

RESEARCH

Open Access



# Psychometric properties of the ultra-brief self-report Patient Health Questionnaire-4 (PHQ-4) to assess anxiety and depression in Arabic-speaking adults

Sahar Obeid<sup>1\*†</sup>, Ali Hemade<sup>2†</sup>, Diana Malaeb<sup>3</sup>, Fouad Sakr<sup>4</sup>, Mariam Dabbous<sup>4</sup>, Jiale Xu<sup>5</sup>, Yiwei Ying<sup>6</sup>, Chen Jiang<sup>5</sup>, Runtang Meng<sup>5,7</sup>, Feten Fekih-Romdhane<sup>8,9†</sup> and Souheil Hallit<sup>10,11,12\*†</sup>

## Abstract

**Background** Anxiety and depression are psychiatric disorders that often coexist and share some features. Developing a simple and cost-effective tool to assess anxiety and depression in the Arabic-speaking population, predominantly residing in low- and middle-income nations where research can be arduous, would be immensely beneficial. The study aimed to translate the four-item composite Patient Health Questionnaire – 4 (PHQ-4) into Arabic and evaluate its psychometric properties, including internal reliability, sex invariance, composite reliability, and correlation with measures of psychological distress.

**Methods** 587 Arabic-speaking adults were recruited between February and March 2023. An anonymous self-administered Google Forms link was distributed via social media networks. We utilized the FACTOR software to explore the factor structure of the Arabic PHQ-4.

**Results** Confirmatory factor analysis (CFA) indicated that fit of the two-factor model of the PHQ-4 scores was modest ( $\chi^2/df = .13/1 = .13$ , RMSEA = .001, SRMR = .002, CFI = 1.005, TLI = 1.000). Internal reliability was excellent (McDonald's omega = .86; Cronbach's alpha = .86). Indices suggested that configural, metric, and scalar invariance were supported across sex. No significant difference was found between males and females in terms of the PHQ-4 total scores, PHQ-4 anxiety scores, and PHQ-4 depression scores. The total score of the PHQ-4 and its depression and anxiety scores were significantly and moderately-to-strongly associated with lower wellbeing and higher Depression Anxiety and Stress Scale (DASS) total and subscales scores.

**Conclusion** The PHQ-4 proves to be a reliable, valid, and cost-effective tool for assessing symptoms related to depression and anxiety. To evaluate the practical effectiveness of the Arabic PHQ-4 and to further enhance the data on its construct validity, future studies should assess the measure in diverse contexts and among specific populations.

**Keywords** Anxiety, Arabic, Depression, Patient Health Questionnaire, Psychometric properties, Validation

<sup>†</sup>Sahar Obeid and Ali Hemade are first coauthors.

<sup>†</sup>Feten Fekih-Romdhane and Souheil Hallit are last coauthors.

\*Correspondence:

Sahar Obeid  
saharobeid23@hotmail.com  
Souheil Hallit  
souheilhallit@usek.edu.lb

Full list of author information is available at the end of the article



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

## Introduction

Anxiety and depression are psychiatric disorders that often coexist and share some features. Both conditions are highly prevalent, with anxiety affecting approximately one in 15 persons annually and depression affecting one in 20 persons annually [1]. The prevalence of anxiety and depression varies across different populations and settings. A meta-analysis encompassing 174 surveys from 63 countries indicated that approximately 1 in 5 respondents (17.6%) experienced a common mental disorder, including mood and anxiety disorders, during the previous 12 months. Regional variations were notable, with lower prevalence rates in North and South East Asia and higher rates in English-speaking countries [1]. Comorbidity of anxiety and depression may be present as the full clinical picture of the two syndromes or as limited symptoms from both two syndromes. Anxiety disorders typically present first [2]. The theoretical framework behind the comorbidity of anxiety and depression integrates several psychological, biological, and environmental factors. This comorbidity is understood through a network theory perspective, which posits that overlapping symptoms, or "bridge symptoms," such as worrying and irritability, facilitate the interaction between anxiety and depression, leading to mutual reinforcement of these disorders [3]. Additionally, early-life stressors and trauma contribute to the development of both disorders, indicating that environmental factors play a critical role in their comorbidity [4]. Anxiety and depression also share other similar risk factors, including female gender, family history, and perinatal factors [5]. Neurobiological research has shown consistent abnormalities across both anxiety and depression, particularly in amygdala hyperactivity [6]. The high comorbidity rate also reflects the interconnected nature of mental health symptoms, where the presence of one disorder significantly increases the likelihood of developing the other, creating a complex interplay of psychological vulnerabilities and stressors that perpetuate both anxiety and depression [7].

