Revista Latinoamericana de Psicología

ORIGINAL

Psychometric properties of the Spanish version of the Patient Health Questionnaire-9 in users of the Ecuadorian public health care system

Carlos Quiñonez-Freire [a,b,*] , Mª Dolores Vara [c,d], José Manuel Tomás [e], Rosa Mª Baños [a,c,d]

[a] Department of Personality, Evaluation and Psychological Treatments, University of Valencia, Spain
[b] Mental Health Unit Hospital Dr. Gustavo Dominguez Z, Santo Domingo, Ecuador
[c] CIBER Fisiopatología Obesidad y Nutrición (CIBEROBN), Instituto Carlos III, Spain
[d] Polibienestar Research Institute, University of Valencia, Spain
[e] Department of Methodology for the Behavioral Sciences, University of Valencia, Valencia, Spain

Abstract **Introduction:** Depression is one of the most prevalent public health problems worldwide and remains an underdiagnosed disease, especially in developing countries. The availability of instruments that permit detecting this problem is essential for reducing the burden of this disorder. The aim of this study was to examine the psychometric properties (factor structure, sex invariance, internal consistency, and convergent and divergent validity) of the Spanish version of the PHQ-9 (Patient Health Questionnaire-9). **Method:** A sample of 366 public hospital users in Ecuador was recruited and accordingly completed a survey that included sociodemographic data, the PHQ-9, and other standards of measurement such as the Beck Depression Inventory-II (BDI-II), the Generalised Anxiety Disorder Scale (GAD-7), the Positive and Negative Affect Scale (PANAS), and the Quality-of-Life Index (QLI). **Results:** The unidimensional factor model of the PHQ-9 provided the best fit for the data. The PHQ-9 assesses depressive symptoms equivalently in both sexes, and presents high internal consistency, as well as good convergent and divergent validity with the other constructs. **Conclusion:** The application of the PHQ-9 could aid in the screening of patients with depressive symptomatology in the Ecuadorian public health system (EPHS).

Propiedades psicométricas de la versión española del Cuestionario de Salud del Paciente-9 en usuarios del sistema público de salud ecuatoriano

Resumen **Introducción:** La depresión es uno de los problemas de salud pública más prevalentes a nivel mundial, y suele ser subdiagnosticada especialmente en países en vías de desarrollo. Contar con instrumentos que permitan realizar un cribado de este problema es fundamental para reducir el costo de este trastorno. El objetivo de este artículo fue examinar

* Corresponding author.
E-mail: carequi@alumni.uv.es

las propiedades psicométricas (estructura factorial, invarianza en función del sexo, consistencia interna, validez convergente y divergente) de la versión española del PHQ-9 (Patient Health Questionnaire-9). **Método:** Se reclutó una muestra de 366 usuarios de un hospital público en Ecuador, los cuales respondieron una encuesta que incluía datos sociodemográficos, el PHQ-9, el inventario de depresión de Beck-II (BDI-II), la Escala de Trastorno de Ansiedad Generalizada-7 (GAD-7), la Escala de Afecto Positivo y Negativo (PANAS), y el índice de calidad de vida (QLI). **Resultados:** El modelo unidimensional del PHQ-9 proporcionó mejor ajuste de los datos. El PHQ-9 evalúa síntomas depresivos de forma equivalente en ambos sexos y presenta una elevada consistencia interna, así como una buena validez convergente y divergente con el resto de los constructos. **Conclusión:** La administración del PHQ-9 podría apoyar el cribado de pacientes con sintomatología depresiva en el sistema público de salud ecuatoriano (SSPE).

Depression is one of the most prevalent diseases in the world (World Health Organisation, 2021). In the Americas, this disorder represents 15% of the global disease burden (World Health Organisation, 2017), having become one of the leading causes of non-fatal health loss (James et al., 2018). Specifically, South America is one of the regions with the highest number of cases of depression (Pan American Health Organisation, 2018). In countries such as Ecuador, the prevalence is 4.6% and it represents 9.2% of the disability adjusted life total (Kohn et al., 2018). Furthermore, it is estimated that only 4.7% of the population in low-and middle-income countries receives adequate treatment (Pan American Health Organisation, 2018).

Depressive disorders are commonly under-recognised in Primary Care (PC) settings (Akincigil & Matthews, 2017), remaining undetected about half the time they are present (Craven & Bland, 2013). There are several barriers to their identification, such as the insufficient training of health professionals, limited time for clinical care (Kroenke & Unützer, 2017), and the fact that many patients more often report somatic or physical complaints without directly referring to their emotional problems related to depression (Haftgoli et al., 2010). The lack of adequate treatment leads to the worsening of symptoms (James et al., 2018), and decreased quality of life for patients (Unützer & Park, 2012).

Reliable, brief, and easy-to-administer depression screening tools are important to help PC professionals identify at risk patients (Nabbe et al., 2017). Specifically, the Patient Health Questionnaire-9 (PHQ-9, Kroenke et al., 2001) is one of the most widely used instruments for screening depression in PC settings (El-Den et al., 2018; Manea et al., 2015) due to its brevity and excellent psychometric properties (Kroenke et al., 2010; Kroenke & Spitzer, 2002). The PHQ-9 is an adjectival type scale derived from the Primary Care Evaluation of Mental Disorders (PRIME-MD) for evaluating depressive symptoms with the criteria of the DSM-IV (Kocalevent et al., 2013).

