

Estudo de validação em Portugal de uma versão reduzida da Escala de Depressão Geriátrica

Ana João Santos* / Baltazar Nunes** / Irina Kislaya*** / Ana Paula Gil**** / Oscar Ribeiro*****

* Instituto de Ciências Biomédicas Abel Salazar, Universidade do Porto, Porto, Portugal / Instituto Nacional de Saúde Doutor Ricardo Jorge, Lisboa, Portugal; ** Instituto Nacional de Saúde Doutor Ricardo Jorge, Lisboa, Portugal / Centro de Investigação em Saúde Pública, Escola Nacional de Saúde Pública, Universidade Nova de Lisboa, Lisboa, Portugal; *** Instituto Nacional de Saúde Doutor Ricardo Jorge, Lisboa, Portugal; **** Centro Interdisciplinar de Ciências Sociais, Faculdade de Ciências Sociais e Humanas, Universidade Nova de Lisboa, Lisboa, Portugal; ***** Centro de Investigação em Tecnologias e Serviços de Saúde, Universidade do Porto, Porto, Portugal / Departamento de Educação e Psicologia, Universidade de Aveiro, Aveiro, Portugal / Instituto Superior de Serviço Social do Porto, Porto, Portugal

As propriedades psicométricas da versão reduzida (5 itens) da Escala de Depressão Geriátrica (GDS), instrumento de avaliação da depressão junto de pessoas idosas, foram examinadas em duas amostras de pessoas com 60 e mais anos.

A validade concorrente, a especificidade e a sensibilidade foram obtidas contra o Inventário de Depressão do Beck (IDB-II) numa amostra de 66 indivíduos (média das idades de 70 anos), enquanto a validade da estrutura fatorial foi observada numa amostra de base populacional de 1023 indivíduos (média das idades de 71 anos). A versão de 5 itens da Escala de Depressão Geriátrica apresenta valores de consistência interna e de correlação com o IDB-II que sugere a sua fiabilidade e valores de sensibilidade e especificidade adequados ao rastreamento de sintomatologia depressiva na população idosa. Os resultados obtidos pela Análise Fatorial Confirmatória (AFC) sugerem que o modelo unifactorial não apresenta as características desejadas, indicando que um dos itens (item 4) poderá ter um menor poder discriminativo, pelo que se observa um melhor ajustamento no modelo obtido pela AFC para 4 itens.

Palavras-chave: Escala de Depressão Geriátrica, Fiabilidade, Validade, Depressão, Pessoas idosas.

A depressão é a desordem psiquiátrica mais comum na população idosa com documentado impacto ao nível da qualidade de vida, constituindo um importante preditor da mortalidade (Baldwin, 2004). A par da importância do diagnóstico, para adequada intervenção e tratamento com reconhecidos ganhos em saúde e qualidade de vida, reconhece-se a dificuldade na identificação da depressão (Baldwin, 2004; Roman & Callen, 2008).

Os instrumentos de autorrelato (e.g., escalas) têm vindo a ganhar importância no rastreio da depressão e dos sintomas depressivos também pela sua fácil administração, número limitado de recursos implicados e tempo necessário (Telles-Correia & Barbosa, 2009). Às reconhecidas vantagens, contrapõe-se as dificuldades de avaliação da depressão em pessoas com outras

Na página oficial da GDS, os autores consideram que se trata de um instrumento no domínio público (<https://web.stanford.edu/~yesavage/GDS.html>).

A correspondência relativa a este artigo deverá ser enviada para: Ana João Santos, Departamento de Epidemiologia, Instituto Nacional de Saúde Doutor Ricardo Jorge, Avenida Padre Cruz, 1600-016 Lisboa, Portugal. E-mail: ana.carvalho@insa.min-saude.pt

comorbilidades, pelo facto de o limiar dos níveis de depressão relevantes para a condição psiquiátrica poderem ser diferentes em pessoas com outras patologias (Telles-Correia & Barbosa, 2009).

Em grupos etários mais avançados, acresce ainda o facto da depressão se apresentar, muitas vezes, através de queixas somáticas e funcionais que podem reproduzir outras patologias médicas (Pocinho, Farate, Dias, Lee, & Yesavage, 2009). Além disso, os sinais e sintomas da depressão manifestados pela população idosa podem ser diferentes dos comumente apresentados na população adulta, aspetos que contribuem para a sua subnotificação (Baldwin, 2004; Roman & Callen, 2008).

O rastreio de sintomatologia depressiva na população idosa com os instrumentos de avaliação desenvolvidos para o efeito poderá auxiliar a identificação e tratamento.

A Escala de Depressão Geriátrica

A Escala de Depressão Geriátrica (Geriatric Depression Scale – GDS) é um instrumento de avaliação da depressão que procura superar as dificuldades de rastreio na população idosa, com vantagens demonstradas relativamente a outros instrumentos do mesmo tipo (Baldwin, 2004; Pocinho et al., 2009). Desenvolvida por Yesavage e colaboradores em 1982, a GDS é, segundo os autores, o único instrumento que se conhece elaborado para ser utilizado especificamente com esta população (Yesavage et al., 1982). Procura eliminar a confusão, geralmente presente em outros instrumentos de avaliação da depressão, entre indicadores somáticos da depressão e manifestações físicas “normais” da velhice.

