



Revista Iberoamericana de Diagnóstico y Evaluación - e  
Avaliação Psicológica

ISSN: 1135-3848

ISSN: 2183-6051

[bgoncalves@psicologia.ulisboa.pt](mailto:bgoncalves@psicologia.ulisboa.pt)

Associação Iberoamericana de Diagnóstico e Avaliação  
Psicológica  
Portugal

Nunes, Daiane; Faro, André  
Estrutura Fatorial, Análise de Invariância e Distribuição Social do Patient Health Questionnaire-9 (PHQ-9)  
Revista Iberoamericana de Diagnóstico y Evaluación - e  
Avaliação Psicológica, vol. 1, núm. 62, 2022, Enero-Marzo, p. 37  
Associação Iberoamericana de Diagnóstico e Avaliação Psicológica  
Portugal

Disponível em: <https://www.redalyc.org/articulo.oa?id=459671925005>

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en [redalyc.org](http://redalyc.org)

[redalyc.org](http://redalyc.org)

Sistema de Información Científica Redalyc

Red de Revistas Científicas de América Latina y el Caribe, España y Portugal  
Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso  
abierto

## Estrutura Fatorial, Análise de Invariância e Distribuição Social do Patient Health Questionnaire-9 (PHQ-9)

### Factor Structure, Invariance Analysis and Social Distribution of the Patient Health Questionnaire-9 (PHQ-9)

Daiane Nunes<sup>1</sup> e André Faro<sup>2</sup>

#### Resumo

Esta pesquisa avaliou a estrutura fatorial do *Patient Health Questionnaire-9* (PHQ-9) em uma população não-clínica no Brasil, através de análises fatoriais exploratória (AFE) e confirmatória (AFC) e de invariância do modelo. Avaliou-se também a distribuição social do diagnóstico de rastreio da sintomatologia depressiva. A amostra total foi composta por 4958 adultos (dividida em duas subamostras). A primeira subamostra ( $n=1000$ ) foi utilizada para a AFE e a segunda ( $n=3958$ ) para a AFC. A AFE indicou uma estrutura interna composta por um único fator, que explicou 78,5% de variância. A AFC atestou unidimensionalidade da medida, com índices de ajuste e indicadores de confiabilidade adequados. O modelo foi invariante em relação às variáveis investigadas em todos os níveis mensurados. Por fim, as variáveis gênero e escolaridade foram preditoras da sintomatologia depressiva. Concluiu-se que os presentes achados suportam a utilidade do PHQ-9, sendo recomendada sua aplicação na população brasileira.

**Palavras-chave:** depressão, PHQ-9, análise fatorial exploratória, análise fatorial confirmatória, análise de invariância

#### Abstract

This research examined the factor structure of the *Patient Health Questionnaire-9* (PHQ-9) with a non-clinical Brazilian population through exploratory (AFE) and confirmatory (AFC) factor analysis and model invariance. The social distribution of the screening diagnosis for depressive symptoms was also assessed. The complete sample consisted of 4958 adults, which was split into two groups. The first group ( $n=1000$ ) was used for the EFA and the second group ( $n=3958$ ) for the CFA. The EFA indicated a single factor structure, which explained 78.5% of variance. The CFA attested to the unidimensionality, with satisfactory fit indices and reliability indicators. The model was invariant in relation to the investigated variables at all measured levels. Finally, the variables sex and school grade were also significant predictors of depressive symptoms. In conclusion, our findings support the utility of the PHQ-9 and it is recommended for usage in research and intervention with the Brazilian population.

**Keywords:** depression, PHQ-9, exploratory factor analysis, confirmatory factor analysis, invariance analysis

<sup>1</sup> Mestre em Psicologia pelo Programa de Pós-Graduação em Psicologia da Universidade Federal de Sergipe (PPGPSI-UFS). Avenida Marechal Rondon, 210. Conjunto Rosa Elze, São Cristóvão – SE, Brasil, CEP 49000-000. Tel.: (+55)79991262173. E-mail: daianenunesufs@gmail.com

<sup>2</sup> Doutor em Psicologia e Docente do Programa de Pós-Graduação em Psicologia da Universidade Federal de Sergipe (PPGPSI-UFS). Pesquisador Bolsista Produtividade em Pesquisa do CNPq - Nível 2. Avenida Marechal Rondon, s/n. Conjunto Rosa Elze, São Cristóvão – SE, Brasil, CEP 49000-000. Tel.: (+55)79981153915. E-mail: andre.faro.ufs@gmail.com

## Introdução

A depressão está entre as principais condições de saúde no mundo, afetando mais de 264 milhões de pessoas globalmente [World Health Organization (WHO; 2017)], sendo a principal causa de anos vividos com incapacidade (*Years Lived with Disability*) e de anos de vida ajustados por incapacidade [*Disability-Adjusted Life Years* (DALY)] mundialmente (GBD 2016 DALYs and HALE Collaborators, 2017). A depressão é uma condição comórbida frequentemente relatada por pacientes com uma variedade de doenças médicas, sendo o diagnóstico precoce, realizado em contextos de atenção primária em saúde, um dos fatores determinantes para curso da doença (Ferenchick et al., 2019; Zimmerman, 2019). Assim, estratégias mais eficazes e de baixo custo para o enfrentamento da depressão na atenção primária e entre a população geral envolvem o rastreamento precoce e adequado da sintomatologia (Siu & US Preventive Services Task Force, 2016).

O subdiagnóstico e baixo reconhecimento dos sintomas depressivos por parte dos profissionais de saúde se configuram como um dos principais desafios ao enfrentamento da doença (Carey et al., 2014). Na atenção primária, a depressão está presente em aproximadamente 14% dos pacientes, mas na maioria das vezes não é diagnosticada. Além disso, quando os instrumentos de rastreamento padronizados não são usados, os profissionais identificam corretamente a presença de sintomas depressivos em apenas 47% dos pacientes com a doença. O uso de questionários de rastreamento da sintomatologia tem sido recomendado por várias sociedades e organizações médicas como estratégia para sanar essa problemática (Hirschtritt & Kroenke, 2017). Há na literatura um número considerado de medidas válidas de avaliação dos sintomas depressivos e que estão à disposição de profissionais de saúde como ferramentas para identificar adequadamente a presença de sintomatologia e auxiliar da intervenção precoce, a exemplo da *Center for Epidemiologic Studies Depression Scale – Revised* e da *Hospital Anxiety Depression Scale* (Fried et al., 2016).

No contexto brasileiro, foram identificados 24 instrumentos para rastreamento de sintomas

depressivos, no entanto, apenas sete avaliam unicamente depressão na população geral, sendo o *Beck Depression Inventory* (BDI; 21 itens) o mais utilizado nas pesquisas conduzidas em amostras nacionais (Baptista & Borges, 2016). Embora haja uma série de instrumentos disponíveis, há uma escassez de informações com base empírica para orientar a seleção e utilização de medidas de depressão por parte dos profissionais de saúde no público geral ou em populações clínicas. Essa problemática culmina em barreiras para a utilização de medidas adequadas à avaliação baseada em evidências da depressão e, consequentemente, traz implicações para o diagnóstico e prognóstico da doença. Um dos principais motivos para não adesão dos profissionais de saúde aos instrumentos de avaliação é a quantidade de itens que, por vezes, dificulta sua aplicação em contextos clínicos diversos, a exemplo da atenção primária. Logo, medidas breves têm sido recomendadas para mensuração baseada em evidências, o que justifica a pertinência de avaliar suas propriedades psicométricas. Com isso, disponibiliza-se aos profissionais de saúde medidas confiáveis para avaliar a sintomatologia depressiva adequadamente e conduzir o tratamento de modo mais assertivo (Weiss et al., 2015).