Subsequent studies in different languages and contexts have also shown that the PHQ-9 is an acceptable standard of measurement with good psychometric properties (Manea et al., 2015). Regarding the PHQ-9 factorial structure, the literature shows mixed results that support both the single-factor (depression) and the two-factor (somatic and cognitive-affective) structure, both in community and clinical settings, as well as for both sexes, multiple age groups and different countries (Lamela et al., 2020).

In the Latin American context, Spanish versions of the PHQ-9 have been reported to be reliable and valid measures in PC from Honduras (Wulsin et al., 2002), Chile, (Baader et al., 2012; Saldivia et al., 2019), Argentina (Urtason et al., 2019), Colombia (Cassiani-Miranda et al., 2021), and Spanish-speaking populations residing in the United States (Huang et al., 2006). In Ecuador, there is only one study that has analysed the tablet-based PHQ-9 as a screening for depression in terms of correlation and feasibility (Grunauer et al., 2014). However, its psychometric properties have not been rigorously validated specifically in Ecuador.

The present study aims to examine the psychometric properties (specifically, the unidimensional and two-dimensional factor models, sex invariance, internal consistency, and convergent and divergent validity) of the Spanish version of the PHQ-9 (Diez-Quevedo et al., 2001) in the Public Health Care setting in Ecuador.

Method

Participants

The sample consisted of 366 patients (127 men and 239 women) of the Dr. Gustavo Domínguez Zambrano Hospital (HGDZ) related to the Public Health Ministry of Santo Domingo (Ecuador). All the participants were between 18 and 71 years of age (M = 32.91, SD = 10.56). The inclusion criteria were: (a) age of 18 years or more; (b) subject to no cognitive impairment (e.g., memory loss, confusion, or problems in understanding the content of the survey); and (c) the granting of voluntary consent for participation in the study. Sociodemographic data appears on Table 1.

Procedure

The study was approved by the Teaching Department and the Ethics Committee of the Hospital Gustavo Domínguez Zambrano (MSP-CZ4-HGDGDZ-DI-2019-0022-M). Participants were recruited from the waiting rooms by an expert psychologist from the Hospital's Mental Health Unit, between August 2019 and January 2020. The psychologist encouraged the patients to participate in the study on the psychometric validation of the PHQ-9, and was present at each application to provide instructions. The survey was administered in paper-and-pencil format. Written informed consent was

requested from all participants after explaining the objectives of the study. Participants were informed of the content of the questionnaires and that participation was completely voluntary. They were also guaranteed anonymity and the confidentiality of their data. On average, it took between 20 and 30 minutes to complete the survey.

Table 1. Sociodemographic characteristics of the sample

Sociodemographic variables	*n* = 366	%
Sex		
Men	127	34.7
Women	239	56.3
Age (years)	32.91(10.56)[a]	
Men	18-66[b]	
Women	18-71[b]	
Marital status		
Single	116	37.7
Free union	137	37.4
Married	69	18.9
Separated	40	10.9
Widowed	4	1.1
Educational level		
No education	6	1.6
Primary school	84	23
Secondary school	48	13.1
High school	157	42.9
Vocational school	13	3.6
University	58	15.8
Work status		
Unemployed	68	18.6
Student	49	13.4
Housekeeper	129	35.2
Full-time job	72	19.7
Part-time job	47	12.8
Retired	1	.3
Income level		
No income	216	59
Lower than minimum I income ($400)	62	16.9
Minimum income	49	13.4
More than minimum I income	39	10.7
Internet access		
Yes	216	59
No	150	40

Note. [a] These values are the mean and standard deviation of age, which are continuous variables. [b] These values correspond to the age range.

Measures

The Patient Health Questionnaire-9 (PHQ-9; Kroenke et al., 2001) is a 9-item self-administered test for screening, diagnosing, monitoring, and measuring depression severity. Participants describe their case, taking into account the two weeks prior to the evaluation. Items are rated from 0 to 3, denoting "*not at all*", "several days", "more than half the days", and "*nearly every day*", respectively. Total scores range from 0-27. Cut-off points of 5, 10, 15, and 20 represent mild, moderate, moderately severe, and severe depression (DSM-IV-TR; American Psychiatric Association [APA], 2002). This study used the Spanish version by Díez-Quevedo et al. (2001), which has been shown to have good psychometric properties (k = .62; overall accuracy, 91%; sensitivity, 84%; specificity, 92%) in the detection of depression in medical and surgical inpatients in the Spanish hospital setting.

The Beck Depression Inventory-II (BDI-II; Beck et al., 1996) is a self-report consisting of 21 items that measure the presence of depressive symptoms. Participants choose the statement that best describes their case in the two weeks leading up to and, including the day of application. Items are rated from 0 ("*not at all*") to 3 ("*extreme form*"), depending on the chosen statement, and the total score ranges from 0 to 63. Standardised score ranges for categorical levels of depressive symptoms are as follows: 0-13 (minimal depression), 14-19 (mild depression), 20-28 (moderate depression), and 29-63 (severe depression). This study employed the Spanish version by Sanz et al. (2003), which has been demonstrated to exhibit good psychometric properties. In the current study the alpha coefficient was excellent (α = 0.90).