Anxiety and depression can lead to functional disability [8], including impaired physical functioning, role limitations due to emotional health problems, and decreased social functioning. Research has shown that both anxiety and depressive symptoms are associated with a reduced quality of life [9]. Compared with those without anxiety, individuals with moderate to severe anxiety were less likely to meet ideal levels of physical activity. Similarly, individuals with moderate to severe depression were less likely than those without depression to meet ideal levels of physical activity. [10]. Furthermore, one study showed that people with low depression and high anxiety, with high depression and low anxiety, and with high depression and high anxiety were at 2.46, 26.32 and 54.77

times more risk for suicide, compared to subjects with low depression and low anxiety [11]. Treatment options including antidepressants and cognitive behavioral therapy are similar for anxiety and depression [12]. Results from 51,547 respondents in a World Health Organization (WHO) study revealed that only 9.8% of individuals with anxiety disorder received adequate treatment, and only 27.6% received any form of treatment at all [13].

Several instruments have been developed for measuring depression, including the Beck Depression Inventory [14], Hamilton Depression Rating Scale [15], Montgomery-Åsberg Depression Rating Scale [16], and Lebanese Depression Scale [17]. A commonly used brief depression severity scale is the Patient Health Questionnaire – 9 (PHQ-9) [18], which has been shortened into the two-item PHQ-2 [19]. Similarly, other scales were developed to measure clinical anxiety, such as the Beck Anxiety Inventory [20], Lebanese Anxiety Scale [21], and Generalized Anxiety Disorder Scale – 7 (GAD-7) [22]. The 7-item GAD-7 scale was shortened into the 2-item GAD-2 [23]. However, these previous scales either measure depression or anxiety separately, but do not measure both. Considering the high comorbidity of these two disorders, there was a need for a validated and reliable scale that measures both depression and anxiety. From this perspective, the four-item composite Patient Health Questionnaire – 4 (PHQ-4) was developed by combining the PHQ-2 and GAD-2 scales [24], which contains two core anxiety items and two core depression items. The PHQ-4 leverages the cognitive-behavioral model to assess both anxiety and depression, rooted in the interconnection between thoughts, feelings, and behaviors. According to this model, mental health conditions are maintained by negative thought patterns and maladaptive behaviors. Depression is characterized by pervasive negative views about oneself, the world, and the future, while anxiety involves overestimating the likelihood of negative events and underestimating one's ability to cope, contributing significantly to emotional distress [25]. Additionally, behavioral avoidance plays a critical role in both conditions; individuals with anxiety avoid situations that trigger their anxiety, and those with depression withdraw from activities that previously brought them joy. By identifying these cognitive and behavioral patterns, the PHQ-4 could effectively screen for anxiety and depression, aiding in early intervention and treatment [25]. In the original study [24], the PHQ-4 demonstrated construct validity, factorial validity, and internal reliability, proving to be an efficient ultra-brief tool for screening for anxiety, depression, or both. Higher scores on the PHQ-4 scale indicate a need for further assessment of anxiety, depression, or both.

