

Propiedades psicométricas del GHQ-28 en pacientes con dependencia a opiáceos

Psychometric properties of the GHQ-28 in opiate-dependent patients

PEDRO PÉREZ MORENO*; OSCAR M. LOZANO ROJAS*;
ANTONIO J. ROJAS TEJADA**

*Universidad de Huelva. Departamento de Psicología Clínica,
Experimental y Social. Área de Metodología de las Ciencias del
Comportamiento. Huelva.

**Universidad de Almería. Departamento de Ciencias Humanas y
Sociales. Área de Metodología de las Ciencias del
Comportamiento. Almería.

Enviar correspondencia a:

Óscar M. Lozano. Departamento de Psicología Clínica, Experimental
y Social. Facultad de Ciencias de la Educación. Campus de 'El
Carmen'. Avda. Fuerzas Armadas, s/n. 21071. Huelva. Tel.: 959
219199 Fax: 959 219201. E-mail: oscar.lozano@dpsi.uhu.es.

recibido: abril 2009
aceptado: noviembre 2009

RESUMEN

Objetivo: Conocer las propiedades psicométricas del GHQ-28 cuando se administra a una muestra de pacientes con dependencia a opiáceos. **Método:** 138 pacientes con diagnóstico de dependencia a opiáceos han participado en el estudio. La fiabilidad se ha estimado por el procedimiento alpha de Cronbach y dos mitades con Spearman-Brown. Se ha aplicado un análisis factorial confirmatorio para conocer la estructura factorial del instrumento. La sensibilidad y especificidad del instrumento para la detección de posibles casos psiquiátricos se ha realizado mediante el estudio de la curva COR (Característica Operativa del Receptor), usando como criterio el diagnóstico DSM-IV aportado por la Mini International Neuropsychiatric Interview. **Resultados:** Los resultados de la estimación de la fiabilidad del instrumento son adecuados. En cuanto a los resultados de validez no se ha podido confirmar la estructura de cuatro factores del GHQ-28 propuesta originariamente por Goldberg y Hillier. Los parámetros de sensibilidad y especificidad muestran unos valores óptimos de 60,7% y 73,7% respectivamente, estando correctamente clasificados el 65,9% de los participantes. **Conclusiones:** El GHQ-28 muestra una adecuada fiabilidad si bien no se ha podido confirmar la estructura factorial original. Los parámetros de sensibilidad y especificidad son apropiados. Dados estos resultados y los proporcionados por otros estudios de validez donde se muestra la relación con la severidad de la dependencia, la salud física y la calidad de vida relacionada con la salud, se puede considerar que el GHQ-28 puede ser un instrumento útil para el screening en la práctica clínica.

Palabras clave: GHQ-28, dependencia a opiáceos, estructura factorial, screening.

ABSTRACT

Objective: To know the psychometric properties of the GHQ-28 when it is applied to a sample of opiate-dependent patients. **Method:** Participants were 138 opiate-dependent patients. Reliability was estimated by means of the Cronbach's alpha and Spearman-Brown coefficients. Confirmatory Factor Analysis was applied to check the four-factor structure proposed by Goldberg and Hillier. The sensitivity and specificity of the instrument for the detection of possible psychiatric cases were measured using ROC (Receiver Operating Characteristic) analysis, using as criterion the DSM-IV diagnosis from the Mini International Neuropsychiatric Interview. **Results:** Reliability estimation was adequate. Validity results could not confirm the four-factor structure of GHQ-28 proposed by Goldberg and Hillier. The sensitivity and specificity parameters show optimum values of 60.7 % and 73.7 %, respectively, 65.9 % of the participants being correctly classified. **Conclusions:** GHQ-28 shows adequate reliability and appropriate sensitivity and specificity parameters when it is applied to opiate-dependent patients. However, the original factorial structure could not be confirmed. Given these results and those from other validity studies in which the relationships with Severity of Dependence, Physical Health and Health-Related Quality of Life were shown, the GHQ-28 can be considered a useful instrument for screening in clinical practice.

Key words: GHQ-28, factorial structure, opiate-dependent patients, screening.

INTRODUCCIÓN

La detección de pacientes con trastorno por consumo de sustancias y otros trastornos mentales concomitantes constituye uno de los principales retos para los profesionales que trabajan en el ámbito de las drogodependencias, aún más si tenemos en cuenta las altas prevalencias encontradas (Mateu, Astals y Torrens, 2005; San, Gutiérrez, Sáiz, González, Bascarán y Bobes, 2000; Torrens, 2008). Diversos estudios han puesto de manifiesto que los pacientes con comorbilidad psiquiátrica tienen un peor pronóstico terapéutico (Herbeck, Hser, Lu, Stark y Paredes, 2006; Kranzler, Del Boca y Rounsaville, 1995), una menor adherencia al tratamiento (Fernández-Montalvo, López, Landa, Illescas, Lorea y Zarzuela, 2004) y demandan más servicios asistenciales (Charney, Palacios-Boix, Negrete, Dobkin y Gill, 2005; Moggi, Hirsbrunner, Brodbeck y Bachmann, 1999).

Los profesionales de este ámbito disponen en la actualidad de una amplia batería de instrumentos de evaluación para la detección de comorbilidad psiquiátrica fundamentados en los actuales sistemas de clasificación nosológica DSM-IV y CIE-10 (para una revisión véase, por ejemplo, Forman, Svikis, Montoya y Blaine, 2004; Sáiz, García-Portilla, Bascarán, Paredes, Martínez, Jiménez et al., 2005). En términos generales, éstos muestran adecuadas propiedades psicométricas y son de gran utilidad en la práctica clínica para la realización de un diagnóstico fiable y válido. No obstante, la aplicación de estos instrumentos es compleja en determinados contextos clínicos. El elevado tiempo de aplicación de algunas de estas entrevistas y la necesidad de tener una formación específica para aplicarlas, constituye algunos de los hándicaps para su uso generalizado en la práctica clínica.