The Generalised Anxiety Disorder Scale (GAD-7; Spitzer et al., 2006) is a 7-item self-reported measure that assesses the symptoms and severity of anxiety based on the DSM-IV diagnostic criteria. Items are rated from 0 ("*not at all*") to 3 ("*nearly every day*"), and the total score ranges from 0 to 21. The Spanish version by García-Campayo et al. (2010), which was used in the present study, has evidenced good internal consistency. In this study the Cronbach's alpha was very satisfactory (α = 0.88).

The Positive and Negative Affect Scale (PANAS; Watson et al., 1988) is a 20-item scale that evaluates two independent dimensions: positive affect (10 items) and negative affect (10 items). Items are rated from 1 ("*very slightly or not at all*") to 5 ("*very much*"), and the total score for each subscale ranges from 5 to 50. This study used the Spanish version by Sandín et al. (1999), which has shown adequate internal consistency. In the current study, the alpha coefficients were quite satisfactory for both subscales (α = 0.90 for positive; α = 0.88 for negative).

The Quality-of-Life Index (QLI, Spanish version, Mezzich et al., 2000) is a 10-item self-reported questionnaire that measures 10 dimensions related to the construct of "quality-of-life" (ranging from physical well-being to spiritual fulfilment, including a global perception of quality-of-life). Items are rated from 0 ("*not at all*") to 10 ("*completely*"). The index presents good psychometric properties. Cronbach's alpha in the present study was excellent (α = 0.90).

Data analysis

First, we performed a confirmatory factor analysis (CFA) to test the factor structure of the PHQ-9. Once the skewness and kurtosis values (see Table 2) in our data were analysed, the unidimensional factor model and a two-dimensional factor model with two latent variables (somatic and cognitive/affective) were estimated by Weighted Least

Squares Mean and Variance Adjusted (WLSMV), a procedure recommended for non-normal ordinal data (Finney & DiStefano, 2006). Moreover, model fit was assessed using the Chi-Square (χ^2) goodness-of-fit test ($p \geq .05$), the Comparative Fit Index (CFI), the Root Mean Square Error of Approximation (RMSEA), and the Standardised Root Mean Square Residual SRMR. Values of at least .90 (excellent over .95) in the CFI, less than .06 in the RMSEA and less than .08 for the SRMR represent a good fit (Hu & Bentler, 1999).

Secondly, in order to examine the measurement invariance with regard to sex, we conducted a multi-group CFA (Gregorich, 2006), also estimated with WLSMV given the ordinal nature of the items. We constructed the following three increasingly restrictive models: where all the parameters were free (model 1: configural invariance, that is, when the same factor structure is freely estimated in both groups); where the loadings were invariant (model 2: metric invariance, that is, when there is equivalence in the meaning of the latent factors studied among groups); and where the loadings and intercepts were invariant (model 3: scalar invariance, that is, when there is equivalence in the factor loadings and the thresholds of the items across sex). The following fit indices were used to evaluate the relative fit of the models: differences in the Chi-Square ($\Delta\chi^2$) and the CFI (ΔCFI). A non-significant $\Delta\chi^2$ and a value of less than .01 in the ΔCFI are indicators of invariance between the groups (Cheung & Rensvold, 2002). Chi-square differences were calculated with DIFFTEST, the procedure for categorical (ordinal) data.

Thirdly, the internal consistency of the PHQ-9 was evaluated using Cronbach's alpha and McDonald's omega coefficients. Finally, to assess the PHQ-9 convergent and divergent validity, bivariate correlations with Pearson's r were performed to test the relationship of the PHQ-9 with the BDI-II, the GAD-7, the PANAS, and the QLI. CFA and invariance analyses were performed with Mplus v. 8.5, and descriptive data, reliability, and correlation analyses with SPSS v. 26.

Results

Descriptive statistics

Table 2 shows the mean, standard deviations, and distribution of the PHQ-9 items in our sample.

Table 2. Descriptive statistics of the PHQ-9 for the Ecuadorian sample

PHQ-9 Items	M	SD	Skewness index	Kurtosis index
Item 1	.978	.996	.680	-.649
Item 2	1.074	1.008	.610	-.712
Item 3	.910	1.028	.805	-.593
Item 4	1.183	.958	.511	-.627
Item 5	.932	1.049	.782	-.672
Item 6	.749	1.022	.109	-.095
Item 7	.691	.957	.256	.471
Item 8	.568	.879	.399	.867
Item 9	.320	.747	.458	.219

Note: *M*, Mean; *SD*, Standard Deviation.