A GDS, na sua versão original e alargada, apresenta-se como uma escala de autorrelato, composta por 30 itens no original, com duas alternativas de resposta (sim ou não), consoante o modo como a pessoa idosa se tem sentido na semana transada.

Esta versão, já traduzida em 30 línguas, apresenta excelentes propriedades psicométricas (a partir de 84% para a sensibilidade e de 95% para a especificidade) (McGivney, Mulvihill, & Taylor, 1994; van Marwijk et al., 1995; Yesavage, 1988). É, segundo Baldwin (2004), o instrumento de rastreio da depressão na população geriátrica mais utilizado e validado em diferentes países. Em Portugal a versão aferida e validada por Pocinho e colaboradores (2009) apresentou uma sensibilidade de 100% (IC=1-1) e uma especificidade de 83% (IC=0.5-1).

Dada a sua extensão, alguns autores indicam que a aplicação da GDS30 pode levar à fadiga das pessoas idosas e interferir no decorrer do atendimento médico (Baldwin, 2004; Rinaldi, Mecocci, Benedetti, & Ercolani, 2003). Além disso a investigação desenvolvida em estudos observacionais obriga à utilização instrumentos de rastreio mais curtos. Para colmatar estas desvantagens, foram desenvolvidas versões mais curtas, nomeadamente de 20, 15, 5, 4 e 1 itens (Hoyl et al., 1999; Kim, DeCoster, Huang, & Bryant, 2013; Pocklington, Gilbody, Manea, & McMillan, 2016; Rinaldi et al., 2003; van Marwijk et al., 1995).

Hoyl e colaboradores (1999) desenvolveram uma das versões curtas da GDS, constituída por cinco itens, mais tarde aferida e validada na população americana (Hoyl et al., 1999; Rinaldi et al., 2003). Os autores, considerando que não se tratava de um diagnóstico de depressão, sugerem que se trata de um bom indicador da presença de sintomas depressivos. Para o desenvolvimento da escala os autores selecionaram, entre os 15 itens da GDS15, os cinco itens que mostraram maior poder discriminativo (Hoyl et al., 1999). Para isso, numa amostra com 74 pessoas idosas procederam à avaliação clínica de depressão e administraram a GDS15. Os cinco itens com uma correlação mais elevada com o diagnóstico clínico de depressão foram selecionados para integrar a versão curta de cinco itens da GDS.

Esta versão reduzida foi posteriormente validada em três contextos distintos: pessoas idosas da comunidade (consultas externas); pessoas idosas hospitalizadas (enfermaria de cuidados agudos) e pessoas idosas em lar residencial (Rinaldi et al., 2003). O estudo incluiu uma amostra

de 181 pessoas sem déficit cognitivo (avaliado clinicamente), às quais foi administrada a GDS15, a GDS5; e que foram alvo de uma avaliação neuropsicológica geriátrica incluindo a avaliação clínica da depressão de acordo com os critérios da 4ª edição do Manual Diagnóstico e Estatístico de Transtornos Mentais (DSM-IV) (Rinaldi et al., 2003).

Esta versão mostrou boa sensibilidade (94% dos indivíduos diagnosticadas com os critérios do DSM-IV apresentaram um score indicativo da presença de sintomas depressivos) e boa especificidade (81% dos indivíduos não diagnosticadas com os critérios do DSM-IV apresentaram um score não indicativo da presença de sintomas depressivos) (Rinaldi et al., 2003). Através da validade concorrente (comparando resultados da GDS5, GDS15 e diagnóstico clínico da depressão), Rinaldi e colaboradores estabeleceram o valor de corte em 1/2 como indicativo da presença de sintomas depressivos. Um estudo inicial de validação foi também conduzido em Espanha e os resultados vão ao encontro do que já tinha sido encontrado na amostra americana, apresentando boa sensibilidade (81%, IC=73.1-89.1) e especificidade (73.2%, IC=66.3-80.1) (de la Iglesia et al., 2005).

A relevância da validação da versão de 5 itens, e do desenvolvimento de versões curtas prende-se com a utilidade de uma ferramenta para a investigação e para o contexto clínico. Se no âmbito da investigação epidemiológica onde o número de perguntas é um aspeto central para a participação dos respondentes, no contexto da prática clínica, escalas de rastreio validadas, mas de fácil aplicação, podem auxiliar os profissionais que trabalham com pessoas idosas a identificar situações que requerem mais atenção para o diagnóstico clínico. Escalas reduzidas, mas que mantenham validade e capacidade discriminativa, cumprem estes objetivos e deverão ser valorizadas (Hagtvet & Sipos, 2016).

Neste âmbito, o presente artigo apresenta os resultados de dois estudos conduzidos para avaliar as propriedades psicométricas da versão de 5 itens da GDS para a população portuguesa. A partir de uma amostra de conveniência analisou-se a sensibilidade e especificidade contra o Inventário de Depressão de Beck (IDB-II). O segundo estudo, conduzido numa amostra de base populacional, avaliou a fiabilidade e a estrutura fatorial.

