

ARTIGO DE INVESTIGAÇÃO (ORIGINAL)

Propriedades psicométricas do Questionário Geral de Saúde 28 itens em enfermeiros: proposta com 24 itens

Psychometric properties of the 28-Item General Health Questionnaire in nurses: a proposal with 24 items

Propiedades psicométricas del Cuestionario de Salud General de 28 ítems en enfermeros: propuesta con 24 ítems

Paulo Seabra^{1,2}
 <https://orcid.org/0000-0001-8296-1021>
Manuel Luís Capelas³
 <https://orcid.org/0000-0002-8993-6854>
Joaquim Oliveira Lopes^{1,2}
 <https://orcid.org/0000-0003-2571-7078>
Mariana Calado⁴
 <https://orcid.org/0000-0003-3409-6058>
Ezequiel Pessoa¹
 <https://orcid.org/0000-0002-0756-2693>

¹ Escola Superior de Enfermagem de Lisboa, Lisboa, Portugal

² Centro de Investigação, Inovação e Desenvolvimento em Enfermagem de Lisboa (CIDNUR), Lisboa, Portugal

³ Universidade Católica Portuguesa, Instituto de Ciências da Saúde, Centro de Investigação Interdisciplinar em Saúde (CIIS), Lisboa, Portugal

⁴ Hospital CUF Descobertas, Lisboa, Portugal

Autor de correspondência

Paulo Seabra

E-mail: pauloseabra@escl.pt

Recebido: 12.08.20

Aceite: 27.02.21

Resumo

Enquadramento: A avaliação das propriedades psicométricas dos instrumentos de medida, quando aplicados em diferentes populações, é essencial.

Objetivo: Analisar as propriedades psicométricas do Questionário de Saúde Geral 28 itens (GHQ28) em enfermeiros.

Metodologia: Estudo metodológico de validação de instrumentos de medida. Análise das propriedades psicométricas, análise fatorial exploratória (AFE) pelo método dos componentes principais, e curva Receiver Operating Characteristic (ROC), com uma amostra não probabilística de 1.264 enfermeiros.

Resultados: A AFE propõe a manutenção de 4 fatores e eliminação de 4 itens em 3 dimensões (item 3,16, 21,26). A variância explicada pelos 4 fatores foi 61,5% e o alfa de Cronbach 0,93. Os resultados de saúde geral não sofreram variabilidade significativa após se retirarem os 4 itens. A curva ROC, comparando a versão 24 itens com a de 28, revelou uma área sob a curva (AUC) de 0,996 ($p < 0,001$) e ponto de corte 20,5, com uma sensibilidade de 96,3% e especificidade de 98,4%.

Conclusão: A reavaliação das propriedades psicométricas do GHQ28 em enfermeiros, sugere redução para 24 itens.

Palavras-chave: psicométrica; questionário geral de saúde; saúde mental; enfermagem

Abstract

Background: The psychometric properties of measurement instruments should be assessed in different populations.

Objective: To analyze the psychometric properties of the 28-Item General Health Questionnaire (GHQ-28) in nurses.

Methodology: Methodological validation study of measurement instruments. Analysis of psychometric properties, exploratory factor analysis (EFA) by the principal component analysis method, and Receiver Operating Characteristic (ROC) curve, with a nonprobability sample of 1,264 nurses.

Results: EFA suggests keeping four factors and deleting four items in three dimensions (items 3, 16, 21, and 26). The variance explained by the four factors was 61.5%, and Cronbach's alpha was 0.93. The general health scores did not change significantly after deleting the four items. The comparison between the 24-item version and the 28-item version revealed an area under the curve (AUC) of 0.996 ($p < 0.001$) and a cutoff point of 20.5, with a 96.3% sensitivity and a 98.4% specificity.

Conclusion: The reassessment of the psychometric properties of GHQ-28 in nurses suggests a reduction to 24 items.

Keywords: psychometrics; general health questionnaire; mental health; nursing

Resumen

Marco contextual: La evaluación de las propiedades psicométricas de los instrumentos de medición, cuando se aplican en diferentes poblaciones, es esencial.

Objetivo: Analizar las propiedades psicométricas del Cuestionario de Salud General de 28 ítems (GHQ28) en enfermeros.

Metodología: Estudio metodológico sobre la validación de instrumentos de medida. Análisis de las propiedades psicométricas, análisis factorial exploratorio (AFE) mediante el método de componentes principales y curva Receiver Operating Characteristic (ROC), con una muestra no probabilística de 1264 enfermeros.

Resultados: La AFE propuso el mantenimiento de 4 factores y la eliminación de 4 ítems en 3 dimensiones (ítem 3, 16, 21, 26). La varianza explicada por los 4 factores fue del 61,5% y el alfa de Cronbach del 0,93. Los resultados de salud general no experimentaron una variabilidad significativa tras eliminar los 4 ítems. La curva ROC, que compara la versión de 24 ítems con la de 28, mostró un área bajo la curva (AUC) de 0,996 ($p < 0,001$) y un punto de corte de 20,5, con una sensibilidad del 96,3% y una especificidad del 98,4%.

Conclusión: La reevaluación de las propiedades psicométricas del GHQ28 en enfermeros sugiere una reducción a 24 ítems.