Otra de las opciones para la evaluación de problemas psiquiátricos es el uso de instrumentos de *screening*. En términos generales, estos instrumentos son fáciles de aplicar y más breves que los anteriores, si bien no ofrecen un diagnóstico psiquiátrico propiamente dicho. Su objetivo es la detección de posibles pacientes psiquiátricos, identificando quienes requieren de un estudio psicopatológico preciso posteriormente.

Probablemente, en el ámbito de las drogodependencias los dos instrumentos de *screening* más utilizados para la detección de posibles pacientes con trastornos en el eje I son el *Symptom Checklist SCL-90-R* (Derogatis, 1988) y el *General Health Questionnaire* (GHQ) (Goldberg, 1972). Además, es frecuente encontrar ambos instrumentos como indicadores de resultado del estado emocional. En este sentido, el GHQ-28 está incorporado dentro de la estructura dimensional del OTI (Darke, Hall, Wodak, Heather y Ward, 1992), mientras el SCL-90-R ha sido utilizado con el EuropASI (Kokkevi y Hartgers, 1995).

El *General Health Questionnaire* (GHQ) es un instrumento desarrollado para la detección de posibles casos con trastornos mentales de tipo no psicótico en los servicios de asistencia primaria (Iraurgi, 2002). Originariamente este instrumento estaba formado por 60 ítems, si bien posteriormente se han desarrollado versiones más cortas de 30, 28,

20 y 12 ítems. Entre estas versiones reducidas, el GHQ-28 presenta propiedades por las que ha resultado interesante su adaptación a varios idiomas (Molina, Andrade-Rosa, González-Parra, Blasco-Fontecilla, Real y Pintor, 2006): su tiempo de aplicación de entre tres y cinco minutos (Lobo, Pérez-Echevarría y Artal, 1986), su facilidad para ser aplicado en servicios de atención primaria (Seva, Magallón, Sarasola y Merino, 1992) y la posibilidad de obtener la interpretación de cuatro factores además de la puntuación general (Goldberg y Hillier, 1979).

Originariamente, Goldberg y Hillier (1979) propusieron la estructura de cuatro factores del GHQ-28 a partir de los resultados de un análisis de componentes principales de la versión originaria de 60 ítems. Los resultados de este estudio mostraron que los ítems se agruparon en cuatro componentes principales con siete ítems cada uno: síntomas somáticos, ansiedad, disfunción social y depresión. No obstante, otros estudios posteriores han mostrado inconsistencias en su estructura factorial. Molina et al. (2006) replicaron el estudio de Goldberg y Hillier (1979) con la versión española de 60 ítems aplicada a una muestra de pacientes de servicios ambulatorios de atención primaria. Estos autores llegaron a una solución final de 28 ítems repartidos en cuatro factores que explicaban el 54,6% de la varianza, si bien los ítems que constituyen esta estructura no son los mismos que los originales. En esta nueva escala aparecen los factores de síntomas somáticos, disfunción social y depresión, aunque con ítems distintos a la versión original. Malakouti, Fatollahi, Mirabzadeh y Zandi (2006) aplicaron la versión iraní del GHQ-28 a una muestra aleatoria de personas mayores seleccionada en un distrito de Teherán (Irak). El estudio muestra la existencia de cuatro factores que explican el 58,9% de la varianza; sin embargo, los ítems asociados a cada factor no son los mismos que en la versión original de Goldberg y Hillier (1979), encontrándose en el primer factor 12 ítems y en el cuarto factor cuatro. Gibbons, Flores de Arévalo y Mónico (2004), en un estudio realizado con estudiantes universitarios de El Salvador que responden a la versión española del GHQ-28 (Muñoz, Vázquez, Pastrana, Rodríguez, y Oneca, 1978), encuentran que tras la rotación varimax los cuatro factores extraídos explican el 54% de la varianza, apareciendo un alto grado de saturación en 9 de los 28 ítems, especialmente entre los de la escala de ansiedad y la de síntomas somáticos. Werneke, Goldberg, Yalcin y Üstün, (2000) analizan la estructura de cuatro factores en un estudio multicéntrico. Tras la rotación varimax encuentran que la extracción de cuatro factores explica el 57,3% de la varianza observada, con las dimensiones disfunción social y depresión estables, aunque no así las de síntomas somáticos y ansiedad.

Los parámetros de sensibilidad y especificidad de este instrumento avalan la aplicación del GHQ-28 como instrumento de *screening* (Andersen, Sestoft, Lillebaek, Gabrielsen, y Hemmingsen, 2002; Goldberg, Gater, Sartorius, Ustun, Piccinelli, Gureje et al., 1997; Likouras, Adrachta, Kalfakis, Oulis, Voulgari, Christodoulou et al., 1996) y, en términos generales, las diferencias de unos estudios a otros pueden deberse a las diferentes muestras utilizadas o a la entrevista diagnóstica utilizada como criterio. Lobo, Pérez-Echevarría y Artal (1986) con la versión española del GHQ-28 informaron de una sen-

sibilidad del 76,9% y especificidad del 90,2% para un punto de corte 6/7 y de una sensibilidad del 84,6% y especificidad del 82% para un punto de corte de 5/6.

A pesar de su uso en el ámbito de las drogodependencias, las propiedades psicométricas del GHQ-28 no han sido estudiadas en pacientes con problemas de adicción. Los puntos de corte utilizados para la detección de posibles casos psiquiátricos han sido adaptados a partir de los estudios realizados con otros grupos poblacionales (Goldberg y Williams, 1991; Taylor, Dal Grande y Parsons, 1997). Este hecho contrasta con las recomendaciones de la *American Psychological Association* (APA, 2002), que establece la necesidad de acumular evidencias sobre la población de interés para realizar una correcta interpretación de las puntuaciones.

