	$r_{i-t}$	$\alpha$	$r_{i-t}$	$\alpha$	$r_{i-t}$
1. ¿Ha podido concentrarse bien en lo que hacía?	0,86	0,52	0,86	0,55	0,89	0,59	0,89	0,54
2. ¿Sus preocupaciones le han hecho perder mucho el sueño?	0,86	0,53	0,86	0,58	0,89	0,57	0,89	0,60
3. ¿Ha sentido que está desempeñando un papel útil en la vida?	0,87	0,44	0,86	0,49	0,89	0,49	0,89	0,55
4. ¿Se ha sentido capaz de tomar decisiones?	0,87	0,43	0,85	0,49	0,89	0,53	0,89	0,56
5. ¿Se ha notado constantemente agobiado y en tensión?	0,86	0,55	0,86	0,57	0,89	0,65	0,89	0,60
6. ¿Ha tenido la sensación de que no puede superar sus dificultades?	0,85	0,65	0,86	0,65	0,88	0,73	0,88	0,69
7. ¿Ha sido capaz de disfrutar de sus actividades normales de cada día?	0,86	0,57	0,85	0,60	0,89	0,65	0,88	0,65
8. ¿Ha sido capaz de hacer frente adecuadamente a sus problemas?	0,86	0,53	0,86	0,58	0,89	0,60	0,88	0,65
9. ¿Se ha sentido poco feliz o deprimido/a?	0,85	0,69	0,86	0,67	0,88	0,73	0,88	0,68
10. ¿Ha perdido confianza en sí mismo/a?	0,86	0,66	0,86	0,61	0,88	0,71	0,88	0,64
11. ¿Ha pensado que usted es una persona que no vale para nada?	0,86	0,56	0,86	0,51	0,89	0,65	0,89	0,57
12. ¿Se siente razonablemente feliz considerando todas las circunstancias?	0,86	0,53	0,86	0,57	0,89	0,60	0,89	0,64

Notas.  $\alpha$  = alfa de Cronbach;  $r_{i-t}$  = correlación ítem-total.

*Estudio de la dimensionalidad del instrumento*

A partir de los resultados del índice de KMO se observa una excelente adecuación de la muestra al análisis factorial (KMO Likert 0,91 y GHQ 0,93) (véase la Tabla 2). Además, a partir del análisis factorial exploratorio se observa que gran parte de la variabilidad de los ítems del GHQ-12, puede ser explicada al reunir los 12 ítems en un solo factor (67% Likert y 73% GHQ). La opción de dos factores explicaría 82% (Likert) y el 84% (GHQ) y la opción de tres factores 91% (Likert) y 89% (GHQ). Además, a partir del análisis factorial exploratorio se observa que ninguno de los ítems presenta valores inferiores a 0,46 (Likert) y 0,55 (GHQ) en el primer factor. A través del análisis factorial forzando la agrupación de dos factores en la puntuación Likert los ítems 2, 5, 6, 9, 10 y 11 se reúnen en el factor 1 y los demás en el factor 2. En la opción de forzar tres factores se observa que en el factor 1 se reúnen 5 ítems (1, 3, 4, 7, 8 y 12), en el factor 2 se reúnen 4 ítems (6, 9, 10 y 11) y en el factor 3 se reúnen 3 ítems (2, 5 y 9). A través de la puntuación GHQ los resultados muestran que los ítems 1, 2, 5, 6, 9, 10 y 11 se reúnen en el factor 1 y los demás en el factor 2. En la opción de forzar tres factores se observa que en el factor 1 se reúnen 6 ítems (1, 2, 5, 6, 7 y 9), en el factor 2 se reúnen 4 ítems (3, 10, 11 y 12) y en el factor 3 se reúnen 2 ítems (4 y 8).

**TABLA 2.** Análisis factorial exploratorio con los valores del factor 1 (datos sin rotar) y matriz de componentes rotados forzando 2 y 3 factores. Porcentaje de varianza explicada por cada factor e índice de adecuación de la muestra (*KMO*) ( $N = 23.760$ ).