Confirmatory factor analysis

The model fit statistics of the unidimensional factor model ($\chi^2[27] = 66.716$, $p < .001$; CFI = .981; RMSEA = .063; SRMR = .039) and the two-dimensional factor model ($\chi^2[26] = 62.186$, $p < .001$; CFI = .983; RMSEA = .062; SRMR = .037) were adequate. However, in the case of the two-dimensional factor model, the correlation between the somatic and cognitive/affective latent factors was very high ($r = .933$), demonstrating non-discrimination between the latent factors. A correlation of .933 per se is not sufficient proof of non-discrimination. However, we also have the relative fit of both models. Following Cheung and Resvold (2002) differences of less than .001 in the CFI (a logic that may be extended to the RMSEA and the SRMR) are meaningless and therefore the more parsimonious model should be retained. These differences in our case are: DCFI = .002; DRMSEA = .001; DSRMR = .001. Additionally, we have calculated a chi-square difference with DIFFTEST, the procedure to compare chi-squares in WLSMV estimation. This difference was not statistically significant Dc2 = 3.85, $p > .05$, thus supporting the one-factor model over the two-factor model. Finally, a 95% confidence interval with regards to the correlation between the two factors included 1 (no discrimination) 95% CI [.859-1.007]. For these reasons, the unidimensional factor model, the most parsimonious, was retained as a better fitting model. Table 3 shows the fit indices of the unidimensional and two-dimensional factor models.

**Table 3.** The fit indices of the unidimensional and two-dimensional factor models

Models	χ^2	df	p	CFI	RMSEA (90% CI)	SRMR
Unidimensional factor model	66.716	27	< .001	.981	.063 (.044; .083)	.039
Two-dimensional factor model	62.186	26	< .001	.983	.062 (.042; .081)	.037

Note: CI, Confidence Interval; CFI, Comparative Fit Index; df, degree of freedom; RMSEA, Root Mean Squared Error of Approximation; SRMR, Standardised Root Mean Square Residual; χ^2, Chi-Square Value.

Figure 1 shows the standardised factor loadings for the unidimensional and two-dimensional factor models using the entire sample. In both models, all factor loadings were significant ($p < .05$), ranging from .664 to .815 for the unidimensional factor model, and from .669 to .822 for the two-dimensional factor model.

Multi-group confirmatory factor analysis (CFA)

Prior to testing for sex measurement invariance, the one-factor model was estimated separately for men and women. Fit for men was adequate: $\chi^2(27) = 38.75$, $p < .001$; RMSEA = .059, 90% CI [.000-.097]; CFI = .981; SRMR = .058.

Figure 1.   Standardised factor loadings for the unidimensional and two-dimensional factor models
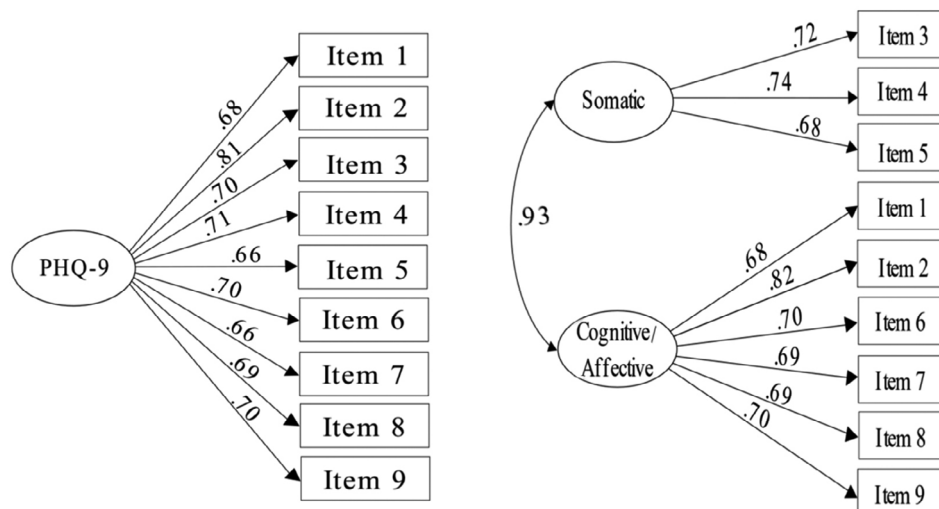
Note: PHQ-9, Patient Health Questionnaire-9.

Table 4.   Summary of goodness of fit indices for tested models in multi-group CFA

Model	χ^2 (df)	p	CFI	RMSEA (90% CI)	SRMR	ΔCFI	$\Delta\chi^2$	Δdf	p
Model 1	97.716 (55)	< .001	.980	.065 (.043; 0.086)	.049	___	___	___	___
Model 2	104.389 (63)	< .001	.981	.060 (.039; .080)	.050	.001	9.222	8	.324
Model 3	116.114 (79)	.004	.983	.051 (.029; .070)	.054	.002	19.124	16	.262

Note: CI, Confidence Interval; CFI, Comparative Fit Index; df, degree of freedom; RMSEA, Root Mean Squared Error of Approximation; SRMR, Standardised Root Mean Square Residual; χ^2, Chi- Square Value.

