

Psychometric Properties of the Patient Health Questionnaire-4 in a Sample of Adults from Ecuador

Propiedades psicométricas del Patient Health Questionnaire-4 en una muestra de adultos de Ecuador

Propriedades psicométricas do Patient Health Questionnaire-4 em uma amostra de adultos do Equador

 Evelyn Cuesta-Andaluz¹
 Rodrigo Moreta-Herrera²
 Adriana Moya-Solís²
 Nazury Santillán-García³
 Esteban Moreno-Montero⁴

¹ Universidad Internacional de La Rioja

² Pontificia Universidad Católica del Ecuador

³ Universidad Estatal de Milagro

⁴ Universidad de La Coruña

Received: 02/15/2025

Accepted: 09/22/2025

Correspondence

Rodrigo Moreta-Herrera
rmoreta@pucesa.edu.ec

How to cite:

Cuesta-Andaluz, E., Moreta-Herrera, R., Moya-Solís, A., Santillán-García, N., & Moreno-Montero, E. (2025). Psychometric Properties of the Patient Health Questionnaire-4 in a Sample of Adults from Ecuador. *Ciencias Psicológicas*, 19(2), e-4481. <https://doi.org/10.22235/cp.v19i2.4481>

Data availability: The data set supporting the results of this study is not available.

Funding: This study did not receive any external funding or financial support.

Conflicts of interest: The authors declare that they have no conflicts of interest.



Abstract: Objective: To analyze the psychometric properties of the Patient Health Questionnaire-4 (PHQ-4) in a sample of Ecuadorian adults. Method: Psychometric study in which factorial validity, measurement equivalence differentiated by gender and internal consistency of the instrument are analyzed. Participants: 406 Ecuadorian adults, 68.47 % women, aged between 18 and 59 years ($M = 27.43$; $SD = 8.20$). Results: The oblique two-factor model of the PHQ is confirmed, with the dimensions of depression and anxiety. Furthermore, the measure is equivalent at the scalar level based on gender. Internal consistency is adequate and presents divergent validity with measures of general well-being. Conclusion: The PHQ-4 is a valid and reliable instrument for rapid assessment of anxiety and depression in Ecuadorian adults.

Keywords: anxiety; depression; measurement equivalence; reliability; validity

Resumen: Objetivo: Analizar las propiedades psicométricas del Patient Health Questionnaire-4 (PHQ-4) en una muestra de adultos de Ecuador. Método: Estudio de carácter psicométrico en el que se analizan la validez factorial, la equivalencia de medida diferenciada por el género y la consistencia interna del instrumento. Participantes: 406 adultos ecuatorianos, el 68.47 % corresponden a mujeres, con edades entre 18 y 59 años ($M = 27.43$; $DE = 8.20$). Resultados: Se confirma el modelo oblicuo de dos factores del PHQ, con las dimensiones de depresión y ansiedad. Además, la medida es equivalente a nivel escalar a partir del género. La consistencia interna es adecuada y presenta una validez divergente con medidas de bienestar general. Conclusión: El PHQ-4 es un instrumento válido y confiable para la evaluación rápida de la ansiedad y la depresión en adultos de Ecuador. **Palabras clave:** ansiedad; depresión; equivalencia de medida; fiabilidad; validez

Resumo: Objetivo: Analisar as propriedades psicométricas do Patient Health Questionnaire-4 (PHQ-4) em uma amostra de adultos equatorianos. Método: Estudo psicométrico no qual se analisa a validade fatorial, a equivalência de medida diferenciada por gênero e a consistência interna do instrumento. Participantes: 406 adultos equatorianos, 68,47 % mulheres, com idades entre 18 e 59 anos ($M = 27,43$; $DP = 8,20$). Resultados: Confirma-se o modelo bifatorial oblíquo do PHQ, com as dimensões depressão e ansiedade. Além disso, a medida mostrou-se equivalente ao nível escalar de acordo com o gênero. A consistência interna é adequada e apresenta validade divergente em relação às medidas de bem-estar geral. Conclusão: O PHQ-4 é um instrumento válido e confiável para a avaliação rápida de ansiedade e depressão em adultos equatorianos.

Palavras-chave: ansiedade; depressão; equivalência de medida; confiabilidade; validade

Mental health is defined as a state of psychological and social well-being that enables individuals to manage daily stress, maintain meaningful relationships, and make sound decisions (World Health Organization [WHO], 2022). It is an essential component of overall health and quality of life, the deficit of which can lead to mental disorders (Cova, 2022). Among these, anxiety and depression stand out due to their prevalence and negative impact on individuals' lives (Fajkowska et al., 2018; Moreta-Herrera et al., 2024) making them relevant for study and intervention.

Depression, characterized by symptoms such as persistent sadness, loss of interest or pleasure in activities, constant fatigue, changes in appetite, concentration difficulties, and recurrent thoughts of death or suicide (American Psychiatric Association [APA], 2014; Bahamón et al., 2023; Lascano-Arias et al., 2025), has multifactorial causes including genetic, biological, psychological, and environmental components (Rutakumwa et al., 2023). Globally, more than 280 million people suffer from this condition, making it a leading cause of social disability (Yildirim et al., 2022), with gender differences as women often report more depression than men (Hyde & Mezulis, 2020; McGuinness et al., 2012). In Latin America, this condition affects 22 % of the population (Huang et al., 2023), while in Ecuador, it is estimated that around 10 % exhibit these symptoms (Lapo-Talledo et al., 2025).

Anxiety, conversely, manifests as a state of excessive worry, restlessness, and neurovegetative symptoms, including palpitations, sweating, and muscle tension (APA, 2014). Like depression, its etiology is multifactorial, encompassing genetic predispositions and stressful life experiences (Chen, 2023; Moreta-Herrera, Jadán-Guerrero, et al., 2025). Gender-based differences are also evident, with women experiencing this condition more frequently than men (Asher & Aderka, 2018; Gao et al., 2020). Globally, anxiety disorders affect approximately 4 % of the population as specific disorders (Anxiety and Depression Association of America, 2023), while in Ecuador, symptomatic presence is noted in 25 % of the population, with an increase observed following the COVID-19 pandemic (Hermosa-Bosano et al., 2021; Sánchez-Vélez & Moreta-Herrera, 2022).