En este sentido, el presente estudio tiene como objetivo mostrar las propiedades psicométricas del GHQ-28 cuando se aplica sobre una muestra de pacientes con dependencia a opiáceos. Más específicamente, los objetivos planteados son: a) estimar la fiabilidad del GHQ-28 cuando es aplicada sobre una muestra de pacientes con dependencia a opiáceos; b) aportar evidencias de validez sobre esta muestra, centrándose en el análisis de la estructura factorial y la detección de un punto de corte óptimo de sensibilidad y especificidad.

MÉTODO

Participantes

La muestra está formada por 138 pacientes con diagnóstico de dependencia a opiáceos, siendo éste uno de los criterios de inclusión en el estudio. El otro criterio para la participación era firmar un consentimiento informado en el que los pacientes mostraban su conformidad para que los datos fueran utilizados con un objetivo exclusivamente estadístico.

El 95,6% son hombres y un 4,4% mujeres. La media de edad de los participantes en el estudio era de 36,1 años (DT=6,6). La mayoría de los pacientes no tenían estudios o como máximo habían realizado los estudios primarios (72,3%), un 16,2% tenía el título de formación profesional y el 11,5% había finalizado BUP/COU o estudios superiores. En relación a la situación laboral, un 13,6% se encontraba trabajando. El 44,2% no tenía incidencias legales en el momento de participar en el estudio.

Un 12,3% de los pacientes realizaba el tratamiento en un centro ambulatorio, mientras que el porcentaje restante estaba en régimen de internamiento en comunidades terapéuticas. Todos los pacientes se encontraban en programa de tratamiento con metadona (PTM), siendo el tiempo medio de 35,9 meses (DT=35,7). La dosis de metadona que tomaban estaba comprendida entre los 10 y los 100 mg., con una media de 46,9 mg. (DT=20,4).

Otras sustancias consumidas durante el mes previo a la entrevista son cocaína (61,3%), alcohol (59,2%), cannabis (52,9%), y opiáceos distintos de la metadona (51,8%).

Instrumentos

Se aplicó una batería de preguntas en la que se recogió información sobre variables sociodemográficas, la historia de consumo y tratamiento de los pacientes, el estado de salud médico, la calidad de vida relacionada con la salud y la presencia de trastornos mentales.

Para el presente estudio, los instrumentos de interés aplicados son el GHQ-28 y la *Mini International Neuropsychiatric Interview* -MINI-.

El GHQ-28 es un test autoadministrado de 28 ítems. Está compuesto de cuatro subescalas que miden: a) síntomas somáticos de origen psicológico (ítems 1 al 7); b) ansiedad (ítems 8 al 14); c) disfunción social (ítems 15 al 21); y, d) depresión (ítems 22 al 28). Los 28 ítems son evaluados con un formato de respuesta tipo Likert con puntuaciones comprendidas entre 0 (mejor que lo habitual) hasta 3 (mucho peor que lo habitual). El sistema de puntuación consiste en otorgar valores de '0' y '1' a las categorías de respuesta siguiendo la secuencia '0,0,1,1'. En la versión adaptada a la población española por Lobo et al. (1986) se recomienda un punto de corte 5/6 para la puntuación total de la escala. Por encima de estos valores estaríamos ante un probable caso psiquiátrico.

La entrevista diagnóstica utilizada para el diagnóstico de trastornos mentales del eje I fue la MINI (Sheehan, Lecrubier, Sheehan, Amorim, Janavs, Weiller et al., 1998). Esta entrevista estructurada puede realizar diagnósticos de acuerdo con los sistemas DSM-IV o CIE-10. En este estudio fue utilizada para la valoración de los trastornos psiquiátricos del eje I según criterios diagnósticos del DSM-IV.

Procedimiento

La recogida de datos fue llevada a cabo por miembros del equipo de investigación. Se realizaron entrevistas individuales desarrolladas en los centros terapéuticos a los que pertenecían los pacientes. Antes de iniciar la recogida de información, a los pacientes se les informaba que la administración de los tests formaba parte de un estudio, explicándoles los objetivos del mismo. Se les decía expresamente que su participación era voluntaria, haciéndoles saber que la información recopilada tendría un tratamiento meramente estadístico, y se les pedía que firmaran un consentimiento informado.

La entrevista MINI fue aplicada a una submuestra de participantes extraída al azar de 50 participantes. De éstos, tres no contestaron a todas las preguntas, por lo que han sido excluidos de los análisis.

Análisis estadístico

El cálculo de la fiabilidad se ha realizado mediante el procedimiento de consistencia interna, calculando el coeficiente Alfa de Cronbach de cada subtest y del conjunto del GHQ-28. Asimismo, se estimó la fiabilidad mediante el procedimiento de dos mitades de Spearman-Brown, considerando una mitad

a los ítems con orden par tras ser graduados en función de su correlación ítem-total corregida, y como otra mitad a los ítems con orden impar. Por último, se estimó también la fiabilidad del test como un compuesto mediante el coeficiente de Raju (1977).

