



ARTIGOS

Análise do questionário de saúde do paciente durante a pandemia da covid-19 via TRI

Analysis of the patient health questionnaire during the covid-19 pandemic using IRT
Análisis del cuestionario de salud del paciente durante la pandemia de covid-19 mediante TRI

Maria Heloísa Santos-Souza¹

orcid.org/0000-0001-5985-7197
mheloisasouzapsi@gmail.com

Brenda Fernanda

Pereira da Silva Ferraz¹

orcid.org/0000-0003-1139-0342
brendafernandapsi@gmail.com

Dandara Barbosa

Palhano¹

orcid.org/0000-0003-1625-6871
palhanodandara@gmail.com

André Faro¹

orcid.org/0000-0002-7348-6297
andrefaro.ufs@gmail.com

Recebido em: 18 dez. 2024

Aprovado em: 13 out. 2025

Publicado em: 24 ago. 2026.

Resumo: Embora o Questionário de Saúde do Paciente-g (PHQ-g) tenha sido amplamente utilizado durante a pandemia da covid-19, para rastreio de sintomas depressivos, ainda há lacunas na literatura sobre suas evidências de validade na população brasileira durante o período. Assim, este estudo teve como objetivo avaliar as propriedades psicométricas da PHQ-g no Brasil, através da Teoria de Resposta ao Item (TRI). A coleta foi feita durante cinco fases da pandemia, entre 2020 e 2023. Os *softwares* SPSS, R e RStudio foram utilizados para análise de dados. Participaram do estudo 9.229 indivíduos, e mais da metade dos respondentes apresentou sintomas de depressão. Foram encontrados bons parâmetros de ajuste no modelo e desempenho satisfatório quanto aos parâmetros de discriminação e dificuldade na escala. O perfil de respostas foi predominantemente dicotômico, e houve funcionamento diferencial dos itens a partir do gênero. Esta pesquisa contribui para o estabelecimento de parâmetros confiáveis para avaliar sintomas depressivos.

Palavras-chave: psicometria; questionário de saúde do paciente; depressão; covid-19.

Abstract: Although the Patient Health Questionnaire-g (PHQ-g) has been widely used during the COVID-19 pandemic to screen for depressive symptoms, gaps remain in the literature regarding its validity in the Brazilian population during this period. This study aimed to evaluate the psychometric properties of the PHQ-g in Brazil using Item Response Theory (IRT). Data were collected during five phases of the pandemic, from 2020 to 2023. SPSS, R, and RStudio software were used for data analysis. The study included 9,229 participants, with more than half of the population presenting clinical symptoms of depression. Good fit parameters were found for the model, with satisfactory performance in the scale's discrimination and difficulty parameters. The response profile was predominantly dichotomous, and differential item functioning was observed by gender. This research contributes to the establishment of reliable parameters for assessing depressive symptoms.

Keywords: psychometrics; patient health questionnaire; depression; COVID-19.

Resumen: Aunque la escala Patient Health Questionnaire-g (PHQ-g) ha sido ampliamente utilizada durante la pandemia de Covid-19 para la detección de síntomas depresivos, aún existen vacíos en la literatura sobre su validez en la población brasileña durante este período. Este estudio tuvo como objetivo evaluar las propiedades psicométricas del PHQ-g en Brasil a través de la Teoría de Respuesta al Ítem (TRI). La recolección de datos se realizó durante cinco fases de la pandemia, entre 2020 y 2023. Se utilizaron los programas SPSS, R y RStudio para los análisis de datos. El estudio incluyó 9229 participantes, y más de la mitad de la población presentó síntomas clínicos de depresión. Se encontraron buenos parámetros de ajuste en el modelo y un desempeño satisfactorio en cuanto a los parámetros de discriminación y dificultad de la escala. El perfil de respuestas fue predominantemente dicotómico y se observó un funcionamiento diferencial de los ítems según el género. Esta investigación contribuye al establecimiento



Este é um artigo de acesso aberto distribuído sob a licença [CC-BY 4.0](https://creativecommons.org/licenses/by/4.0/), que permite a cópia e redistribuição do material em qualquer formato e para qualquer finalidade, desde que a autoria original e os créditos de publicação sejam mantidos.

¹ Universidade Federal do Sergipe, São Cristóvão, SE, Brasil.

de parâmetros confiáveis para avaliar los síntomas depresivos.

Palabras clave: psicometria; cuestionario de salud del paciente; depresión; Covid-19.

A depressão é uma preocupação pública no Brasil, já que houve aumento generalizado nos índices do transtorno no país entre 2013 e 2019 (Brito et al., 2022). No que concerne à população mundial, a Organização Mundial de Saúde (OMS) estimou que cerca de 320 milhões de pessoas conviviam com a depressão até 2015 (Organização Pan-Americana de Saúde [OPAS], 2017). Além disso, houve um aumento de 25% nos níveis de depressão em todo o planeta durante o primeiro ano da pandemia da covid-19 (OPAS, 2022).

A pandemia foi uma crise global ocasionada pela disseminação do novo coronavírus e propagação de uma pneumonia contagiosa e letal, que afetou vários países (Tassara et al., 2021). A inexistência inicial de vacinas e tratamentos eficazes instaurou um cenário de medo e incertezas (Moura et al., 2022). Medidas restritivas, como o isolamento social, foram adotadas para diminuir o contágio (Tassara et al., 2021), o que gerou impactos significativos na qualidade de vida e na saúde mental das pessoas (Santana et al., 2020).

Após o início da emergência sanitária, a elevação dos níveis de depressão foi amplamente identificada na literatura (Badaró et al., 2021; Musse et al., 2022; Tassara et al., 2021). Observou-se que a depressão e as emoções negativas estavam associadas à ampla divulgação de informações sobre a situação da saúde pública (Santana et al., 2020). Também houve maior prevalência de sintomas de depressão em mulheres durante a pandemia, do mesmo modo que fora dela (Barros et al., 2020; Moura et al., 2022). Outro estudo mostrou que pessoas com diagnósticos anteriores de depressão apresentaram a saúde mental ainda mais vulnerável do que pessoas que nunca foram diagnosticadas previamente (Barros et al., 2020).

Os transtornos depressivos são caracterizados pelo humor triste, irritável ou vazio e por prejuízos no funcionamento do indivíduo, em decorrência de modificações somáticas e cognitivas (*American Psychiatric Association [APA]*,

2014). Cognitivamente, a depressão é composta de uma tríade: a visão negativa que a pessoa tem de si, do mundo e do futuro (Beck & Alford, 2011). Pacientes com sintomas depressivos tendem a ver a si mesmos como indignos ou inadequados, a interpretar as experiências como derrotas ou depreciações e a acreditar que seus sofrimentos e suas frustrações permanecerão ao longo do tempo.