Given the high incidence of these symptoms in the general population, which necessitates professional assistance, and the limitations and difficulties encountered in accessing timely public healthcare (Berry et al., 2020), the evaluation and diagnostic process is crucial. Consequently, the role of psychometric research is significant, not only in generating and adapting measures to diverse contexts and situations but also in enabling precise assessments of mental attributes of sanitary priority, as well as being rapid and efficient in terms of resource availability (personal, time, costs, among others). In the case of Ecuador, there is a marked limitation of available measures for psychological assessment, including conditions of anxiety and depression.

Depression and anxiety assessment measures

Various assessment tools have been developed globally to evaluate anxiety and depression, each employing distinct evaluative approaches. Among the most frequently utilized specialized measures are the Beck Anxiety Inventory (BAI; Beck et al., 1988), the State-Trait Anxiety Inventory (STAI; Spielberger et al., 1970), the Beck Depression Inventory (BDI; Beck et al., 1996), and the Zung Self-Rating Depression Scale (Zung, 1986). Furthermore, multi-faceted assessment measures for mental health conditions, such as the General Health Questionnaire (GHQ-28; Goldberg, 1972) which includes a 28-item adapted version in Ecuador (GHQ-28; Moreta-Herrera et al., 2021), the Symptom Checklist-90-Revised (SCL-90 R; Derogatis et al., 1973) and the Depression Anxiety Stress Scales-21 (DAS-21; Lovibond & Lovibond, 1995), are also prominent. These instruments are designed for rapid screening processes and brief segmentation of individuals in primary care, facilitating improved classification of patients requiring specialized care.

Within the group of measures that jointly assess anxiety and depression, the Patient Health Questionnaire (PHQ; Spitzer et al., 1999) initially developed in English as a 15-item self-report, is notable. This tool aims to identify depression and anxiety criteria in primary care, not for specific diagnoses, but rather for subsequent referral. Its versatility has led to widespread use globally, with shorter versions such as the PHQ-9 (Kroenke et al., 2001) and the PHQ-4 (Kroenke et al., 2009) developed to expedite the identification of indicators of disturbance.

The Patient Health Questionnaire-4 (PHQ-4) is an ultra-brief version of the PHQ, comprising four items: two related to anxiety (GAD-2) and two to depression (PHQ-2), facilitating its application across diverse clinical and population settings. Its psychometric properties have been extensively tested and validated in numerous international studies with adults, including research in Austria, Croatia, Georgia,

Germany, Lithuania, Portugal, and Sweden (Kazlauskas et al., 2023), Greece (Christodoulaki et al., 2022), Poland (Larionow & Mudło-Głagolska, 2023), and Philippines (Mendoza et al., 2024) which also verifies measurement equivalence (ME) based on gender, among other factors. Within Spanish-speaking contexts, validation and adaptation studies have been conducted in Latin American populations in the United States (Mills et al., 2015), Paraguay (Caycho-Rodríguez et al., 2024) and Perú (Carranza Esteban et al., 2024). In Ecuador, the analysis of its psychometric properties and validation has been reported, albeit with samples of university students (Guerra et al., 2022).

Reviewed studies confirm the two-factor correlated (oblique) structure of the measure, in addition to ratifying adequate internal consistency within its items according to each factor. Convergent validity with measures of well-being, acceptance, and commitment is maintained, as is divergent validity with measures of stress, anxiety, and depression, and finally, measurement equivalence (ME) based on gender.

The present study

The COVID-19 post-pandemic era has witnessed increased levels of anxiety and depression in both the region (Moreno-Montero et al., 2025) and Ecuador (Moreta-Herrera, Caycho-Rodríguez, et al., 2025; Sánchez-Vélez & Moreta-Herrera, 2022), placing strain on the public health system. This necessitates the availability of rapid screening instruments such as the Patient Health Questionnaire-4 (PHQ-4), duly validated and calibrated for the general population and adapted to the evolving post-pandemic circumstances, to facilitate their application in efficient screening processes. Furthermore, from a psychometric perspective, while a study of the PHQ-4 exists within an Ecuadorian sample that may serve as a reference (Guerra et al., 2022), it is imperative to note that confirmatory factor analyses (CFA) were not conducted utilizing robust estimation methods, despite the ordinal nature of the items, a factor that may distort the ultimate criteria regarding the adjustments and suitability of the measure, potentially compromising the internal validity of the study and leading to measurement bias (Li, 2016) and an elevated risk of Type I error. Therefore, it is essential to verify the psychometric properties, including factorial validity, measurement invariance (MI) based on gender, given the accumulating evidence suggesting gender-based differences in anxiety and depression (Asher & Aderka, 2018; Gao et al., 2020; Hyde & Mezulis, 2020; McGuinness et al., 2012) and validity based on relationships with other associated variables through the application of robust estimators, particularly with measures such as well-being, considering its typically close association with mental health conditions (Kazlauskas et al., 2023; Mendoza et al., 2024) as well as internal consistency to minimize any factors that might compromise the measure's precision.

This study aims to (1) evaluate the dimensionality of the PHQ-4 in an Ecuadorian adult sample, (2) determine its internal consistency, (3) analyze measurement invariance based on gender, and (4) explore validity based on its relationship with psychological well-being, hypothesizing adequate internal validity (H_1), satisfactory internal consistency (H_2), adequate measurement invariance based on gender (H_3), and a significant relationship with psychological well-being (H_4).

Method

Design

The present study employed a quantitative, instrumental approach (Ato et al., 2013) to confirm the internal structure and internal consistency reliability of the PHQ-4 in a sample of Ecuadorian adults. Furthermore, the study examined differential item functioning (DIF) by gender and assessed validity based on relationships with other variables.

Participants

Participants included 406 Ecuadorian individuals (68.47 % female), aged 18 to 59 years ($M = 27.43$; $SD = 8.20$), with 81.42 % residing in urban areas and 18.58 % in rural areas. The sample consisted of 51.97 % university students and 48.03 % general population, with 73.89 % of the total sample having attained a university education. The majority of participants were single (74.87 %). A non-probabilistic snowball sampling method was used, along with convenience sampling, and inclusion criteria comprised voluntary participation, being over 18 years of age, and an educational level appropriate for their age to ensure comprehension of the instruments. Exclusion criteria included

intellectual disability, effects of substances or drugs affecting consciousness, and lack of Spanish language proficiency.