A escolha do IDB-II decorreu da necessidade de selecionar uma escala de autorrelato para avaliação da sintomatologia depressiva, dado não haver possibilidade, no âmbito do estudo, de inclusão de uma avaliação para diagnóstico clínico. Esta escala, desenvolvida por Aaron Beck em 1961, é um instrumento de medida construído segundo a metodologia proposta por Likert, com vinte e um itens referentes a sintomas ou grupos sintomatológicos medidos numa escala de quatro pontos (de 0 a 3 consoante o grau de gravidade) (Campos & Gonçalves, 2004; Canavarro, Pereira, Simões, Pintassilgo, & Ferreira, 2008; Oliveira-brochado & Oliveira-brochado, 2008).

Apesar de ser preferencial a utilização de diagnóstico clínico enquanto *gold standard* na avaliação da validade concorrente de doenças e síndromes, a utilização de um instrumento que avalia o mesmo constructo é também comum quando se deseja substituir um instrumento por outro mais simples, mais barato ou menos invasivo (Streiner & Norman, 1995). Adicionalmente, alguns estudos têm vindo a demonstrar a aplicabilidade do IDB-II junto da população idosa (Argimon, Paloski, Farina, & Irigaray, 2016). A sua adequabilidade para o objetivo proposto, prende-se também com o facto de ser uma das escalas de autorrelato que melhor reflete os critérios de diagnóstico do DSM-V, sendo considerada como o *gold standard* entre as escalas de autorrelato para avaliação da depressão (Wang & Gorenstein, 2013). Utilizada anteriormente em estudos de validade da GDS, na sua versão original, os resultados mostraram correlações elevadas entre o IDB e a GDS (Olin, Schneider, Eaton, Zemansky, & Pollock, 1992), bem como valores de sensibilidade e de especificidade aceitáveis (Ferraro & Chelminski, 1996; Kiernan, Wilson, Suter, Naquin & Meltzen, 1986; Scogin, 1987). Estes valores tendem a decrescer em populações clínicas, sugerindo que ambas as escalas funcionarão melhor junto da população em geral (Olin et al., 1992).

Os dois estudos apresentam resultados relativos ao ponto de corte mais favorável da GDS5, de acordo com os valores de sensibilidade e de especificidade, relativos à fiabilidade e à estrutura fatorial da escala.

Método

Desenvolveram-se dois estudos, que se apresentam em seguida. O Estudo 1 teve como objetivo avaliar a validade concorrente da GDS5 utilizando o IDB-II como *gold standard*. Junto de uma amostra de conveniência (Amostra A) avaliou-se o ponto de corte mais favorável, a sensibilidade e a especificidade. O Estudo 2, conduzido numa amostra de base populacional (Amostra B) permitiu examinar a fiabilidade e a estrutura fatorial da GDS5.

Participantes

A população-alvo dos dois estudos realizados foi constituída por pessoas com idade igual ou superior a 60 anos residentes em Portugal, em domicílios particulares, há pelo menos 12 meses. A amostra A (Estudo 1) foi constituída por pessoas com 60 e mais anos e incluiu 71 participantes (Tabela 1), sendo mais de metade mulheres (70% mulheres e 30% homens), com idades compreendidas entre os 60 e os 82 anos ($M=70.16$, $DP=6.24$). Considerando apenas os indivíduos sem valores omissos em todos os itens das 2 escalas avaliadas obteve-se um total de 66 participantes.

Tabela 1

Distribuição do sexo e idade nas duas amostras

		Amostra A (N=71)	Amostra B (N=1123)
Sexo	Mulheres % (n)	69.6% (49)	67.3% (693)
	Homens % (n)	30.4% (22)	32.7% (337)
Intervalo de idades		60-82	60-92
(M, Dp)		(70.16; 6.24)	(70.92; 7.06)

A amostra B (Estudo 2) obteve-se a partir do inquérito via telefone a uma amostra representativa da população portuguesa de pessoas a residir na comunidade com 60 ou mais anos. A amostra foi composta por 1123 participantes, dos quais 67% eram mulheres e 32% eram homens, com idades compreendidas entre os 60 e os 92 anos ($M=70.92$, $DP=7.06$). Considerando apenas os participantes sem valores omissos na GDS5, obteve-se um total de 1030 participantes.

Instrumento

Para o desenvolvimento da versão reduzida da GDS5 para Portugal, os autores utilizaram os itens que integraram a versão original de 5 itens em língua inglesa (Hoyl et al., 1999). Os cinco itens correspondentes em português (itens 1, 4, 10, 12 e 17) à versão desenvolvida em língua inglesa foram selecionados a partir da versão portuguesa da GDS30 já traduzida e validada para o contexto nacional (Barreto, Leuschner, Santos, & Sobral, 2008). Aos participantes foi lida a seguinte introdução: “Por favor responda Sim ou Não a cada uma das seguintes frases quanto à forma como se tem sentido de há uma semana para cá”: (1) Satisfeito(a) com a sua vida?; (2) Muitas vezes aborrecido(a)?; (3) A sentir-se muitas vezes desamparado(a)?; (4) A preferir ficar

em casa, em vez de sair e fazer coisas novas?; e (5) A sentir-se inútil?. O score total corresponde à soma de cada um dos itens, sendo 1 ponto atribuído para a resposta “sim” e 0 pontos para a resposta “não”, exceto nos itens invertidos.