Model fit in the case of women was also adequate: $\chi^2(27) = 60.82$, $p < .001$; RMSEA = .072, 90% CI [.048-.097]; CFI = .977; SRMR = .044. Table 4 shows the summary of goodness of fit indices for the tested models in multi-group CFA. Model 1, where all parameters were free showed adequate fit indices (CFI = .980; RMSEA = .065; SRMR = .049), determining configural variance. There were no significant differences between this unconstrained model and the more restrictive models according to the significance of the Chi-Square indices ($p > .05$) and the differences found for the CFIs (i.e., they were less than the .01 cut-off point). Hence, metric ($p = .324$; ΔCFI = .001) and scalar ($p = .262$; ΔCFI = .002) invariance was established for women and men.

Internal consistency

Overall scale reliability was good (α = .852; ω = .855), in both women (α = .862; ω = .864) and men (α = .824; ω = .828) too. The reliability indices for the other scales were also adequate: BDI-II (α = .904; ω = .905); GAD-7 (α = .883; ω = .884); QLI (α = .902; ω = .903), and positive and negative affect dimensions of the PANAS respectively (α = .896; ω = .897, α = .881; ω = .886).

Convergent and divergent validity

The correlations between the PHQ-9 total score and the measuring scales of depression (BDI-II), anxiety (GAD-7), positive and negative affect (PANAS), and quality-of-life (QLI) are shown on Table 5. A strong correlation was found between the PHQ-9 total score and the BDI-II ($r = .758$), as well as a moderate correlation with the GAD-7 ($r = .631$) and the negative dimension of the PANAS ($r = .631$). In addition, a less strong correlation was found with the positive dimension of the PANAS ($r = -.399$) and the QLI ($r = -.570$).

Table 5.   Correlations between the PHQ-9 and BDI-II, GAD-7, PANAS, and QLI scales

Scales	M	SD	r^a
PHQ-9	7.41	5.876	--
BDI-II	11.67	9.593	.758***
GAD-7	7.72	5.379	.631***
PANAS (+)	31.92	8.568	-.399***
PANAS (-)	20,17	7,949	.610***
QLI	73.15	16.002	-.570***

Note: BDI-II, Beck Depression Inventory-II; GAD-7, General Anxiety Disorder-7; M, Mean; PANAS; Positive and Negative Affect Schedule; (-), negative; (+), positive; PHQ-9, Patient Health Questionnaire-9; QLI, Quality-of-Life Index; SD, Standard Deviation.

Discussion

The main goal of our study was to provide evidence regarding the psychometric properties of the Spanish version of the PHQ-9 (Diez-Quevedo et al., 2001) in the public health context of the Ecuadorian population. The factor structure (unidimensional and two-dimensional factor model), sex invariance, reliability, and convergent and divergent validity were tested.

The results of our CFA support the unidimensional factor model with excellent fit and adequate factor loadings for the all PHQ-9 items as conceived in the original design (Kroenke et al., 2001) and replicated in a series of subsequent studies (e.g., Aslan et al., 2020; González-Blanch et al., 2018). In contrast to other studies (Granillo, 2012; Krause et al., 2010), the high correlation between somatic and cognitive/affective factors in the two-dimensional factor model suggests that these factors are not well differentiated, thus providing additional support for a single-factor solution.

According to the measurement invariance results, our data demonstrate satisfactory fit on the configural, metric and scalar invariance models with regards to sex, showing an equivalence in the PHQ-9 structure for men and women. This finding is consistent with other studies carried out with a sample of primary care patients in Spain (González-Blanch et al., 2018), American university students (Keum et al., 2018), and Peruvian citizens (Villarreal-Zegarra et al., 2019). In addition, a large section of the literature supports using the unidimensional model in the general population in the comparison between men and women (Doi et al., 2018).

The internal consistency coefficient reported in this study was good for the whole sample (α = 0.852; ω = .855) and for both groups (α = 0.862; ω = .864, women; α = 0.824; ω = .828, men), being only slightly lower than that obtained with the samples of the primary care and gynecology patients in the original study by Kroenke et al. (2001), but higher than in other studies with similar health settings (Baader et al., 2012; Hanlon et al., 2015), showing adequate reliability for our sample.

The results regarding the correlation of the PHQ-9 with measures of depression, anxiety, positive and negative affect, and quality of life reinforce the evidence of its construct validity. As in previous studies, the PHQ-9 global score showed a strong positive correlation with the BDI-II (Diez-Quevedo et al., 2001; Lamela et al., 2020), demonstrating good convergent validity. Moreover, moderate correlations between the PHQ-9 with the GAD-7 and the affect dimension of the PANAS, as well as weaker correlations with the positive affect dimension of the PANAS and the QLI were interpreted as evidence of discriminant validity as in previous studies (Díaz-García et al., 2020; Lee & Kim, 2019; Mira et al., 2019).

A strong point of this work is related to the characteristics of our sample, and also the context in which it has been carried out. So far, only one study exists in the Ecuadorian population that analyses various psychometric properties of the PHQ-9 (Grunauer et al., 2014). However, that study took place in the private health sector, which hinders transferring results to other contexts, such as the public health setting. Therefore, the present study is spearheading in that it provides evidence regarding the application of the

Spanish version of the PHQ-9 in the Ecuadorian population, and in the public sector.