Instruments

Patient Health Questionnaire-4 (PHQ-4; Kroenke et al., 2009) a concise tool designed for assessing symptoms of anxiety and depression in pre-clinical and research settings, comprises four items divided into two subscales: the PHQ-2 and the GAD-2. Responses are recorded on a 4-point Likert scale ranging from never (0) to every day (3). The total score reflects the intensity of general symptoms, while the subscales allow for the identification of anxiety and depression levels separately, with higher scores indicating greater symptom severity. Psychometric properties reported for the original version show reliability values of $\alpha = .86$ and $.89$ for the PHQ-2 and the GAD-2, respectively, while the Ecuadorian sample yielded $\omega = .780$ for the PHQ-2 and $\omega = .838$ for the GAD-2 (Guerra et al., 2022).

World Health Organization Five Well-Being Index (WHO-5; WHO, 2024) in its Spanish-translated version (Simancas-Pallares et al., 2016), includes five items, with higher scores representing higher levels of perceived well-being over the preceding 2 weeks. Items are rated on a Likert scale ranging from none of the time (0) to all the time (5), resulting in a total score ranging from 0 to 25 points. The instrument exhibits a Cronbach's alpha of $\alpha = .852$, indicating adequate reliability (Simancas-Pallares et al., 2016).

Procedure

For the procedures, a survey was designed on Google Forms and disseminated across various higher education institutions in Ecuador to recruit a significant number of participants. Potential participants received detailed information regarding the study's objectives, security measures, anonymity of responses, voluntary participation, and the option to withdraw without repercussions. Those who agreed to participate completed the assessment in approximately 15 minutes. Upon completion, participants were asked to share the evaluation link among their personal contacts to expand the accessibility to the general population. Subsequently, the information was digitally cleaned and systematized for statistical management, results formulation, hypothesis testing, and the preparation of research reports. Ethical considerations adhered to the ethical criteria and standards for research involving human subjects as outlined in the Declaration of Helsinki. The project from which this study is derived was approved by the institutional review board of the Pontificia Universidad Católica del Ecuador Sede Ambato (Pucesa).

Data analysis

The statistical management of the results was divided into five blocks. Initially, a preliminary analysis of the PHQ-4 items was conducted, employing measures such as the arithmetic mean (M), standard deviation (SD), skewness (g_1), and kurtosis (g_2). Univariate normality was assessed, adhering to the criterion that g_1 and g_2 values should fall within ~ 1.5 (Ferrando & Anguiano-Carrasco, 2010). Subsequently, multivariate normality was analyzed using Mardia's test (Mardia, 1970), where non-significance ($p > .05$) in g_1 and g_2 was considered evidence of multivariate normality. The second block involved Confirmatory Factor Analysis (CFA) based on a matrix of polychoric correlations, employing Diagonally Weighted Least Squares (DWLS) estimation due to the absence of multivariate normality and the categorical nature of the items (Li et al., 2022; Moreta-Herrera et al., 2025). This phase analyzed three models: (1) unidimensional, (2) oblique two-factor, and (3) hierarchical with two first-order factors and one second-order factor. Assessment of the oblique model, as proposed by Kroenke et al. (2009) considered seven indices: Chi-square (χ^2), normed Chi-square (χ^2/df), Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), Standardized Root Mean Square Residual (SRMR), Root Mean Square Error of Approximation (RMSEA), and item factor loadings (λ). Adequate fit was defined as $\chi^2 p > .05$ or $\chi^2/df < 4$; CFI and TLI $> .95$; SRMR and RMSEA $< .06$ (with $.08$ as an acceptable tolerance) (Browne & Cudeck, 1992; Byrne, 2008; Wolf et al., 2013; Yang-Wallentin et al., 2010). Item factor loadings (λ) were deemed adequate if exceeding $\lambda > .40$ (Dominguez-Lara, 2018).

In the third section, measurement equivalence (ME) of the PHQ-4 was analyzed based on gender using a multigroup confirmatory factor analysis (MG-CFA) with DWLS estimation to rule out differences between groups in the internal structure (Asparouhov & Muthén, 2014; Moreta-Herrera et al., 2022). During this process, restrictions were imposed on the factor loadings (metric), intercepts (scalar or strong), and residuals (strict) to determine whether changes in fit indices (χ^2 , CFI, and RMSEA) were

significant. ME was assumed if the change (Δ) in χ^2 was not significant ($p > .05$), although this may not be met due to the sample size (Luong & Flake, 2023); furthermore, it was required that $\Delta CFI \leq .015$ and $\Delta RMSEA \leq .01$ for each stage of restriction (Rutkowski & Svetina, 2017).

The third section also evaluated internal consistency using Cronbach's alpha (α). Although McDonald's omega coefficient (ω) has been recognized as a more robust metric for assessing scale reliability, especially in models with heterogeneous factor loadings, its calculation requires well-identified models (Şimşek & Noyan, 2013). In this case, due to the oblique structure of the PHQ-4 and the problems associated with estimating ω in individual dimensions, α was chosen as the primary measure of reliability.

Accepted reference values for interpreting internal consistency indicate that an $\alpha \geq .70$ is considered acceptable. It should be noted that, although with caution, reliability analysis with two items per factor is fully viable when the correlations between items are high or extremely high, being recommended from .60 (Lapo-Talledo et al., 2025). As the internal consistency is broken down by groups, the differences in reliability between groups are also presented, for which both the values of the coefficient ($\Delta\alpha$) and the confidence intervals (CI) are subtracted. Differences in internal consistency between groups are estimated to exist when the values of the CIs do not include 0.

Finally, the fifth section addressed validity based on the relationship with other variables through Structural Equation Modeling (SEM), identifying latent covariances between the PHQ-4 and a measure of well-being (WHO-5). Relationships greater than $\rho > .20$ are expected, with fit indices like those of the CFA in the general fit model.

It should be noted that the statistical analyses were performed using the R programming language, version 4.3.1 (R Core Team, 2024) and the packages used were lavaan for the CFA, semTools for the EM, MNV for multivariate normality and MBESS for the reliability indices.

Results

Preliminary analysis of items

Table 1 presents the preliminary analysis of the items comprising the PHQ-4. The means for the items ranged from $M(\text{Item3}) = 0.66$, $SD = 0.80$ to $M(\text{Item1}) = 0.77$, $SD = 0.79$. Regarding distribution indices, skewness values ranged from 0.93 to 1.06 and kurtosis values ranged from 0.47 to 0.68, suggesting a slight deviation from normality in the item distributions. Furthermore, results from the multivariate normality analysis indicated significant value for both skewness and kurtosis, confirming the violation of multivariate normality assumptions. Finally, the correlation matrix demonstrated positive associations between the items, with coefficients ranging from .528 to .632, supporting the internal consistency of the instrument. Given these findings, the application of robust estimators for subsequent factor analyses is deemed essential to ensure valid and reliable results.