Na prática clínica, os questionários de autorrelato são preferíveis às escalas avaliadas pelo profissional de saúde, pois requerem menos envolvimento para a administração e tem baixo custo associado. Essas medidas têm sido usadas, principalmente, para triagem, monitoramento de resultados e estabelecimento de níveis de gravidade da sintomatologia, orientando o tratamento e intervenção adequados da assistência em saúde. O *Patient Health Questionnaire-9* [Questionário de Saúde do Paciente (PHQ-9; Kroenke et al., 2001)] está entre as medidas de depressão mais breves e replicáveis no mundo, contendo cerca da metade dos itens das escalas em uso no Brasil (e.g., BDI-I, 21 itens; CES-D, 20 itens; Escala de Hamilton, 21 itens). PHQ-9 é recomendado, ainda, pelo *National Quality Forum* (Fórum Nacional de Qualidade) como um resultado clínico e indicador dos níveis de sintomas na população (Zimmerman, 2019).

O PHQ-9 consiste em nove itens que representam os critérios diagnósticos para

depressão maior do DSM-IV, incluindo humor deprimido, problemas com o sono, falta de energia, mudança no apetite ou peso, sentimento de culpa ou inutilidade, entre outros. Trata-se de uma medida de autorrelato em que os respondentes são solicitados a classificar, em uma escala de quatro pontos, cada um dos itens com base em quanto um sintoma experimentado os incomodou nas duas últimas semanas (de “nada” a “quase todos os dias”). Um décimo item foi adicionado, em que se pergunta aos pacientes o quão difícil os problemas identificados tornavam para eles gerenciar o trabalho, a vida diária e os relacionamentos (Kroenke et al., 2001). A medida é amplamente utilizada no mundo e as vantagens de sua aplicação incluem seu formato simples e breve de autorrelato, fornece pontuação e interpretação rápidas, sendo acessível a profissionais e não-profissionais de saúde mental e adequada à administração remota, por exemplo, via telessaúde (Gunzler et al., 2020; Hirschtritt & Kroenke, 2017; Mitchell et al., 2016; Zhen et al., 2020).

O PHQ-9 já foi traduzido para 49 idiomas, incluindo espanhol, japonês, russo, português, entre outros e adaptado para 32 culturas diferentes, estando disponível gratuitamente (<https://www.phqscreeners.com>; acessado em 16 de outubro de 2020; Zimmerman, 2019; Lamela et al., 2020). A estrutura fatorial da medida já foi extensivamente avaliada e algumas estruturas foram suportadas empiricamente. Dentre elas, foram identificados na literatura quatro modelos distintos da medida: modelo unidimensional (e.g., Bélanger et al., 2019; Boothroyd et al., 2019; Cameron et al., & Reid, 2008; Yu et al., 2012), de dois (e.g., Chilcot et al., 2013; Patel et al., 2019) e três fatores (e.g., Hinz et al., 2016), além de modelo *bifactor* (e.g., Doi et al., 2018). Outro estudo sugeriu, ainda, um modelo fatorial de segunda ordem com três fatores de primeira ordem, a saber: somático, cognitivo e afetivo (Monteiro et al., 2019). Em recente revisão sistemática, verificou-se que ambos os modelos de um e dois fatores foram apoiados em amostras gerais, clínicas e psiquiátricas. Isso indica que a heterogeneidade nas estruturas fatoriais do PHQ-9 não pode ser explicada exclusivamente pelas diferenças culturais ou pela diversidade sociodemográfica das amostras, requerendo

cautela em suas interpretações e que investigações adicionais sejam conduzidas para fins de verificação de sua estrutura interna (Lamela et al., 2020).

No Brasil, embora o PHQ-9 já tenha sido traduzido para o contexto nacional (Copyright © 2005 Pfizer Inc., New York, NY), não há nenhum estudo de avaliação de sua estrutura interna por meio de Análise Fatorial Exploratória (AFE), Confirmatória (AFC) e análise de invariância da medida. Identificou-se apenas um estudo de avaliação da validade discriminativa da medida, estabelecendo apenas um ponto de corte ( $\geq 10$  pontos) para rastreamento da depressão (de Lima et al., 2009) e outro estudo de determinação dos parâmetros de sensibilidade e especificidade do instrumento (Santos et al., 2013). Além disso, análises mais sofisticadas são possíveis para avaliação das propriedades psicométricas da escala, a exemplo da AFE, AFC e invariância da medida (Damásio, 2013). Tais análises são consideradas etapas essenciais em processos de avaliação das características psicométricas de um instrumento de medição, principalmente quando se tratam de medidas amplamente utilizadas em cenários internacionais (Flora & Flake, 2017; Milfont & Fisher, 2010).

Cumprido ressaltar que à análise de invariância da medida é pouco realizada inclusive em estudos internacionais, visto que menos de um terço dos trabalhos que analisaram evidências de validade do PHQ-9 realizaram análise de invariância da medida (Lamela et al., 2020). Dentre os estudos, as variáveis gênero (González-Blanch et al., 2018; Keum et al., 2018; Patel et al., 2019), condições clínicas (Chung et al., 2015; Doi et al., 2018; Schuler et al., 2018) e nível de escolaridade (González-Blanch et al., 2018; Patel et al., 2019) foram os grupos sociodemográficos mais utilizados para análise multigrupo.

Em face destas questões em torno da escala PHQ-9, a análise das propriedades psicométricas do PHQ-9 em uma amostra brasileira parece ser uma contribuição relevante para o campo de estudos sobre depressão no contexto nacional. Tal investigação se mostra pertinente por fornecer um instrumento de fácil aplicação e interpretação por diferentes profissionais da saúde, sendo uma ferramenta útil para identificação precoce da presença de sintomatologia e assegurando

adequadamente a atenção em saúde. Além disso, contribuiria com o esclarecimento acerca da inconsistência dos modelos disponíveis mundialmente, apresentando evidências adicionais de sua estrutura interna. Assim, este estudo teve como objetivo examinar a estrutura fatorial do PHQ-9 em uma amostra de base populacional não-clínica. Para tanto, foram realizadas AFE, AFC e análise de invariância da medida por gênero, nível de escolaridade e doenças crônicas. Por fim, verificou-se também a distribuição social do diagnóstico de rastreio da sintomatologia depressiva, buscando-se avaliar a presença de diferenças em termos de proporção de ocorrência da depressão segundo variáveis de interesse.

## Método

### Participantes

Este foi um estudo transversal, tendo como público-alvo adultos, de ambos os gêneros, residentes de todos os estados brasileiros, incluindo Distrito Federal. A amostra inicial contou com 4.976 indivíduos. Após exclusão dos participantes com idade acima de 70 anos e estrangeiros ( $n=9$ ; 0,18%, cada), restaram 4.958 respondentes, que compuseram a amostra final deste estudo. Posteriormente, dividiu-se o banco de dados em duas subamostras para realização de cada etapa de análise [Análise Fatorial Exploratória (AFE) e Análise Fatorial Confirmatória (AFC)].