O transtorno depressivo maior (TDM), que contém os elementos clássicos desse conjunto de transtornos, é caracterizado pelo humor deprimido quase todos os dias, pela perda do interesse ou prazer em atividades, pelo aumento ou pela perda de peso, pela insônia ou hipersonia frequente, pela fadiga, pela agitação psicomotora, pela diminuição da concentração e pelos pensamentos de morte (APA, 2014). O TDM está associado à alta mortalidade, sobretudo por suicídio, sendo que o risco de comportamentos suicidas é constante em indivíduos em episódio depressivo maior. Portanto, a mensuração de sintomas da depressão é necessária para a identificação de panoramas sobre a saúde mental e de comportamentos de risco durante e após o período pandêmico.

A *patient health questionnaire-g* (PHQ-g) é uma escala de autorrelato desenvolvida por Kroenke, Spitzer e Williams para a triagem e avaliação da gravidade de sintomas depressivos (Kroenke et al., 2001). O instrumento é composto de nove itens, que se baseiam nos critérios do Manual Diagnóstico e Estatístico dos Transtornos Mentais (DSM) e avaliam o humor, o sono, o apetite, a anedonia, a concentração, as mudanças psicomotoras e os pensamentos de morte. A PHQ-g apresentou robustez psicométrica na população brasileira e evidências de validade no período anterior à pandemia (Nunes & Faro, 2022). A escala também apresentou evidências de validade em vários países, como na Argentina (Urtasun et al., 2019), na Coreia do Sul (Shin et al., 2020), na China (Ye et al., 2020), na Espanha (Gómez-Gómez et al., 2023), nos Estados Unidos (Kalkbrenner et al., 2024), no México (Arrieta et al., 2017) e em Moçambique (Cumbe et al., 2020).

O instrumento tem sido utilizado em diferentes contextos, na literatura nacional e internacional, a exemplo da população quilombola (Barroso et al., 2019), dos pacientes da atenção primária (Cumbe et al., 2020; Damiano et al., 2023), dos estudantes universitários (Lingán-Huamán et al., 2023) e do público adolescente (Gao & Liu, 2024). Durante a pandemia, a PHQ-9 foi utilizada em diferentes países, como no Brasil (Almeida et al., 2023; Musse et al., 2022), na Itália, na Irlanda, no Reino Unido, na Espanha (Shevlin et al., 2022) e no Peru (Lingán-Huamán et al., 2023).

O uso de escalas de rastreio validadas possibilita uma melhor avaliação da prevalência de sintomas de transtornos, reduzindo dúvidas acerca de diagnósticos relatados (Brito et al., 2022). Logo, para identificar aspectos de saúde mental, profissionais e pesquisadores precisam de instrumentos de medida confiáveis (Kalkbrenner et al., 2024). Apesar de a confiabilidade e a validade da PHQ-9 terem sido avaliadas em diferentes contextos, a literatura ainda carece de estudos sobre sua validade no contexto pandêmico no Brasil.

A Teoria de Resposta ao Item (TRI) é uma linha teórica da psicometria que tem sido utilizada em estudos de saúde pública para verificar a confiabilidade e validade de instrumentos (Gomes et al., 2018). Seu uso possibilita a identificação da relação entre os traços latentes (teta), isto é, as variáveis hipotéticas, e os comportamentos observáveis através de equações logísticas (Pascual & Primi, 2003). Os parâmetros produzidos por meio da equação logística e da curva característica dos itens (CCI) indicam se os itens de um teste são bons e confiáveis (Bichi & Talib, 2018).

Dois dos parâmetros que são utilizados pela TRI são a discriminação (α) e a dificuldade (b) (Bichi & Talib, 2018). O parâmetro α avalia a capacidade do item de diferenciar pessoas a partir dos seus níveis de teta. Um item com bom valor de discriminação indica que pessoas com maior nível do traço latente possuem maior probabilidade de endosso, ao passo que as pessoas com menor teta possuem menos chances. O inverso ocorre quando o item obtém valores

insatisfatórios de discriminação. O α dos itens de um bom teste se encontra em um intervalo entre 0,5 e 2,0 (Bichi & Talib, 2018).

A dificuldade é um parâmetro que avalia se pessoas com maior valor de teta têm maior probabilidade de endosso e se participantes com menor valor de teta têm menor probabilidade de endosso. Valores satisfatórios de dificuldade são encontrados em itens com b maior que um (Bichi & Talib, 2018). Além disso, para avaliar a existência de possíveis vieses na escala, a TRI dispõe de um mecanismo denominado "funcionamento diferencial do item" (DIF). O DIF é um elemento que identifica se as características das pessoas exercem influência sobre suas respostas (Gomes et al., 2018).

Assim, a TRI possibilita a transformação de dados sobre informações subjetivas em probabilidades matemáticas para indicar a confiabilidade dos testes (Gomes et al., 2018). Estudos recentes já avaliaram a PHQ-9, a partir da TRI (Cumbe et al., 2020; Gao & Liu, 2024; Jiraniramai et al., 2021), inclusive na população brasileira (Barroso et al., 2019; Damiano et al., 2023) e durante a pandemia no cenário peruano (Lingán-Huamán et al., 2023). Apesar disso, ainda não há estudos que verifiquem a validade do instrumento na população brasileira, no período da pandemia, por meio da TRI.

Diante desses aspectos, entende-se a crescente importância da TRI na avaliação de escalas psicométricas. Dados o uso da escala PHQ-9 no Brasil, durante a pandemia (Almeida et al., 2023; Musse et al., 2022), e a necessidade de se ter instrumentos confiáveis de avaliação da saúde mental (Brito et al., 2022; Kalkbrenner et al., 2024), é fundamental avaliar a validade da PHQ-9 em uma amostra brasileira durante o período. Logo, o objetivo deste estudo é avaliar as propriedades psicométricas da escala PHQ-9 em uma amostra brasileira, durante a pandemia, e verificar se há funcionamento diferencial do(s) item(ns) a partir do gênero dos participantes.