Table 1
Preliminary analysis of PHQ-4 items

Ítems	Descriptive statistics				Correlation matrix			
	<i>M</i>	<i>DE</i>	<i>g</i> ₁	<i>g</i> ₂	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>
PHQ-2 (Depression)								
Ítem 1	0.77	0.79	0.93	0.58	1			
Ítem 2	0.70	0.79	1.01	0.63	.618	1		
GAD-2 (Anxiety)								
Ítem 3	0.66	0.80	1.06	0.47	.528	.551	1	
Ítem 4	0.70	0.80	1.06	0.68	.574	.623	.632	1
Mardia			374.91***	19.65***				

Note. M: mean; SD: standard deviation; g1: asymmetry; g2: kurtosis.
*** $p < .001$

Confirmatory factor analysis

Table 2 presents the results of the confirmatory factor analyses (CFA) of the PHQ-4, evaluating various models of fitness. Both the oblique and hierarchical models demonstrated the best fit, with near-maximum values. In accordance with previous literature, the oblique model is frequently utilized in

psychometric studies of the PHQ-4, as it assumes a structure of two correlated factors reflecting the dimensions of depression and anxiety. This model is congruent with the results obtained, highlighting its superiority in terms of parsimony and theoretical fit. Therefore, it is concluded that the oblique model is the most appropriate for describing the factorial structure of the PHQ-4 in the sample studied.

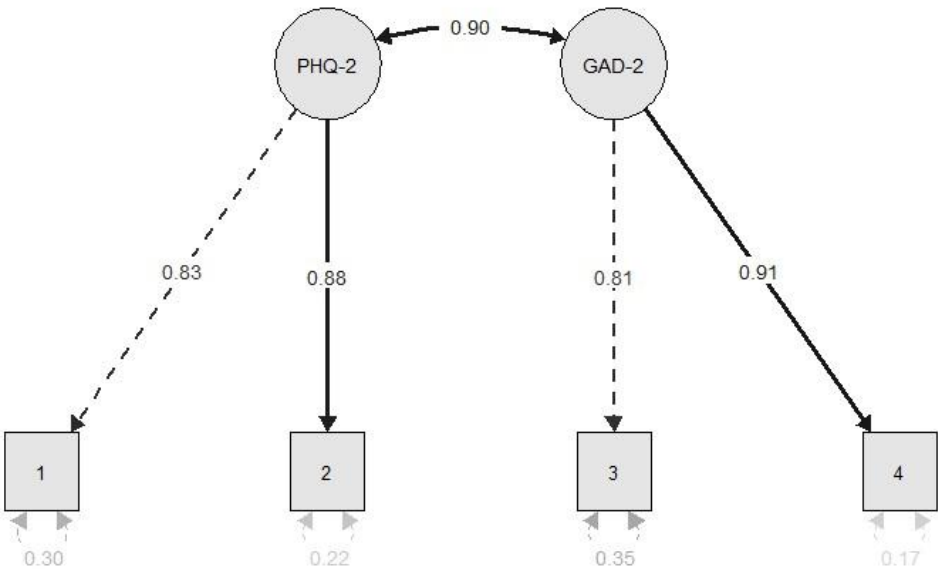
Table 2
Confirmatory factor analysis of PHQ-4

Model	χ^2	gl	χ^2/gl	CFI	TLI	SRMR	RMSEA
A factor	9.255	2	4.620	.998	.993	.029	.135 [.121 – .238]
Oblique	0.203	1	0.203	.999	.999	.005	.000 [.000 – .126]
Hierarchical++	0.203	1	0.203	.999	.999	.005	.000 [.000 – .126]

Note. χ^2 : chi-square; df: degrees of freedom; χ^2/df : normative chi-square; CFI: Comparative Adjustment Index; TLI: Tucker-Lewis Index; SRMR: Standardized Mean Square Residual; RMSEA: Mean Square Error of Approximation.
++ Two first-order factors and one second-order factor

As depicted in Figure 1, the oblique model emerges as the most suitable configuration for the PHQ-4, demonstrating results that validate its appropriate fit within the adult sample in Ecuador. From a theoretical standpoint, this approach facilitates the correlation of variables such as depression and anxiety, acknowledging their interrelationships without resorting to the simplification inherent in integrating them into a single component. Furthermore, the magnitude of the λ coefficients, ranging from $\lambda_{item1} = .83$ to $\lambda_{item4} = .91$, is deemed appropriate in relation to the factor, indicating that these items are pertinent within the proposed model and contribute significantly to explaining the variance of the measure.

Figure 1
Factor structure of PHQ-4



Gender-based measurement equivalence

Table 3 presents the results of the Measurement Equivalence (ME) based on gender for the PHQ-4. The fit indices for the configural model were adequate, indicating that the factorial structure of the instrument is consistent across groups. With the addition of metric constraints, the fit remained high with minimal change in the comparative indices, suggesting that the factor loadings are equivalent. Upon imposing scalar invariance, the results continued to demonstrate excellent fit, and changes in the indices were irrelevant, indicating that the item intercepts are also equivalent. However, when evaluating strict invariance, although the global fit indices remained acceptable, a notable deterioration in fit was

observed compared to the scalar model, suggesting that the variances of the measurement errors are not equivalent between groups. Thus, it is estimated that the PHQ-4 is invariant only at the scalar level.

Table 3
Gender-based measurement equivalence of PHQ-4

<i>Restrictions</i>	<i>(df) χ^2</i>	<i>CFI</i>	<i>RMSEA</i>	<i>(df) $\Delta\chi^2$</i>	<i>ΔCFI</i>	<i>ΔRMSEA</i>
Men's base	(1) 0.501	.995	.077	-	-	-
Women's base	(1) 0.005	.999	.010	-	-	-
Configural	(2) 2.964	.998	.014	-	-	-
Metric	(4) 3.584	.999	.000	(2) 0.619	.001	.014
Scalar	(6) 3.833	.999	.000	(2) 0.249	.000	.000
Strict	(10) 24.95	.977	.085	(4) 21.12	.022	.085

Note. χ^2 : chi-squared; (DF): degrees of freedom; CFI: Comparative Fit Index; RMSEA: Root Mean Square Error of Approximation; Δ : change.