Palabras clave: psicométrica; cuestionario de salud general; personal de salud; salud mental; enfermería



Como citar este artigo: Seabra, P., Capelas, M. L., Lopes, J. O., Calado, M., & Pessoa, E. (2021). Propriedades psicométricas do Questionário Geral de Saúde 28 itens em enfermeiros: proposta com 24 itens. *Revista de Enfermagem Referência*, 5(6), e20136. <https://doi.org/10.12707/RV20136>



Introdução

A avaliação em saúde assume uma grande variabilidade de possibilidades, baseada nas inúmeras necessidades associadas à multidimensionalidade e complexidade do conceito. Uma das dimensões é a saúde mental, com fatores subjetivos e de difícil monitorização. A saúde mental pode ser definida como um “estado de bem-estar no qual o indivíduo realiza as suas capacidades, [podendo] fazer face ao stress normal da vida, trabalhar de forma produtiva e frutífera, e contribuir para a comunidade em que se insere” (World Health Organization [WHO], 2007, p.1), ou seja, um estado que não é oposto ao de doença mental, mas que é influenciado por fatores do dia-a-dia, contribuindo para enfrentar situações stressantes.

Em Portugal, estudos já realizados, com amostras reduzidas de profissionais de saúde e de difícil comparabilidade, também pela utilização de instrumentos diferentes, evidenciam níveis elevados de *burnout* (25%-60% dos enfermeiros; Marôco et al., 2016) e sofrimento mental (15%-61,5% sofrem de ansiedade e entre 14%-18% de depressão; Gomes & Oliveira, 2013). Neste sentido, têm surgido muitos instrumentos para a avaliação da saúde mental ou para clarificar um possível diagnóstico em termos de doença mental (Pais-Ribeiro, 2011; Ali et al., 2016), sendo que adquirem aceitabilidade internacional pela sua transversalidade em indicadores de sintomatologia específica, mas não necessariamente patológica, como o *General Health Questionnaire 28* (GHQ28). Sobre o GHQ28, Jackson (2007), citado por Pais-Ribeiro et al. (2015), referia que este é um dos questionários de rastreio de saúde mental mais utilizados, particularmente pela adequação a qualquer tipo de população. No entanto, ainda é pouco utilizado com profissionais de saúde. Assim, o objetivo do presente estudo é analisar as propriedades psicométricas do GHQ28 em enfermeiros portugueses.

Enquadramento

O GHQ de 28 itens foi desenvolvido por Goldberg e Hillier (1979) com o objetivo de avaliar a capacidade de dar resposta a algumas atividades e ao surgimento de fenómenos stressantes. Não tem como objetivo avaliar traços estáveis, mas sim alterações no funcionamento. Neste sentido, é um instrumento para avaliar a saúde mental e o bem-estar psicológico (Pais-Ribeiro & Antunes, 2003). É constituído por quatro subescalas - Sintomas somáticos; Ansiedade e insónia; Disfunção social; e Depressão grave - cada uma com 7 itens. Os *scores* apresentam-se numa escala de medida ordinal tipo *Likert* com 4 pontos (0-3), que possibilita, através da soma dos itens, um intervalo de resposta entre 0-21 nas subescalas e 0-84 na escala global. Valores mais elevados correspondem a pior saúde mental. Quanto ao ponto de corte, os estudos de validação em Portugal têm evidenciado, para as subescalas, o valor de 4/5, e para a escala global 23/24, sendo que valores superiores indicam a necessidade da pessoa ser avaliada clinicamente por um profissional de saúde. As diferentes dimensões reportam-se à sintomatologia e

não a distintos diagnósticos, embora os valores possam ser relevantes e independentes em cada subescala (Pais-Ribeiro et al., 2015).

Sendo a GHQ28 uma escala de uso internacional, apresentando atualmente versões mais reduzidas (12 ou 20 itens), vários estudos têm sido desenvolvidos para a sua validação em diferentes populações de Portugal. Em 2003, com 30 pessoas com tuberculose e 30 pessoas saudáveis, obteve-se um alfa de Cronbach de 0,94 (Pais-Ribeiro & Antunes, 2003). Em 2011, num estudo com 171 pessoas em situação não-clínica apurou-se um alfa de Cronbach de 0,92 (Monteiro, 2011) e em 2015, com 384 pessoas que recorriam a serviços de saúde no geral e a serviços sociais, obteve-se um alfa de Cronbach de 0,94 (Pais-Ribeiro et al., 2015). Quanto à versão com 12 itens, foi validada para Portugal com uma população de 790 estudantes do ensino secundário, tendo-se verificado uma consistência interna de 0,91 (Laranjeira, 2008). É possível verificar que, em todos os estudos, os valores de fiabilidade são aceitáveis (Marôco, 2018).

A realização de estudos de validação da escala especificamente em profissionais de saúde ainda é escassa, embora já tenha sido feita em diferentes países. Na Índia, com 448 profissionais em unidades de saúde, verificou-se que 41,1% apresentavam sinais de saúde mental prejudicada (Vinod et al., 2017), embora nenhuma medida de consistência interna tenha sido apresentada. No Japão, com profissionais emigrados da Indonésia, apurou-se uma consistência interna de 0,89 (Sato et al., 2016), sendo que 22,5% dos profissionais apresentavam menor perceção da sua saúde mental.

Num estudo recente, com aplicação do GHQ28 a uma população de enfermeiros provenientes da rede de cuidados nacional, procedeu-se à verificação da confiabilidade do instrumento (consistência interna de 0,94) e da validade de construto (KMO = 0,946 e teste de esfericidade de Bartlett com $p < 0,001$). A análise evidenciou que 60,6% dos participantes apresentavam perceção negativa da sua saúde mental, 71,6% percecionavam sintomas somáticos, 76% ansiedade significativa, 94,1% alguma forma de disfunção social e 22,2% sintomas de depressão severa (Seabra et al., 2019). Pela falta de estudos em Portugal acerca deste tópico, justifica-se um maior aprofundamento das propriedades psicométricas da escala, assim como da sua estrutura de forma (Marôco, 2018), para contribuir para a adaptação deste instrumento nesta população.