O Quadro 1 descreve todas as variáveis sociodemográficas das amostras referentes à AFE e AFC. A primeira amostra foi utilizada para realização da AFE e contou com 1000 sujeitos, dispondo, assim, com pouco mais de 100 participantes por item da escala. A idade média foi de 30,5 anos [ $DP=11.4$ ; *Mínimo* (*Mín.*)=18; *Máximo* (*Max.*)=70], em sua maioria do gênero feminino (87.9%) e com nível superior de escolaridade (74.1%). A maior parte da amostra foi constituída por participantes da região Nordeste (52.0%), seguida pela região Sudeste (29.2%). Quanto às condições clínicas, 76.1% relataram não possuir nenhuma doença crônica diagnosticada. A segunda amostra ( $n=3958$ ) foi utilizada para AFC, contando com, aproximadamente, 450 sujeitos por parâmetro. A

idade média foi de 31,4 anos ( $DP=11.8$ ; *Mín.*=18; *Máx.*=70), predominantemente do gênero feminino (86,9%;  $n=3.439$ ) e de nível superior de escolaridade (76,6%;  $n=3033$ ). A região Nordeste também foi a mais frequente (51,9%;  $n=2054$ ), seguida novamente pela região Sudeste (30,4%;  $n=1205$ ). A maioria da amostra informou não possuir nenhuma doença crônica (77,5%;  $n=3066$ ). Ao serem comparadas as proporções das variáveis sociodemográficas por tipo de análise, verificou-se que não houve diferença estatisticamente significativa em todas as comparações, ou seja, as amostras da AFE e AFC foram equivalentes nessas variáveis.

### Instrumentos

O *Patient Health Questionnaire-9* (PHQ-9; Kroenke et al., 2001), traduzido e adaptado para o português brasileiro por de Lima Osório et al. (2009), avalia a presença de cada um dos sintomas de depressão maior descritos pelo Manual Diagnóstico e Estatístico dos Transtornos Mentais (DSM-IV). O questionário conta com nove perguntas dispostas em uma escala tipo *likert* (0 “nunca” a 3 “em quase todos os dias”), que mensura a presença de sinais e sintomas de depressão nas últimas duas semanas. O escore final é obtido a partir da soma dos itens, variando entre 0 e 27 pontos. Uma pontuação igual ou superior a 10 pontos indica presença de sintomatologia depressiva. O instrumento possibilita, ainda, avaliar níveis de severidade dos sintomas [leve (10 a 14 pontos), moderado (15 a 19 pontos) ou grave (acima de 20 pontos)].

Para análise de evidências de validade baseada na relação com outras variáveis, utilizou-se a Escala de Estresse Percebido – 10 itens [*Perceived Stress Scale* (PSS-10; Cohen, Kamarck, & Mermelstein, 1983)] e a Escala de Rastreamento do Transtorno Geral de Ansiedade [Generalized Anxiety Disorder Screener (GAD-7)]. Na PSS-10, o participante é questionado sobre com que frequência (no último mês) tenha, por exemplo, “Sentido que as coisas estão sob seu controle” (item 10) ou “Se sentido nervoso e estressado” (item 3), numa escala variando entre 1 “nunca” e 5 “sempre”. A versão adaptada e validada para o contexto brasileiro apresentou confiabilidade satisfatória da medida ( $\alpha=.79$ ; Faro,

Quadro 1. Caracterização das Amostras para Análises Fatoriais Exploratória e Confirmatória

Variáveis		F% (n=1000) <sup>1</sup>	F% (n=3958) <sup>2</sup>	p-valor <sup>3</sup>
Gênero	Masculino	12.1 (121)	13.1 (519)	.393
	Feminino	87.9 (879)	86.9 (3439)	
Escolaridade	Até Nível Médio	25.9 (259)	23.4 (925)	.094
	Nível Superior	74.1 (741)	76.6 (3033)	
Doenças Crônicas	Não	76.1 (761)	77.5 (3066)	.359
	Sim	23.9 (239)	22.5 (892)	
	Norte	3.8 (38)	4.0 (158)	
Região	Nordeste	52.0 (520)	51.9 (2054)	.821
	Centro-Oeste	6.3 (63)	5.6 (223)	
	Sudeste	29.2 (292)	30.4 (1205)	
	Sul	8.7 (87)	8.0 (318)	
Idade [Média (desvio-padrão)]		30.5 (11,48)	31.4 (11,81)	.067

Nota. <sup>1</sup>Amostra utilizada para Análise Fatorial Exploratória; <sup>2</sup>Amostra utilizada para Análise Fatorial Confirmatória; <sup>3</sup>p-valor=significância estatística. F%=frequência percentual; n=número de sujeitos

2013). Neste estudo, o instrumento apresentou alfa de Cronbach de .90.

A GAD-7, por sua vez, consiste em um instrumento breve para avaliação da presença de sintomas ansiosos nas duas últimas semanas, disposto em uma escala com possibilidades de resposta de 0 “nenhuma vez” a 3 “quase todos os dias”. O escore final pode variar entre 0 e 21 pontos, sendo considerado indicador positivo de sinais e sintomas ansiosos, valor igual ou superior a 10 pontos (Kroenke et al., 2007). Estudo de evidências de validade da escala apontou confiabilidade satisfatória em amostra brasileira ( $\alpha=.91$ ; Moreno et al., 2016). Na presente investigação, a medida apresentou alfa de Cronbach de .90.

Por fim, aplicou-se também um questionário sociodemográfico, contendo informações para caracterização da amostra, incluindo gênero (feminino ou masculino), idade (em anos), escolaridade (até o ensino médio ou ensino superior), região geográfica (Norte, Nordeste, Centro-Oeste, Sudeste e Sul) e presença de doenças crônicas (sim ou não).

## Procedimentos

Os dados foram coletados após aprovação da Comissão Nacional de Ética em Pesquisa (CONEP, registro: informações omitidas), através de formulário online. Os participantes foram abordados por meio de convites digitais encaminhados pelos pesquisadores. O Termo de Consentimento Livre e Esclarecido (TCLE), contendo informações sobre a pesquisa e aspectos éticos, foi incluindo na primeira página do formulário, de modo que o participante só tinha acesso às perguntas do questionário após

confirmação de consentimento.

## Análise dos dados

A AFE foi efetuada no *software* Factor (v. 10.8, Lorenzo-Seva & Ferrando, 2006), utilizando o *Robust Diagonally Weighted Least Squares* (DWLS) como método de estimação dos dados e Análise Paralela (AP) para retenção do número de fatores. O *Closeness to Unidimensionality Assessment* foi indicado para avaliar a plausibilidade de a medida ser unidimensional (UniCO>.95; ECV .85; MIREAL .30). Foram aplicados o *Comparative Fit Index* (CFI), o *Goodness of Fit Index* (GFI), *Non-normed Fit Index* (NNFI) (desejável >.95) e o *Root Mean Square Error of Approximation* (RMSEA) (desejável <.08) como índices de ajuste do modelo (Hair et al., 2014). Os indicadores de confiabilidade [Alfa de Cronbach ( $\alpha$ ; desejável >.60) e Ômega de McDonald ( $\Omega$ ; desejável >.70)] e de replicabilidade (*H-Latent*; desejável >.80; Hancock & Mueller, 2001) também foram mensurados.