Ultra brief assessment tools such as the PHQ-4 are essential for busy clinicians, patient follow-up after treatment, and time-restricted research. Numerous studies have shown that ultra-short two- or three-question tests perform comparably to their longer counterparts [26–28]. In fact, the PHQ-4 has shown comparable performance to longer depression and anxiety measures in terms of correlating well with measures of quality of life, disability days, and healthcare utilization [24]. The PHQ-4 has been validated and proved to be reliable for use in various populations, including the general population [25], college students [29], pregnancy [30] and attention-deficit/hyperactivity disorder (ADHD) patients [31]. Multiple languages including Spanish [32], Portuguese [33], Greek [34], Korean [35], and Colombian [36] has been translated and validated. Numerous studies have proven a cross-cultural validation of the PHQ-4 [37, 38]. The PHQ-4 has been translated to Arabic and validated on a group of Syrian refugees in Germany [39]. While the authors' efforts are valuable in expanding the cross-cultural adaptability of the PHQ-4, we believe that validating the Arabic PHQ-4 scale in Lebanon, despite its prior validation in the German study, is essential due to several reasons. Cultural differences significantly influence how mental health symptoms are expressed and perceived, necessitating a culturally relevant validation within the local context and population, to ensure accuracy. Additionally, the unique stressors and challenges faced by the specific population of Syrian refugees in Germany, such as different levels of access to resources and support, may impact mental health differently compared to the Lebanese population. Thus, the PHQ-4 still needs to be validated in Arabic-speaking adults from the general population more broadly. Psychometric properties like reliability and validity need testing within the specific population to confirm the scale's effectiveness. Finally, for seamless integration into the Lebanese healthcare system, local validation ensures that healthcare providers can trust and confidently use the scale for screening and diagnosing mental health conditions among Lebanese individuals. Thus, validating the PHQ-4 in Lebanon ensures it is culturally, linguistically, and contextually appropriate, providing accurate and reliable measurements for diagnosing psychological distress.

In Lebanon, research has shown that 17% of individuals met the criteria for a 12-month mental disorder, and 27% were classified as serious [40], and half of the respondents had a history of exposure to war-related traumatic events. Lebanon has recently experienced a series of profound tragedies, including the COVID-19 pandemic and the devastating explosion at the Beirut port on August fourth, which stands as the world's

most powerful non-nuclear explosion [41]. Additionally, Lebanon is grappling with its worst economic crisis in modern history, characterized by the rapid devaluation of the national currency, one of the highest inflation rates globally, and severe shortages of essential resources such as electricity and fuel. All those factors are significantly impacting the well-being and contributing to an increase in psychological disorders among the Lebanese population. Consequently, interventions aimed at early detection and treatment may play a crucial role in reducing the persistence or severity of primary anxiety and depressive disorders and preventing the onset of secondary disorders.

The availability of a psychometrically validated ultra-brief scale such as the PHQ-4 in Arabic will assist clinicians in screening for anxiety and depression, provide a valuable follow-up tool to assess intervention efficacy, augment mental health research in the Arab world, and contribute to a broader cross-cultural validation of the PHQ-4. Therefore, the aim of this study was to translate the PHQ-4 into Arabic and determine the psychometric properties of this translation, including internal reliability, sex invariance, composite reliability, and its correlation with measures of psychological distress and well-being. We hypothesize that our translation of the PHQ-4 will demonstrate a fit for a two-factor solution similar to the original scale [24], good internal consistency reliability, adequate convergent and concurrent validity, as well as cross-sex invariance.

## Methods

### Participants and procedures

All data were collected via a Google Form link, between February and March 2023. After being trained by the research team, five university students were asked to collect data via the snowball sampling technique. Students were instructed to forward the link to acquaintances, who were asked to forward the link to other family members and friends. Inclusion criteria for participation included being of a resident and citizen of Lebanon of adult age. Exclusion criteria were those who refused to fill out the questionnaire. Internet protocol (IP) addresses were monitored to prevent duplicate survey responses. Participants provided digital informed consent before completing the survey instruments, which were presented in a pre-randomized order to control for order effects. The survey was anonymous and participation was voluntary and without remuneration. A total of 587 participants completed the survey (mean age of  $34.48 \pm 15.06$  years, 69.4% females, 42.7% married and 74.0% with a university level of education).

### Translation procedure

According to Beaton's guidelines [42], the forward-backward translation approach was employed for the scale. Initially, the English version was translated into Arabic by two Lebanese translators who were unaffiliated with the study. Subsequently, two Lebanese psychologists who were proficient in English, back-translated the Arabic version back into English. To ensure the accuracy of the translation, the original English version and the translated one were compared, and any inconsistencies were identified and corrected by a committee of experts comprising the research team and the translators [43]. Furthermore, an adaptation of the measure to the Arab context was conducted to ascertain any potential misunderstanding of the item wordings and the ease of item interpretation, ensuring the conceptual equivalence between the original and Arabic scale in both contexts [44]. Following translation and adaptation of the scale, a pilot study was conducted with 20 participants to confirm comprehension of all questions; no alterations were made after the pilot study.