Metodologia

Estudo metodológico de validação de instrumentos de medida (Marôco, 2018; Hair et al., 2009). Foram incluídos no estudo 1.264 enfermeiros, numa amostra de conveniência, a partir de um universo de 69486 enfermeiros inscritos na Ordem dos Enfermeiros Portugueses (OE) que disponibilizou um *link* no seu *website* entre abril e julho de 2017. Os enfermeiros foram convidados a responder por correio eletrónico. Através deste *link* os participantes acediam a um questionário de auto-preenchimento, para caracterização sociodemográfica,

profissional e laboral e ainda uma avaliação de saúde mental pelo GHQ-28 (Pais-Ribeiro & Antunes, 2003). A necessidade de preenchimento da totalidade do GHQ minimizou as não respostas.

Os dados foram depois extraídos da base de dados em formato Excel para Windows. A análise estatística foi feita através do software IBM SPSS Statistics, versão 25 (SPSS Inc., Chicago, IL, USA) para Windows. A análise da estrutura relacional dos itens foi realizada através da análise fatorial exploratória (AFE) sobre a matriz das correlações, com extração dos fatores pelo método das componentes principais, seguida de uma rotação *Varimax*, à semelhança de estudos prévios com o mesmo instrumento (Pais-Ribeiro et al., 2015). A validade da AFE foi avaliada através de: matrizes de correlações e correlações parciais, medida Kaiser-Meyer-Olkin (KMO) e teste de esfericidade de Bartlett, assumindo-se quanto à adequação do modelo (KMO) a referência: adequação média valores entre 0,7-0,8, boa entre 0,8-0,9 e excelente entre 0,9-1,0, assim como a necessidade de um *p value* no teste de Bartlett < 0,05 (Marôco, 2018; Hair et al., 2009). Os fatores comuns retidos foram aqueles que apresentaram um valor próprio (*eigenvalue*) superior a 1, em consonância com o *scree plot* e a percentagem da variância retida de acordo com Marôco (2018). No que se refere ao peso fatorial dos itens, e quando se observa uma carga cruzada em mais que um fator, a opção foi que o peso mínimo tem de ser superior a 0,3 (para amostras maiores que 350 participantes) e a diferença entre estes têm que ser maior que 0,1 (Marôco, 2018), a que se associou uma avaliação qualitativa quando o mesmo respeitava estes critérios e estava alocado a mais que um fator. Procedeu-se à avaliação da consistência interna do instrumento, através do alfa de Cronbach, que nos dá o grau com que os diferentes itens que constituem um instrumento, e que medem o mesmo construto, produzem pontuações semelhantes (Oliveira, 2014) e avaliou-se a sensibilidade psicométrica e de diagnóstico como sugerido por Marôco (2018). Para a determinação dos pontos de corte da escala, a sua sensibilidade e especificidade, no geral e nas subescalas, utilizou-se a curva ROC e o Índice de Youden, tendo como padrão apenas a avaliação comparativa com a GHQ28.

Foi solicitada e obtida autorização para utilização do GHQ28 aos autores da validação do instrumento para a população portuguesa. O estudo foi aprovado pela Comissão de Ética do Instituto de Ciências da Saúde da Universidade Católica Portuguesa, parecer nº 23/2017 e encontra-se publicado (Seabra et al., 2019), fazendo-se neste estudo um aproveitamento secundário dos dados.

Resultados

A caracterização sociodemográfica da amostra revelou que 83,2% (1052) dos participantes eram mulheres, com uma mediana de idades de 36 anos (AIQ = 14) com idade mínima de 22 e máxima de 64, que 87,7% (1108) viviam acompanhados e que 55,0% (695) tinham dependentes a seu cargo. A nível profissional, 60,0% (759) exerciam em contexto hospitalar e 22,2% (280) nos cuidados de saúde primários, 58,5% (740) trabalhavam por turnos, com a mediana de 40 (AIQ = 15) horas de trabalho por semana e o tempo médio de experiência profissional encontrado foi de 15,2 ($\pm 9,5$) anos.

A AFE foi precedida do diagnóstico da fatorabilidade da matriz de correlação do GHQ-28. A inspeção visual da matriz de correlações revelou um número substancial de correlações superiores a 0,30, sugerindo a existência de uma estrutura de agrupamento dos itens, e a inspeção das correlações parciais (matriz anti-imagem), com valores inferiores a 0,7, sugerindo que os itens, individualmente, podem ser explicados pelos outros itens que compõem o fator (Hair et al., 2009). O teste de esfericidade de Bartlett revelou significância estatística para a matriz de correlações ($p < 0,001$) e a medida KMO um valor de 0,946 que traduz uma fatorabilidade da matriz de correlações excelente, de acordo com os critérios definidos por Marôco (2018). Confirmou-se, assim, a validade da AFE para a matriz de correlações do GHQ-28. Seguidamente procedeu-se à AFE, conduzida pelo método das componentes principais, seguida de uma rotação *Varimax* para extração dos fatores. A decisão sobre o número de fatores a ser extraído foi tomada com base na regra do *eigenvalue* superior a 1 e o *scree plot* (Marôco, 2018), definindo uma estrutura relacional dos itens do GHQ-28 explicada por quatro fatores latentes.