	Puntuación Likert						Puntuación GHQ					
	2 factores		3 factores		3 factores		2 factores		3 factores		3 factores	
	Factor	Factor	Factor	Factor	Factor	Factor	Factor	Factor	Factor	Factor	Factor	Factor
1. ¿Se ha podido concentrar bien en lo que hacía?	0,56	0,33	0,67	0,38	0,58	0,67	0,65	0,67	0,65	0,52	0,50	
2. ¿Sus preocupaciones le han hecho perder mucho el sueño?	0,60	0,68			0,60							
3. ¿Ha sentido que juega un papel útil en la vida?	0,47	0,58		0,53	0,55	0,56	0,42					
4. ¿Se ha sentido capaz de tomar decisiones?	0,46	0,65		0,62	0,56	0,55	0,53					0,53
5. ¿Se ha notado constantemente cargado/a y en tensión?	0,63	0,74	0,74		0,60		0,70			0,71	0,70	
6. ¿Ha tenido la sensación de que no puede superar sus dificultades?	0,71	0,70	0,54		0,69		0,49			0,52	0,49	
7. ¿Ha sido capaz de disfrutar de sus actividades de cada día?	0,62	0,43		0,48	0,65	0,48	0,46			0,48	0,46	
8. ¿Ha sido capaz de hacer frente, adecuadamente, a sus problemas?	0,56	0,61		0,61	0,65	0,58	0,53					0,53
9. ¿Se ha sentido poco feliz y deprimido/a?	0,76	0,76	0,58		0,72		0,57			0,58	0,57	
10. ¿Ha perdido confianza en sí mismo/a?	0,75	0,63		0,80	0,69	0,63	0,68					0,68
11. ¿Ha pensado que es una persona que no sirve para nada?	0,63	0,49		0,73	0,60	0,64	0,65					0,65
12. ¿Se siente razonablemente feliz, considerando todas las circunstancias?	0,56	0,54		0,51	0,64	0,54	0,46					0,46
Varianza explicada (%) por los factores	67	82		91	73	84				GHQ 0,93		89
KMO	Likert 0,91						GHQ 0,93					



*Análisis de la fiabilidad*

A partir de la constatación de que los 12 ítems del GHQ ofrecen información complementaria y que ningún de ellos compromete la fiabilidad del instrumento se ha analizado la fiabilidad del instrumento con todos los ítems. El GHQ-12 presenta una elevada consistencia interna en población general ( $\alpha = 0,86$ ) y también en población mayor de 65 años ( $\alpha = 0,90$ ) en ambas puntuaciones (Likert y GHQ). No existen casi diferencias entre la consistencia del instrumento al ser analizado por separado en hombres y mujeres (véase la Tabla 3).

**TABLA 3.** Análisis del alfa de Cronbach de toda la escala. La muestra esta dividida en < 65 años ( $n = 23.760$ ) y  $\geq 65$  años ( $n = 5.718$ ).

	<i>Likert</i>		<i>GHQ</i>	
	< 65 años	$\geq 65$ años	< 65 años	$\geq 65$ años
	$\alpha$	$\alpha$	$\alpha$	$\alpha$
Toda la muestra	0,86	0,90	0,86	0,90
Hombres	0,86	0,90	0,86	0,89
Mujeres	0,86	0,89	0,86	0,89

*Nota.*  $\alpha$  = alfa de Cronbach.

*Evidencias de validez externa del GHQ-12*

Al realizar el análisis de *Chi-cuadrado de Pearson* entre la variable depresión, ansiedad u otro trastorno mental y el GHQ-12 se observa que hay una asociación estadísticamente significativa ( $\chi^2$  significativo;  $p < 0,001$ ) entre estas dos medidas. El 62% de los entrevistados que afirmaron haber sufrido depresión, ansiedad u otro trastorno mental fueron positivos en el cribado GHQ-12, lo que indica la posible presencia de un trastorno mental. En la misma dirección, se observa una asociación significativa entre el estado de salud autopercebido y el GHQ-12 ( $\chi^2$  significativo;  $p < 0,001$ ). El 64% de los que declararon que su estado de salud era regular, malo o muy malo fueron positivos en el GHQ-12.