### Measures

#### *Patient Health Questionnaire (PHQ-4)*

The PHQ-4 is a concise 4-item questionnaire designed to assess anxiety and depressive symptoms experienced over the past two weeks [24]. It comprises two subscales: anxiety (e.g. "Feeling nervous, anxious or on edge") and depression (e.g. "Little interest or pleasure in doing things"), each consisting of two items. Each item is rated on a 4-point Likert scale ranging from 0 (not at all) to 3 (nearly every day). To calculate the total PHQ-4 score, the scores from all four items are summed. The cut-off score for the PHQ-4's subscales is greater than or equal to 3. A score between 3 and 5 suggests mild psychological distress, between 6 and 8 suggests moderate psychological distress, and between 9 and 12 suggests severe psychological distress. A score of 0 to 2 indicates the absence of psychological distress.

#### *The Depression Anxiety and Stress Scale-8 items (DASS-8)*

This instrument has been developed and validated in Arabic by Ali et al. [45]. It is composed of eight items and three dimensions: (1) stress (two items; e.g. "I felt that I was using a lot of nervous energy"), (2) anxiety (three items; e.g. "I felt scared without any good reason"), and (3) depression (three items; "I was unable to become enthusiastic about anything"). Higher scores reflect higher level of symptom affirmation (McDonald's  $\omega=0.88$ ; Cronbach's  $\alpha=0.88$ ).

#### *WHO-5 wellbeing index*

Validated in Lebanon [46], formed of 5 items scored on a 6-point Likert scale with anchors ranging from "at no time" to "all the time" (e.g. "In the last two weeks, I have felt cheerful in good spirits"). Items are summed on a scale from 0 to 25, with higher scores reflecting higher wellbeing [47] (McDonald's  $\omega=0.93$ ; Cronbach's  $\alpha=0.93$ ).

### Analytic strategy

#### *Data treatment*

There were no missing responses in the dataset. To examine the factor structure of the PHQ-4, we conducted a confirmatory factor analysis (CFA) using the data from the total sample via SPSS and AMOS (version 29) software. A minimum sample size of 80 participants was needed, based on 20 participants per item on the scale was deemed necessary to conduct the CFA [48]. Parameter estimates were obtained using the maximum likelihood method. Calculated fit indices were the normed model chi-square ( $\chi^2/df$ ), the Steiger-Lind root mean square error of approximation (RMSEA), standardized root mean square residual (SRMR), the Tucker-Lewis index (TLI) and the comparative fit index (CFI). Values  $\leq 5$  for  $\chi^2/df$ ,  $\leq 0.08$  for RMSEA,  $\leq 0.05$  for SRMR and  $\geq 0.95$  for CFI and TLI indicate good fit of the model to the data [49]. Multivariate normality was verified (Bollen-Stine  $p=0.752$ ). Convergent validity was confirmed via an average extracted variance (AVE) value  $> 0.5$ .

#### *Sex invariance*

To examine sex invariance of the PHQ-4 scores, we conducted multi-group CFA using the total sample [50]. Measurement invariance was assessed at the configural, metric and scalar levels [51]. We accepted  $\Delta CFI \leq 0.010$  and  $\Delta RMSEA \leq 0.015$  or  $\Delta SRMR \leq 0.010$  as evidence of invariance [50]. Comparison between males and females was done using the *Student t-test* only if scalar or partial scalar invariance held.

#### *Further analyses*

Composite reliability in both subsamples was assessed using McDonald's omega ( $\omega$ ) and Cronbach's alpha ( $\alpha$ ), with values greater than 0.70 reflecting adequate composite reliability [52]. The normality of the PHQ-4 score was verified, since the skewness and kurtosis values for each item of the scale varied between -1 and +1 [53]. To assess concurrent validity, Pearson's correlation coefficient was used to correlate the PHQ-4 scores with DASS-8 and WHO-5. Correlation coefficients values  $\leq 0.10$  were considered weak,  $\sim 0.30$  were considered

**Table 1** Sociodemographic and other characteristics of the participants (n = 592)

<b>Age (year)</b>	<b>34.48 ± 15.06</b>
Gender	
Male	181 (30.6%)
Female	411 (69.4%)
Marital status	
Single	339 (57.3%)
Married	253 (42.7%)
Education	
Primary	15 (2.5%)
Complementary	50 (8.4%)
Secondary	89 (15.0%)
University	438 (74.0%)
PHQ4 total (mean/SD)	4.69 ± 3.05
PHQ4 depression (mean/SD)	2.41 ± 1.60
PHQ4 anxiety (mean/SD)	2.28 ± 1.69
PHQ4 categories	
None (scores between 0–2)	142 (24.0%)
Mild (scores between 3–5)	220 (37.2%)
Moderate (scores between 6–8)	179 (30.2%)
Severe (scores between 9–12)	51 (8.6%)

moderate, and ~ 0.50 were considered strong correlations [54].