Na avaliação da matriz fatorial, os itens 3, 16, 21 e 26 apresentaram pesos fatoriais (correlação item-fator) elevados, quase sobreponíveis, em mais que um fator e com diferenças entre eles < 0,10 (Tabela 1). Este facto pode ocorrer se, segundo Marôco (2018), um item for explicado em simultâneo por mais do que um fator, o que não contribui para a ortogonalidade (não correlação) das dimensões da escala. Verificando-se esta condição, pondera-se a eliminação desses itens. Antes de se tomar essa decisão, procedeu-se à análise adicional com métodos rotacionais ortogonais *Quartimax* e *Equimax*, procurando definir uma estrutura simples, em que cada variável se associasse a um único fator (Hair et al., 2009). No entanto, a análise com estes métodos manteve esse problema fundamental.

Tabela 1*Matriz fatorial rodada a 28 itens (resumida)*

ITENS	Pesos Fatoriais				KMO
	Fator 1	Fator 2	Fator 3	Fator 4	
GHQ (28 itens incluídos)					0,946
3 - Tem-se sentido em baixo de forma e mal-disposto?	0,501			0,508	
16 - Tem levado mais tempo a fazer as tarefas normais?			0,494	0,437	
21 - Tem tido prazer nas suas atividades normais do dia-a-dia?	0,405		0,476		
26 - Acha que às vezes não consegue fazer nada por causa dos nervos?	0,401	0,491			
<i>Eigenvalue</i>	11,122	2,557	1,930	1,105	
Variância explicada	39,7%	9,1%	6,9%	3,9%	
Variância cumulativa explicada	39,7%	48,8%	55,7%	59,7%	

Nota. GHQ = Questionário de Saúde Geral; KMO = Medida Kaiser-Meyer-Olkin. A negrito apresentam-se os pesos fatoriais, conservando os valores de correlação item-fator acima de 0,32 (Tabachnick & Fidel, 2007).

Procedeu-se então à eliminação sucessiva de cada item (3, 16, 21 e 26, por esta ordem) e respetiva avaliação da matriz fatorial gerada em cada passo. Os resultados da AFE realizada com os itens 27, 26 e 25, continuaram a mostrar pesos fatoriais cruzados elevados, com diferenças de $< 0,10$, procedendo-se, assim, à eliminação desses itens da escala. Posteriormente, o diagnóstico da fatorabilidade da matriz de correlação do GHQ com 24 itens revelou um teste de esfericidade de Bartlett com significância estatística ($p < 0,001$) e uma medida de KMO = 0,931

que traduz uma fatorabilidade da matriz de correlações excelente (Marôco, 2018), confirmando a validade da AFE para a matriz de correlações dos itens. Seguidamente, conduziu-se a AFE pelo método das componentes principais, seguida de uma rotação Varimax para extração dos fatores. A estrutura relacional mais estável da escala com 24 itens foi explicada por quatro fatores latentes, extraídos com base nos pesos fatoriais de cada item, nas comunalidades, na regra do *eigenvalue* > 1 , as covariâncias e o *scree plot* (Marôco, 2018; Tabela 2).

Tabela 2*Matriz fatorial rodada a 24 itens*

ITENS	Pesos Fatoriais				h ²
	Fator 1	Fator 2	Fator 3	Fator 4	
1- Tem-se sentido perfeitamente bem de saúde?	0,366		0,320	0,519	0,513
2- Tem sentido necessidade de um tónico?	0,306			0,406	0,337
4- Tem-se sentido doente?	0,414			0,563	0,612
5- Tem tido dores de cabeça?				0,816	0,733
6- Tem tido uma sensação de aperto na cabeça ou pressão na cabeça?				0,797	0,723
7- Tem tido ataques de frio ou calor?				0,479	0,377
8- Tem perdido o sono devido a preocupações?	0,663			0,331	0,587
9- Depois de adormecer acorda várias vezes?	0,557			0,365	0,457
10- Tem-se sentido constantemente sob tensão?	0,771				0,700
11- Tem-se sentido irritável e de mau humor?	0,701				0,643
12- Tem-se sentido assustado ou tem entrado em pânico sem razão?	0,597	0,337			0,536
13- Tem tido a sensação de que está tudo a cair em cima de si?	0,745				0,684
14- Tem-se sentido permanentemente nervoso e tenso?	0,766				0,743
15- Tem conseguido manter-se ativo e ocupado?			0,535		0,394

17- Acha que, de um modo geral, tem trabalhado bem?		0,773		0,638
18- Sente-se satisfeito com a maneira como tem cumprido as suas tarefas?		0,777		0,668
19- Tem-se sentido útil no que faz?		0,751		0,607
20- Tem-se sentido capaz de tomar decisões?		0,733		0,603
22- Tem-se considerado uma pessoa sem valor nenhum?		0,592	0,343	0,591
23- Tem sentido que já não há nada a esperar da vida?		0,772		0,720
24- Tem sentido que a vida já não vale a pena?		0,797		0,740
25- Já pensou na hipótese de um dia vir a acabar consigo?		0,824		0,695
27- Tem dado consigo a pensar estar morto e longe de tudo?		0,803		0,722
28- Acha que a ideia de acabar com a sua vida está sempre a vir-lhe à cabeça?		0,847		0,743
<i>Eigenvalue</i>	9,391	2,451	1,861	1,062
Variância explicada	39,1%	10,2 %	7,8 %	4,4 %
Variância cumulativa explicada	39,1%	49,3%	57,1%	61,5%

Nota. h^2 = Comunalidades. A negrito apresentam-se os pesos fatoriais, conservando os valores de correlação item-fator acima de 0,32 (Tabachnick & Fidel, 2007).