Métodos

Participantes

O presente estudo teve como público-alvo pessoas residentes de todas as regiões do Brasil. No total, participaram 9.929 pessoas, dentre as quais 88,2% ($n = 8756$) eram mulheres, 55,0% brancos ($n = 5841$) e 44,8% pardos ou pretos ($n = 4448$). Cerca de 82,0% possuía Ensino Superior ($n = 8098$). A maior parte da amostra residia nas regiões Nordeste (39,3%; $n = 3902$) e Sudeste (38,6%; $n = 3834$). Em menor proporção, houve participantes do Sul (13,8%; $n = 1368$), Centro-Oeste (5,2%; $n = 519$) e Norte (3,1%; $n = 306$). A amostra foi composta apenas de pessoas adultas, com média de 36,07 anos ($DP = 13,68$).

Instrumentos

Patient health questionnaire-g (Kroenke et al., 2001). Os sintomas depressivos foram mensurados por meio da PHQ-9. O instrumento possui nove itens baseados nos critérios do DSM-5 para depressão. Cada item apresenta quatro opções de resposta, que representam a frequência dos sintomas nas últimas duas semanas. As categorias são: nenhuma vez (0), vários dias (1), mais da metade dos dias (2) e quase todos os dias (3). Os escores podem variar de 0 a 27 pontos, e os sintomas foram avaliados em cinco níveis: sem depressão (0-4), depressão leve (5-9), depressão moderada (10-14), depressão moderadamente grave (15-19) e depressão grave (acima de 20). A validação inicial do instrumento encontrou bons resultados de confiabilidade, cujo valor do α de Cronbach foi de 0,89 em uma aplicação com profissionais da atenção primária e 0,86 com pacientes (Kroenke et al., 2001). O mesmo estudo identificou sensibilidade e especificidade de 88% na escala. Em 2022, a PHQ-9 também apresentou indicadores satisfatórios de confiabilidade ($\alpha = 0,93$; $\Omega = 0,93$) e replicabilidade (H -latent = 0,94) em estudo recente no Brasil (Nunes & Faro, 2022).

A caracterização da amostra foi feita por meio de um questionário sociodemográfico. Foram incluídas perguntas sobre o gênero, por meio das categorias (feminino, masculino, outros), entretanto a categoria "outros" foi excluída, em virtude do baixo número de participação. Tam-

bém foram solicitadas informações sobre a idade (em anos), cor de pele (amarela, branca, indígena, parda, preta ou outra) e região (Centro-Oeste, Nordeste, Norte, Sudeste ou Sul).

Procedimento e aspectos éticos

Trata-se de um estudo *on-line*, de delineamento transversal, com amostragem por bola de neve. A coleta foi dividida em cinco fases. Duas fases aconteceram em 2020, sendo a primeira delas entre 18 e 21 de março, e a segunda entre 2 e 8 de junho. As demais fases ocorreram nos anos seguintes. A terceira fase aconteceu entre 18 de março e 4 de abril de 2021, e a quarta e a quinta fase aconteceram na segunda quinzena de março de 2022 e 2023, respectivamente. O convite de participação da pesquisa foi feito por meio de redes sociais (Instagram, Facebook e WhatsApp).

O estudo foi aprovado pela Comissão Nacional de Ética em Pesquisa (CONEP, CAAE: 30485420.6.0000.0008), assegurando o cumprimento dos requisitos éticos necessários em pesquisas com seres humanos. Os participantes assentiram ao Registro de Consentimento Livre e Esclarecido (RCLE) para garantir o consentimento na participação do estudo.

Análise estatística

As estatísticas descritivas foram obtidas e ajustadas por meio do *software* SPSS (versão 25), sendo obtidas frequências, porcentagens, médias e desvios-padrão. Os *softwares* R 4.3.2. e RStudio foram usados para obter os parâmetros e pressupostos da TRI. Utilizou-se do pacote "Mirt" para gerar o modelo para itens politômicos. Ainda, optou-se pela escolha do modelo de resposta gradual (GRM), uma vez que esse é considerado um modelo robusto (Dai et al., 2021). Os índices Root Mean Square Error of Approximation (RMSEA) e Tucker-Lewis Index (TLI) foram utilizados para atestar a qualidade de ajuste do modelo. A partir do modelo escolhido, foram obtidos os parâmetros de discriminação e dificuldade e a curva característica dos itens, bem como foram calculados a informação e o erro.

Resultados

Caracterização da amostra

A média do escore total do PHQ-9 na amostra foi de 13,25 ($DP = 7,51$). Mais da metade dos participantes apresentou sintomas clínicos para a depressão (64,1%; acima de 10 pontos). Quanto aos níveis de pontuação da escala, 14,7% ($n = 1464$) se encaixou no nível sem depressão, 21,1% possuía depressão leve ($n = 2098$), 19,3%, depressão moderada ($n = 1918$), 19,8%, depressão moderadamente grave ($n = 1961$), e 25,1%, depressão grave ($n = 2488$).

Análise da PHQ-9 por meio da TRI

O modelo apresentou índices de ajuste considerados satisfatórios (CFI = 0,94; RMSEA = 0,056; SRMR = 0,059; TLI = 0,93). A Tabela 1 descreve os parâmetros a e b , além dos achados de DIF da PHQ-9. O parâmetro de discriminação de cada item apresentou variações, já que a pontuação mínima foi de 1,68 e a máxima de 3,69. A menor discriminação foi identificada no item 9, sobre pensar em se autolesionar e ter desejo de morte, seguida pelo item 8, sobre lentidão/agitação. O item 2, sobre estar deprimido ou sem perspectiva, obteve a maior discriminação, seguido pelo item 4, sobre se sentir cansado.

Tabela 1 – Discriminação, dificuldade e função diferencial do item da escala PHQ-9

Itens	a	b			DIF		
		b1	b2	b3	M	X2	p
1. Ter pouco interesse ou prazer com as coisas.	2,928	-1,156	0,102	0,629	-0,106	11,75	0,022
2. Sentir-se "para baixo", deprimido/a ou sem perspectiva.	3,696	-1,138	-0,002	0,474	-0,166	28,454	0,000
3. Demonstrar dificuldade para pegar no sono ou permanecer dormindo ou dormir mais do que de costume.	2,104	-1,272	-0,176	0,295	-0,288	22,102	0,000
4. Sentir-se cansado/a ou com pouca energia.	2,972	-1,584	-0,252	0,206	-0,407	62,787	0,000
5. Apresentar falta de apetite ou comer demais.	1,933	-1,125	-0,097	0,419	-0,2	41,953	0,000
6. Sentir-se mal consigo mesmo/a – ou achar que é um fracasso ou que decepcionou a família ou a si mesmo.	2,417	-0,730	0,116	0,523	-0,022	28,045	0,000
7. Ter dificuldade para se concentrar nas coisas, como ler o jornal ou ver a televisão.	2,102	-0,948	0,189	0,713	-0,011	19,079	0,001
8. Apresentar lentidão para se movimentar ou falar, a ponto de as outras pessoas perceberem, ou o oposto, isto é, estar tão agitado/a ou inquieto/a, que fica andando de um lado para o outro mais do que de costume.	1,758	-0,103	0,803	1,416	0,529	0,465	0,977
9. Pensar em se ferir de alguma maneira ou que seria melhor estar morto.	1,686	0,688	1,464	1,918	1,017	19,887	0,001

Notas. a = parâmetro de discriminação; b = parâmetro de dificuldade.