Para finalizar, se ha realizado también el análisis de la distribución en percentiles de la puntuación del GHQ-12 en la muestra (véase el Anexo 1). La muestra fue dividida en dos grupos, la población menor y mayor de 65 años, y analizada la distribución utilizando la puntuación Likert y GHQ. A partir de estos resultados es posible concluir que la población mayor de 65 años tiene una distribución diferente de la población más joven, teniendo puntuaciones más elevadas, lo que puede confirmar la necesidad de utilizar puntos de corte más elevados en esta población.

**Discusión y conclusiones**

Es importante destacar que por primera vez se publican datos psicométricos del GHQ-12 a partir de una muestra representativa de toda la población española, incluyendo además un subanálisis para la población mayor. Asimismo, por primera vez se exploran las diferencias entre los diferentes métodos de puntuación del GHQ-12 en

España (Likert y GHQ). En este sentido se observa que en ambos métodos de puntuación, los análisis muestran que cada uno de los ítems 12 del GHQ tienen un elevado poder discriminatorio. Además, el GHQ-12 presenta una elevada consistencia interna en población general, en ambos métodos de puntuación y también en personas mayores de 65 años. El análisis factorial exploratorio muestra que gran parte de la variabilidad del GHQ-12 puede ser explicada al reunir los 12 ítems en un solo factor, lo que refuerza que el GHQ-12 puede ser utilizado como una escala unidimensional de cribado. Cabe destacar que la puntuación GHQ consigue explicar más variabilidad al reunir todos los ítems en el primer factor que la puntuación Likert, lo que refuerza que al utilizar el instrumento como escala unidimensional la puntuación GHQ es la más adecuada. El análisis de validez externa confirma que existe asociación entre el GHQ-12 y el hecho de declarar que se padece depresión, ansiedad u otro trastorno mental en el año anterior a la entrevista, en ambas puntuaciones, lo que refuerza la sensibilidad del GHQ-12 para detectar problemas de salud mental.

Respecto a las limitaciones del presente estudio, cabe destacar que la variable utilizada para analizar la validez externa que es referir tener depresión, ansiedad u otro trastorno mental, no puede ser considerada una variable *gold estándar*, como podría ser una prueba diagnóstica de trastorno mental. Otro aspecto a considerar es que no se realizaron análisis factoriales confirmatorios para analizar la estructura multidimensional del GHQ-12.

Por otro lado, una de las principales fortalezas del presente estudio es que al trabajar con una muestra representativa de toda población española se evitan algunos de los errores que se podían producir en función de la selección de la muestra. Ello posibilita tener datos de las características psicométricas del GHQ-12 de gran poder estadístico. Además, todos los análisis fueron realizados con dos métodos diferentes de puntuación (Likert y GHQ), lo que permite contrastarlos. Un dato a destacar es la elevada consistencia interna y el adecuado poder discriminatorio de cada uno de los ítems del GHQ-12 entre la población mayor de 65 años, lo que coincide con los resultados de otros estudios que utilizan el GHQ-12 en población mayor de 65 años (Costa *et al.*, 2006; Salama-Younes *et al.*, 2009). No hay datos específicos de sensibilidad, especificidad, valores predictivos positivo y negativo que puedan contrastar los resultados de estudios anteriores que pusieron en evidencia el elevado porcentaje de falsos positivos en el GHQ-12 en población mayor de 65 años (Papassotiropoulos y Hein, 1999). Futuros estudios que utilicen pruebas diagnósticas pueden explorar la adopción de puntos de cortes más elevados para reducir el porcentaje de falsos positivos en población española mayor de 65 años, estrategia que ya fue adoptada en otros países (Papassotiropoulos, Heun y Maier, 1997). El GHQ-12 también presenta una elevada consistencia interna en población general, lo que confirma resultados de estudios anteriores con el GHQ-12 (Goldberg y Williams, 1988; López-Castedo y Fernández, 2005).

Los resultados del presente estudio indican que los 12 ítems del GHQ-12 aportan informaciones complementarias, presentando un elevado poder discriminatorio. Estudios anteriores referían que el ítem 11 del GHQ-12 («¿Ha pensado que usted es una persona que no vale para nada?») presenta un bajo poder discriminatorio (Sánchez-López y Dresh, 2008), lo que no fue confirmado con los resultados del presente estudio.