A matriz fatorial revelada após a eliminação dos quatro itens, minimiza o número de pesos fatoriais significantes sobre cada linha, fazendo com que cada variável se associe com um único fator e, esta matriz, mostra-nos a estrutura de fatores latentes (correlações item-fator) amplamente conhecida e divulgada na literatura, em que o primeiro fator se relaciona com Ansiedade e insónia (mantendo sete itens originais), o segundo com Depressão grave (passa a ter seis itens), o terceiro com Disfunção social (passa a ter cinco itens) e o quarto com Sintomas somáticos (passa a ter seis itens). O modelo fatorial retido reproduz bem a estrutura correlacional observada a partir do critério das comunalidades (Hair et al., 2009) e a matriz dos resíduos (Marôco, 2018). As comunalidades, superiores a 0,3 demonstram que os quatro fatores retidos são apropriados para descrever a estrutura correlacional latente entre os itens. Por sua vez, a matriz de resíduos no modelo, que

resulta da diferença entre a matriz das correlações observadas e a matriz das correlações estimadas pelo modelo, revela 79% de resíduos com valores absolutos inferiores a 0,05. Uma percentagem elevada de resíduos (>50%) inferiores a 0,05 é indicador de um modelo fatorial com bom ajustamento (Marôco, 2018).

A avaliação da fiabilidade do instrumento (com 24 itens) revelou um alfa de Cronbach de 0,93 e para as diferentes dimensões: Sintomas somáticos $\alpha = 0,820$; Ansiedade e insónia $\alpha = 0,890$; Disfunção social $\alpha = 0,810$; Depressão grave $\alpha = 0,900$. Os valores observados sugerem homogeneidade dos itens em cada uma das dimensões do instrumento. Na proposta que apresentamos (24 itens), a eliminação de itens não revelou alterações significativas na fiabilidade da escala total, comparativamente à versão original (28 itens), nem nas suas subescalas, agora com menos itens em três delas (Tabela 3).

Tabela 3

Comparação das medidas de consistência interna da GHQ a 28 e a 24 itens

	Consistência interna (alfa de Cronbach)				
	Total	Sintomas somáticos	Ansiedade e insónia	Disfunção social	Depressão grave
GHQ28	0,940	0,850	0,890	0,842	0,900
GHQ24	0,930	0,820	0,890	0,810	0,900

Nota. GHQ28 = Questionário de Saúde Geral 28 itens; GHQ24 = Questionário de Saúde Geral 24 itens.

Outro passo dado que reforça a análise da estrutura da escala é a correlação entre fatores, que obrigatoriamente tem de ser significativa, e a correlação entre fatores deve

ser < 0,70, demonstrando a sua independência. Na Tabela 4 apresentam-se os valores com base nos coeficientes de correlação de Spearman.

Tabela 4

Matriz de correlações entre os fatores da escala GHQ a 24 itens

Fatores	Sintomas somáticos	Ansiedade e insônia	Disfunção social	Depressão grave
Sintomas somáticos	1,000			
Ansiedade	0,717*	1,000		
Disfunção social	0,485*	0,505*	1,000	
Depressão	0,428*	0,501*	0,342*	1,000

Nota. * A correlação é significativa no nível 0,01 (2 extremidades); GHQ = Questionário Geral de Saúde.

Procedeu-se também à avaliação da sensibilidade psicométrica dos itens, através das medidas de forma (Marôco, 2018). A análise dos coeficientes de assimetria e achatamento dos itens revelaram valores de 2,191 e 3,959 respetivamente, o que de acordo com Kline (2011), indica que os itens da escala têm a capacidade de discriminar indivíduos que são estruturalmente diferentes.

Quanto à avaliação da sensibilidade de diagnóstico/clínica, o GHQ com 24 itens mantém a capacidade de discriminar diferenças estatisticamente significativas na percepção da saúde mental. Por exemplo, face ao género (*homens*: $m_e = 20,00$, $AIQ = 14,75$; *mulheres*: $m_e = 24$, $AIQ = 16$; $U = 131039,50$; $p \leq 0,001$) e face à prática de atividade desportiva/*hobbie* (*sim*: $m_e = 21$, $AIQ = 14$; *não*: $m_e = 26$, $AIQ = 16$; $U = 241145,00$; $p \leq 0,001$). Em relação às subescalas (medida ajustada ao número de itens) e face às mesmas variáveis, verificam-se diferenças face ao género na subescala Sintomas Somáticos (*homens*: $m_e = 5$, $AIQ = 5$; *mulheres*: $m_e = 7$, $AIQ = 6$; $U =$

$132447,50$; $p \leq 0,001$), na subescala Ansiedade (*homens*: $m_e = 7$, $AIQ = 7$; *mulheres*: $m_e = 9$, $AIQ = 7$; $U = 130420,50$; $p \leq 0,001$) e na subescala Disfunção Social (*homens*: $m_e = 5$, $AIQ = 2$; *mulheres*: $m_e = 6$, $AIQ = 3$; $U = 122997,00$; $p < 0,05$); face à prática de atividade desportiva/*hobbie* verificam-se diferenças na subescala Sintomas Somáticos (*sim*: $m_e = 6$, $AIQ = 5$; *não*: $m_e = 7$, $AIQ = 6$; $U = 236115,50$; $p \leq 0,001$), na subescala Ansiedade (*sim*: $m_e = 8$, $AIQ = 6$; *não*: $m_e = 10$, $AIQ = 7,75$; $U = 239517,50$; $p \leq 0,001$), na subescala Disfunção Social (*sim*: $m_e = 5$, $AIQ = 2$; *não*: $m_e = 6$, $AIQ = 3$; $U = 228251,50$; $p \leq 0,001$) e na subescala Depressão Grave (*sim*: $m_e = 1$, $AIQ = 3$; *não*: $m_e = 1$, $AIQ = 5$; $U = 220532,00$; $p < 0,05$).