A AFC e Análise de Invariância da medida foram conduzidas no *software* JASP (versão 0.12.2). Utilizou-se o DWLS robusto, como método de estimação dos dados e os índices de ajuste do modelo foram: o CFI, o GFI, o TLI, com valores acima de .95 considerados desejáveis (Hair et al., 2014). Adotou-se, ainda, o RMSEA (desejável <.08) e o *Standardized Root Mean Square Residual* (SRMR; desejável <.08) (Hu & Bentler, 1999). A invariância do modelo foi avaliada nos níveis configural, métrico, escalar e estrito, sendo empregado Delta CFI ( $\Delta$ CFI; esperado  $\leq .01$ ) e Delta RMSEA ( $\Delta$ RMSEA; esperado  $\leq .015$ ) como parâmetro para rejeição de invariância (Chen, 2007; Hair et al., 2014).

**Quadro 2.** Resultados das Análises Fatoriais Exploratória e Confirmatória do Patient Health Questionnaire-9 (PHQ-9) no Brasil

Itens	$\lambda^1$	$\lambda^2$	%	$M(DP)$	Skew	Kurt
1. Tive pouco interesse ou prazer em fazer coisas.	.82	.78	85.1	1.6 (1.06)	-.03	-1.28
2. Senti desânimo, desalento ou falta de esperança.	.89	.84	85.0	1.6 (1.07)	-.11	-1.32
3. Tive dificuldade em adormecer ou em dormir sem interrupções ou dormi demais.	.74	.70	84.0	1.8 (1.14)	-.31	-1.40
4. Senti cansaço ou falta de energia.	.86	.80	90.6	2.0 (1.05)	-.33	-1.26
5. Tive falta ou excesso de apetite.	.74	.68	82.3	1.7 (1.15)	-.25	-1.44
6. Senti que não gosto de mim mesmo ou que sou um (a) fracassado/a ou desiluído/a a mim mesmo ou à minha família.	.80	.76	70.0	1.4 (1.22)	.03	-1.53
7. Tive dificuldade em me concentrar nas coisas, como ao ler o jornal ou ver televisão.	.77	.72	78.0	1.5 (1.13)	.04	-1.41
8. Movimentei-me ou falei tão lentamente que outras pessoas puderam ter notado. Ou o oposto: estive agitado (a) a ponto de andar de um lado para o outro muito mais do que é habitual.	.68	.62	50.6	1.0 (1.08)	.82	-.65
9. Pensei que seria melhor estar morto/a, ou em magoar-me a mim próprio/a de alguma forma.	.68	.51	29.2	.05 (.91)	.74	1.91

Nota. <sup>1</sup>Análise Fatorial Exploratória; <sup>2</sup>Análise Fatorial Confirmatória.  $\Lambda$ =cargas fatoriais;  $M$ =média;  $DP$ =Desvio Padrão; Mínimo=0; Máximo=3; %=percentual de sujeitos que sinalizaram as opções de resposta 1, 2 ou 3 (presença de sintomatologia com ponto de corte 10); Skew=Assimetria; Kurt=Achatamento (centrado em zero)

A associação entre os escores do PHQ-9 e suas categorias em relação aos escores da PSS-10 e GAD-7 (correlações de Pearson) foi executada no SPSS (versão 25). As análises de distribuição social (gênero, escolaridade, doenças crônicas) da sintomatologia depressiva foram realizadas no JASP. Para tal, conduziu-se análise de regressão logística (método *enter*), em que o diagnóstico de depressão (presença ou ausência) foi inserido como variável dependente (VD). Os indicadores utilizados para avaliação do modelo foram: o Akaike (AIC) e o Critério de Informação Bayesiano (BIC), sendo que valores menores em relação à hipótese nula ( $H_0$ ) são preferíveis; o Qui-Quadrado e graus de liberdade [ $X^2$  (gl)]; o Nagelkerke  $R^2$  e o Tjur  $R^2$ .

## Resultados

### Análise Fatorial Exploratória

O teste de esfericidade de Bartlett [ $\chi^2_{(36)}=4421,5$ ;  $p<.001$ ] e o índice Kaiser-Meyer-Olkin (KMO=.92) indicaram adequação para a realização da AFE. O resultado da AP apontou uma estrutura interna com apenas um fator (78,5% de variância explicada), sendo reafirmada pelos outros indicadores (UniCO=.99, IC95%=.97-.99; ECV=.93, IC95%=.92-.95; MIREAL=.18, IC95%=.15-.20). As cargas fatoriais variaram entre 0,68 (Itens 8 e 9) e .89 (Item 2), mantendo a mesma estrutura de nove itens da escala. Todos os índices de ajuste do modelo foram satisfatórios (CFI=.99; GFI=1,0; NNFI=.99; RMSEA=.07). Os

indicadores de confiabilidade ( $\alpha=.93$ ;  $\Omega=.93$ ) e replicabilidade ( $H\text{-latent}=.94$ ) também apresentaram valores acima do esperado (Quadro 2).

### Análise Fatorial Confirmatória e Invariância da Medida

A AFC confirmou a unidimensionalidade do PHQ-9, apresentando bom ajuste do modelo ( $\chi^2=301,63$ ;  $gl=79$ ; CFI=.99; GFI=.99; TLI=.99; RMSEA=.05, IC95%=.04-.05; SRMR=.04). O menor valor de carga fatorial foi visto no item 9 ( $\lambda=.51$ ) e o maior no item 2 ( $\lambda=.84$ ) (Quadro 1). Quanto à invariância da medida, observou-se que o modelo foi invariante por gênero (feminino ou masculino), escolaridade (até o ensino médio ou nível superior) e doenças crônicas (sim ou não) em todos os níveis de análise (configural, métrico, escalar e estrito) (Quadro 3).

### Evidências de validade baseada na relação com outras variáveis

As correlações entre os escores do PHQ-9, da PSS-10 e da GAD-7 foram estatisticamente significativas, positivas e fortes. A correlação entre os escores do PHQ-9 e da PSS-10 foi de .751 ( $p<.001$ ). Por sua vez, a correlação entre os escores do PHQ-9 e da GAD-7 foi de 0,746 ( $p<.001$ ).