One limitation of this study is that a gold standard criterion for diagnosing depression was not included, and therefore, we cannot determine the sensitivity and specificity or the cut-off points for the diagnosis nor the severity of depression by means of the PHQ-9 in our sample. Another important limitation of our study is that non-probability sampling was used, in consequence the results cannot be generalised to the total population and should be interpreted with caution. Future studies should examine the potential of the PHQ-9 in terms of its diagnostic efficacy in the Ecuadorian population, including a larger and more representative sample (e.g., by sex, gender, educational level, work status) and considering other regions of the country. In addition, another aspect that needs to be taken into account in future studies is the use of cross-cultural adaptations of existing instruments so that they can be used effectively in a wide range of cultural settings and languages.

In conclusion, the results of this study support the unidimensional factor model and the sex invariance of the PHQ-9 scale. Moreover, this instrument is reliable and presents good convergent and divergent validity with other related constructs. The quick and easy application of the PHQ-9 would allow correct identification of depressive symptoms in the population attending public health services in Ecuador.

References

Akincigil, A., & Matthews, E. B. (2017). National rates and patterns of depression screening in primary care: Results from 2012 and 2013. *Psychiatric Services*, 68(7), 660-666. https://doi.org/10.1176/appi.ps.201600096

American Psychiatric Association (2002). *Manual diagnóstico y estadístico de los trastornos mentales.* Cuarta edición, texto revisado (DSM-IV-TR). Masson.

Aslan, J., Cova, F., Saldivia, S., Bustos, C., Inostroza, C., Rincón, P., Ortiz, C., & Bühring, V. (2020). Psychometric properties of the patient health questionnaire-9 in elderly Chilean primary care users. *Frontiers in Psychiatry, 11*, 555011. https://doi.org/10.3389/fpsyt.2020.555011

Baader, M., Molina, J., Venezian, S., Rojas, C., Farías, R., Fierro-Freixeneta, C., Backenstrass, M., & Mundt, C. (2012). Validación y utilidad de la encuesta PHQ-9 (Patient Health Questionnaire) en el diagnóstico de depresión en pacientes usuarios de atención primaria en Chile. *Revista Chilena de Neuro-Psiquiatría, 50*(1), 10-22. https://doi.org/10.4067/S0717-92272012000100002

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for the Beck Depression Inventory-II.* Psychological Corporation.

Cassiani-Miranda, C. A., Cuadros-Cruz, A. K., Torres-Pinzón, H., Scoppetta, O., Pinzón-Tarrazona, J. H., López-Fuentes, W. Y., Paez, A., Cabanzo-Arenas, D. F., Ribero-Marulanda, S., & Llanes-Amaya, E. R. (2021). Validez del cuestionario de salud del paciente-9 (PHQ-9) para cribado de depresión en adultos usuarios de atención primaria en Bucaramanga, Colombia. *Revista Colombiana de Psiquiatría, 50*(1), 11-21. https://doi.org/10.1016/j.rcp.2019.09.001

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness of fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 9(2)*, 233-255. https://doi.org/10.1207/S15328007SEM0902_5

Craven, M. A., & Bland, R. (2013). Depression in primary care: Current and future challenges. *The Canadian Journal of Psychiatry, 58*(8), 442-448. https://doi.org/10.1177/070674371305800802

Díaz-García, A., González-Robles, A., Mor, S., Mira, A., Quero, S., García-Palacios, A., Baños, R.M, & Botella, C. (2020). Positive and Negative Affect Schedule (PANAS): Psychometric properties of the online Spanish version in a clinical sample with emotional disorders. *BMC Psychiatry, 20*(1), 56. https://doi.org/10.1186/s12888-020-2472-1

Diez-Quevedo, C., Spitzer, R. L., Kroenke, K., Rangil, T., & Sanchez-Planell, L. (2001). Validation and utility of the patient health questionnaire in diagnosing mental disorders in 1003 general hospital Spanish inpatients. *Psychosomatic Medicine, 63*(4), 679-686. https://doi.org/10.1097/00006842-200107000-00021

Doi, S., Ito, M., Takebayashi, Y., Muramatsu, K., & Horikoshi, M. (2018). Factorial validity and invariance of the patient health questionnaire (PHQ)-9 among clinical and non-clinical populations. *PLoS ONE, 13*(7), e0199235. https://doi.org/10.1371/journal.pone.0199235

El-Den, S., Chen, T. F., Gan, Y. L., Wong, E., & O'Reilly, C. L. (2018). The psychometric properties of depression screening tools in primary healthcare settings: A systematic review. *Journal of Affective Disorders, 225*, 503-522. https://doi.org/10.1016/j.jad.2017.08.060

Finney, S. J., & DiStefano, C. (2006). Non-normal and categorical data in structural equation modeling. In *Structural equation modeling: A second course* (pp. 269-314). Information Age Publishing.