El estudio de la estructura latente del GHQ-28 se ha realizado mediante análisis factorial confirmatorio. Se comprobó la estructura unidimensional de cada subtest así como el modelo de cuatro factores de Goldberg y Hillier (1979) y un modelo de tres factores basado en Gibbons et al. (2004) y Werneke et al. (2000). La estimación se realizó mediante el método ML robusto (Máxima Verosimilitud corregido para distribuciones donde no puede ser asumida la normalidad multivariada). Como prueba de bondad de ajuste del modelo se empleó la Chi-Cuadrado robusta de Satorra-Bentler (SB^2 , Satorra y Bentler, 1994). Además, debido a la sensibilidad de dicha prueba a factores como el tamaño de la muestra y las desviaciones con respecto al cumplimiento de los supuestos de linealidad, normalidad multivariada y aditividad (Jöreskog y Sörbom, 1989), se interpretaron también la razón entre el valor SB^2 y sus grados de libertad (SB^2/df), considerando que el ajuste es adecuado en caso de contar con un valor igual o inferior a 3 (Carmines y McIver, 1981). Otros indicadores de ajuste interpretados fueron el NFI (Normed Fit Index, índice de ajuste normativo; Bentler y Bonnet, 1980) aunque subestima el valor de ajuste en muestras pequeñas, NNFI (Non-Normed Fit Index, índice de ajuste no-normativo; Bentler y Bonnet, 1980), CFI (Comparative Fit Index, índice de ajuste comparativo; Bentler, 1990) e IFI (Incremental Fit Index, índice de ajuste incremental, Bollen, 1989), todos ellos índices comparativos, así como el GFI (Goodness of Fit Index, Índice de Bondad de Ajuste, Jöreskog y Sörbom, 1989), índice de ajuste global, aunque presenta sensibilidad a las desviaciones de la normalidad multivariada. Para estos cinco índices se consideraron valores superiores a 0,90 como indicadores de un ajuste adecuado. También se interpretó el valor RMSEA (Root Mean Square Error of Approximation, raíz media cuadrática del error de aproximación; Byrne, 2001; Steiger, 1990), considerando indicador de un ajuste adecuado todo valor igual o inferior a 0,08 (Browne y Cudeck, 1993).

El análisis de las curvas COR (Característica Operativa del Receptor) se ha realizado utilizando como criterio la presencia de trastornos mentales diagnosticados con la MINI, y bajo el supuesto de una distribución no paramétrica y un nivel de confianza del 95%.

El software estadístico utilizado para la estimación de la fiabilidad y el análisis de las curvas COR ha sido el SPSS versión 15.0. El análisis factorial confirmatorio se ha realizado con el programa EQS 6.0 (Bentler, 2005).

RESULTADOS

Estimación de la fiabilidad

La fiabilidad se ha estimado mediante el coeficiente alfa de Cronbach y por el procedimiento de dos mitades. La fiabilidad estimada a través del alfa de Cronbach aportó valo-

res de 0,826 para la subescala de síntomas somáticos, 0,859 para la subescala de ansiedad, 0,862 para la subescala de disfunción social y 0,834 para la subescala de depresión. La fiabilidad del instrumento se calculó como fiabilidad de un compuesto (Raju, 1977) aportando un valor de 0,938.

Para la estimación de la fiabilidad por el procedimiento de dos mitades, los ítems fueron ordenados según su valor de correlación ítem-total corregido. Se utilizó como primera mitad los ítems pares, y la otra mitad los ítems impares. Los valores aportados por el coeficiente de Spearman-Brown fueron de 0,725 para la subescala de síntoma somáticos, 0,857 para la de ansiedad, 0,851 para la de disfunción social y 0,919 para la de depresión. La estimación para el GHQ-28 fue de 0,948.

A la vista de estos resultados debemos considerar que la fiabilidad de las subescalas y del instrumento en su conjunto es adecuada.

Evidencias de validez

Estructura factorial

En la comprobación de la estructura unidimensional de cada subtest se obtuvieron indicadores de un ajuste inadecuado en las dimensiones de Síntomas Somáticos, Ansiedad y Depresión y una mayoría de indicadores adecuados en el caso de la dimensión de Disfunción Social (ver Tabla 1, modelo inicial). Las dimensiones con un mal ajuste se reespecificaron (Byrne, 2001) incluyendo en sus modelos como parámetros a estimar las covarianzas entre errores de medida de los ítems cuyos residuos se habían mostrado estadísticamente significativos, hasta alcanzar un ajuste adecuado en la mayoría de indicadores (ver Tabla 1, modelo final). En este proceso se encontraron residuos de covarianzas entre errores de medida estadísticamente significativos en los pares de ítems "¿Ha padecido dolores de cabeza?" y "¿Ha tenido sensación de opresión en la cabeza, o de que la cabeza le va a estallar?" del subtest de Síntomas Somáticos, "¿Sus preocupaciones le han hecho perder mucho sueño?" y "¿Ha tenido dificultades para seguir durmiendo toda la noche?" en el subtest de Ansiedad y entre "¿Ha pensado en la posibilidad de "quitarse de en medio"?" y "¿Ha notado que la idea de quitarse la vida le viene repetidamente a la cabeza?", entre éste y "¿Ha notado que desea estar muerto y lejos de todo?", y entre éste y "¿Ha pensado en la posibilidad de "quitarse de en medio"?" en el subtest de Depresión. Puede apreciarse que en los cuatro modelos resultantes, salvo la significación estadística de los contrastes $SB\chi^2$ de cada uno de ellos (prueba que, por otra parte se muestra sensible ante factores como el tamaño de la muestra o las desviaciones de los supuestos) y el valor RMSEA de la dimensión de Depresión (superior a 0,10) todos los índices reflejan un ajuste adecuado. Se comprobó si la estimación de estos parámetros mejoraba significativamente la bondad de ajuste de los modelos mediante el contraste del incremento en el valor Chi-cuadrado ($\Delta SB\chi^2$), resultando una mejora estadísticamente significativa ($p < 0,05$) en todos los casos.