### Distribuição social da sintomatologia depressiva

O modelo obtido pela regressão logística apresentou indicadores satisfatórios ( $X^2 = 92,57$ ;  $p<.001$ ). Viu-se que as variáveis gênero e escolaridade foram preditoras da sintomatologia

**Quadro 3.** Análise de Invariância da Medida do Modelo Unidimensional do PHQ-9 por Gênero, Escolaridade e Doenças crônicas

Parâmetros	Gênero	Escolaridade	Doenças Crônicas
$\chi^2(gl)$	-	-	-
Configural	362.533 (70)	403.498 (70)	349.546 (70)
Métrico	350.600 (61)	375.187 (61)	325.194 (61)
Escalar	362.533 (70)	403.498 (70)	349.546 (70)
Estrito	407.884 (79)	453.315 (79)	363.079 (79)
CFI	-	-	-
Configural	.99	.99	.99
Métrico	.99	.99	.99
Escalar	.99	.99	.99
Estrito	.99	.99	.99
$\Delta CFI$	-	-	-
Configural	-	-	-
Métrico	0	0	0
Escalar	0	0	0
Estrito	0	0	0
RMSEA	-	-	-
Configural	.046	.051	.046
Métrico	.049	.051	.046
Escalar	.046	.049	.046
Estrito	.046	.049	.043
$\Delta RMSEA$	-	-	-
Configural	-	-	-
Métrico	.003	0	0
Escalar	.003	.002	0
Estrito	0	0	.003

*Nota.* Gênero: Feminino ou Masculino; Escolaridade: Até o ensino médio ou ensino superior; e Doenças Crônicas: Sim ou Não.  $\chi^2$ =Qui-Quadrado;  $gl$ =graus de liberdade; CFI=Comparative Fit Index; RMSEA=Root Mean Square Error of Approximation;  $\Delta$ =Delta.  $\Delta CFI$  (esperado  $\leq .01$ );  $\Delta RMSEA$  (esperado  $\leq .015$ ).

**Quadro 4.** Resultados da Regressão Logística para Distribuição Social da Sintomatologia Depressiva com o PHQ-9

Variáveis	F% (n = 4958)	OR (IC 95%)	p-valor
Gênero			
Masculino	12,9 (640)	1	
Feminino	87,1 (4318)	1,9 (1,2 - 2,1)	< .001
Escolaridade			
Nível Superior	76,1 (3774)	1	-
Até Nível Médio	23,9 (1184)	1,5 (1,4 - 2,0)	< .001
Doenças Crônicas			
Não	77,2 (3827)	1	-
Sim	22,8 (1131)	1,1 (1,0 - 1,2)	.162

*Nota.* F%=frequência percentual; n=número de sujeitos. 2. OR=Odds Ratio; IC 95%=intervalo de confiança de 95%; p-valor=significância estatística. Estatísticas de ajuste do modelo: AIC=6348,652; BIC=6374,687;  $\chi^2$ =92.575;  $gl$ =4954;  $p < .001$ ; Nagelkerke  $R^2$ =.018; Tjur  $R^2$ =.084; Cox & Snell  $R^2$ =.018.

depressiva (presença ou ausência). As mulheres apresentaram cerca de duas vezes mais chances de estar no grupo com sintomatologia depressiva ( $OR=1.9$ ;  $IC95\%=1.2-2.1$ ;  $p < .001$ ). Já os participantes com até o nível médio de escolaridade exibiram 50% mais chances de pertencerem ao grupo com presença de sintomas ( $OR=1.5$ ;  $IC95\%=1.4-2.0$ ;  $p < .001$ ). A variável doença crônica foi excluída do modelo final por não apresentar significância estatística (Quadro 4).

## Discussão

O presente estudo objetivou avaliar as propriedades psicométricas do PHQ-9 numa amostra populacional brasileira. Os resultados

indicaram uma estrutura unidimensional, com boas evidências de validade da escala, sustentando a pertinência de sua utilização para mensurar a ocorrência de sintomatologia depressiva no Brasil. Tais achados estão em conformidade com outros estudos com o PHQ-9 no mundo (Amtmann et al., 2014; Bélanger et al., 2019; Boothroyd et al., 2019; González-Blanch et al., 2018; Keum et al., 2018).

A estrutura fatorial do PHQ-9 já foi extensivamente avaliada em outros países e diferentes estruturas (um, dois e três fatores, além do modelo bifactor) receberam suporte empírico (Boothroyd et al., 2019; Chilcot et al., 2013; Doi et al., 2018). Em recente revisão sistemática, foram identificados 33 estudos que examinaram a



estrutura interna da PHQ-9 (Lamela et al., 2020). A maioria dos estudos (57.6%) encontrou suporte para uma estrutura de um único fator, 12 (36.4%) para uma estrutura de dois fatores e dois (6.0%) para modelos bifactor ou de três fatores. Dentre os que predominaram, os estudos que confirmaram um modelo com um fator foram realizados na população em geral, em pacientes de cuidados primários e em grupos clínicos (pacientes com esclerose múltipla e pacientes com infecção por HIV). No que se refere à adequação do modelo unidimensional, a maioria (68.4%) apresentou um ajuste dos dados aceitável (Lamela et al., 2020). Portanto, os resultados do presente estudo seguem a tendência observada na referida revisão sistemática, reforçando a maior plausibilidade para o modelo unidimensional da medida e os achados desta pesquisa.

Com os dados da presente amostra, constatou-se que o PHQ-9 é mais bem representado pelo cálculo de um escore total, implicando considerar sua unidimensionalidade e satisfatória saturação dos itens em um fator primário (Vitória et al., 2006). Isso corrobora o fato de que a maior parte dos estudos acerca da estrutura interna de medidas de depressão resume os itens em uma única pontuação, interpretando essa pontuação em termos de severidade da sintomatologia (Fried et al., 2016). A unidimensionalidade de uma medida é considerada desejável num teste, pois implica no conceito de homogeneidade. Sua avaliação tem sido ponderada como central no processo de desenvolvimento de instrumentos psicológicos e deve ser entendida em termos de grau (Vitória et al., 2006). Por exemplo, na mensuração da depressão, esse escore total pode ser interpretado em termos de gravidade da sintomatologia, similar ao que ocorre com escalas próximas ao PHQ-9, tais como a CESD-R (Haroz et al., 2014; Rababah et al., 2020) e a Escala Hamilton de Depressão (Vindbjerg et al., 2019). Logo, os achados desta pesquisa apontam para a alternativa de interpretação da PHQ-9 segundo um único escore total.

Ainda acerca das evidências de validade do PHQ-9, viu-se que o instrumento apresentou correlação estatisticamente significativa com a PSS-10, que mensura o estresse percebido, e com a GAD-7, que avalia a presença e intensidade de sintomas ansiosos. Em resumo, quanto mais grave

a presença de sintomas depressivos, maiores os níveis de estresse e de ansiedade. Na literatura, há evidências robustas de que essas variáveis estão altamente associadas, sendo o estresse relatado como um dos principais fatores de risco para o desenvolvimento de sintomatologia depressiva (Campos, 2018; Gao et al., 2020; Maharaj et al., 2019; Moutinho et al., 2017; Suzuki et al., 2018). Também no caso da ansiedade, tal resultado está em conformidade com outros achados na literatura (Dominguez-Lara & Merino-Soto, 2018; Dugré et al., 2020; Kircanski et al., 2017; Salari et al., 2020). Os transtornos depressivos e de ansiedade têm uma das maiores taxas de co-ocorrência de todas as categorias de diagnóstico psiquiátrico, sendo associada ao curso clínico mais grave, maior incapacidade e maior risco de suicídio (Kircanski et al., 2017).