No caso do IDB-II considerou-se o valor de corte superior a 13 para a presença de sintomatologia depressiva, uma vez que em estudos de validação em Portugal, os resultados sugerem valores inferiores ou iguais a 13 para os indivíduos sem sintomatologia depressiva e valores superiores a 13 para indivíduos com sintomatologia depressiva (Campos & Gonçalves, 2004; Oliveira-brochado & Oliveira-brochado, 2008).

Procedimento

Os dois estudos foram realizados enquanto parte do projeto de investigação Envelhecimento e violência (Gil et al., 2015), que decorreu entre 2011 e 2014. Os participantes que aceitaram colaborar na fase piloto do projeto foram convidados a preencher um pequeno questionário onde se incluíram a GDS e o IDB-II (Estudo 1). Esta amostra, a partir da qual se desenvolveu o Estudo 1, foi recrutada através de pontos focais de uma universidade sénior, de um centro de dia e de uma junta de freguesia da área metropolitana de Lisboa.

O Estudo 2 obteve-se a partir do inquérito via telefone a uma amostra representativa da população portuguesa de pessoas a residir na comunidade com 60 ou mais anos (Amostra B), tendo-se administrado um questionário em que se incluiu também a versão curta de 5 itens da GDS (Gil et al., 2015). A amostra utilizada no estudo foi estratificada pelas 7 regiões NUTS II e regiões autónomas, com alocação homogénea das unidades de amostragem. A recolha de dados teve em conta a existência de telefones fixos e de telefones móveis nos alojamentos e a utilização de cada uma das redes (fixa e móvel) pela população idosa (60% da amostra recrutada através de telefones fixos e 40% recrutada através de telefones móveis). O questionário foi aplicado através de entrevistas telefónicas assistidas por computador (CATI – Computer Assisted Telephone Interview) a 1123 pessoas com 60 e mais anos, que forneceram o consentimento verbal após terem sido informadas sobre o estudo. De acordo com as definições normalizadas pela American Association for Public Opinion Research (American Association For Public Opinion Research, 2011) a taxa de resposta ao inquérito telefónico foi de 74.03% e de recusa 25.97%.

Resultados

Estudo 1

Validade concorrente

O estudo sobre a validade de critério, conduzido na Amostra A ($N=66$), foi obtido através do IDB-II. O poder discriminativo foi avaliado a partir da área sob a curva ROC (AUC – Area Under the Curve). Considerou-se valores aceitáveis quando se obteve curvas ROC com valores superiores a .70 (Aguiar, 2007; Hosmer, Lemeshow & Sturdivant, 2013). A sensibilidade e a especificidade foram calculadas considerando o nível de confiança de 95%, utilizando o IDB-II como critério (*gold standard*). O ponto de corte foi selecionado de acordo com a maximização dos valores obtidos tanto para a especificidade, como para a sensibilidade (Hosmer et al., 2013).

Os resultados indicam que, com uma precisão de 85%, o valor de corte mais adequado em termos de sensibilidade (70.0%) e especificidade (85.3%) é de igual ou superior a 2 (Tabela 2).

Tabela 2

Sensibilidade e especificidade da GDS5 face ao IDB-II

Valor de corte	GDS5 * IDB-II	
	Sensibilidade	Especificidade
(≥ 0)	100.00%	0.00%
(≥ 1)	90.00%	52.94%
(≥ 2)	70.00%	85.29%
(≥ 3)	50.00%	97.06%
(≥ 4)	30.00%	100.00%
(≥ 5)	20.00%	100.00%
(>5)	0.00%	100.00%
Área ROC (IC95%)	84.41 (69.9-93.4)	

Nota. Estudo 1, Amostra A (N=66).

Estudo 2

Fiabilidade

A fiabilidade da GDS5, através da consistência interna foi obtida junto da Amostra B (N=1030). O Alfa de Cronbach foi de .69, valor abaixo do usualmente utilizado para classificar a fiabilidade como adequada ($\alpha \geq .70$). Note-se, ainda, no que respeita aos diferentes itens que compõem a escala, o valor do Alfa de Cronbach para escala é ligeiramente superior ($\alpha > .71$) no caso de eliminação do item 4 (“A preferir ficar em casa, em vez de sair e fazer coisas novas?”).

Validade

A avaliação da validade fatorial do instrumento foi realizada com uma Análise Fatorial Confirmatória (AFC), considerando o modelo unifatorial, uma vez que a unidimensionalidade da escala pode ser encontrada consistentemente em estudos da GDS5 para outros países (de la Iglesia et al., 2005; Hoyl et al., 1999; Rinaldi et al., 2003). Optou-se pela análise fatorial das correlações tetracóricas por ser a mais apropriada para avaliar variáveis dicotómicas e pelo grande número de participantes na amostra (N=1030) (Savalei, Bonett, & Bentler, 2015).