Apurando ainda a sensibilidade clínica em termos da saúde mental com os 28 itens, comparativamente ao instrumento com 24 itens, ajustando a média ao número de itens, verifica-se que os valores apurados são sensivelmente os mesmos (Tabela 5).

Tabela 5

Comparação da sensibilidade clínica da GHQ a 28 e a 24 itens

	Total Média (\pm DP) / Mediana (AIQ)	Sintomas somáticos Média (\pm DP) / Mediana (AIQ)	Ansiedade e insônia Média (\pm DP) / Mediana (AIQ)	Disfunção social Média (\pm DP) / Mediana (AIQ)	Depressão grave Média (\pm DP) / Mediana (AIQ)
GHQ28	30,09 (\pm 13,86) 28 (18)	8,35 (\pm ,30) 8 (6)	9,13 (\pm 4,83) 9 (7)	9,23 (\pm 3,13) 4 (4)	3,38 (\pm 4,34) 2 (5)
GHQ24	25 (\pm 11,92) 24 (16)	6,86 (\pm 3,70) 6 (5)	9,13 (\pm 4,84) 9 (7)	6,32 (\pm 2,28) 6 (2)	2,70 (\pm 3,83) 1 (4)
GHQ28 ^a /24 ^a	Média 1,07 / 1,04	Média 1,19 / 1,14	Média 1,30 / 1,30	Média 1,31 / 1,26	Média 0,48 / 0,45

Nota. ^a = Ajustado ao número de itens; GHQ28 = Questionário de Saúde Geral 28 itens; GHQ24 = Questionário Geral de saúde 24 itens; DP = Desvio padrão; AIQ = Amplitude interquartilica.

Por último, perspetivando a validade clínica e tendo em conta os valores de referência para se interpretar os resultados, ou seja, os valores a partir dos quais os avaliados devem ser aconselhados a uma avaliação clínica

mais precisa, realizamos a análise da curva ROC, tanto para a escala no total como para as subescalas. As análises revelaram áreas sob a curva (AUC), que variaram entre 0,993 e 1 com $p < 0,001$ (Tabela 6).

Tabela 6*AUC da Curva ROC*

Escala	AUC	p
GHQ24-Total	0,996	
GHQ24-Sintomas somáticos	0,993	
GHQ24-Ansiedade	1,000	<0,001
GHQ24-Disfunção social	0,990	
GHQ24-Depressão	0,996	

Nota. GHQ24 = Questionário de Saúde Geral 24 itens; AUC = Área sob a curva; *p* = teste de significância.

Procedendo-se, com recurso ao Índice de Youden, à determinação dos pontos de corte e respetivas sensibilidades e especificidades da escala no total, obteve-se o valor de

20,5, a que corresponde sensibilidade de 96,3% e especificidade de 98,4%. Os valores relativos às subescalas podem ser visualizados na Tabela 7.

Tabela 7*Pontos de corte, sensibilidade e especificidade da GHQ com 24 itens*

Escala	Ponto de corte	Sensibilidade	Especificidade
GHQ24-Total	20,5	96,3%	98,4%
GHQ24-Sintomas somáticos	4,5	96,2%	98,1%
GHQ24-Ansiedade	5,5	100,0%	100,0%
GHQ24-Disfunção social	4,5	94,3%	98,6%
GHQ24-Depressão	4,5	95,4%	98,5%

Nota. GHQ24 = Questionário de Saúde Geral 24 itens.

Discussão

O desafio de reavaliar as propriedades psicométricas e a factorialidade de uma escala com vasto uso nacional e internacional, é seguramente necessário, principalmente, por ser pouco estudada a sua aplicabilidade em profissionais de saúde. O tamanho da amostra permitiu este aprofundamento de modo a estarem garantidas a fiabilidade das análises estatísticas e a tornar mais robustas as conclusões e a proposta apresentada. Reavaliámos e foram mantidos os quatro fatores originais da escala nomeadamente, Ansiedade e insónia (mantendo sete itens originais), Depressão grave (passa a ter seis itens), Disfunção social (passa a ter cinco itens) e Sintomas somáticos (passa a ter seis itens), com a garantia dos *eigenvalue* e variância explicada por fator sustentada nas diferentes recomendações consultadas e garantindo a significância estatística dos pesos fatoriais (Hair et al., 2009; Marôco 2018). A necessidade de propor uma nova estrutura, discutir a eliminação dos itens e a sua distribuição factorial (saturação dos itens em mais de um fator, não ortogonalidade da escala, independência de fatores) deve seguir as recomendações estatísticas consensualizadas e

manter a integridade teórica dos construtos (Hair et al., 2009), o que se considera ter sido conseguido, seguindo as recomendações de outras versões mais reduzidas e desenvolvidas a nível internacional e nacional (Laranjeira, 2008; Kashyap & Singh, 2017). Neste caso, a opção foi partir da sua versão mais extensa pela especificidade da amostra.