Cabe destacar que el elevado poder discriminatorio de cada uno de los ítems del GHQ-12, sumado a que gran parte de la variabilidad del instrumento puede ser explicada al reunir todos los ítems en un único factor, refuerzan la utilización del GHQ-12 como un instrumento unidimensional de cribado, especialmente al utilizar la puntuación GHQ.

El análisis factorial muestra un incremento de la variabilidad explicada al utilizar dos factores. Al analizar los datos de la puntuación Likert, los resultados del presente estudio se asemejan a la estructura factorial descrita en la literatura forzando dos factores (Werneke *et al.*, 2000) y tiene una distribución idéntica a la estructura descrita en la literatura al forzar tres factores (Martin, 1999). Sin embargo, al analizar las diferencias existentes se observa que en nuestro estudio en el factor 1 se reúnen todos los ítems cuya opción de respuesta son sentencias negativas y en el factor 2 se reúnen todos los ítems cuya opción de respuesta son sentencias positivas. Lo mismo pasa en el factor 1 (*Estrategias de afrontamiento*) que reúne todos los ítems positivos de la escala. Los resultados del presente estudio, de manera general, corroboran los hallazgos de Hankins (2008), que concluye que los factores encontrados se deben en realidad a sesgos entre las opciones de respuestas con sentencias positivas (6 ítems) y negativas (6 ítems) y que el GHQ-12 es un instrumento que debe ser utilizado prioritariamente como una escala unidimensional, lo que también es corroborado por otros estudios (Hankins, 2008; Hu *et al.*, 2007; Werneke *et al.*, 2000; Ye, 2009). En la puntuación GHQ, los resultados factoriales no se asemejan a los descritos en la literatura y tampoco hay una división tan clara de los ítems positivos y negativos en factores aislados.

Un estudio que ha realizado una comparación entre las soluciones factoriales de los tres diferentes métodos de puntuación del GHQ-12 (Likert, GHQ y C-GHQ) concluye que la puntuación GHQ tiene la más alta correlación entre los diferentes factores lo que indica una baja capacidad de diferenciación entre diferentes factores, y refuerza su utilización como una puntuación unidimensional (Campbell y Knowles, 2007). En el estudio de validación del GHQ-12 los autores describen que la puntuación GHQ presenta los mejores resultados, siendo la peor opción el método Likert (Goldberg *et al.*, 1997).

El GHQ-12 tiene como ventaja ser un instrumento corto, de fácil comprensión y auto-administrado, lo que facilita su utilización como instrumento de cribado. Los resultados del presente estudio refuerzan que el GHQ-12 es un instrumento con buenas características psicométricas en el contexto español y que puede ser utilizado como un instrumento unidimensional de cribado, en este caso utilizando prioritariamente la puntuación GHQ.

Sería conveniente asimismo realizar una validación actualizada en España, que permita obtener datos de sensibilidad, especificidad y capacidad discriminatoria del GHQ-12 en una muestra poblacional representativa, en la que se pueda explorar más acerca de las diferentes puntuaciones del GHQ-12 (Likert, GHQ y C-GHQ), así como las implicaciones de la adopción de diferentes puntos de corte en población general y también en población mayor de 65 años.