Tabla 1. Indicadores de ajuste de los cuatro subtest del GHQ-28

	Síntomas Somáticos		Ansiedad		Disfunción Social		Depresión	
	Inicial	Final	Inicial	Final	Inicial	Inicial	Final	
SB χ^2	58,964	23,811	43,867	22,768	25,572	61,071	26,493	
g.l.	14	13	14	13	14	14	11	
Significación	0,000	0,033	0,000	0,045	0,029	0,000	0,005	
(SB χ^2)/gl	4,212	1,832	3,133	1,751	1,827	4,362	2,408	
NFI	0,815	0,925	0,896	0,946	0,929	0,912	0,962	
NNFI	0,773	0,941	0,888	0,961	0,949	0,895	0,956	
GFI	0,862	0,945	0,909	0,952	0,930	0,805	0,926	
CFI	0,849	0,964	0,925	0,976	0,966	0,930	0,977	
IFI	0,852	0,965	0,927	0,976	0,967	0,931	0,977	
RMSEA	0,153	0,078	0,125	0,074	0,078	0,157	0,101	
90% RMSEA	0,113	0,022	0,084	0,011	0,024	0,017	0,052	
	0,193	0,126	0,167	0,123	0,124	0,196	0,151	

SB χ^2 : Chi-cuadrado de Satorra-Bentler; NFI: Normed Fit Index; NNFI: Non-Normed Fit Index;

GFI: Goodness of Fit Index; CFI: Comparative Fit Index; IFI: Incremental Fit Index;

RMSEA: Root Mean Square Error of Approximation; 90% RMSEA: Intervalo de confianza al 90% del RMSEA

Tras la comprobación de la estructura unidimensional de cada subescala se procedió a contrastar dos modelos entre sí: el propuesto por Goldberg y Hillier (1979) en el que cada subtest es una puntuación independiente, estando todos correlacionados entre sí, y la propuesta por Gibbons et al. (2004) y Werneke et al. (2000) donde los factores de Síntomas Somáticos y Ansiedad forman un único subtest, estando compuesta la escala, por lo tanto, por tres puntuaciones con sentido propio. Al haberse detectado la existencia de covarianzas entre errores de medida estadísticamente significativas, se liberaron dichos parámetros para que participasen en ambas estructuras.

En la Tabla 2 se muestran los indicadores de ajuste de ambos modelos. Podemos apreciar que ambos cuentan con indicadores de ajuste poco adecuados (contraste SB χ^2 , GFI y NFI, siendo los dos primeros sensibles a las desviaciones de los supuestos estadísticos, mientras que el tercero infraestima el valor de ajuste en muestras pequeñas) así como una

mayoría de indicadores adecuados (SB χ^2 /gl, NNFI, CFI, IFI y RMSEA). Por lo tanto, podemos considerar que existe la suficiente evidencia empírica favorable sobre la validez de estas estructuras. No obstante, la comparación entre ambos valores SB χ^2 muestran que la ganancia en ajuste del modelo de cuatro factores es estadísticamente significativa (Δ SB $\chi^2_{(3)} = 20,48$; $p < 0,05$). Por lo tanto, nada se opone a afirmar que el ajuste del modelo de cuatro factores es significativamente mayor que el de tres factores. En este modelo todos los ítems cuentan con saturaciones estadísticamente significativas ($p < 0,05$) estando cada uno relacionado con el factor propuesto por Goldberg y Hillier (1979).

Sensibilidad y especificidad

Respecto a los trastornos mentales de los pacientes (Tabla 3), el 59,6% de los pacientes entrevistados presentaban al menos el diagnóstico de un trastorno mental del eje

Tabla 2. Indicadores de ajuste del modelo del GHQ-28 con cuatro (Goldberg y Hillier, 1979) y con tres factores (Gibbons et al., 2004; y Werneke et al., 2000)

Modelo	SB χ^2	g.l.	Sig	(SB χ^2)/gl	NFI	NNFI	GFI	CFI	IFI	RMSEA	90% RMSEA
Cuatro factores	504,07	339	0,000	1,487	0,797	0,912	0,771	0,921	0,923	0,060	0,048 0,070
Tres factores	524,55	342	0,000	1,534	0,788	0,904	0,764	0,913	0,915	0,062	0,051 0,073

SB χ^2 : Chi-cuadrado de Satorra-Bentler; NFI: Normed Fit Index; NNFI: Non-Normed Fit Index;

GFI: Goodness of Fit Index; CFI: Comparative Fit Index; IFI: Incremental Fit Index;

RMSEA: Root Mean Square Error of Aproximation; 90% RMSEA: Intervalo de confianza al 90% del RMSEA

I. Los trastornos con mayor prevalencia fueron el trastorno distímico actual, el riesgo de suicidio (leve o alto), y el trastorno de ansiedad generalizada, presentes en un 29,8% de la muestra.

Tabla 3. Distribución de frecuencias de los trastornos psiquiátricos del Eje I encontrados en los pacientes (n=47).

<i>Diagnóstico psiquiátricos del eje I</i>	
• Trastorno Distímico Actual	14 (29,8%)
• Riesgo de Suicidio (leve + alto)	14 (29,8%)
• Trastorno de Angustia actual	4 (8,5%)
• Trastorno de Ansiedad Generalizada	14 (29,8%)
• Agorafobia	10 (21,3%)
• Fobia Social	5 (10,6%)
• Trastorno de Estrés Postraumático	1 (2,1%)
• Bulimia Nerviosa	1 (2,1%)

La capacidad del GHQ-28 de discriminar entre pacientes con alteraciones psiquiátricas y sin alteraciones se ha estimado mediante la curva COR. El área bajo la curva resultante es de 0,684 (IC 95%: 0,529 a 0,839), siendo este valor estadísticamente significativo. El punto de corte más adecuado se sitúa en el valor 7/8 con una sensibilidad de 60,7% y una especificidad de 73,7% (Tabla 4).

Tabla 4. Parámetros de sensibilidad y especificidad

Punto de Corte	Sensibilidad (%)	Especificidad (%)	Correctamente clasificados (%)
0/1	89,3	5,3	55,3
1/2	89,3	15,8	59,6
2/3	85,7	21,1	59,6
3/4	85,7	26,3	61,7
4/5	82,1	31,6	61,7
5/6	78,6	42,1	63,8
6/7	71,4	57,9	65,9
7/8	60,7	73,7	65,9
8/9	57,1	78,9	65,9
9/10	53,6	84,2	65,9
10/11	53,6	84,2	65,9
11/12	50,0	84,2	63,8
12/13	46,4	84,2	61,7
13/14	42,9	84,2	59,6
14/15	39,3	84,2	57,5
15/16	35,7	84,2	55,3

Si establecemos como punto de corte el valor 8, el porcentaje de pacientes en la muestra con probables casos psiquiátricos representa un 46,8%, mientras que con el valor 7 los probables casos psiquiátricos se incrementan a un 59,6%. Esta última prevalencia coincide con la aportada por la MINI, si bien el porcentaje de diagnósticos coincidentes entre ambos instrumentos es del 65,9%.