Internal consistency analysis

Table 4 presents the α values for evaluating the internal consistency of the PHQ-4 in the total sample and by gender. Prior verification confirmed that correlations between items within each dimension exceed 0.60, rendering this type of analysis viable. The total scale exhibited adequate reliability for both the depression and anxiety factors. When classified by gender, the male group demonstrated higher levels of internal consistency compared to the female group; however, reported levels were acceptable for both groups. Significant differences were observed in the depression factor reliability coefficients, as the confidence intervals (CI) did not include 0, suggesting greater internal consistency among males than females. Conversely, the anxiety factor showed similar internal consistency between both groups.

Table 4
Internal consistency of PHQ-4

Factors	Total		Men		Women		$\Delta\alpha$
	α	IC95 %	α	IC95 %	α	IC95 %	
Depression	.77	[.73 - .81]	.89	[.84 - .92]	.70	[.62 - .77]	.19 [.10 - .28]
Anxiety	.77	[.72 - .81]	.78	[.69 - .85]	.77	[.70 - .81]	.01 [-.09 - .11]

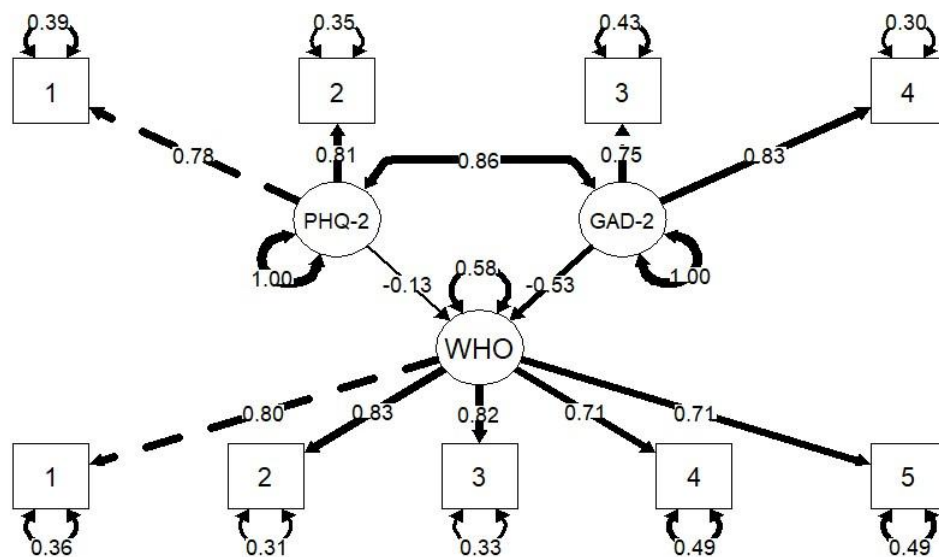
Note. α : Cronbach's coefficient; Δ : change/difference.

Validity analysis based on the relationship with other variables

Figure 2 provides a representation of the validity analysis based on the interaction of diverse variables, the purpose of which is to clarify the covariance that manifests between mental health, specifically anxiety and depression, and general well-being, considering these elements as interrelated constructs with relevant theoretical proximity. The WHO-5 presents an exceptionally low negative relationship with the PHQ-2 and a moderate negative relationship with the GAD-2, suggesting that an increase in general well-being correlates with a slight reduction in depression levels and a significant decrease in anxiety levels. These findings not only provide empirical evidence on the connection between these constructs but also underscore, through the adjustment indices, the existence of a robust and stable relationship that is fundamental for the evaluation of these phenomena in the Ecuadorian adult population.

Figure 2

Model of general fit of the latent relationship between PHQ-4 and WHO-5



$$\chi^2 = 9.429; p < .001; gl = 24; CFI = .999; TLI = .999; SRMR = .024; RMSEA = .000 [0.000 - .000]$$

Note. χ^2 : Chi-square; df: Degrees of freedom; χ^2/df : Normative chi-square; CFI: Comparative Adjustment Index; TLI: Tucker-Lewis Index; SRMR: Standardized mean square residual; RMSEA: Mean square error of approximation. The circles represent the latent variables, and the rectangles represent the observable ones. The values of the arrows pointing to the latent variables are the factor loads (λ).

Discussion

The objectives of the present study were to examine the factorial structure, measurement equivalence (ME) across gender, validity based on relationships with other variables, and reliability of the PHQ-4 in a sample of Ecuadorian adults.

Regarding factorial validity, findings indicated that the most suitable model for the Ecuadorian sample was the oblique two-factor (correlated) model, which distinctly differentiates the assessment of both anxiety and depression. This model alignment is consistent with results from similar studies conducted in other countries (Carranza Esteban et al., 2024; Caycho-Rodríguez et al., 2024; Christodoulaki et al., 2022; Kazlauskas et al., 2023; Larionow & Mudło-Głagolska, 2023; Mendoza et al., 2024; Mills et al., 2015) and within Ecuador (Guerra et al., 2022). In this context, the utilization of robust estimators such as DWLS for factor analyses enhances the precision and objectivity of inferences regarding the internal validity of the measure and reduces the likelihood of measurement bias (Li, 2016).

Concerning ME based on gender, the study concluded that the PHQ-4 exhibits invariance across gender only in the factor loadings and intercepts, signifying scalar (strong) invariance. Consequently, observed differences between groups are attributable to the inherent characteristics of those groups rather than variations in the configuration of the measure's internal structure (loadings and intercepts) (Asparouhov & Muthén, 2014). These results facilitate comparative studies by gender within the adult population of Ecuador. Notably, these findings agree with previous research that reported similar conclusions (Mendoza et al., 2024) identifying that gender is not a source of variability in the measure.

Regarding internal consistency, the PHQ-4 has demonstrated acceptable and adequate reported consistency for assessment in Ecuadorian adults, both overall and within subgroups. These findings align with previous studies conducted globally (Christodoulaki et al., 2022; Kazlauskas et al., 2023; Larionow & Mudło-Głagolska, 2023) and within Ecuador (Guerra et al., 2022). However, a notable observation is the difference in internal consistency ($\Delta\alpha$) between genders within our sample, particularly in the depression factor, where men exhibit greater internal consistency than women, with significant differences (the CIs did not include 0). While this does not fundamentally alter the interpretation of the results, it indicates that men are more consistent in the reliability of the measure

compared to women. Notably, no prior reference studies highlight this characteristic, making it a potentially relevant advancement.