Na Tabela 3 são apresentadas as intercorrelações entre os itens. As correlações variam entre .37 e .80 sendo o item 4 aquele que apresenta valores de correlação menos elevados, o que vai ao encontro da análise da fiabilidade obtida através do Alfa de Cronbach. Os valores de correlação dos itens da GDS5 são todos positivos, uma vez que o único item com uma cotação invertida (“Satisfeito com a sua vida?”) foi corrigido antes da análise.

Tabela 3

Correlações tetracóricas dos itens da GDS5

	GDS_1	GDS_2	GDS_3	GDS_4	GDS_5
GDS_1 Satisfeito(a) com a sua vida?	1				
GDS_2 Muitas vezes aborrecido(a)?	.8004	1			
GDS_3 A sentir-se muitas vezes desamparado(a)?	.5514	.7351	1		
GDS_4 A preferir ficar em casa, em vez de sair e fazer coisas novas?	.3738	.3763	.3982	1	
GDS_5 A sentir-se inútil?	.4590	.5786	.4684	.4305	1

Nota. Estudo 2, Amostra B (N=1030).

Na análise da estrutura fatorial, foram utilizados os 5 itens do instrumento, numa AFC, recorrendo à versão 15 do programa estatístico Stata. Conforme indicado por alguns autores, foram utilizados três índices para verificar o ajustamento do modelo (Bentler & Bonnet, 1980; Pilati & Laros, 2007). O índice absoluto – χ^2 do ajustamento (χ^2/gl); índice relativo – CFI (*Comparative Fit Index*) e índice de discrepância populacional – RMSEA (*Root Mean Square Error of Approximation*). Os indicadores de ajustamento do modelo permitem perceber um fraco ajustamento do modelo ($\chi^2/\text{gl}=10.824$, CFI=.942 e RMSEA=.117). Os coeficientes de ponderação fatorial dos itens para o modelo original (GDS5) mostram o menor peso do item 4, que apresenta uma variância única de 84%, sendo que os restantes apresentam coeficientes elevados (Tabela 4). Considerando-se o potencial menor poder discriminativo deste item e de modo a encontrar um modelo com ajustamento adequado, procedeu-se à sua eliminação.

Tabela 4

Coefficientes de ponderação fatorial e variância de dois modelos confirmatórios da Escala de Depressão Geriátrica (GDS5 e GDS4)

	GDS5		GDS4	
	β (ep)	Var	β (ep)	Var
GDS_1	1	.341	1	.127
GDS_2	1.211(.034)***	.034	1.152(.075)***	.001
GDS_3	.919(.038)***	.442	.723(.053)***	.135
GDS_4	.492(.036)***	.839		
GDS_5	.729(.032)***	.649	.492(.045)***	.128
	$\chi^2(5)=54.12, p<.0001$		$\chi^2(5)=8.42, p<0.05$	

Nota. Estudo 2, Amostra B (N=1030); β =coeficiente de ponderação fatorial; *** $p<.0001$; ** $p<.005$; Var= variância única explicada.

O novo modelo apresentou valores de ajustamento adequados ($\chi^2/\text{gl}=4.210$, CFI=.991 e RMSEA=.055), sobretudo considerando o tamanho amostral. Este modelo sugere a utilização de 4 itens, com valores de variância única de cada um dos itens menores, sendo o item 5 aquele que apresenta um coeficiente de ponderação fatorial mais baixo.

Discussão

Existem, atualmente, uma grande quantidade de escalas de avaliação de depressão disponíveis quer para efeitos de investigação, quer para efeitos de rastreamento na prática clínica. A necessidade de um instrumento rápido, sensível, específico e o mais possível destituído de interferência cultural são fatores potenciadores do contínuo estudo no âmbito das escalas de avaliação de depressão ou da sintomatologia depressiva (Almeida & Almeida, 1999). A Escala de Depressão Geriátrica tem sido amplamente utilizada, tanto em contextos clínicos, como de investigação e está disponível em vários idiomas de países europeus, países latino-americanos e países asiáticos, o que sugere uma consistência entre diferentes países (Almeida & Almeida, 1999) e a facilidade da sua utilização para estudos internacionais comparativos.

Este estudo pretendeu examinar as características psicométricas da versão curta da Escala de Depressão Geriátrica (GDS5), utilizando duas amostras distintas para avaliar a validade concorrente, a fiabilidade e a validade da estrutura fatorial.

A sensibilidade (70%) e especificidade (85%) da versão com 5 itens não sendo elevada é aceitável considerando não só o número de itens, como o número de participantes da amostra ($N=66$). Outras versões obtiveram, contra diagnóstico clínico (DSM-IV ou CID-10) valores de sensibilidade entre os 81% e os 94% para a GDS5 e de especificidade entre 73% e 81% (de la Iglesia et al., 2005; Rinaldi et al., 2003). Estes resultados sugerem uma capacidade discriminativa semelhante entre as versões curtas desta escala em diferentes países.