Outro objetivo do presente estudo consistiu na avaliação da invariância do modelo unidimensional do PHQ-9 por meio de algumas variáveis independentes de interesse. Concluiu-se que o modelo foi invariante por gênero, escolaridade e doenças crônicas nos níveis configural (equivalência na configuração do modelo), métrico (equivalência de cargas fatoriais), escalar (equivalência nos níveis dos interceptos dos itens) e estrito (equivalência dos erros da medida) (Damásio, 2013; Milfont & Fisher, 2010). Em revisão sistemática, dentre 33 estudos, pouco menos que um terço dos trabalhos (30,3%;  $n=10$ ) realizaram análise de invariância da medida em diferentes grupos (Lamela et al., 2020), o que mostra que esse tipo de análise com a PHQ-9 – mesmo que recomendada na validação de medidas (Flora & Flake, 2017; Milfont & Fisher, 2010) – ainda é escassa. De acordo com essa revisão, as variáveis gênero (González-Blanch et al., 2018; Keum et al., 2018; Patel et al., 2019) e condições clínicas (Chung et al., 2015; Doi et al., 2018; Schuler et al., 2018) foram os grupos sociodemográficos mais utilizados para avaliar a invariância do PHQ-9. Dois estudos abordaram a invariância da medida por nível de escolaridade (González-Blanch et al., 2018; Patel et al., 2019). Portanto, os achados do presente estudo quanto à invariância da medida sustentam que os escores do PHQ-9 podem ser comparados de forma válida e confiável nessas variáveis investigadas.

Diferenças na sintomatologia depressiva por gênero estão entre as descobertas mais robustas da pesquisa em psicopatologia (Salk et al., 2017). Os resultados desta pesquisa corroboram estimativas globais, tendo as mulheres maiores chances de estarem no grupo com diagnóstico de rastreamento positivo para depressão. Similarmente, níveis mais baixos de escolaridade também estão associados à ocorrência de depressão (Hakulinen et al., 2019), o que reiterou os achados desta investigação, em que os participantes com até nível médio de escolaridade tinha mais chances de compor o grupo com sintomatologia. Neste estudo, a presença de doenças crônicas, aqui utilizada como um indicador simples de condições clínicas dos participantes, não foi fator de risco para ocorrência de depressão. Cumpre ressaltar que tais resultados são psicometricamente válidos e não se devem a erro de medida, considerando que a unidimensionalidade do PHQ-9 foi invariante entre os grupos avaliados.

Dentre as limitações deste estudo, aponta-se que foi utilizado o ponto de corte do PHQ-9 indicado na literatura, ou seja, não foi realizada a normatização da medida em termos de classificação de severidade, pois isso é executado mediante a análise de validade de critério; o que não foi objetivo desta pesquisa. Outro aspecto é que os resultados encontrados foram obtidos com uma amostra composta basicamente de adultos brasileiros, já que a participação de idosos foi reduzida e outros grupos etários, como o de adolescentes, não foram incluídos. Sendo assim, é preciso ter em consideração esse perfil amostral para a compreensão destes resultados. Além disso, a maioria da amostra foi composta por participantes com nível superior de escolaridade, o que não representa a população brasileira que é majoritariamente composta por indivíduos com até nível médio de escolaridade (51%) [Instituto Brasileiro de Geografia e Estatística (IBGE), 2016]. Por fim, destaca-se que houve maior concentração de participantes da região nordeste, com menor representatividade das outras regiões do Brasil. Sugere-se que futuras investigações busquem sanar as limitações sinalizadas por este estudo.

Como uma característica peculiar desta investigação, ressalta-se que os dados foram

coletados durante a pandemia pelo novo coronavírus. Portanto, há a possibilidade de que os escores médios de depressão medidos pelo PHQ-9 tenham sido afetados em decorrência do contexto atípico de saúde vivenciado mundialmente. No entanto, os achados desta pesquisa referentes à estrutura interna da medida – em termos de hipótese – não devem ter sido afetados, considerando que esses se assemelharam a outros achados na literatura em contextos distintos ao atual (Bélanger et al., 2019; Boothroyd et al., 2019; González-Blanch et al., 2018). Cumpre ressaltar, ainda, que se trata de uma amostra populacional, a nível nacional, coletada durante o contexto de surto da COVID-19, sendo, então, uma investigação pioneira nos estudos de estrutura interna e de evidências de validade do PHQ-9.

É preciso também salientar as contribuições desta pesquisa. Primeiramente, este foi o primeiro estudo de avaliação da estrutura interna do PHQ-9 no país, verificando sua adequação para investigação da sintomatologia depressiva na população brasileira. Em segundo lugar, o tamanho da amostra; alcance regional e de quantidade de municípios; como já mencionado acima, durante a pandemia; bons indicadores em todas as análises executadas, exibindo robustez psicométrica. Finalmente, recomenda-se a utilização da PHQ-9 em pesquisas sobre depressão, principalmente devido à facilidade de administração, por se tratar de um instrumento curto, de simples entendimento, aplicação e correção. As evidências de validade reunidas e os achados concernentes à invariância da medida suportaram o modelo unidimensional do PHQ-9, reforçando a aplicabilidade e uso da PHQ-9 como medida de rastreamento da depressão no Brasil.

## Referências

- Amtmann, D., Kim, J., Chung, H., Bamer, A. M., Askew, R. L., Wu, S., Cook, & Johnson, K. L. (2014). Comparing CESD-10, PHQ-9 and PROMIS depression instruments in individuals with multiple sclerosis. *Rehabilitation Psychology, 59*(2), 220-229.  
<https://doi.org/10.1037/a0035919>