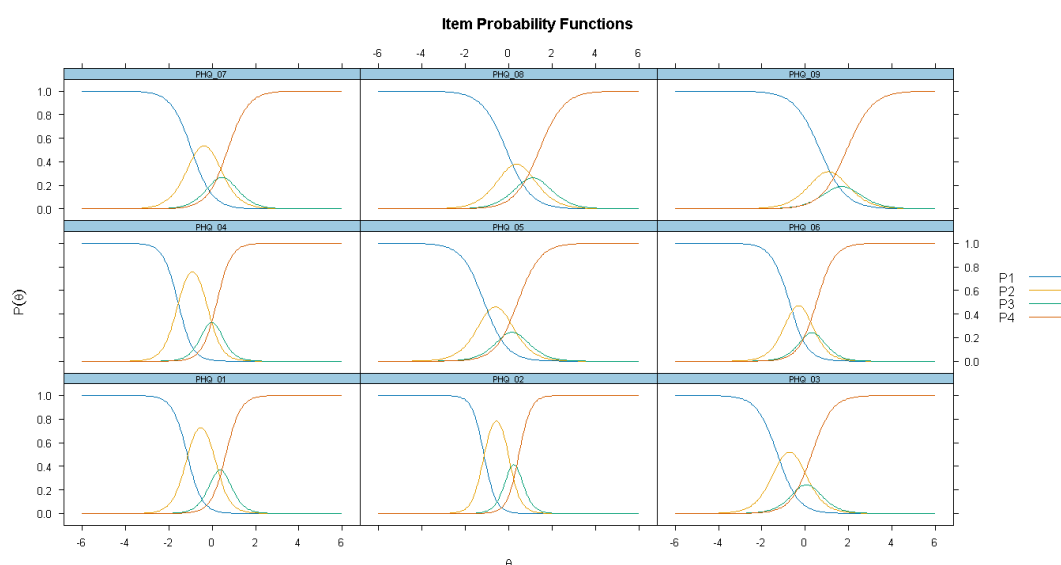
De modo geral, os itens apresentaram probabilidade moderada de endosso. A questão 9

apresentou maior dificuldade de endosso entre os itens, com valores de 0,68, 1,46 e 1,91, para b1, b2 e b3, respectivamente, e média de 1,017. O item 4, sobre se sentir cansado ou com pouca energia, apresentou maior probabilidade de endosso, com os valores de -1,58, -1,25 e 0,2 para b1, b2 e b3, respectivamente, e média de -0,407.

A Figura 1 apresenta a curva característica de cada item. Predominantemente, os participantes marcaram categorias nas extremidades, as quais representam maior frequência de sintomas depressivos ou ausência de sintomas. Os gráficos apresentaram disposições similares de teta em relação às categorias, com variações maiores nos itens 8 e 9. Na categoria 1 (nenhuma vez),

houve predominância de respostas em participantes que pontuaram até a faixa de teta de até -2. A categoria 2 apresentou níveis de teta no intervalo entre -1 e 0. A categoria 4 (quase todos os dias) se sobrepôs à categoria 3 (mais da metade dos dias), em participantes com teta igual ou superior a 1, sendo que, nos itens 3, 5, 6 e 7, a sobreposição foi mais expressiva. Os itens 8 e 9 apresentaram maior incidência de respostas nas categorias extremas, com nível de teta até -1 e 0, respectivamente, para a categoria 1. Houve sobreposição das categorias 2 e 3 pela categoria 4, sobretudo no item 9, em participantes com teta superior ou igual a 2,5.

Figura 1 – Curva característica dos itens da PHQ-9

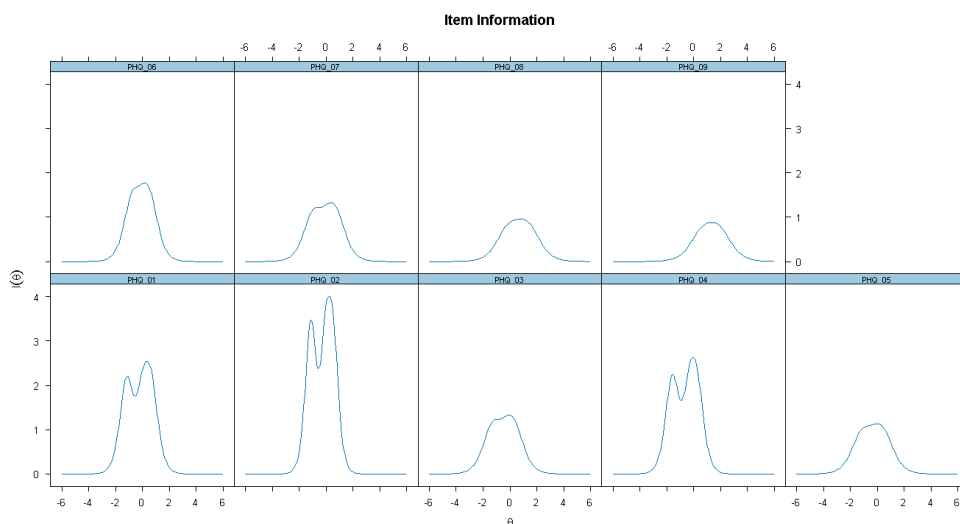


Fonte: Dados da pesquisa.