García-Campayo, J., Zamorano, E., Ruiz, M. A., Pardo, A., Pérez-Páramo, M., López-Gómez, V., Freire, O., & Rejas, J. (2010). Cultural adaptation into Spanish of the generalized anxiety disorder-7 (GAD-7) scale as a screening tool. *Health and Quality of Life Outcomes, 8*(1), 8. https://doi.org/10.1186/1477-7525-8-8

González-Blanch, C., Medrano, L. A., Muñoz-Navarro, R., Ruíz-Rodríguez, P., Moriana, J. A., Limonero, J. T., Limonero, J., Schmitz, F., & Cano-Vindel, A. (2018). Factor structure and measurement invariance across various demographic groups and over time for the PHQ-9 in primary care patients in Spain. *PLoS ONE, 13*(2), e0193356. https://doi.org/10.1371/journal.pone.0193356

Granillo, M. T. (2012). Structure and function of the Patient Health Questionnaire-9 among Latina and non-Latina white female college students. *Journal of the Society for Social Work and Research, 3*(2), 80-93. https://doi.org/10.5243/jsswr.2012.6

Gregorich, S. E. (2006). Do self-report instruments allow meaningful comparisons across diverse population groups? Testing measurement invariance using the confirmatory factor analysis framework. *Medical Care, 44*(11 Suppl 3), S78-S94. https://doi.org/10.1097/01.mlr.0000245454.12228.8f

Grunauer, M., Schrock, D., Fabara, E., Jimenez, G., Miller, A., Lai, Z., Kilbourn, A., & McInnis, M. G. (2014). Tablet-based screening of depressive symptoms in Quito, Ecuador: Efficiency in primary care. *International Journal of Family Medicine, 2014*, 1-7. https://doi.org/10.1155/2014/845397

Haftgoli, N., Favrat, B., Verdon, F., Vaucher, P., Bischoff, T., Burnand, B., & Herzig, L. (2010). Patients presenting with somatic complaints in general practice: Depression, anxiety and somatoform disorders are frequent and associated with psychosocial stressors. *BMC Family Practice, 11*(1), 67. https://doi.org/10.1186/1471-2296-11-67

Hanlon, C., Medhin, G., Selamu, M., Breuer, E., Worku, B., Hailemariam, M., Lund, C., Prince, M., & Fekadu, A. (2015). Validity of brief screening questionnaires to detect depression in primary care in Ethiopia. *Journal of Affective Disorders, 186*, 32-39. https://doi.org/10.1016/j.jad.2015.07.015

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1-55. https://doi.org/10.1080/10705519909540118

Huang, F. Y., Chung, H., Kroenke, K., Delucchi, K. L., & Spitzer, R. L. (2006). Using the patient health questionnaire-9 to measure depression among racially and ethnically diverse primary care patients. *Journal of General Internal Medicine, 21*(6), 547-552. https://doi.org/10.1111/j.1525-1497.2006.00409.x

James, S. L., Abate, D., Abate, K. H., Abay, S. M., Abbafati, C., Abbasi, N., A., Abbastabar, H., Abd-Allah, F., Abdela, J., Abdelalim, A., Abdollahpour, I., Abdulkader, R. S., Abebe, Z., Abera, S. F., Abil, O. Z., Abraha, H. N., Abu-Raddad, L. J., Abu-Rmeileh, N., Accrombessi, M., Acharya, D ... Murray, C. J. L. (2018). Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990-2017: A systematic analysis for the Global Burden of Disease Study 2017. *Lancet, 392*(10159), 1789-1858. https://doi.org/10.1016/S0140-6736(18)32279-7

Keum, B. T., Miller, M. J., & Inkelas, K. K. (2018). Testing the factor structure and measurement invariance of the PHQ-9 across racially diverse U.S. college students. *Psychological Assessment, 30*(8), 1096-1106. https://doi.org/10.1037/pas0000550

Kocalevent, R. D., Hinz, A., & Brähler, E. (2013). Standardization of the depression screener patient health questionnaire (PHQ-9) in the general population. *General Hospital Psychiatry, 35*(5), 551-555.

Kohn, R., Ali, A. A., Puac-Polanco, V., Figueroa, C., López-Soto, V., Morgan, K., Saldivia, S., & Vicente, B. (2018). Mental health in the Americas: An overview of the treatment gap. *Revista Panamericana de Salud Pública, 42*. https://doi.org/10.26633/RPSP.2018.165

Krause, J. S., Reed, K. S., & McArdle, J. J. (2010). A structural analysis of health outcomes after spinal cord injury. *The Journal of Spinal Cord Medicine, 33*(1), 22-32. https://doi.org/10.1080/10790268.2010.11689671

Kroenke, K., & Spitzer, R. L. (2002). The PHQ-9: A new depression diagnostic and severity measure. *Psychiatric Annals, 32*(9), 509-515. https://doi.org/10.3928/0048-5713-20020901-06

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: Validity of a Brief Depression Severity Measure. *Journal of General Internal Medicine, 16*(9), 606-613. https://doi.org/10.1046/j.1525-1497.2001.016009606.x

Kroenke, K., Spitzer, R. L., Williams, J. B. W., & Löwe, B. (2010). The patient health questionnaire somatic, anxiety, and depressive symptom scales: A systematic review. *General Hospital Psychiatry, 32*(4), 345-359. https://doi.org/10.1016/j.genhosppsych.2010.03.006

Kroenke, K., & Unutzer, J. (2017). Closing the false divide: Sustainable approaches to integrating mental health services into primary care. *Journal of General Internal Medicine, 32*(4), 404-410. https://doi.org/10.1007/s11606-016-3967-9