- Baptista, M. N., & Borges, L. (2016). Revisão integrativa de instrumentos de depressão em crianças/adolescentes e adultos na população brasileira. *Avaliação Psicológica*, 15, 19-32. <https://doi.org/10.15689/ap.2016.15ee.03>
- Bélanger, E., Thomas, K. S., Jones, R. N., Epstein-Lubow, G., & Mor, V. (2019). Measurement validity of the Patient-Health Questionnaire-9 in US nursing home residents. *International Journal of Geriatric Psychiatry*, 34(5), 700-708. <https://doi.org/10.1002/gps.5074>
- Boothroyd, L., Dagnan, D., & Muncer, S. (2019). PHQ-9: One factor or two?. *Psychiatry Research*, 271, 532-534. <https://doi.org/10.1016/j.psychres.2018.12.048>
- Cameron, I. M., Crawford, J. R., Lawton, K., & Reid, I. C. (2008). Psychometric comparison of PHQ-9 and HADS for measuring depression severity in primary care. *British Journal of General Practice*, 58(546), 32-36. <https://doi.org/10.3399/bjgp08X263794>
- Campos, R. C. (2018). Dados psicométricos adicionais da Versão Portuguesa do Questionário de Experiências Depressivas. *Revista Iberoamericana de Diagnóstico y Evaluación – e Avaliação Psicológica*, 46(1), 121-131. <https://doi.org/10.21865/RIDEP46.1.09>
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, 14, 464-504. <https://doi.org/10.1080/10705510701301834>
- Chilcot, J., Rayner, L., Lee, W., Price, A., Goodwin, L., Monroe, B., Sykes, N., Hansford, P., Hotopf, M. (2013). The factor structure of the PHQ-9 in palliative care. *Journal of Psychosomatic Research*, 75(1), 60-64. <https://doi.org/10.1016/j.jpsychores.2012.12.012>
- Chung, H., Kim, J., Askew, R. L., Jones, S. M. W., Cook, K. F., Amtmann, D., 2015. Assessing measurement invariance of three depression scales between neurologic samples and community samples. *Quality of Life Research*, 24, 1829-1834. <https://doi.org/10.1007/s11136-015-0927-5>
- Damásio, B. F. (2013). Contribuições da Análise Fatorial Confirmatória Multigrupo (AFCMG) na avaliação de invariância de instrumentos psicométricos. *Psico-USF*, 18(2), 211-220. <https://doi.org/10.1590/S1413-82712013000200005>
- de Lima Osório, F., Vilela Mendes, A., Crippa, J. A., & Loureiro, S. R. (2009). Study of the discriminative validity of the PHQ-9 and PHQ-2 in a sample of Brazilian women in the context of Primary Health care. *Perspectives in Psychiatric Care*, 45(3), 216-227. <https://doi.org/10.1111/j.1744-6163.2009.00224.x>
- Doi, S., Ito, M., Takebayashi, Y., Muramatsu, K., & Horikoshi, M. (2018). Factorial validity and invariance of the Patient Health Questionnaire (PHQ)-9 among clinical and non-clinical populations. *PloS One*, 13(7), 1-9. <https://doi.org/10.1371/journal.pone.0199235>
- Dominguez-Lara, S., & Merino-Soto, C. (2018). Cognitive Emotional Regulation Questionnaire-18 en universitarios: Evidencias de validez convergente y discriminante. *Revista Iberoamericana de Diagnóstico y Evaluación – e Avaliação Psicológica*, 47(2), 171-184. <https://doi.org/10.21865/RIDEP47.2.12>
- Dugré, J. R., Dumais, A., Dellazizzo, L., & Potvin, S. (2020). Developmental joint trajectories of anxiety-depressive trait and trait-aggression: Implications for co-occurrence of internalizing and externalizing problems. *Psychological Medicine*, 50(8), 1338-1347. <https://doi.org/10.1017/S0033291719001272>
- Faro, A. (2013). Estresse e estressores na pós-graduação: Estudo com mestrandos e doutorandos no Brasil. *Psicologia: Teoria e Pesquisa*, 29, 51-60. <https://doi.org/10.1590/S0102-37722013000100007>
- Ferenchick, E. K., Ramanuj, P., & Pincus, H. A. (2019). Depression in primary care: Part 1 – Screening and diagnosis. *BMJ*, 365, 1-10. <https://doi.org/10.1136/bmj.l794>
- Flora, D. B., & Flake, J. K. (2017). The purpose and practice of exploratory and confirmatory factor analysis in psychological research: Decisions for scale development and validation. *Canadian Journal of Behavioral Science*, 49(2), 78-88.

- <https://doi.org/10.1037/cbs0000069>.
- Fried, E. I., van Borkulo, C. D., Epskamp, S., Schoevers, R. A., Tuerlinckx, F., & Borsboom, D. (2016). Measuring depression over time... Or not? Lack of unidimensionality and longitudinal measurement invariance in four common rating scales of depression. *Psychological Assessment*, 28(11), 1354-1367.  
<https://doi.org/10.1037/pas0000275>
- Gao, W., Ping, S., & Liu, X. (2020). Gender differences in depression, anxiety, and stress among college students: A longitudinal study from China. *Journal of Affective Disorders*, 263, 292-300.  
<https://doi.org/10.1016/j.jad.2019.11.121>
- GBD 2016 DALYs and HALE Collaborators (2017). Global, regional, and national disability-adjusted life-years (DALYs) for 333 diseases and injuries and healthy life expectancy (HALE) for 195 countries and territories, 1990-2016: A systematic analysis for the Global Burden of Disease Study 2016. *The Lancet*, 390(10100), 1260-1344.  
[https://doi.org/10.1016/S0140-6736\(17\)32130-X](https://doi.org/10.1016/S0140-6736(17)32130-X)
- González-Blanch, C., Medrano, L. A., Muñoz-Navarro, R., Ruíz-Rodríguez, P., Moriana, J. A., Limonero, J. T., Schmitz, F., & Cano-Vindel, A. (2018). Factor structure and measurement invariance across various demographic groups and over time for the PHQ-9 in primary care patients in Spain. *PloS One*, 13(2), 1-16.  
<https://doi.org/10.1371/journal.pone.0193356>
- Gunzler, D., Sehgal, A. R., Kauffman, K., Davey, C. H., Dolata, J., Figueroa, M., Huml, A., Pencak, J., & Sajatovic, M. (2020). Identify depressive phenotypes by applying RDOC domains to the PHQ-9. *Psychiatry Research*, 286(112872), 1-10.  
<https://doi.org/10.1016/j.psychres.2020.112872>
- Hair, J., Black, W., Babin, B., Anderson, R., & Tatham, R. (2014). *Multivariate data analysis*. Pearson New International Edition/Prentice Hall.
- Hakulinen, C., Musliner, K. L., & Agerbo, E. (2019). Bipolar disorder and depression in early adulthood and long-term employment, income, and educational attainment: A nationwide cohort study of 2,390,127 individuals. *Depression and Anxiety*, 36(11), 1080-1088. <https://doi.org/10.1002/da.22956>
- Hancock, G. R., & Mueller, R. O. (2001). Rethinking construct reliability within latent variable systems. In R. Cudeck, S. du Toit, & D. Sörbom (Eds.), *Structural Equation Modeling: Present and Future* (pp. 195-216). Scientific Software International, Inc.
- Haroz, E. E., Ybarra, M. L., & Eaton, W. W. (2020). Psychometric evaluation of a self-report scale to measure adolescent depression: The CESDR-10 in two representative adolescent samples in the United States. *Journal of Affective Disorders*, 158, 154-160.  
<https://doi.org/10.1016/j.jad.2014.02.009>
- Hinz, A., Mehnert, A., Kocalevent, R. D., Brähler, E., Forkmann, T., Singer, S., & Schulte, T. (2016). Assessment of depression severity with the PHQ-9 in cancer patients and in the general population. *BMC Psychiatry*, 16(1), 22. <https://doi.org/10.1186/s12888-016-0728-6>
- Hirschtritt, M. E., & Kroenke, K. (2017). Screening for depression. *JAMA*, 318(8), 745-746. <https://doi.org/10.1001/jama.2017.9820>
- Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, 3, 424-453.  
<https://doi.org/10.1037/1082-989X.3.4.424>
- Keum, B. T., Miller, M. J., & Inkelas, K. K. (2018). Testing the factor structure and measurement invariance of the PHQ-9 across racially diverse US college students. *Psychological Assessment*, 30(8), 1096-1106.  
<https://doi.org/10.1037/pas0000550>
- Kircanski, K., LeMoult, J., Ordaz, S., & Gotlib, I. H. (2017). Investigating the nature of co-occurring depression and anxiety: Comparing diagnostic and dimensional research approaches. *Journal of Affective Disorders*, 216, 123-135.  
<https://doi.org/10.1016/j.jad.2016.08.006>
- Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine*, 16(9), 606-613.  
<https://doi.org/10.1046/j.1525-1497.2001.016009606.x>