**Results**

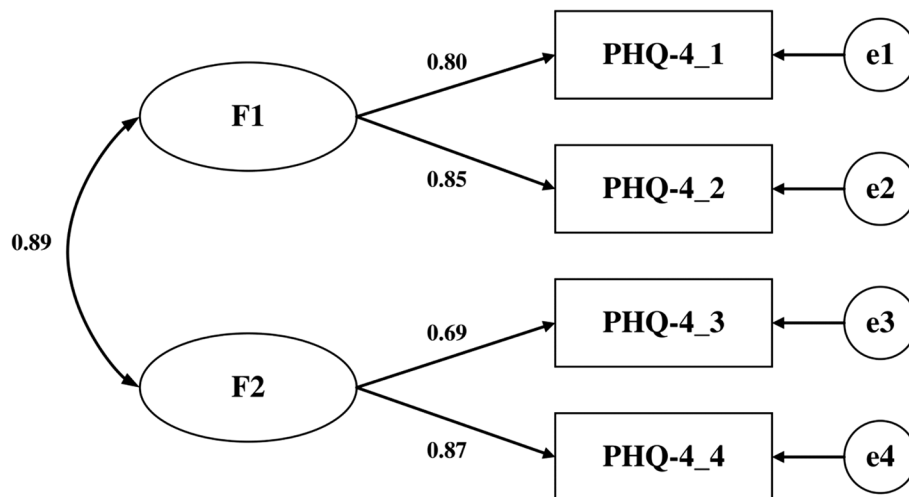
Five hundred eighty-seven participants completed the survey, with a mean age of 34.48 ± 15.06 years and 69.4% females. Other characteristics of the participants are summarized in Table 1.

**Confirmatory factor analysis**

CFA indicated that the fit of the two-factor model of the PHQ-4 scores was modest:  $\chi^2/df=0.13/1=0.13$ , R MSEA=0.001 (90% CI<0.001, 0.078), SRMR=0.002, CFI=1.005, TLI=1.000. The standardized estimates of factor loadings were all adequate (Fig. 1). Internal reliability was excellent (McDonald’s  $\omega=0.86$ ; Cronbach’s  $\alpha=0.86$ ). The AVE value was satisfactory at =0.65.

**Sex invariance**

The indices suggested that configural, metric, and scalar invariance were supported across sex (Table 2). No significant difference was found between males and females in terms of the PHQ-4 total scores (4.75 ± 3.16



**Fig. 1** Standardized Estimates of Factor Loadings from the Confirmatory Factor Analysis in the total sample

**Table 2** Measurement invariance across sex in the total sample

Model	CFI	RMSEA	SRMR	Model Comparison	ΔCFI	ΔRMSEA	ΔSRMR
Configural	1.000	< .001	.004				
Metric	.995	.047	.025	Configural vs metric	.005	.047	.021
Scalar	.991	.053	.025	Metric vs scalar	.004	.006	< .001

CFI Comparative fit index, RMSEA Steiger-Lind root mean square error of approximation, SRMR Standardized root mean square residual

vs  $4.67 \pm 3.00$ ,  $t(590) = 0.30$ ,  $p = 0.762$ ), PHQ-4 anxiety scores ( $2.31 \pm 1.66$  vs  $2.27 \pm 1.71$ ,  $t(590) = 0.24$ ,  $p = 0.807$ ), and PHQ-4 depression scores ( $2.44 \pm 1.68$  vs  $2.40 \pm 1.57$ ,  $t(590) = 0.32$ ,  $p = 0.751$ ).

### Convergent and concurrent validity

The PHQ-4 total score and the PHQ-4 depression and anxiety scores were significantly and moderately-to-strongly associated with lower wellbeing and higher DASS total and subscales scores (Table 3).