A Figura 2 apresenta a curva de informação de cada item, e a Figura 3, a curva da informação e do erro. Em geral, os itens foram mais informativos, com teta entre -3 e 3, e, nos demais intervalos de teta, há queda da informação e maior erro. Os itens 8 e 9 se apresentaram de modo

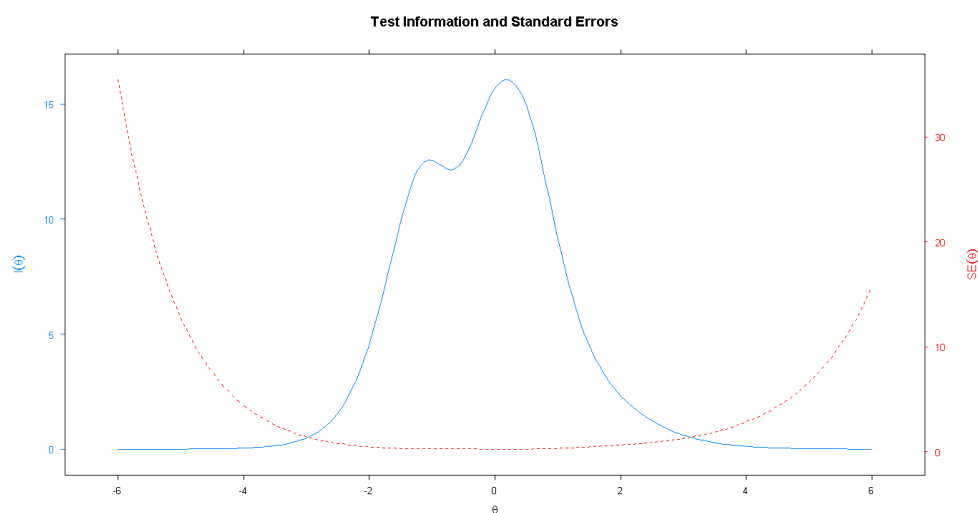
distinto, já que foram mais informativos entre os tetras -1 e 3. O item mais informativo foi o item 2, sobre se sentir deprimido ou sem perspectiva. Os itens 1, 2 e 3 apresentaram curvas similares, com informação entre 3 e 4 e desnível entre teta -1 e 0. A área da informação total foi de 45,43.

Figura 2 – Curva da informação do item da escala PHQ-9



Fonte: Dados da pesquisa.

Figura 3 – Curva da informação e erro da escala PHQ-9



Fonte: Dados da pesquisa.

Por fim, foi avaliada a função diferencial do item (DIF) de gênero, a fim de identificar se há viés nos itens da PHQ-9. Foi utilizado o valor de $p < 0,001$ como critério para o DIF em virtude do tamanho amostral. Assim, a PHQ-9 mensurou de modo distinto homens e mulheres com mesmo nível de teta neste estudo, pois foram identifica

das diferenças estatisticamente significativas nos valores de p em sete itens da escala. Apenas os itens 1 e 8 mensuraram igualmente pessoas de ambos os gêneros com mesmo traço latente. A Tabela 1 apresenta os valores encontrados nas análises do DIF.

Discussão

Este estudo teve como objetivo analisar os itens da escala PHQ-9 na população brasileira, durante a pandemia da Covid-19, por meio da TRI. Foram obtidos os parâmetros da discriminação e dificuldade, bem como a informação e o erro dos itens, e avaliada a função diferencial do item (DIF) para gênero. De modo geral, foram encontrados valores satisfatórios para o parâmetro de discriminação, e a maioria dos itens obteve dificuldade moderada. A curva da informação identificou que os participantes marcaram as categorias extremas, que representam ausência de sintomas depressivos ou maior frequência de sintomas depressivos.

Cerca de 64% dos participantes apresentaram sintomas moderados ou severos para depressão. De modo similar, outros estudos brasileiros encontraram que mais da metade dos participantes possuíam depressão moderada ou severa durante a pandemia (Almeida et al., 2023; Moura et al., 2022; Musse et al., 2022). Em contrapartida, um estudo pré-pandêmico encontrou que apenas 10,9% dos participantes possuíam indicativos de depressão (Barros et al., 2021). Esses dados indicam que a pandemia foi um evento que acentuou os sintomas depressivos no Brasil.

A PHQ-9 é uma escala que apresentou boas propriedades psicométricas e consistência interna na população brasileira, sendo uma ferramenta importante no rastreio de sintomas depressivos na avaliação clínica (Damiano, 2023). No presente estudo, os parâmetros de discriminação apresentaram pontuações mínimas e máximas entre 1,68 e 3,69, nos itens 9 e 2, respectivamente, sendo classificados entre bons e satisfatórios (Bichi & Talib, 2018). Tais resultados indicam que os itens apresentam boa capacidade de distinguir pessoas com diferentes níveis de sintomas depressivos.

A literatura aponta que a discriminação dos itens da PHQ-9 também tem sido adequada em diferentes contextos (Barroso et al., 2019; Cumbe et al., 2020; Damiano et al., 2023; Gao & Liu, 2024). Um estudo sobre a PHQ-9, em uma amostra de adolescentes chineses, encontrou pontuações

próximas a esses achados para a discriminação dos itens, sendo que o item 2 também apresentou o escore mais discriminativo (Gao & Liu, 2024). A capacidade de discriminação do item 2 pode ser determinada pelo seu conteúdo, que representa um dos sintomas centrais do TDM, tendo em vista que a presença do humor deprimido quase todos os dias e em uma parte significativa do dia é critério diagnóstico do transtorno (APA, 2014).

Predominantemente, os itens apresentaram dificuldade moderada (Bichi & Talib, 2018), com valor mínimo de -0,011 e máximo de 0,529 entre os itens de 1 a 8. Apenas o item 9 foi classificado como difícil, pois encontrou valor superior a um ($b = 1,017$). O achado corrobora com outros estudos, que também encontraram a menor probabilidade de endosso no item sobre autolesão e pensamento de morte (Cumbe et al., 2020; Gao & Liu, 2024). Nesse sentido, um estudo internacional identificou a gravidade de sintomas cognitivos, afetivos e somáticos da depressão associada ao risco de suicídio (Cui & Fiske, 2020), o que pode explicar o maior endosso do item 9 apenas em pessoas com valor teta mais elevado. Apesar de ter o maior escore de dificuldade, o último item apresentou a menor pontuação de discriminação, o que indica que ele possui baixa probabilidade de endosso e baixa capacidade para distinguir pessoas com diferentes níveis de teta.

Sobre os gráficos da CCI, notou-se que as categorias 1, 2 e 4 apresentaram boa cobertura entre os tetras, enquanto a categoria 3 (mais da metade dos dias) foi sobreposta pela 4 (quase todos os dias). Um estudo brasileiro sobre a PHQ-9 encontrou menor informação na categoria 3 em todos os itens (Damiano et al., 2023). A sobreposição pode ser decorrente da dificuldade do participante em diferenciar ambas as categorias, devido à semelhança semântica entre elas. Logo, a eficácia da escala pode ser aumentada por meio da modificação da escrita das categorias, de modo a tornar a diferença entre ambas mais compreensível. Sugere-se que as categorias sejam escritas da seguinte forma: "não ocorre" (categoria 1), "ocorre em poucos dias" (categoria

2), "ocorre em vários dias" (categoria 3) e "ocorre quase todos os dias" (categoria 4).