Confirmaram-se os fatores determinados previamente, o que permite uma avaliação imediata da significância prática dos resultados, ou seja, confirmou-se a identificação de variáveis-chave que neste caso são as próprias dimensões (Hair et al., 2009). Na análise destes dados é de notar que os itens 3, 16, 21 e 26 saturavam em dois fatores, 1 e 4, 3 e 4, 1 e 3, 1 e 2, respetivamente, refletindo uma aproximação concetual e de interação entre construtos teóricos (quadros psicológicos) com pontos comuns, nomeadamente entre sintomatologia somática e ansiedade, entre a sensação de disfunção social e a sintomatologia somática, entre a ansiedade e a sensação de disfunção social e, por último, entre a depressão e ansiedade, tal como assinalado pela literatura (Lu et al.,

2019). Sabemos que são sintomas ou expressões do estado mental das pessoas, com fronteiras que se influenciam mutuamente, até porque os seus indicadores objetivos e subjetivos são frequentemente similares. Os domínios do instrumento têm tido uma organização com base na literatura (Pais-Ribeiro et al., 2015), com a capacidade de apontar indicações clínicas independentes, verificado pela independência das correlações entre construtos (entre 0,342 e 0,717). Tal como identificado anteriormente, apurou-se uma correlação maior entre as subescalas Ansiedade e insónia e Sintomas somáticos (Pais-Ribeiro & Antunes, 2003). Deste modo, os itens mencionados foram retirados após a exploração de soluções ortogonais e não ortogonais. Os itens retirados são efetivamente de ligação aos construtos teóricos das suas subdimensões, mas os resultados de fiabilidade, tendo como elemento comparativo o GHQ28, revelam que esses construtos conseguem ser medidos sem a presença dos itens eliminados, como demonstram os resultados na escala total e nas subescalas, quando se ajusta a média obtida ao número de itens, e quando se avalia a sua sensibilidade clínica face a algumas variáveis sociodemográficas ou socioprofissionais (Hair et al., 2009; Mâroco, 2018; Pais-Ribeiro et al., 2015). Poder-se-á justificar a realização de um estudo metodológico com o objetivo primário de validação desta versão reduzida com 24 itens, comparando com instrumentos padrão de análise do *outcome* geral e das subdimensões. Por exemplo, o Short Form-36, questionário de avaliação de saúde, usado por Pais-Ribeiro e Antunes (2003) e outros clinicamente relevantes para os construtos. Destaque-se, ainda, que os itens 3 e 26, já anteriormente tinham sido identificados, como assumindo carga elevada em mais de um fator, tendo-se questionado a sua discriminação (Pais-Ribeiro et al., 2015).

Para a sua operacionalidade, na escala total com 28 itens a literatura propõe uma pontuação de 23/24 como valor fronteira para se estar em presença de um caso para posterior avaliação clínica, e o valor acima de 4/5 nas subescalas (com sete itens cada; Pais-Ribeiro et al., 2015). Estes resultados, apontam a possibilidade de se fazer uma adaptação para o número de itens quer na escala global, quer nas subescalas, com a particularidade de estas ficarem com um número de itens entre cinco e sete. A escala global ficará com intervalo possível de respostas entre 0-72 e as subescalas com cinco itens entre 0-15, as subescalas com seis itens entre 0-18 e a que se mantém com sete itens entre 0-21. Neste processo, realizou-se a análise tendo-se principalmente em conta o que os dados estatísticos foram apresentando, evitando a variabilidade que é proporcionada pelos vários métodos para extração, rotação e pontuação de fatores. Considera-se que cada método produz resultados diferentes, pelo que se evitou o caminho de experimentar diversos métodos até que os resultados concordassem com alguma convicção prévia do grupo de investigadores, tal como referido na literatura (Oliveira, 2014). Uma opção mais psicométrica e menos teórica-clínica (Pais-Ribeiro et al., 2015), mas que se apresenta fiável, uma vez que a análise qualitativa manteve os itens nos fatores com maior peso.

No que se refere à aplicabilidade prática, a análise da curva

ROC permite considerar que o GHQ24 mantém excelentes características psicométricas, quando comparado com o GHQ28. Será necessário, no futuro, providenciar a validação de critério com instrumentos *gold standard* para cada uma das dimensões. É necessário considerar que este processo de reavaliação das propriedades psicométricas beneficiaria da utilização de outro instrumento, com um construto teórico aproximado, para se poder reforçar com uma análise de validação concorrente, o que pode ser considerado uma limitação deste estudo. O processo de seleção dos participantes, pode ser apontado também como uma limitação do estudo pois, apesar do acesso ser pelo site da OE, não se podia garantir que o acesso fosse realizado exclusivamente por estes.

Conclusão

O resultado da análise das propriedades psicométricas do GHQ28 sugere uma redução do número de itens de 28 para 24 (perdendo os itens 3, 16, 21 e 26, que eram explicados em simultâneo por mais do que um fator). Mantém uma estrutura fatorial ortogonal (não correlacionada), teoricamente significativa e fácil de interpretar num instrumento que produz uma pontuação única, resultante da soma das diferentes dimensões.

Avaliando o comportamento do GHQ a 24 itens, confirma-se boa consistência interna da escala e das suas dimensões, e o contributo homogêneo e significativo dos itens para a avaliação do conceito e das dimensões que o constituem. Confirma-se também a capacidade da escala em discriminar indivíduos estruturalmente diferentes e com perceções distintas, numa manifestação da sensibilidade do instrumento na avaliação de indicadores de sintomatologia em saúde mental e bem-estar psicológico. A redução do número de itens, permite uma redução do tempo de resposta dos participantes, sem perda de fiabilidade nos resultados.