### Referencias

- Artazcoz, L., Borrell, C. y Benach, J. (2001). Gender inequalities in health among workers: the relation with family demands. *Journal Epidemiological Community Health*, 55, 639-647.
- Artazcoz, L., Borrell, C., Benach, J., Cortès, I. y Rohls, I. (2004). Women, family demands and health: The importance of employment status and socio-economic position. *Social Science & Medicine*, 59, 263-274.
- Borrell, C., Artazcoz, L., Gil-González, D., Pérez, G., Rohlfs, I. y Pérez, C. (2010). Perceived sexism as a health determinant in Spain. *Journal of Women's Health*, 19, 741-750.
- Borrell, C., Muntaner, C., Gil-González, D., Artazcoz, L., Rodríguez-San, M., Rohlfs, I., Pérez, K., García-Calvente, M., Villegas, R. y Álvarez-Dardet, C. (2010). Perceived discrimination and health by gender, social class, and country of birth in a Southern European country. *Preventive Medicine*, 50, 86-92.
- Bowling, A. (2007). Gender-specific and gender-sensitive associations with psychological health and morbidity in older age. Baseline findings from a British population survey of ageing. *Aging & Mental Health*, 11, 301-309.
- Campbell, A. y Knowles, S. (2007). A confirmatory factor analysis of the GHQ-12 using a large Australian Sample. *European Journal of Psychological Assessment*, 28, 2-8.
- Carretero-Dios, H. y Pérez, C. (2007). Standards for the development and review of instrumental studies: Considerations about test select in psychological research. *International Journal of Clinical and Health Psychology*, 7, 863-882.
- Costa, E., Barreto, S.M., Uchoa, E., Firmo, J.O.A. y Lima-Costa, F. (2006). Is the GDS-30 better than the GHQ-12 for screening depression in elderly people in the community? The Bambui Health Aging Study (BHAS). *International Psychogeriatrics*, 18, 493-503.
- Fogel, J. e Israel, S. (2009). Consumer attitudes regarding internet health information and communication: Gender, locus of control, and stress information. *International Journal of Clinical and Health Psychology*, 9, 275-286.
- Gill, S.C., Butterworth, P., Rodgers, B. y Mackinnon, A. (2007). 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 Research*, 152, 63-71.
- Goldberg, D.P., Gater, R., Sartorius, N., Ustun, T.B., Piccinelli, M., Gureje, O. y Rutter, C. (1997). The validity of two versions of the GHQ in the WHO study of mental health illness in general health care. *Psychological Medicine*, 27, 191-197.
- Goldberg, D.P. y Williams P. (1988). *A user's guide to the General Health Questionnaire*. UK: NFER-NELSON.
- Graetz, B. (1991). Multidimensional properties of the General Health Questionnaire. *Social Psychiatric and Psychiatric Epidemiology*, 26, 132-138.
- Hamer, M., Stamatakis, E. y Steptoe, A. (2009). Dose-response relationship between physical activity and mental health: The Scottish Health Survey. *British Journal of Sports Medicine*, 42, 1111-1114.
- Hankins, M. (2008). The reliability of the twelve-item general health questionnaire (GHQ-12) under realistic assumptions. *BMC Public Health*, 8, 355.
- Hu, Y., Stewart-Brown, S., Twigg, L. y Weich, S. (2007). Can the 12-item General Health Questionnaire be used to measure positive mental health? *Psychological Medicine*, 37, 1005-1013.
- Hewitt, C.E., Perry, A.E., Adams, B. y Gilbod, M. (2010). Screening and case finding for depression in offender populations: A systematic review of diagnostic properties. *Journal Affective Disorders* (2010). doi:10.1016/j.jad.2010.06.029
- Laaksonen, E., Martikainen, P., Lahelma, E., Lallukka, T., Rahkonen, O., Head, J. y Marmot, M. (2007). Socioeconomic circumstances and common mental disorders among Finnish