A CCI evidenciou a presença de maior índice de respostas nas categorias extremas, o que pode representar um perfil quase dicotômico quanto aos sintomas depressivos na amostra. No entanto, optou-se por não dicotomizar a escala, pois manter os níveis originais do instrumento possibilita uma melhor compreensão sobre a distribuição social dos sintomas depressivos durante a pandemia. Além disso, o estabelecimento de pontos de corte arbitrários, durante a dicotomização, pode afetar o agrupamento, de modo que indivíduos com pontuações próximas sejam categorizados em grupos distintos (Purgato & Barbui, 2013).

A escala foi satisfatória para os níveis de teta entre -3 e 3, o que representa que a PHQ-9 avaliou adequadamente brasileiros durante a pandemia, independentemente do traço latente. Uma informação satisfatória é produzida quando os itens apresentam variações na dificuldade conforme o nível de teta do indivíduo, a ótima discriminação e o pouco acerto devido a chutes (Pasquali, 2007). Quanto à informação dos itens e do teste, foram encontradas diferenças em relação a uma pesquisa brasileira de Damiano et al. (2023). A curva da informação do presente estudo evidenciou que o item 2, sobre se sentir deprimido, foi mais informativo. De modo dispar, Damiano et al. (2023) encontraram que os itens 3 e 7, sobre dificuldades com o sono e concentração, respectivamente, foram mais informativos. As diferenças possivelmente são decorrentes do período de coleta de dados. Enquanto o outro estudo utilizou uma amostra anterior à pandemia, a amostra do presente estudo estava inserida na crise sanitária, momento em que foram acentuados sintomas negativos de saúde mental na população, como sentimentos de tristeza e depressão (Barros et al., 2020).

A PHQ-9 apresentou viés de gênero na maioria dos itens, com exceção do item 8. Os achados do presente estudo divergem da literatura internacional. Uma pesquisa tailandesa com profissionais da saúde na pandemia encontrou

uma predominância diferente de DIF de sexo na escala (Jiraniramai et al., 2021). Apenas o item 6, sobre se sentir mal consigo, mesmo apresentou maior probabilidade de endosso entre os homens. De modo similar, um estudo multicêntrico com pacientes que sofreram traumatismo cranioencefálico não encontrou viés de gênero na maioria dos itens, sendo que somente o item 1 apresentou DIF (Teymoori et al., 2020). A literatura brasileira sobre o uso da TRI para análise da PHQ-9 ainda é insuficiente, de forma que não foram encontrados estudos de DIF na escala. Logo, é necessário que novas pesquisas explorem a existência de DIF nesse instrumento aplicado na população brasileira.

Os achados desta pesquisa contribuem para o estabelecimento de parâmetros confiáveis para a avaliação de sintomas depressivos, por meio da PHQ-9, na população brasileira durante o contexto pandêmico. De modo geral, identificou-se que a escala apresentou desempenho satisfatório quanto à discriminação e dificuldade dos itens e foi informativa para avaliar indivíduos com diferentes traços latentes. Além disso, a PHQ-9 apresentou funcionamento distinto em pessoas com mesmo nível de teta; logo, sugere-se a produção de dois protocolos de interpretação a partir do gênero.

Este estudo apresenta limitações. Optou-se por não dicotomizar a escala, a fim de possibilitar uma melhor identificação da distribuição social da depressão na população. No entanto, a escala apresentou perfil predominantemente dicotômico de respostas, dadas às respostas nas categorias extremas da PHQ-9. Outra limitação diz respeito à predominância de mulheres (88,02%) e pessoas com Ensino Superior completo (82%) na constituição da amostra. A existência de DIF para gênero pode ter sido impactada pela desproporcionalidade entre homens e mulheres no estudo. Assim, apesar dos bons resultados quanto aos parâmetros de dificuldade e discriminação, as propriedades e qualidades de um instrumento não são universais e precisam ser continuamente avaliadas (Barroso et al., 2019). Os presentes achados não são generalizáveis para

a população brasileira, em qualquer contexto temporal, já que foram encontrados especificamente no cenário da pandemia da covid-19.

Conclusões

A PHQ-9 tem sido utilizada para o rastreio de sintomas depressivos em diferentes contextos e nacionalidades. A escala contribuiu para a identificação do panorama de saúde mental da população durante a pandemia, tendo em vista o cenário de incertezas e medidas restritivas demandadas. O presente estudo contribuiu para a identificação de parâmetros de validade do instrumento, de acordo com a TRI. A escala foi informativa para mensurar pessoas com diferentes níveis de teta e apresentou valores satisfatórios de discriminação e dificuldade. Diferentemente do que foi encontrado em outros estudos, a PHQ-9 mensurou, de modo distinto, homens e mulheres com o mesmo nível de teta, o que pode ser indicativo da necessidade de fazer dois protocolos de interpretação a partir do gênero.

A TRI tem sido usada para atestar a validade e a confiabilidade de instrumentos de rastreio (Gomes et al., 2018). Este é o primeiro estudo que avaliou as propriedades psicométricas da PHQ-9, por meio da TRI, utilizando dados coletados de diferentes fases da pandemia da covid-19. Os achados contribuem como parâmetros para a avaliação de estudos brasileiros que utilizaram a PHQ-9 durante o cenário pandêmico. Sugere-se que estudos posteriores avaliem a função diferencial do item para gênero e outras categorias sociodemográficas, como raça, local de origem, idade e escolaridade na população brasileira.