Este estudo reforça a necessidade de reavaliar periodicamente as propriedades de um instrumento de utilização global, nomeadamente quando se utiliza numa população diferente e específica. No futuro, sugere-se dar continuidade a este estudo com uma população de profissionais de saúde não enfermeiros, de forma a expandir o processo de validação junto de profissionais de saúde.

Contribuição de autores

Conceptualização: Seabra, P., Lopes, J., Calado, M.

Tratamento de dados: Seabra, P., Lopes, J., Calado, M., Capelas, M., Pessoa, E.

Metodologia: Seabra, P., Lopes, J., Calado, M., Capelas, M., Pessoa, E.

Redação - preparação do rascunho original: Seabra, P., Lopes, J., Calado, M., Capelas, M., Pessoa, E.

Redação - revisão e edição: Seabra, P., Lopes, J., Calado, M., Capelas, M., Pessoa, E.

Referências bibliográficas



- Ali, G.-C., Ryan, G., & De Silva, M. J. (2016). Validated screening tools for common mental disorders in low and middle income countries: A systematic review. *PLoS ONE*, *11*(6), e0156939. <https://doi.org/10.1371/journal.pone.0156939>
- Goldberg, D., & Hillier, V. (1979). A scaled version of the General Health Questionnaire. *Psychology Medicine*, *9*(1), 139-145. <https://doi.org/10.1017/S0033291700021644>
- Gomes, R. K., & Oliveira, V. B. (2013). Depressão, ansiedade e suporte social em profissionais de enfermagem. *Boletim de Psicologia*, *68*(138), 23-33. http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S0006-59432013000100004
- Hair, J. F., Jr., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2009). *Análise multivariada de dados* (6ª ed.). Bookman.
- Kashyap, G. C., & Singh, S. K. (2017). Reliability and validity of general health questionnaire (GHQ-12) for male tannery workers: A study carried out in Kanpur, India. *BMC Psychiatry*, *17*, 102. <https://doi.org/10.1186/s12888-017-1253>
- Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). The Guilford Press.
- Laranjeira, C. A. (2008). General health questionnaire - 12 items: Adaptation study to the Portuguese population. *Epidemiology and Psychiatric Sciences*, *7*(2), 148-151. <https://doi.org/10.1017/S1121189X00002840>
- Lu, Y., Li, J., & Liu, Y. (2019). Depression as a mediator of quality of life in patients with neuropathic pain: A cross-sectional study. *Journal of Advanced Nursing*, *75*(11), 2719-2726. <https://doi.org/10.1111/jan.14111>
- Marôco, J. (2018). *Análise estatística com o SPSS Statistics 25* (7ª ed.). Report Number.
- Marôco, J., Marôco, A. L., Leite, E., Bastos, C., Vazão, M. J., & Campos J. (2016). Burnout in Portuguese healthcare professionals: An analysis at the national level. *Acta Médica Portuguesa*, *29*(1), 24-30. <https://doi.org/10.20344/amp.6460>
- Monteiro, A. P. (2011). Assessment of the factor structure and reliability of the Portuguese version of the General Health Questionnaire-28 among adults. *Journal of Mental Health*, *20*(1), 15-20. <https://doi.org/10.3109/09638237.2010.492414>
- Oliveira, A. G. (2014). *Bioestatística decodificada: Bioestatística, epidemiologia e investigação* (2ª ed.). Lidel.
- Pais-Ribeiro, J. L. (2011). *Inventário de saúde mental*. Placebo.
- Pais-Ribeiro, J. L., & Antunes, S. (2003). Contribuição para o estudo de adaptação do Questionário de Saúde Geral de 28 itens. *Revista Portuguesa de Psicossomática*, *5*(1), 37-45. <https://www.redalyc.org/articulo.oa?id=28750105>
- Pais-Ribeiro, J. L., Silva, M., Abrantes, C., Coelho, M., & Nunes, J. (2015). Ulterior validação do Questionário de Saúde Geral de Goldberg de 28 itens. *Psicologia, Saúde e Doenças*, *16*(3), 278-286. http://www.scielo.mec.pt/scielo.php?script=sci_abstract&pid=S1645-00862015000300001&lng=en&nrm=iso
- Sato, F., Hayakawa, K., & Kamide, K. (2016). Investigation of mental health in Indonesian health workers immigrating to Japan under the Economic Partnership Agreement. *Nursing and Health Sciences*, *18*, 342-349. <https://doi.org/10.1111/nhs.12275>
- Seabra, P., Lopes, J., Calado, M. E., & Capelas, M. L. (2019). A national survey of the nurses' mental health: The case of Portugal. *Nursing Forum*, *54*(3), 425-433. <https://doi.org/10.1111/nuf.12350>
- Tabachnick, B., & Fidell, L. (2007). *Using multivariate statistics* (5th ed.). Pearson/Allyn & Bacon.
- Vinod, A., Choudhari, S. B., Waghmare, S., Rajshri, Dewani, K., & Moni, S. (2017). Study of psychiatric morbidity among health professionals of different groups using the General Health Questionnaire-28: A cross-sectional study. *International Journal of Indian Psychology*, *4*(2). <https://doi.org/10.25215/0402.186>
- World Health Organization. (2007). *Strengthening mental health promotion*. Geneva, WHO Fact sheet. <https://mindyourmindproject>