Finally, concerning validity based on the relationship with other variables, the PHQ-4 demonstrates an extremely low negative convergence (PHQ-2) and a moderate negative convergence (GAD-2) with the WHO-5, which assesses general well-being. Furthermore, the model fit from which the latent relationship is derived indicates adequate fit for comprehension within the adult population of Ecuador. These findings are consistent with prior work concluding divergent validity with the construct of well-being through the WHO-5 (Kazlauskas et al., 2023; Mendoza et al., 2024) suggesting the proximity of these attributes.

Regarding the implications of the study, it is noted that the PHQ-4 is a valid and precise measure for evaluating symptoms of anxiety and depression in Ecuadorian adults. This will allow its practical application in primary care assessment processes, aiding in the segmentation of individuals according to their care needs. Furthermore, the confirmed measurement equivalence (ME) facilitates comparative studies by gender, assisting in the detection of inherent differences between men and women. From a theoretical perspective, these results contribute to the development of psychometric research on the measure, expanding its generalizability for evaluation processes globally, particularly in the South American context where validation studies are scarce.

It is necessary to acknowledge a limitation related to the age characteristics of the participants, as the study focused solely on Ecuadorian adults, excluding other segments such as adolescents or older adults, thereby reducing its validation capacity for these groups, and thus these results cannot be generalized. Future studies validating the measure in other age groups of interest are recommended. In the analyses of validity evidence based on the relationship with other variables, the WHO-5, which evaluates general well-being, was used. While its use is pertinent for this process, future confirmatory studies employing measures more closely aligned with the PHQ-4, such as specific anxiety and depression tests, may be advisable.

References

- American Psychiatric Association. (2014). *Manual diagnóstico y estadístico de los trastornos mentales. DSM-5* (5ª ed.). Editorial Médica Panamericana.
- Anxiety and Depression Association of America (2023). *Understand the Fact*. <https://adaa.org/understanding-anxiety>
- Asher, M., & Aderka, I. M. (2018). Gender Differences in Social Anxiety Disorder. *Journal of Clinical Psychology*, 74(10). <https://doi.org/10.1002/jclp.22624>
- Asparouhov, T., & Muthén, B. (2014). Multiple-group Factor Analysis Alignment. *Structural Equation Modeling*, 21(4), 495-508. <https://doi.org/10.1080/10705511.2014.919210>
- Ato, M., López, J. J., & Benavente, A. (2013). Un sistema de clasificación de los diseños de investigación en psicología. *Anales de Psicología*, 29(3), 1038-1059. <https://doi.org/10.6018/analesps.29.3.178511>
- Beck, A. T., Epstein, N., Brown, G., & Steer, R. (1988). An Inventory for Measuring Clinical Anxiety: Psychometric Properties. *Journal of Consulting and Clinical Psychology*, 56(6), 893-897. <https://doi.org/10.1037/0022-006X.56.6.893>
- Beck, A., Steer, R., & Brown, G. (1996). *Manual for the Beck Depression Inventory—II*. Psychological Corporation.
- Berry, K., Sheardown, J., Pabbineedi, U., Haddock, G., Cross, C., & Brown, L. J. E. (2020). Barriers and Facilitators to Accessing Psychological Therapies for Severe Mental Health Difficulties in Later Life. *Behavioural and Cognitive Psychotherapy*, 48(2). <https://doi.org/10.1017/S1352465819000596>
- Browne, M. W., & Cudeck, R. (1992). Alternative Ways of Assessing Model Fit. *Sociological Methods & Research*, 21(2). <https://doi.org/10.1177/0049124192021002005>
- Byrne, B. M. (2008). Testing for Multigroup Equivalence of a Measuring Instrument: A Walk Through the Process. *Psicothema*, 20(4). <https://www.psicothema.com/pdf/3569.pdf>
- Carranza Esteban, R. F., Mamani-Benito, O., Cjuno, J., Tito-Betancur, M., Caycho-Rodríguez, T., Vilca, L. W., Torales, J., & Barrios, I. (2024). Adaptation and Validation of the Patient Health Questionnaire for Depression and Anxiety (PHQ-4) in a Sample of Quechua-speaking Peruvians. *Medicina Clínica y Social*, 8(1). <https://doi.org/10.52379/mcs.v8i1.361>