Em relação ao valor de corte escolhido (2), de acordo com o valor máximo de sensibilidade e especificidade, este é igual ao que tem sido proposto por outros estudos (Almeida & Almeida, 1999; de la Iglesia et al., 2005; Hoyl et al., 1999; Pocklington et al., 2016; Rinaldi et al., 2003).

O resultado de consistência interna ($\alpha=.69$) não vai ao encontro do desejado, ainda que se aproxime do que é usualmente considerado um valor aceitável na literatura em termos de confiabilidade da escala ($\alpha>.70$). Porém, importa referir que, dado que o valor da correlação poderá ser influenciado pelo número de itens, se a escala for constituída por um número reduzido os valores acima dos .60 podem ser considerados aceitáveis (Ribeiro, 2010). Do mesmo modo, na mensuração de constructos complexos, um Alfa de Cronbach de 0.60 pode ser considerado aceitável, ainda que requerendo uma interpretação cautelosa (DeVellis, 2003).

Outros estudos de validação tendem a apresentar valores mais elevados nas versões com um maior número de itens (GDS 30, 15 e 10), tanto em Portugal (Apóstolo et al., 2014; Pocinho et al., 2009) como noutros países (Kim et al., 2013; Pocklington et al., 2016). No caso das versões mais curtas, os valores de fiabilidade obtidos são, usualmente, mais baixos (Kim et al., 2013; Pocklington et al., 2016).

Com o recurso a procedimentos de análise confirmatória, os resultados de validade sugerem um modelo de ajustamento fraco ($\chi^2/gl=10.824$, CFI=.942 e RMSEA=.117), ainda que os valores de correlações entre os itens sejam aceitáveis. Também os coeficientes de ponderação fatorial mostram a elevada contribuição da grande maioria dos itens para a dimensão, sendo o item 4 (*A preferir ficar em casa, em vez de sair e fazer coisas novas*) aquele que apresenta uma saturação menos elevada e um valor de variância única mais elevado. Assim, um novo modelo eliminando o item 4 apresentou valores de ajustamento mais aceitáveis ($\chi^2/gl=4.210$, CFI=.991 e RMSEA=.055). Este item apresentou valores de correlação menos elevados e na escala, o valor de consistência interna é mais elevado com a sua eliminação. Apesar de estes resultados irem ao encontro do que já foi observado num estudo espanhol, em que das duas versões avaliadas (de 4 e 5 itens), a GDS de 4 itens sugeriu melhores resultados (de la Iglesia et al., 2005), há que considerar que existem sempre fatores aleatórios nos dados que podem promover o comportamento flutuante de alguns itens. Além disso, é importante considerar que estes resultados podem decorrer de algumas limitações do próprio estudo.

Em primeiro lugar os itens utilizados no desenvolvimento da versão curta foram extraídos da versão original e não validados através de um estudo exploratório de uma versão portuguesa mais longa do instrumento. De facto, têm sido desenvolvidas várias versões diferentes relativamente aos itens incluídos, sendo que nenhuma versão é exatamente igual entre si, relativamente aos itens selecionados a partir da GDS15 ou da GDS10 (Pocklington et al., 2016). Além disso os diversos estudos utilizam diferentes tipos de populações (por exemplo, clínicas e não clínicas) o que dificulta a comparabilidade dos resultados e tem levado alguns autores a afirmarem maior evidência da GDS15 face a versões mais reduzidas e a necessidades de futuros estudos para as versões mais curtas (Pocklington et al., 2016).

Também a amostra de conveniência obtida além das implicações implícitas, é pequena e requer alguma cautela na leitura dos resultados. Ainda assim, a análise do comportamento da escala junto de uma amostra representativa desta população permite analisar a adequabilidade dos resultados encontrados, nomeadamente, no que respeita ao valor de corte aceitável e que vai ao encontro do que tem sido observado noutros países.

Conclusões

A versão de 5 itens da Escala de Depressão Geriátrica apresenta valores de consistência interna e de correlação com o IDB-II que sugere a sua fiabilidade e valores de sensibilidade e especificidade adequados ao rastreamento de sintomatologia depressiva na população idosa. Os resultados obtidos pela AFC apesar de não sugerirem um modelo bem ajustado, não põe em causa as propriedades da medida. Os resultados indicam ainda que um dos itens (item 4) poderá ter um menor poder discriminativo e que se observa um melhor ajustamento no modelo obtido pela AFC para 4 itens. Tendo em conta a existência de versões mais curtas (de 4 e 1 itens) noutros países, serão necessários mais estudos, utilizando diagnósticos clínicos como *gold standard*, para se determinar se as propriedades psicométricas de versão de 4 itens são melhores que as da versão de 5 itens.

Em suma, os estudos mostram que em Portugal, a versão curta da GDS (5 itens), apresenta resultados semelhantes aos obtidos noutros países e noutras versões. Por ser de fácil administração, simples de entender e requerer tempo mínimo, esta versão poderá ser uma ferramenta de grande utilidade na prática clínica, mas sobretudo, para a investigação na população idosa portuguesa.