DISCUSIÓN

El objetivo del presente estudio ha sido aportar evidencias de las propiedades psicométricas del GHQ-28 cuando se aplica sobre una muestra de pacientes con diagnóstico de adicción a opiáceos, centrándonos en el análisis de la fiabilidad, la estructura factorial del GHQ-28 y la sensibilidad y especificidad para la detección de posibles casos psiquiátricos.

Los coeficientes de fiabilidad estimados con los distintos procedimientos han aportado resultados positivos, tanto para cada una de las subescalas como para el instrumento en su conjunto.

En cuanto a las evidencias de validez centradas en la estructura factorial del instrumento, hemos optado por someter a contraste la estructura de cuatro factores propuesta por Goldberg y Hillier (1979) y la de tres factores basada en los resultados de Gibbons et al. (2004) y Werneke et al. (2000). Como se ha podido apreciar, los resultados reflejan que ninguna de las dos estructuras factoriales se confirman sobre esta muestra de pacientes, si bien la estructura de cuatro factores presenta un ajuste significativamente mejor que la de tres factores.

Una de las explicaciones a esta falta de ajuste es la existencia de covariación entre el error de algunos ítems. Esto ocurre en el factor Síntomas Somáticos con los ítems '*¿Ha padecido dolores de cabeza?*' y '*¿Ha tenido sensación de opresión en la cabeza, o de que la cabeza le va a estallar?*', en el factor Ansiedad con los ítems '*¿Sus preocupaciones le han hecho perder mucho sueño?*' y '*¿Ha tenido dificultades para seguir durmiendo de un tirón toda la noche?*', y en el de Depresión con los ítems '*¿Ha pensado en la posibilidad de quitarse de en medio?*', '*¿Ha notado que desea estar muerto y lejos de todo?*' y '*¿Ha notado que la idea de quitarse la vida le viene repetidamente a la cabeza?*'.

Para entender porqué se produce esta covariación debemos fijarnos en el contenido de estos ítems, apreciándose que los enunciados son similares dentro de cada factor. Si, como se ha visto en la descripción de la muestra la mayoría de los participantes tienen un bajo nivel de estudios, es posible que se esté produciendo confusión en la comprensión de los ítems. Así, si los pacientes en distintos ítems están entendiendo un mismo contenido, lo lógico es que las respuestas sean iguales y por tanto aparezca covariación entre los errores.

Esta hipótesis tiene aún más sentido si apreciamos los resultados encontrados en otros estudios, ya que son otros los ítems con resultados no acordes a la propuesta de Goldberg y Hillier (1979). Así, por ejemplo, la muestra empleada por Werneke et al. (2000) registra, en sus palabras, un amplio rango de clases sociales. Por otra parte, la muestra empleada por Gibbons et al. (2004) es representativa de la población de estudiantes de la Universidad Tecnológica de El Salvador. Por el contrario, la muestra empleada en este estudio se compone de un 72,3% de personas con un nivel educativo máximo de estudios primarios, un 16,2% de personas

con título de formación profesional y un 11,5% de personas con BUP o COU terminados.

En relación al uso del GHQ-28 como instrumento de *screening*, los resultados encontrados reflejan que el mayor equilibrio entre la sensibilidad y la especificidad se encuentra en el valor comprendido entre 7/8. Con este valor se diagnóstica correctamente (siguiendo el diagnóstico de la MINI) a un 65,9% de los pacientes que han participado en este estudio. En este sentido, los autores consideran que este porcentaje justifica adecuadamente la utilidad del GHQ-28 como instrumento de *screening* en la práctica clínica. Los casos que pueden presentar más dudas acerca del diagnóstico serían aquellos a los que habría que someter a pruebas diagnósticas complementarias.

Al comparar el punto de corte y el número de casos identificados correctamente en este estudio con los resultados de otros autores, se puede apreciar que el funcionamiento del GHQ-28 es más deficiente en este estudio. Como se ha dicho anteriormente, en el estudio de originario de Lobo et al. (1986) encontraron que el valor 5/6 detectaba correctamente el 83% de los casos, mientras que el valor 6/7 a un 85%. Aspectos metodológicos pueden ocasionar las diferencias encontradas en este estudio con respecto a los citados anteriormente. No obstante, otra posibilidad es que pueda existir un mimetismo de síntomas entre los ocasionados por el consumo de drogas y los síntomas de otros trastornos psiquiátricos independientes, pues se aprecian correlaciones estadísticamente significativas entre las puntuaciones del GHQ-28 y el consumo de sustancias (Lozano, Rojas, Pérez, González-Sáiz, Ballesta, y Bilbao, 2008). Este motivo justificaría que, en términos estadísticos, la precisión para detectar un punto de corte válido resulte más compleja, redundando negativamente sobre esta. Igualmente, ello justificaría que el punto de corte más adecuado de sensibilidad y especificidad se sitúe por encima del hallado en otras poblaciones.