Referências

- Almeida, M., Costa, D., Carvalho, L., Carvalho, M., Oliveira, C., Neves, J., Leite, A., Nicolau, L., Costa, N., Lopes, G., Fernandes, C., & Haikal, D. (2023). Impactos da COVID-19 na saúde mental entre universitários da saúde. *Psicologia, Saúde & Doenças*, 24(1), 80-92. https://scielo.pt/scielo.php?script=sci_arttext&pid=S1645-00862023000100080
- American Psychiatric Association. (2014). *Manual diagnóstico e estatístico de transtornos mentais: DSM-5*. Artmed.
- Arrieta, J., Aguerrebere, M., Raviola, G., Flores, H., Elliott, P., Espinosa, A., Reyes, A., Ortiz-Panozo, E., Rodriguez-Gutierrez, E. G., Mukherjee, J., Palazuelos, D., & Franke, M. F. (2017). Validity and utility of the Patient Health Questionnaire (PHQ)-2 and PHQ-9 for screening and diagnosis of depression in rural Chiapas, Mexico: A crosssectional study. *Journal of Clinical Psychology*, 73(9), 1076-1090. <https://doi.org/10.1002/jclp.22390>
- Badaró, A. F. B., Fonseca, T. R., & dos Santos, M. F. R. (2021). Transtorno de ansiedade generalizada (TAG) e pandemia por Covid-19: Uma abordagem cognitivo comportamental/Generalized anxiety disorder (GAD) and the Covid-19 pandemic: A cognitive behavioral approach. *Brazilian Journal of Development*, 7(6), 57729-57739. <https://doi.org/10.34117/bjdv7n6-262>
- Barros, M. B. de A., Lima, M. G., Malta, D. C., Szwarcwald, C. L., Azevedo, R. C. S. de, Romero, D., Souza, P. R. B., Júnior, Azevedo, L. O., Machado, Í. E., Damacena, G. N., Gomes, C. S., Werneck, A. O., Silva, D. R. P., Pina, M. F., & Gracie, R. (2020). Relato de tristeza/depressão, nervosismo/ansiedade e problemas de sono na população adulta brasileira durante a pandemia de COVID-19. *Epidemiologia e Serviços de Saúde*, 29, 1-12. <https://doi.org/10.1590/S1679-49742020000400018>
- Barros, M. B. de A., Medina, L. P. B., Lima, M. G., Azevedo, R. C. S., Sousa, N. F. S., & Malta, D. C. (2021). Associação entre comportamentos de saúde e depressão: resultados da Pesquisa Nacional de Saúde de 2019. *Revista Brasileira de Epidemiologia*, 24, 1-15. <https://doi.org/10.1590/1980-549720210010.supl.2>
- Barroso, S. M., Melo, A. P. S., Silva, M. A. da, & Guimarães, M. D. C. (2019). Evaluation of the Brazilian version of Patient Health Questionnaire (PHQ-9) in quilombola population using the Item Response Theory. *Salud mental*, 42(1), 43-50. https://www.scielo.org.mx/scielo.php?pid=S0185-33252019000100043&script=sci_arttext&tlng=en
- Beck, A. T., & Alford, B. A. (2011). *Depressão: Causas e tratamento*. Artmed.
- Bichi, A. A., & Talib, R. (2018). Item response theory: An introduction to latent trait models to test and item development. *International Journal of Evaluation and Research in Education*, 7(2), 142-151. <http://doi.org/10.11591/ijere.v7i2.12900>
- Brito, V. C. de A., Bello-Corassa, R., Stopa, S. R., Sardinha, L. M. V., Dahl, C. M., & Viana, M. C. (2022). Prevalência de depressão autorreferida no Brasil: Pesquisa Nacional de Saúde 2019 e 2013. *Epidemiologia e Serviços de Saúde*, 31, 1-13. <https://doi.org/10.1590/SS2237-9622202200006.especial>
- Cui, R., & Fiske, A. (2022). Relation between depression symptoms and suicide risk in adults and older adults: A brief report. *Journal of Applied Gerontology*, 41(1), 176-180. <https://doi.org/10.1177/0733464820970849>
- Cumbe, V. F., Muanido, A., Manaca, M. N., Fumo, H., Chiruca, P., Hicks, L., Mari, J. J., & Wagenaar, B. H. (2020). Validity and item response theory properties of the Patient Health Questionnaire-9 for primary care depression screening in Mozambique (PHQ-9-MZ). *BMC Psychiatry*, 20, 1-15. <https://doi.org/10.1186/s12888-020-02772-0>

Dai, S., Vo, T. T., Kehinde, O. J., He, H., Xue, Y., Demir, C., & Wang, X. (2021). Performance of polytomous IRT models with rating scale data: An investigation over sample size, instrument length, and missing data. *Frontiers in Education*, 6, Article 721963. <https://doi.org/10.3389/educ.2021.721963>

Damiano, R. F., Hoffmann, M. S., Gosmann, N. P., Pan, P. M., Miguel, E. C., & Salum, G. A. (2023). Translating measurement into practice: Brazilian norms for depressive symptom assessment with the Patient Health Questionnaire (PHQ-9). *Brazilian Journal of Psychiatry*, 45(4), 310-317. <https://doi.org/10.47626/1516-4446-2022-2945>

Gao, X., & Liu, Z. (2024). Analyzing the psychometric properties of the PHQ-9 using item response theory in a Chinese adolescent population. *Annals of General Psychiatry*, 23(1), Article 2. <https://doi.org/10.1186/s12991-024-00492-3>

Gomes, D. E., Santos, J. L. G. dos, Borges, J. W. P., Alves, M. P., Andrade, D. F. de, & Erdmann, A. L. (2018). Teoria da resposta ao item nas pesquisas em saúde pública. *Revista de Enfermagem da UFPE on line*, 12(6), 1800-1812. <https://doi.org/10.5205/1981-8963-v12i6a234740p1800-1812-2018>

Gómez-Gómez, I., Benítez, I., Bellón, J., Moreno-Peral, P., Oliván-Blázquez, B., Clavería, A., Zabaleta-del-Olmo, E., Llobera, J., Serrano-Ripoll, M. J., Tamayo-Morales, O., & Motrico, E. (2023). Utility of PHQ-2, PHQ-8 and PHQ-9 for detecting major depression in primary health care: a validation study in Spain. *Psychological Medicine*, 53(12), 5625-5635. <https://doi.org/10.1017/S0033291722002835>

Jiraniramai, S., Wongpakaran, T., Angkurawaranon, C., Jiraporncharoen, W., & Wongpakaran, N. (2021). Construct validity and differential item functioning of the PHQ-9 among health care workers: Rasch analysis approach. *Neuropsychiatric Disease and Treatment*, 17, 1035-1045. <https://doi.org/10.2147/NDT.S271987>

Kalkbrenner, M. T., Hunt, A. J., Ryan, A. F., & Rahman, S. R. (2024). Factorial invariance of scores on the Patient Health Questionnaire-9 and Generalized Anxiety Disorder-7 with adults in the United States. *Journal of Mental Health Counseling*, 46(1), 74-94. <https://doi.org/10.17744/mehc.46.1.05>

Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (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>

Lingán-Huamán, S. K., Dominguez-Lara, S., Alarcón-Almeyda, M. S., de Moura, G. B., & Paiva, T. T. (2023). Psychometric properties of the Patient Health Questionnaire-9 in Peruvian university students. *Psychological Topics*, 32(3), 451-470. <https://doi.org/10.31820/pt.32.3.3>