Referências

- Aguiar, P. (2007). *Guia prático de estatística em investigação epidemiológica: SPSS*. Lisboa: Climepsi.
- Almeida, O. P., & Almeida, S. A. (1999). Short versions of the Geriatric Depression Scale: A study of their validity for the diagnosis of a major depressive episode according to ICD-10 and DSM-IV. *International Journal of Geriatric Psychiatry*, 14, 858-865. doi: 10.1002/(SICI)1099-1166(199910)14:10<858::AID-GPS35>3.0.CO;2-8
- American Association for Public Opinion Research. (2011). *Standard definitions: Final dispositions of case codes and outcome rates for RDD telephone surveys*. Lenexa, KS: American Association for Public Opinion Research.
- Apóstolo, J., Loureiro, L., Reis, I., Silva, I., Cardoso, D., & Sfectcu, R. (2014). Contribuição para a adaptação da Geriatric Depression Scale-15 para a língua portuguesa. *Revista de Enfermagem Referência, Série IV*(3), 65-73.
- Argimon, I. I. de L., Paloski, L. H., Farina, M., & Irigaray, T. Q. (2016). Aplicabilidade do Inventário de Depressão de Beck-II em idosos: Uma revisão sistemática. *Avaliação Psicológica*, 15, 11-17. doi: 10.15689/ap.2016.15ee.02
- Baldwin, R. (2004). Management of depression in later life. *Advances in Psychiatric Treatment*, 10, 131-139. doi: 10.1192/apt.10.2.131
- Barreto, J., Leuschner, A., Santos, F., & Sobral, M. (2008). Escala de depressão geriátrica: Tradução portuguesa da Geriatric Depression Scale. In A. Mendonça & M. Guerreiro (Eds.), *Escala e testes na demência* (pp. 37-43). Lisboa.
- Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness-of-fit in the analysis of covariance structures. *Psychological Bulletin*, 88, 588-606. doi: 10.1037/0033-2909.88.3.588
- Campos, R. C., & Gonçalves, B. (2004). Alguns dados sobre a prevalência de sintomatologia depressiva na população universitária portuguesa. In C. Machado, L. Almeida, M. Gonçalves, & V. Ramalho (Eds.), *Actas da X Conferência Internacional de Avaliação Psicológica: Formas e contextos* (pp. 50-53). Braga: Psiquilibrios Edições.

- Canavarro, M. C., Pereira, M., Simões, M. R., Pintassilgo, A. L., & Ferreira, A. (2008). Estudos psicométricos da versão portuguesa (de Portugal) do instrumento de avaliação da qualidade de vida na infecção VIH da Organização Mundial de Saúde (WHOQOL-HIV). *Psicologia, Saúde & Doenças*, 9, 15-28.
- de la Iglesia, J. M., Vilches, M. C. O., Herrero, R. D., Taberné, C. A., Colomer, C. A., & Blanco, M. C. A. (2005). Abreviar lo breve. Aproximación a versiones ultracortas del cuestionario de Yesavage para el cribado de la depresión. *Atención Primaria*, 35, 14-21. doi: 10.1157/13071040
- DeVellis, R. F. (2003). *Scale development: Theory and applications*. Thousand Oaks: Sage Publications.
- Ferraro, F. R., & Chelminski, I. (1996). Preliminary normative data on the Geriatric Depression Scale-Short Form (GDS-SF) in a young adult sample. *Journal of Clinical Psychology*, 52, 443-447. doi: 10.1002/(SICI)1097-4679(199607)52:4<443::AID-JCLP9>3.0.CO;2-Q
- Gil, A. P. M., Kislaya, I., Santos, A. J., Nunes, B., Nicolau, R., & Fernandes, A. A. (2015). Elder abuse in Portugal: Findings from the first national prevalence study. *Journal of Elder Abuse & Neglect*, 27, 174-195. doi: 10.1080/08946566.2014.953659
- Hagtvet, K. A., & Sipos, K. (2016). Creating short forms for construct measures: The role of exchangeable forms. *Pedagogika*, 66, 689-713. doi: 10.14712/23362189.2016.346
- Hosmer, J. D., Lemeshow, S., & Sturdivant, R. (2013). *Applied Logistic Regression*. Berlin: Wiley.
- Hoyl, M. T., Alessi, C. A., Harker, J. O., Josephson, K. R., Pietruszka, F. M., Koelfgen, M., . . . Rubenstein, L. Z. (1999). Development and testing of a five-item version of the Geriatric Depression Scale. *Journal of the American Geriatrics Society*, 47, 873-878. doi: 10.1111/j.1532-5415.1999.tb03848.x
- Kiernan, B. U., Wilson, D., Suter, N., Naquin, A., & Meltzen, J. (1986). Comparison of the Geriatric Depression Scale and the Beck Depression Inventory in a nursing home. *Clinical Gerontologist*, 6, 54-56.
- Kim, G., DeCoster, J., Huang, C.-H., & Bryant, A. N. (2013). A meta-analysis of the factor structure of the Geriatric Depression Scale (GDS): The effects of language. *International Psychogeriatrics*, 25, 71-81. doi: 10.1017/S1041610212001421
- McGivney, S. A., Mulvihill, M., & Taylor, B. (1994). Validating the GDS depression screen in the nursing home. *Journal of the American Geriatrics Society*, 42, 490-492. doi: 10.1111/j.1532-5415.1994.tb04969.x
- Olin, J. T., Schneider, L. S., Eaton, L. M., Zemansky, M. F., & Pollock, V. E. (1992). The Geriatric Depression Scale and the Beck Depression Inventory as screening instruments in an older adult outpatient population. *Psychological Assessment*, 4, 190-192.
- Oliveira-brochado, F., & Oliveira-brochado, A. A. (2008). Estudo da presença de sintomatologia depressiva na adolescência. *Revista Portuguesa de Saúde Pública*, 26(2), 27-36.
- Pilati, R., & Laros, J. A. (2007). Modelos de equações estruturais em psicologia: Conceitos e aplicações. *Psicologia: Teoria e Pesquisa*, 23, 205-216.
- Pocinho, M. T. S. S., Farate, C., Dias, C. A., Lee, T. T., & Yesavage, J. A. (2009). Clinical and psychometric validation of the Geriatric Depression Scale (GDS) for Portuguese elders. *Clinical Gerontologist*, 32, 223-236. doi: 10.1080/07317110802678680
- Pocklington, C., Gilbody, S., Manea, L., & McMillan, D. (2016). The diagnostic accuracy of brief versions of the Geriatric Depression Scale: A systematic review and meta-analysis. *International Journal of Geriatric Psychiatry*, 31, 837-857. doi: 10.1002/gps.4407
- Ribeiro, J. L. P. (2010). *Investigação e avaliação em psicologia e saúde*. Lisboa: Placebo Editora, Lda.
- Rinaldi, P., Mecocci, P., Benedetti, C., & Ercolani, S. (2003). Validation of the five-item Geriatric Depression Scale. *Journal of the American Geriatrics Society*, 51, 694-698. doi: 10.1034/j.1600-0579.2003.00216.x
- Roman, M. W., & Callen, B. L. (2008). Screening instruments for older adult depressive disorders: Updating the evidence-based toolbox. *Issues in Mental Health Nursing*, 29, 924-941. doi: 10.1080/01612840802274578