- Caycho-Rodríguez, T., Travezaño-Cabrera, A., Torales, J., Barrios, I., Vilca, L. W., Samaniego-Pinho, A., Moreta-Herrera, R., Reyes-Bossio, M., Barria-Asenjo, N. A., Ayala-Colqui, J., & Garcia-Cadena, C. H. (2024). Psychometric Network Analysis of the Patient Health Questionnaire-4 (PHQ-4) in Paraguayan General Population. *Psicologia: Reflexão e Crítica*, 37(1), 1-13. <https://doi.org/10.1186/S41155-024-00299-X/METRICS>
- Chen, Q. (2023). Causes and Treatment of Anxiety Disorder. *Lecture Notes in Education Psychology and Public Media*, 9(1). <https://doi.org/10.54254/2753-7048/9/20230230>
- Christodoulaki, A., Baralou, V., Konstantakopoulos, G., & Touloumi, G. (2022). Validation of the Patient Health Questionnaire-4 (PHQ-4) to Screen for Depression and Anxiety in the Greek General Population. *Journal of Psychosomatic Research*, 160. <https://doi.org/10.1016/j.jpsychores.2022.110970>
- Cova, F. (2022). ¿Problemas de salud mental o trastornos mentales? Una distinción ilusoria. *Praxis Psy*, 22(36). <https://doi.org/10.32995/praxispsy.v22i36.173>
- Derogatis, L. R., Lipman, R. S., & Covi, L. (1973). SCL-90: An Outpatient Psychiatric Rating Scale--Preliminary Report. *Psychopharmacology Bulletin*, 9(1).
- Dominguez-Lara, S. (2018). Propuesta de puntos de corte para cargas factoriales: una perspectiva de fiabilidad de constructo. *Enfermería Clínica*, 28(6). <https://doi.org/10.1016/j.enfcli.2018.06.002>
- Fajkowska, M., Domaradzka, E., & Wytykowska, A. (2018). Types of Anxiety and Depression: Theoretical Assumptions and Development of the Anxiety and Depression Questionnaire. *Frontiers in Psychology*, 8. <https://doi.org/10.3389/fpsyg.2017.02376>
- Ferrando, P. J., & Anguiano-Carrasco, C. (2010). El análisis factorial como técnica de investigación en psicología. *Papeles del Psicólogo*, 31(1). <https://www.papelesdelpsicologo.es/resumen?pii=1793>
- 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. <https://doi.org/10.1016/j.jad.2019.11.121>
- Goldberg, D. P. (1972). *The Detection of Psychiatric Illness by Questionnaire*. Oxford University Press.
- Guerra, V. L., Mejia, Á. J. A., & Alcedo, J. M. G. (2022). Psychometric Properties and Factorial Structure of the Patient Health Questionnaire (PHQ-4) in Ecuadorian University Students. *Revista Cubana de Enfermería*, 38(3).
- Hermosa-Bosano, C., Paz, C., Hidalgo-Andrade, P., García-Manglano, J., Sádaba-Chalezquer, C., López-Madrigal, C., & Serrano, C. (2021). Síntomas de depresión, ansiedad y estrés en la población general ecuatoriana durante la pandemia por covid-19. *Revista Ecuatoriana de Neurología*, 30(2). <https://doi.org/10.46997/revecuatneurol30200040>
- Huang, Y., Li, Y., Pan, H., & Han, L. (2023). Global, Regional, and National Burden of Neurological Disorders in 204 Countries and Territories Worldwide. *Journal of Global Health*, 13. <https://doi.org/10.7189/JOGH.13.04160>
- Hyde, J. S., & Mezulis, A. H. (2020). Gender Differences in Depression: Biological, Affective, Cognitive, and Sociocultural Factors. *Harvard Review of Psychiatry*, 28(1), 4-13. <https://doi.org/10.1097/hrp.0000000000000230>
- Kazlauskas, E., Gelezelyte, O., Kvedaraite, M., Ajdukovic, D., Johannesson, K. B., Böttche, M., Bondjers, K., Dragan, M., Figueiredo-Braga, M., Grajewski, P., Anastassiou-Hadjicharalambous, X., Javakhishvili, J. D., Lioupi, C., Lueger-Schuster, B., Mouthaan, J., Bagaric, I. R., Sales, L., Schäfer, I., Soydas, S., ... Lotzin, A. (2023). Psychometric Properties of the Patient Health Questionnaire-4 (PHQ-4) in 9,230 Adults across Seven European Countries: Findings from the ESTSS ADJUST Study. *Journal of Affective Disorders*, 335, 18-23. <https://doi.org/10.1016/j.jad.2023.05.007>
- 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). <https://doi.org/10.1046/j.1525-1497.2001.016009606.x>
- Kroenke, K., Spitzer, R. L., Williams, J. B. W., & Löwe, B. (2009). An Ultra-brief Screening Scale for Anxiety and Depression: The PHQ-4. *Psychosomatics*, 50(6). [https://doi.org/10.1016/s0033-3182\(09\)70864-3](https://doi.org/10.1016/s0033-3182(09)70864-3)

- Lapo-Talledo, G. J., Montes-Escobar, K., Rodrigues, A. L. S., & Siteneski, A. (2025). Hospitalizations for Depressive Disorders in Ecuador: An Eight-year Nationwide Analysis of Trends and Demographic Insights. *Journal of Affective Disorders*, 374, 433-442. <https://doi.org/10.1016/J.JAD.2025.01.074>
- Larionow, P., & Mudło-Głagolska, K. (2023). The Patient Health Questionnaire-4: Factor Structure, Measurement Invariance, Latent Profile Analysis of Anxiety and Depressive Symptoms and Screening Results in Polish Adults. *Advances in Cognitive Psychology*, 19(2). <https://doi.org/10.5709/acp-0391-2>
- Li, C. H. (2016). Confirmatory Factor Analysis with Ordinal Data: Comparing Robust Maximum Likelihood and Diagonally Weighted Least Squares. *Behavior Research Methods*, 48(3). <https://doi.org/10.3758/s13428-015-0619-7>
- Li, L., Niu, Z., Mei, S., & Griffiths, M. D. (2022). A Network Analysis Approach to the Relationship Between Fear of Missing Out (Fomo), Smartphone Addiction, and Social Networking Site Use among a Sample of Chinese University Students. *Computers in Human Behavior*, 128. <https://doi.org/10.1016/j.chb.2021.107086>
- Lovibond, S. H., & Lovibond, P. F. (1995). *Manual for the Depression, Anxiety and Stress Scale-21 Items (DASS-21)*. Psychology Foundation of Australia.
- Luong, R., & Flake, J. K. (2023). Measurement Invariance Testing using Confirmatory Factor Analysis and Alignment Optimization: A Tutorial for Transparent Analysis Planning and Reporting. *Psychological Methods*, 28(4), 905-924. <https://doi.org/10.1037/met0000441>
- Mardia, K. V. (1970). Measures of Multivariate Skewness and Kurtosis with Applications. *Biometrika*, 57(3), 519. <https://doi.org/10.1093/biomet/57.3.519>
- McGuinness, T. M., Dyer, J. G., & Wade, E. H. (2012). Gender Differences in Adolescent Depression. *Journal of Psychosocial Nursing and Mental Health Services*, 50(12). <https://doi.org/10.3928/02793695-20121107-04>
- Mendoza, N. B., Frondoza, C. E., Dizon, J. I. W. T., & Buenconsejo, J. U. (2024). The Factor Structure and Measurement Invariance of the PHQ-4 and the Prevalence of Depression and Anxiety in a Southeast Asian Context amid the covid-19 Pandemic. *Current Psychology*, 43(25). <https://doi.org/10.1007/s12144-022-02833-5>
- Mills, S. D., Fox, R. S., Pan, T. M., Malcarne, V. L., Roesch, S. C., & Sadler, G. R. (2015). Psychometric Evaluation of the Patient Health Questionnaire-4 in Hispanic Americans. *Hispanic Journal of Behavioral Sciences*, 37(4). <https://doi.org/10.1177/0739986315608126>
- Moreno-Montero, E., Moreta-Herrera, R., Rodas, J. A., Oriol-Granado, X., Puerta-Cortés, D. X., Ferrufino-Borja, D., Diaz, R. G., Rivera, M. E. L., Samaniego-Pinho, A., Buenahora-Bernal, M., Rojas-Jara, C., & Vega-Arce, M. (2025). Cross-cultural Measurement Equivalence of the Seven-item General Anxiety Disorder Scale (GAD-7) in College Students of Six Countries of Latin American. *Journal of Affective Disorders*, 380, 598-606. <https://doi.org/10.1016/J.JAD.2025.03.188>
- Moreta-Herrera, R., Caycho-Rodríguez, T., Salinas, A., Jiménez-Borja, M., Gavilanes-Gómez, D., & Jiménez-Mosquera, C. J. (2025). Factorial Validity, Reliability, Measurement Invariance and the Graded Response Model for the covid-19 Anxiety Scale in a Sample of Ecuadorians. *Omega*, 90(3), 1078-1093. <https://doi.org/10.1177/00302228221116515>
- Moreta-Herrera, R., Dominguez-Lara, S., Vaca-Quintana, D., Zambrano-Estrella, J., Gavilanes-Gómez, D., Ruperti-Lucero, E., & Bonilla, D. (2021). Psychometric Properties of the General Health Questionnaire (GHQ-28) in Ecuadorian College Students. *Psihologijske Teme*, 30(3). <https://doi.org/10.31820/pt.30.3.9>
- Moreta-Herrera, R., López-Calle, C., Caycho-Rodríguez, T., Cabezas Guerra, C., Gallegos, M., Cervigni, M., Martino, P., Barés, I., & Calandra, M. (2022). Is it Possible to Find a Bifactor Structure in the Fear of covid-19 Scale (FCV-19S)? Psychometric Evidence in an Ecuadorian Sample. *Death Studies*, 46(9). <https://doi.org/10.1080/07481187.2021.1914240>
- Moreta-Herrera, R., Rodríguez-Lorenzana, A., Mascialino, G., Castro-Ochoa, F., Narváez-Pillco, V., Caycho-Rodríguez, T., & Mayorga-Lascano, M. (2024). Psychometric Properties of the 28-item General Health Scale (GHQ-28). Analysis from the Classical Test and Item Response Theories in Ecuadorian University Students. *Psychology Hub*, 41(3), 43-54. <https://doi.org/10.13133/2724-2943/18460>