Moura, A. A. M. D. de, Bassoli, I. R., Silveira, B. V. de, Diehl, A., Santos, M. A. dos, Santos, R. A. dos, Wagstaff, C., & Pilon, S. C. (2022). Is social isolation during the COVID-19 pandemic a risk factor for depression? *Revista Brasileira de Enfermagem*, 75, e202 10594. <https://doi.org/10.1590/0034-7167-2021-0594>

Musse, F. C. C., Castro, L. de S., Mestre, C. G., Pelloso, S. M., Poyares, L. L., Lozinski Musse, A., & Carvalho, T. (2022). Violência mental: Ansiedade e depressão durante a pandemia de COVID-19 no Brasil. *Saúde e Pesquisa*, 15(1), e9684. <https://doi.org/10.17765/2176-9206.2022v15n1.e9684>

Nunes, D., & Faro, A. (2022). Estrutura fatorial, análise de invariância e distribuição social do Patient Health Questionnaire-9 (PHQ-9). *Revista Iberoamericana de Diagnóstico y Evaluación - e Avaliação Psicológica*, 62(1), 37-49. <https://doi.org/10.21865/RIDEP62.1.04>

Organização Pan-Americana de Saúde. (2017). *Aumenta o número de pessoas com depressão no mundo*. <https://www.paho.org/pt/noticias/23-2-2017-aumenta-numero-pessoas-com-depressao-no-mundo>

Organização Pan-Americana de Saúde. (2022). *Pandemia de COVID-19 desencadeia aumento de 25% na prevalência de ansiedade e depressão em todo o mundo*. <https://www.paho.org/pt/noticias/2-3-2022-pandemia-covid-19-desencadeia>

Pasquali, L. (2007). TRI: Teoria de resposta ao item: Teoria, procedimentos e aplicações. LabPAM.

Pasquali, L., & Primi, R. (2003). Fundamentos da teoria da resposta ao item: TRI. *Avaliação Psicológica: Interamerican Journal of Psychological Assessment*, 2(2), 99-110. <https://dialnet.unirioja.es/servlet/articulo?codigo=5115864>

Purgato, M., & Barbui, C. (2013). Dichotomizing rating scale scores in psychiatry: A bad idea? *Epidemiology and Psychiatric Sciences*, 22(1), 17-19. <https://doi.org/10.1017/S2045796012000613>

Santana, V. V. R., do Nascimento, R. Z., Lima, A. A., & Nunes, I. C. M. (2020). Alterações psicológicas durante o isolamento social na pandemia de covid-19: revisão integrativa. *Revista Família, Ciclos de Vida e Saúde no Contexto Social*, 8(2), 754-762. <https://www.redalyc.org/journal/4979/497965721011/497965721011.pdf>

Shevlin, M., Butter, S., McBride, O., Murphy, J., Gibson-Miller, J., Hartman, T. K., Levita, L., Mason, L., Martinez, A. P., McKay, R., Stocks, T. V. A., Bennett, K. M., Hyland, P., Vallieres, F., Valiente, C., Vazquez, C., Contreras, A., Peinado, V., Trucharte, A., ... Bentall, R. P. (2022). Measurement invariance of the Patient Health Questionnaire (PHQ-9) and Generalized Anxiety Disorder scale (GAD-7) across four European countries during the COVID-19 pandemic. *BMC Psychiatry*, 22, Article 154. <https://doi.org/10.1186/s12888-022-03787-5>

Shin, C., Ko, Y. H., An, H., Yoon, H. K., & Han, C. (2020). Normative data and psychometric properties of the Patient Health Questionnaire-9 in a nationally representative Korean population. *BMC Psychiatry*, 20, Article 194. <https://doi.org/10.1186/s12888-020-02613-0>

Tassara, I. G., Okabayashi, N. Y. T., Casaca, M. C. G., & de Souza Veronez, F. (2021). Prevalência de sintomas psicológicos em tempos de isolamento social. *Brazilian Journal of Health Review*, 4(1), 1295-1309. <https://doi.org/10.34119/bjhrv4n1-112>

Teymoori, A., Real, R., Gorbunova, A., Haghish, E. F., Andelic, N., Wilson, L., Asendorf, T., Menon, D., & von Steinbüchel, N. (2020). Measurement invariance of assessments of depression (PHQ-9) and anxiety (GAD-7) across sex, strata and linguistic backgrounds in a European-wide sample of patients after traumatic brain injury. *Journal of Affective Disorders*, 262, 278-285. <https://doi.org/10.1016/j.jad.2019.10.035>

Urtasun, M., Daray, F. M., Teti, G. L., Coppolillo, F., Herlax, G., Saba, G., Rubinstein, A., Araya, R., & Irazola, V. (2019). Validation and calibration of the patient health questionnaire (PHQ-9) in Argentina. *BMC Psychiatry*, 19, Article 291. <https://doi.org/10.1186/s12888-019-2262-9>

Ye, X., Shu, H. L., Feng, X., Xia, D. M., Wang, Z. Q., Mi, W. Y., Yu, B., Zhang, X., & Li, C. (2020). Reliability and validity of the Chinese version of the Patient Health Questionnaire-9 (C-PHQ-9) in patients with psoriasis: a cross-sectional study. *BMJ Open*, 10(7), e033211. <https://doi.org/10.1136/bmjopen-2019-033211>

Maria Heloísa Santos-Souza

Graduanda em psicologia pela Universidade Federal de Sergipe.

Brenda Fernanda Pereira da Silva Ferraz

Mestre em Psicologia pela Universidade Federal de Sergipe.

Dandara Barbosa Palhano

Doutora em Psicologia pela Universidade Federal de Sergipe.

André Faro

Pós-doutor em Psicologia pela Universidade Federal de Sergipe.

Endereço para correspondência

Maria Heloísa Santos Souza

Universidade Federal de Sergipe
Avenida Marcelo Deda Chagas
Rosa Elze, 49107-230
São Cristóvão/SE, Brasil

Disponibilidade de dados

Entrar em contato com os autores.

Conflito de interesse

Não há conflitos de interesse.

Como citar este artigo

Santos-Souza, M. H., Ferraz, B. F. P. da S., Palhano, D. B., & Faro, A. (2026). Análise do Questionário de Saúde do Paciente durante a pandemia da covid-19 via TRI. *Psico*, 57(1), eID47284. <https://doi.org/10.15448/1980-8623.2026.1.47284>

Os textos deste artigo foram revisados pela Texto Certo Assessoria Linguística e submetidos para validação dos autores antes da publicação.