- and British public sector employees: Evidence from the Helsinki Health Study and the Whitehall II Study. *International Journal of Epidemiology*, 36, 776-786.
- López-Castedo, A. y Fernández, L. (2005). Psychometric properties of the Spanish version of the 12-item General Health Questionnaire in adolescents. *Perceptual & Motor Skills*, 100, 676-680.
- Martin, A.J. (1999). Assessing the multidimensionality of the 12-Item General Health Questionnaire. *Psychological Reports*, 84, 927-935.
- McDowell, I. (2006). *Measuring Health – a guide to rating scales and questionnaires*. 3<sup>a</sup>.ed. New York: Oxford.
- Ministerio de Sanidad y Consumo (2006). *Metodología de la Encuesta Nacional de Salud de 2006*. España. Recuperado de <http://www.msc.es/estadEstudios/estadisticas/encuestaNacional/encuesta2006.htm>.
- Molloy, G.J., Stamatakis, E., Randall, G. y Hamer, M. (2009). Marital status, gender and cardiovascular mortality: Behavioural, psychological distress and metabolic explanations. *Social Science & Medicine*, 69, 223-228.
- Montero, I. y León, O.G. (2007). A guide for naming research studies in Psychology. *International Journal of Clinical and Health Psychology*, 7, 847-862.
- Muntaner, C., Borrell, C., Benach, J., Pasarín, M. y Fernández, E. (2003). The associations of social class and social stratification with patterns of general and mental health in Spanish population. *International Journal of Epidemiology*, 32, 950-958.
- Navarro, P., Ascaso, C., Garcia-Esteve, L., Aguado, J., Torres, A. y Martin-Santos, R. (2007). Postnatal psychiatric morbidity: A validation study of the GHQ-12 and the EPDS as screening tools. *General Hospital Psychiatry*, 29, 1-7.
- Netuveli, G., Wiggins, R.D., Montgomery, S. M., Hildon, Z. y Blane, D. (2008). Mental health and resilience at older ages: Bouncing back after adversity in the British Household Panel Survey. *Journal of Epidemiology & Community Health*, 62, 987-991.
- Oviedo, H.C. y Campo-Arias, A. (2005). Aproximación al uso del coeficiente alfa de Cronbach. *Revista Colombiana de Psiquiatría*, 24, 572-580.
- Papassotiropoulos, A. y Hein, R. (1999). Screening for depression in the elderly: A study on misclassification by screening instruments and improvement of scale performance. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 23, 431-446.
- Papassotiropoulos, A., Heun, R. y Maier, W. (1997). Age and cognitive impairment influence the performance of the General Health Questionnaire. *Comprehensive Psychiatry*, 38, 335-340.
- Patel, V., Araya, R., Chowdhary, N., King, M., Kirkwood, B., Nayak, S., Simon, G. y Weiss, A. (2008). Detecting common mental disorders in primary care in India: A comparison of five questionnaires. *Psychological Medicine*, 38, 221-228.
- Ramos, P., Moreno, C., Rivera, F. y Pérez, P.J. (2010). Integrated analysis of the health and social inequalities of Spanish adolescents. *International Journal of Clinical and Health Psychology*, 10, 477-498.
- Rocha, K.B., Pérez, K., Rodríguez-Sanz, M., Borrell, C. y Obiols, E.J. (2010). Prevalencia de problemas de salud mental y su asociación con variables socioeconómicas, de trabajo y salud: Resultados de la Encuesta Nacional de Salud de España. *Psicothema*, 22, 389-395.
- Salama-Younes, M., Montazeri, L., Ismaïl, A. y Roncin, C. (2009). Factor structure and internal consistency of the 12-item General Health Questionnaire (GHQ-12) and the Subjective Vitality Scale (VS), and the relationship between them: A study from France. *Health and Quality of Life Outcomes*, 7:22.
- Sánchez-López, M.P. y Dresch, V. (2008). The 12-Item General Health Questionnaire (GHQ-12): Reliability, external validity and factor structure in the Spanish population. *Psicothema*, 20, 839-843.

- Werneke, U., Goldberg, D.P., Yalcin, I. y Ustun, B.T. (2000). The stability of the factor structure of the General Health Questionnaire. *Psychological Medicine*, 30, 823-829.
- Ye, S. (2009). Factor structure of the General Health Questionnaire (GHQ-12): The role of wording effects. *Personality and Individual Differences*, 46, 197-201.

Recibido 10 de septiembre 2009

Aceptado 8 de abril 2010

**ANEXO 1.** Distribución de la puntuación en los diferentes percentiles utilizando la puntuación Likert y GHQ.

<i>Percentil</i>	<i>Puntuación Likert</i>		<i>Puntuación GHQ</i>	
	<i>&lt; 65 años</i> <i>(n = 23.760)</i>	<i>≥ 65 años</i> <i>(n = 5.718)</i>	<i>&lt; 65 años</i> <i>(n = 23.760)</i>	<i>≥ 65 años</i> <i>(n = 5.718)</i>
1°	0	0	0	0
5°	4	4	0	0
10°	6	6	0	0
25°	7	7	0	0
50°	9	11	0	0
75°	12	14	2	3
90°	16	19	5	7
95°	19	22	7	9
99°	26	29	11	12
<i>M</i>	10,0	11,4	1,41	2,05
<i>DS</i>	4,89	5,79	2,42	3,06
<i>Varianza</i>	23,9	33,5	5,86	9,37
<i>Skeweness</i>	1,10	0,89	2,21	1,63
<i>Curtosis</i>	6,16	4,77	7,48	4,82

*Nota.* *P:* percentil; *M:* media en la muestra; *DS:* desviación estándar.