- Kroenke, K., Spitzer, R. L., Williams, J. B., Monahan, P. O., & Löwe, B. (2007). Anxiety disorders in primary care: Prevalence, impairment, comorbidity, and detection. *Annals of Internal Medicine*, 146(5), 317-325. <https://doi.org/10.7326/0003-4819-146-5-200703060-00004>
- Lamela, D., Soreira, C., Matos, P., & Morais, A. (2020). Systematic review of the factor structure and measurement invariance of the patient health questionnaire-9 (PHQ-9) and validation of the Portuguese version in community settings. *Journal of Affective Disorders*, 276, 220-233. <https://doi.org/10.1016/j.jad.2020.06.066>
- Maharaj, S., Lees, T., & Lal, S. (2019). Prevalence and risk factors of depression, anxiety, and stress in a cohort of Australian nurses. *International Journal of Environmental Research and Public Health*, 16(1), 61-71. <https://doi.org/10.3390/ijerph16010061>
- Milfont, T. L. & Fisher, R. (2010). Testing measurement invariance across groups: Applications in cross-cultural research. *International Journal of Psychological Research*, 3, 111-121. <https://doi.org/10.21500/20112084.857>
- Mitchell, A. J., Yadegarfar, M., Gill, J., & Stubbs, B. (2016). Case finding and screening clinical utility of the Patient Health Questionnaire (PHQ-9 and PHQ-2) for depression in primary care: a diagnostic meta-analysis of 40 studies. *BJ Psych Open*, 2(2), 127-138. <https://doi.org/10.1192/bjpo.bp.115.001685>
- Monteiro, S., Bártolo, A., Torres, A., Pereira, A., & Albuquerque, E. (2019). Examining the construct validity of the Portuguese version of the Patient Health Questionnaire-9 among college students. *Revista PSICOLOGIA*, 33(2), 1-8. <https://doi.org/10.17575/rpsicol.v33i2.1421>
- Moutinho, I. L. D., Maddalena, N. C. P., Roland, R. K., Lucchetti, A. L. G., Tibiriçá, S. H. C., Ezequiel O. S. (2017). Depression, stress and anxiety in medical students: A cross-sectional comparison between students from different semesters. *Revista da Associação Médica Brasileira*, 63(1), 21-28. Recuperado de [https://www.scielo.br/scielo.php?script=sci\\_ar](https://www.scielo.br/scielo.php?script=sci_ar) ttext&pid=S0104-42302017000100021&lng=en&tlng=en
- Patel, J. S., Oh, Y., Rand, K. L., Wu, W., Cyders, M. A., Kroenke, K., & Stewart, J. C. (2019). Measurement invariance of the patient health questionnaire-9 (PHQ-9) depression screener in US adults across sex, race/ethnicity, and education level: NHANES 2005-2016. *Depression and Anxiety*, 36(9), 813-823. <https://doi.org/10.1002/da.22940>
- Rababah, J., Al-Hammouri, M. M., Drew, B. L., Alawawdeh, A., Dawood, Z., & Jawhar, W. E. (2020). Validation of the Arabic version of the Center for Epidemiologic Studies Depression-Revised: A comparison of the CESD-R and CESDR-12. *Journal of Affective Disorders*. <https://doi.org/10.1016/j.jad.2020.05.124>
- Salari, N., Hosseini-Far, A., Jalali, R., Vaisi-Raygani, A., Rasoulpoor, S., Mohammadi, M., Rasoulpoor, S., & Khaledi-Paveh, B. (2020). Prevalence of stress, anxiety, depression among the general population during the COVID-19 pandemic: A systematic review and meta-analysis. *Globalization and Health*, 16(1), 1-11. <https://doi.org/10.1186/s12992-020-00589-w>
- Salk, R. H., Hyde, J. S., & Abramson, L. Y. (2017). Gender differences in depression in representative national samples: Meta-analyses of diagnoses and symptoms. *Psychological Bulletin*, 143(8), 783. <https://doi.org/10.1037/bul0000102>
- Santos, I. S., Tavares, B. F., Munhoz, T. N., Almeida, L. S. P. D., Silva, N. T. B. D., Tams, B. D., Patella, A. M., & Matijasevich, A. (2013). Sensibilidade e especificidade do Patient Health Questionnaire-9 (PHQ-9) entre adultos da população geral. *Cadernos de Saúde Pública*, 29, 1533-1543. <https://doi.org/10.1590/0102-311X00144612>
- Schuler, M., Strohmayer, M., Mühlig, S., Schwaighofer, B., Wittmann, M., Faller, H., & Schultz, K. (2018). Assessment of depression before and after inpatient rehabilitation in COPD patients: Psychometric properties of the German version of the Patient Health Questionnaire (PHQ-9/PHQ-2). *Journal of Affective Disorders*, 232, 268-275. <https://doi.org/10.1016/j.jad.2018.02.037>

- Siu, A. L., & US Preventive Services Task Force (2016). Screening for depression in adults: US preventive services task force recommendation statement. *JAMA*, 315(4), 380-387.  
<https://doi.org/10.1001/jama.2015.18392>
- Suzuki, M., Furihata, R., Konno, C., Kaneita, Y., Ohida, T., & Uchiyama, M. (2018). Stressful events and coping strategies associated with symptoms of depression: A Japanese general population survey. *Journal of Affective Disorders*, 238, 482-488.  
<https://doi.org/10.1016/j.jad.2018.06.024>
- Vindbjerg, E., Makranksy, G., Mortensen, E. L., & Carlsson, J. (2019). Cross-cultural psychometric properties of the Hamilton Depression Rating Scale. *The Canadian Journal of Psychiatry*, 64(1), 39-46.  
<https://doi.org/10.1177/0706743718772516>
- Vitória, F., Almeida, L. S., & Primi, R. (2006). Unidimensionalidade em testes psicológicos: Conceito, estratégias e dificuldades na sua avaliação. *Psic: Revista da Vetor Editora*, 7(1), 01-07. Recuperado de [http://pepsic.bvsalud.org/scielo.php?script=sci\\_arttext&pid=S1676-73142006000100002](http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1676-73142006000100002)
- Weiss, R. B., Aderka, I. M., Lee, J., Beard, C., & Björgvinsson, T. (2015). A comparison of three brief depression measures in an acute psychiatric population: CES-D-10, QIDS-SR, and DASS-21-DEP. *Journal of Psychopathology and Behavioral Assessment*, 37(2), 217-230.  
<https://doi.org/10.1007/s10862-014-9461-y>
- WHO, World Health Organization. Depression and other common mental disorders: Global health estimates. Geneva: WHO, 2017.
- Yu, X., Tam, W. W., Wong, P. T., Lam, T. H., & Stewart, S. M. (2012). The Patient Health Questionnaire-9 for measuring depressive symptoms among the general population in Hong Kong. *Comprehensive Psychiatry*, 53(1), 95-102.  
<https://doi.org/10.1016/j.comppsy.2010.11.002>
- Zhen, L., Wang, G., Xu, G., Xiao, L., Feng, L., Chen, X., Liu, M., & Zhu, X. (2020). Evaluation of the paper and smartphone versions of the Quick Inventory of Depressive Symptomatology-Self-Report (QIDS-SR16) and the Patient Health Questionnaire-9 (PHQ-9) in Depressed Patients in China. *Neuropsychiatric Disease and Treatment*, 16, 993-1001.  
<https://doi.org/10.2147/NDT.S241766>
- Zimmerman, M. (2019). Using the 9-Item Patient Health Questionnaire to Screen for and Monitor Depression. *JAMA*, 322(21), 2125-2126.  
<https://doi.org/10.1001/jama.2019.15883>