- Savalei, V., Bonett, D. G., & Bentler, P. M. (2015). CFA with binary variables in small samples: A comparison of two methods. *Frontiers in Psychology*, 5, 1-11. doi: 10.3389/fpsyg.2014.01515
- Scogin, F. (1987). The concurrent validity of the Geriatric Depression Scale with depressed older adults. *Clinical Gerontologist: The Journal of Aging and Mental Health*, 7, 23-31. doi: 10.1300/J018v07n01_04
- Streiner, D. L., & Norman, G. R. (1995). *Health measurement scales: A practical guide to their development and use*. Oxford: Oxford University Press.
- Telles-Correia, D., & Barbosa, A. (2009). Ansiedade e depressão em medicina: Modelos teóricos e avaliação. *Acta Medica Portuguesa*, 22, 89-98.
- van Marwijk, H. W., Wallace, P., de Bock, G. H., Hermans, J., Kaptein, A. A., & Mulder, J. D. (1995). Evaluation of the feasibility, reliability and diagnostic value of shortened versions of the geriatric depression scale. *The British Journal of General Practice: The Journal of the Royal College of General Practitioners*, 45(393), 195-199.
- Wang, Y., & Gorenstein, C. (2013). Assessment of depression in medical patients: A systematic review of the utility of the Beck Depression Inventory-II. *Clinics*, 68, 1274-1287. doi: 10.6061/clinics/2013(09)15
- Yesavage, J. A. (1988). Geriatric Depression Scale. *Psychopharmacology Bulletin*, 24, 709-711.
- Yesavage, J. A., Brink, T. L., Rose, T. L., Lum, O., Huang, V., Adey, M., & Leirer, V. O. (1982). Development and validation of a geriatric depression screening scale: A preliminary report. *Journal of Psychiatric Research*, 17, 37-49.

Validation study of a reduced version of the Geriatric Depression Scale in Portugal

The validity of the short Geriatric Depression Scale (GDS) version (5 items) was assessed in two samples of older adults aged 60 or more years. Concurrent validity, specificity, and sensitivity were obtained in a convenience sample of 66 individuals (mean of 70 years) by comparison to the Beck Depression Inventory (IDB-II). Factorial analysis was conducted in a population-based sample of 1023 participants (mean of 71 years). The 5-item version of the Geriatric Depression Scale presented adequate internal consistency values and good correlation results with the IDB-II. These results suggest its reliability and adequate sensitivity, and specificity values for screening for depressive symptoms in the elderly population. The results obtained by the Confirmatory Factor Analysis (CFA) suggest that the one-factor model does not have the desired characteristics – one of the items (item 4) may have a lower discriminative power and a better fit in the model obtained by the CFA for 4 items was observed.

Key words: Geriatric Depression Scale, Reliability, Validity, Depression, Elderly.

Submissão: 23/11/2017

Aceitação: 04/09/2018