En cuanto a las limitaciones del estudio, por una parte consideramos que la extrapolación de los resultados encontrados sobre esta muestra a otros grupos de consumidores debe realizarse con precaución. Debemos tener en cuenta que se trata de pacientes consumidores de opiáceos que, en buena medida, presentan características distintas en cuanto al perfil de pacientes con dependencia a otras sustancias psicoactivas (García, Secades, Fernandez, Carballo, Errasti y Al-Halabi, 2005; Le Bon, Basiaux, Streel, Tecco, Hanak, Hansenne et al., 2004). En cualquier caso, los pacientes con dependencia a opiáceos constituyen un elevado porcentaje de pacientes admitidos cada año a tratamiento en España (OED, 2007), de ahí la relevancia de este estudio. Además, cabe decir que el estudio tomado como referencia de Lobo et al. (1986) estaba realizado sobre pacientes que acudían a atención primaria, cuyo perfil generalmente difiere de los pacientes atendidos en los centros especializados de drogodependencias.

Por otro lado, consideramos que el tamaño de la muestra empleado para la detección del punto de corte representa otra limitación del estudio. Por ello, para salvar esta limitación, se ha optado por usar un procedimiento de estimación no paramétrico. Este procedimiento no hace asunción alguna

sobre la distribución de los datos y es más robusto a tamaños de muestras pequeños.

Teniendo en cuenta los resultados encontrados, pensamos que el GHQ-28 es un instrumento que puede ser utilizado en el ámbito de las drogodependencias con ciertas garantías métricas. Por un lado, en este estudio se han aportado evidencias de que el instrumento presenta una adecuada fiabilidad para esta población. En cuanto a la validez del instrumento, otros estudios reflejan relaciones entre las puntuaciones del GHQ-28 y otros constructos relacionados teóricamente como la severidad de la dependencia, la salud física y la calidad de vida relacionada con la salud (Lozano et al., 2008). Sobre la estructura factorial, la no confirmación de la estructura factorial lleva a que los autores recomienden prudencia en la interpretación de las puntuaciones correspondientes a cada subescala, siendo necesaria más evidencias (con mayores tamaños muestrales y distintos diagnósticos de dependencia) para rechazar o no la estructura dimensional teórica propuesta por los autores del instrumento. Sin embargo, los parámetros de sensibilidad y especificidad encontrados para la puntuación total, si bien son más bajos que los vistos en otras poblaciones, permiten la clasificación correcta de más del 65% de los pacientes. Si tenemos en cuenta que el tiempo de aplicación de este instrumento es breve, consideramos que el GHQ-28 puede ser un instrumento útil para la práctica clínica, siendo necesario complementarlos con otras pruebas en aquellos casos que existan dudas.

REFERENCIAS

- American Educational Research Association, American Psychological Association, y National Council on Measurement in Education. (2002). *Standards for educational and psychological testing*. Washington DC: American Educational Research Association.
- Andersen, H. S., Sestoft, D., Lillebaek, T., Gabrielsen, G. y Hemmingsen, R. (2002). Validity of the General Health Questionnaire (GHQ-28) in a prison population: data from a randomized sample of prisoners on remand. *International Journal of Law and Psychiatry*, 25, 573-580.
- Bentler, P. M. (2005). *EQS 6 Structural equations program manual*. Encino, CA: Multivariate Software.
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107, 238-246.
- Bentler, P. M. y Bonnet, D. G. (1980). Significance test and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88, 588-606.
- Bollen, K. A. (1989). A new incremental fit index for general structural equation models. *Sociological Methods & Research*, 17, 303-316.
- Browne, M. W. y Cudeck, R. (1993). Alternative ways of assessing model fit. En K.A. Bollen y J.S. Long (Eds.), *Testing structural equation models* (pp. 136-162). Newbury Park: Sage.
- Byrne, B. M. (2001). *Structural equation modeling with AMOS: basic concepts, applications, and programming*. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Carmine, E. G. y McIver, J. P. (1981). Analyzing models with unobserved variables. En G. W. Bohrnstedt y E. F. Borgatta (Eds.). *Social measurement: Current issues* (pp. 122-133). Beverly Hills: Sage.