Lamela, D., Soreira, C., Matos, P., & Morais, A. (2020). Systematic review of the factor structure and measurement invariance of the patient health questionnaire-9 (PHQ-9) and validation of the Portuguese version in community settings. *Journal of Affective Disorders, 276*, 220-233. https://doi.org/10.1016/j.jad.2020.06.066

Lee, B., & Kim, Y. E. (2019). The psychometric properties of the generalized anxiety disorder scale (GAD-7) among Korean university students. *Psychiatry and Clinical Psychopharmacology, 29*(4), 864-871. https://doi.org/10.1080/24750573.2019.1691320

Manea, L., Gilbody, S., & McMillan, D. (2015). A diagnostic meta-analysis of the patient Health Questionnaire-9 (PHQ-9) algorithm scoring method as a screen for depression. *General Hospital Psychiatry, 37*(1), 67-75. https://doi.org/10.1016/j.genhosppsych.2014.09.009

Mezzich, J. E., Ruipérez, M. A., Pérez, C., Yoon, G., Liu, J., & Mahmud, S. (2000). The Spanish Version of the Quality of Life Index:

Presentation and validation. *The Journal of Nervous and Mental Disease, 188*(5), 301–305. https://doi.org/10.1097/00005053-200005000-00008

Mira, A., González-Robles, A., Molinari, G., Miguel, C., Díaz-García, A., Bretón-López, J., García-Palacios, A., Quero, S., Baños, R., & Botella, C. (2019). Capturing the severity and impairment associated with depression: The Overall Depression Severity and Impairment Scale (ODSIS) validation in a Spanish clinical sample. *Frontiers in Psychiatry, 10,* 180. https://doi.org/10.3389/fpsyt.2019.00180

Nabbe, P., Le Reste, J. Y., Guillou-Landreat, M., Munoz Perez, M. A., Argyriadou, S., Claveria, A., Fernández San Martín, M.I., Czachowski, S., Lingner, H., Lygidakis, C., Sowinska, A., Chiron, B., Derriennic, J., Le Prielec, A., Le Floch, B., Montier, T., Van Marwijk, H., & Van Royen, P. (2017). Which DSM validated tools for diagnosing depression are usable in primary care research? A systematic literature review. *European Psychiatry, 39,* 99–105. https://doi.org/10.1016/j.eurpsy.2016.08.004

Pan American Health Organization. (2018). *The Burden of Mental Disorders in the Region of the Americas.* PAHO. https://iris.paho.org/handle/10665.2/49578

Saldivia, S., Aslan, J., Cova, F., Vicente, B., Inostroza, C., & Rincón, P. (2019). Propiedades psicométricas del PHQ-9 (Patient Health Questionnaire) en centros de atención primaria de Chile. *Revista Médica de Chile, 147*(1), 53-60. https://doi.org/10.4067/S0034-98872019000100053

Sandín, B., Chorto, P., Lostao, L., Joiner, T. E., Santed, M., & Valiente, R. M. (1999). Escalas PANAS de afecto positivo y negativo: Validación factorial y transcultural. *Psicothema, 11*(1), 37-51. http://www.psicothema.com/psicothema.asp?id=229

Sanz, J., Perdigón, A.L., & Vázquez, C. (2003). Adaptación española del inventario para la depresión de Beck-II (BDI-II): 2. Propiedades psicométricas en población general. *Clínica y Salud, 14*(3), 249-280.

Spitzer, R. L., Kroenke, K., Williams, J. B. W., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder. *Archives of Internal Medicine, 166(*10), 1092. https://doi.org/10.1001/archinte.166.10.1092

Unützer, J., & Park, M. (2012). Strategies to improve the management of depression in primary care. *Primary Care: Clinics in Office Practice, 39*(2), 415-431. https://doi.org/10.1016/j.pop.2012.03.010

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*(1), 291. https://doi.org/10.1186/s12888-019-2262-9

Villarreal-Zegarra, D., Copez-Lonzoy, A., Bernabé-Ortiz, G., Melendez-Torres, J., & Bazo-Alvarez, J. C. (2019). Valid group comparisons can be made with the Patient Health Questionnaire (PHQ-9): A measurement invariance study across groups by demographic characteristics. *PLoS ONE, 14*(9), e0221717. https://doi.org/10.1371/journal.pone.0221717

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of Brief Measures of Positive and Negative Affect: The PANAS scales. *Journal of Personality and Social Psychology, 54*(6), 1063-1070. https://doi.org/10.1037/0022-3514.54.6.1063

World Health Organization (2017). Depression and Other Common Mental Disorders. https://apps.who.int/iris/handle/10665/254610

World Health Organization (2021, March 28). Depression. https://www.who.int/news-room/fact-shEets/detail/depression

Wulsin, L., Somoza, E., & Heck, J. (2002). The feasibility of using the Spanish PHQ-9 to screen for depression in primary care in Honduras. *The Primary Care Companion to The Journal of Clinical Psychiatry, 04*(5), 191-195. https://doi.org/10.4088/PCC.v04n0504