- R Core Team. (2024). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*. <https://www.r-project.org/>
- Rutakumwa, R., Tusiime, C., Mpango, R. S., Kyohangirwe, L., Kaleebu, P., Patel, V., & Kinyanda, E. (2023). A Qualitative Exploration of Causes of Depression among Persons Living with HIV Receiving Antiretroviral Therapy in Uganda: Implications for Policy. *Psychiatry Journal*. <https://doi.org/10.1155/2023/1986908>
- Rutkowski, L., & Svetina, D. (2017). Measurement Invariance in International Surveys: Categorical Indicators and Fit Measure Performance. *Applied Measurement in Education*, 30(1). <https://doi.org/10.1080/08957347.2016.1243540>
- Sánchez-Vélez, H., & Moreta-Herrera, R. (2022). Fear and Anxiety of covid-19, Stress and Health Perception. A Predictive Model in Ecuadorian Hospital Patients. *Anales de Psicología*, 38(3), 439-447. <https://doi.org/10.6018/analesps.489761>
- Simancas-Pallares, M., Díaz-Cárdenas, S., Barbosa-Gómez, P., Buendía-Vergara, M., & Arévalo-Tovar, L. (2016). Propiedades psicométricas del Índice de Bienestar General-5 de la Organización Mundial de la Salud en pacientes parcialmente edéntulos. *Revista de la Facultad de Medicina*, 64(4). <https://doi.org/10.15446/revfacmed.v64n4.52235>
- Şimşek, G. G., & Noyan, F. (2013). McDonald's ω , Cronbach's α , and Generalized θ for Composite Reliability of Common Factors Structures. *Communications in Statistics - Simulation and Computation*, 42(9). <https://doi.org/10.1080/03610918.2012.689062>
- Spielberger, C. D., Gorsuch, R. L., & Lushene, R. E. (1970). *Manual for the State-Trait Anxiety Inventory (Self-Evaluation Questionnaire)*. Consulting Psychologists Press.
- Spitzer, R. L., Kroenke, K., & Williams, J. B. W. (1999). Validation and Utility of a Self-report Version of PRIME-MD: The PHQ Primary Care Study. *Journal of the American Medical Association*, 282(18). <https://doi.org/10.1001/jama.282.18.1737>
- Wolf, E. J., Harrington, K. M., Clark, S. L., & Miller, M. W. (2013). Sample Size Requirements for Structural Equation Models: An Evaluation of Power, Bias, and Solution Propriety. *Educational and Psychological Measurement*, 73(6). <https://doi.org/10.1177/0013164413495237>
- World Health Organization. (2022). *World Mental Health Report: Transforming Mental Health for all. Executive Summary*. <https://www.who.int/publications/i/item/9789240049338>
- World Health Organization (2024). *The World Health Organization-Five Well-Being Index (WHO-5)*. <https://www.who.int/publications/m/item/WHO-UCN-MSD-MHE-2024.01>
- Yang-Wallentin, F., Jöreskog, K. G., & Luo, H. (2010). Confirmatory Factor Analysis of Ordinal Variables with Misspecified Models. *Structural Equation Modeling*, 17(3). <https://doi.org/10.1080/10705511.2010.489003>
- Yildirim, M., Gaynes, B. N., Keskinocak, P., Pence, B. W., & Swann, J. (2022). DIP: Natural history model for major depression with incidence and prevalence. *Journal of Affective Disorders*, 296. <https://doi.org/10.1016/j.jad.2021.09.079>
- Zung, W. W. K. (1986). Zung Self-Rating Depression Scale and Depression Status Inventory. *Assessment of Depression*. https://doi.org/10.1007/978-3-642-70486-4_21

Authors' contribution (CRediT Taxonomy): 1. Conceptualization; 2. Data curation; 3. Formal Analysis; 4. Funding acquisition; 5. Investigation; 6. Methodology; 7. Project administration; 8. Resources; 9. Software; 10. Supervision; 11. Validation; 12. Visualization; 13. Writing: original draft; 14. Writing: review & editing.

E. C. A. has contributed in 1, 2, 3, 6, 9, 11, 12, 13, 14; R. M. H. in 1, 2, 3, 5, 6, 9, 10, 12, 13, 14; A. M. S. in 3, 5, 6, 8, 11, 13, 14; N. S. G. in 3, 5, 6, 8, 11, 13, 14; E. M. M. in 1, 2, 3, 6, 9, 11, 12, 13, 14.

Scientific editor in charge: Dra. Cecilia Cracco.