- Charney, D. A., Palacios-Boix, J., Negrete, J. C., Dobkin, P. L. y Gill, K. J. (2005). Association between concurrent depression and anxiety and six-month outcome of addiction treatment. *Psychiatric Services*, 56, 927-933.
- Darke, S., Hall, W., Wodak, A., Heather, N. y Ward, J. (1992). Development and validation of a multidimensional instrument for assessing outcome of treatment among opiate users: the Opiate Treatment Index. *British Journal of Addictions*, 87, 733-742.
- Derogatis L. R. (1998). *SCL-90-R: administration scoring and procedures manual*. Towson: Clinical Psychometric Research.
- Fernández-Montalvo, J., López, J. J., Landa, N., Illescas, C., Lorea, I. y Zarzuela, A. (2004). Trastornos de personalidad y abandonos terapéuticos en pacientes adictos: resultados en una comunidad terapéutica. *International Journal of Clinical and Health Psychology*, 4, 271-283.
- Forman, R., Svikis, D., Montoya, I. D. y Blaine, J. (2004). Selection of a substance use disorder diagnostic instrument by the National Drug Abuse Treatment Clinical Trials Networks. *Journal of Substance Abuse Treatment*, 27, 1-8.
- García, O., Secades, R., Fernández, J. R., Carballo, J. L., Errasti, J. M. y Al-Halabi, S. (2005). Comparación de pacientes cocaínómanos y heroínómanos en el EuropASI. *Adicciones*, 17, 33-42.
- Gibbons, P., Flores de Arévalo, H., y Mónico, M. (2004). Assessment of the factor structure and reliability of the 28 item version of the General Health Questionnaire (GHQ-28) in El Salvador. *International Journal of Clinical and Health Psychology*, 4, 389-398.
- Goldberg, D. P. (1972). *The detection of psychiatric illness by questionnaire*. London: Oxford University Press.
- Goldberg, D. P., Gater, R., Sartorius, N., Ustun, T. B., Piccinelli, M., Gureje, O. y Rutter, C. (1997). Validity of two versions of the GHQ in the WHO study of mental illness in general health care. *Psychological Medicine*, 27, 191-197.
- Goldberg, D. P. y Hillier, V. F. (1979). A scaled version of the General Health Questionnaire. *Psychological Medicine*, 9, 139-145.
- Goldberg, D.P. y Williams, P. (1991). *A user's guide to the General Health Questionnaire*. NFER-Nelson: Windsor.
- Herbeck, D. M., Hser, Y., Lu, T. H., Stark, M. E. y Paredes, A. (2006). A 12-year follow-up study of psychiatric symptomatology among cocaine-dependent men. *Addictive Behaviors*, 31, 1974-1987.
- Iraurgi, I. (2002). Instrumentos de Evaluación de la Calidad de Vida Relacionada con la Salud en Toxicomanías. En I. Iraurgi y F. González-Saiz (Eds.). *Instrumentos de evaluación en drogodependencias* (pp. 481-511). Madrid: Aula Médica.
- Jöreskog, K. G. y Sörbom, D. (1989). *LISREL 7 User's Reference Guide*. Chicago: Scientific Software, Inc.
- Kokkevi, A., y Hartgers, C. (1995). European Adaptation of a multidimensional assessment instrument for drug and alcohol dependence. *European Addiction Research*, 1, 208-210.
- Kranzler, H., Del Boca, F. y Rousanville, B. (1995). Comorbid psychiatric predicts three-year outcomes in alcoholics: A posttreatment natural history study. *Journal of Studies on Alcohol*, 3, 619-626.
- Le Bon, O., Basiaux, P., Streel, E., Tecco, J., Hanak, C., Hansenne, M., et al. (2004). Personality profile and drug of choice; a multivariate analysis using Cloninger's TCI on heroin addicts, alcoholics and a random population group. *Drug and Alcohol Dependence*, 73, 175-182.
- Lobo, A., Pérez-Echevarría, M. J. y Artal, J. (1986). Validity of the scaled version of the General Health Questionnaire (GHQ-28) in a Spanish population. *Psychological Medicine*, 16, 135-140.
- Lozano, O. M., Rojas, A. J., Pérez, C., González-Sáiz, F., Ballesta, R. y Bilbao, A. (2008). Evidencias de validez del test para la evaluación de la calidad de vida en adictos a sustancias psicoactivas a partir del modelo biaxial de la adicción. *Psicothema*, 20, 311-317.
- Likouras, L., Adrachta, D., Kalfakis, N., Oulis, P., Voulgari, A., Christodoulou, G. N., et al. (1996). GHQ-28 as an aid to detect mental disorders in neurological inpatients. *Acta Psychiatrica Scandinavica*, 93, 212-216.
- Malakouti, S. K., Fatollahi, P., Mirabzadeh, A. y Zandi, T. (2006). Reliability, validity and factor structure of the GHQ-28 used among elderly Iranians. *International Psychogeriatrics*, 19, 623-634.
- Mateu, G., Astals, M. y Torrens, M. (2005). Comorbilidad psiquiátrica y trastorno por dependencia de opiáceos: del diagnóstico al tratamiento. *Adicciones*, 17, 111-121.
- Moggi, F., Hirsbrunner, H. P., Brodbeck, J. y Bachmann, K. M. (1999). One-year outcome of an integrative inpatient treatment for dual diagnosis patients. *Addictive Behaviors*, 24, 589-592.
- Molina, J. D., Andrade-Rosa, C., González-Parra, S., Blasco-Fontecilla, H., Real, M. A. y Pintor, C. (2006). The factor structure of the General Health Questionnaire (GHQ): a scaled version for general practice in Spain. *European Psychiatry*, 21, 478-486.
- Muñoz, P. E., Vázquez, J. L., Pastrana, E., Rodríguez, F. y Oneca, C. (1978). Study of the validity of Goldberg's 60-item GHQ in its Spanish version. *Social Psychiatry*, 13, 99-104.
- Observatorio Español sobre Drogas. (2007). *Informe 2007 del Observatorio Español sobre Drogas. Situación y tendencias de los problemas de drogas en España*. Madrid: Ministerio de Sanidad y Consumo.
- Raju, N. S. (1977). A generalization of Alpha Coefficient. *Psychometrika*, 42, 549-565.
- Sáiz, P. A., García-Portilla, P. M^a, Bascarán, M^a. T., Paredes, M^a. B., Martínez, S., Jiménez, L. y Bobes, J. (2005). Instrumentos de evaluación de la dependencia de heroína. *Adicciones*, 17 (supl. 2), 81-94.
- San, G. I., Gutiérrez, E., Sáiz, P.A., González, M^a.P., Bascarán, M^a. T. y Bobes, J. (2000). Evaluación de trastornos de la personalidad en pacientes heroínómanos mediante el International Disorders Examination (IPDE). *Adicciones*, 12, 43-56.
- Satorra, A. y Bentler, P. M. (1994). Corrections to test statistics and standard errors in covariance structure analysis. In A.V. Eye y C.C. Clogg (Eds.). *Latent variables analysis: Applications for developmental research*, 399-419. Thousand Oaks, CA: Sage.
- Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., et al. (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I.): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *Journal of Clinical Psychiatry*, 59 (suppl. 20), 22-33.
- Seva, A., Magallón, R., Sarasola, A. y Merino, J. A. (1992). GHQ-28 validation in a Spanish general urban population. *European Journal of Psychiatry*, 6, 147-153.
- Steiger, J. H. (1990). Structural model evaluation and modification: an interval estimation approach. *Multivariate Behavioral Research*, 25, 173-180.
- Taylor, A., Del Grande, E. y Parsons, J. (1997). *Mental Health Status of South Australians*. South Australia: National Library of Australia Cataloging Publication.
- Torrens, M. (2008). Patología dual: situación actual y retos del futuro. *Adicciones*, 20, 315-320.
- Werneke, U., Goldberg, D. P., Yalcin, I. y Üstün, B. T. (2000). The stability of the factor structure of the General Health Questionnaire. *Psychological Medicine*, 30, 823-829.