### Discussion

The findings from this study suggest that an ultra-brief 4-item measure can reliably and validly measure depression and anxiety in the general population. Overall, the results support the reliability and validity of the instrument, as well as its suitability for use in Arabic-speaking adults from the general population. It is important to note the PHQ-4 serves only as a screening tool, and individuals with elevated PHQ-4 scores should undergo further assessment to determine whether they meet the full diagnostic criteria for either disorder or if intervention is warranted.

Though depression and anxiety often coexist [55, 56], assessment for both conditions seems necessary. Consistent with prior research conducted in various countries, such as Germany [25], Colombia [36], the United States [24, 32], Spain [57], and Iran [58], our study reaffirmed the two-factor structure of the PHQ-4, indicating distinct subscales for anxiety and depression. Hence, our findings support the differentiation between the two scales, PHQ-2 and GAD-2, rather than relying solely on one of them or the total PHQ-4 score. However, these findings contrast with a study conducted among a sample of Quechua speakers, which supported a one-dimensional model, where anxiety and depression items combined to form a single latent variable of emotional problems [59].

Given the brevity of the PHQ-4, its reliability was notably high. In our study, internal reliability was

excellent (McDonald's  $\omega = 0.86$ ; Cronbach's  $\alpha = 0.86$ ), slightly surpassing those reported in the German validation study for the PHQ-2 ( $\alpha = 0.75$ ) and GAD-2 ( $\alpha = 0.82$ ) [25], and to some extent similar to the values from previous studies conducted across various other populations [24, 25, 29, 30, 32, 35, 36]. This suggests that the PHQ-4 is equally effective in measuring symptoms of depression and anxiety in the Arabic-speaking population as it is in other demographic groups.

Another finding of our research is that the factor loadings of the Arabic PHQ-4 remained consistent across sex at the three levels (configural, metric, and scalar). Consequently, comparisons between sex groups showed no statistically significant variances in PHQ-4 total scores and two subscores within our sample. Put differently, individuals of both genders comprehend and interpret the significance of PHQ-4 items similarly. Consistent with our findings, studies conducted in Colombia [36], Germany [25], and Greece [34] have also provided evidence supporting the consistency of measurement across sexes.

Our findings also revealed that the PHQ-4 depression and anxiety scores were significantly and moderately-to-strongly associated with depression, anxiety and stress scores as measured using another brief scale, i.e. the DASS-8, thus confirming the convergent validity of the Arabic PHQ-4. Furthermore, the PHQ-4 total score and the PHQ-4 depression and anxiety scores were linked to lower levels of well-being, in line with a previous study [60]. Indeed, the presence of both anxiety and depressive disorders has been shown to have negative impacts on various aspects of an individual's life, including perceived well-being, satisfaction in relationships [61–63], poorer occupational outcomes [64], and more loneliness [65]. These negative impacts may be attributed to the common characteristics observed in individuals experiencing depressive and anxiety symptoms, such as fatigue, loss of energy, feeling slowed down or agitated, poor attention and concentration, slow thinking, distractibility, impaired

**Table 3** Pearson correlation matrix

	Mean $\pm$ SD	1	2	3	4	5	6	7
1. PHQ-4 total	4.69 $\pm$ 3.05	1						
2. PHQ-4 anxiety	2.28 $\pm$ 1.69	.93***	1					
3. PHQ-4 depression	2.41 $\pm$ 1.60	.92***	.71***	1				
4. DASS total	10.42 $\pm$ 6.07	.56***	.53***	.51***	1			
5. DASS depression	3.79 $\pm$ 2.47	.52***	.47***	.49***	.92***	1		
6. DASS anxiety	3.52 $\pm$ 2.68	.50***	.48***	.43***	.91***	.76***	1	
7. DASS stress	3.11 $\pm$ 1.69	.46***	.44***	.41***	.80***	.65***	.59***	1
8. Well-being	13.68 $\pm$ 5.60	-.50***	-.46***	-.47***	-.36***	-.36***	-.24***	-.39***

\*\*\*  $p < .001$

memory, and indecisiveness [66], all of which contribute to a diminished well-being.

Our study showed that, according to the PHQ-4 cut-off, 24.0% of the participants had no depression, while 37.2%, 30.2% and 8.6% of the participants exhibited mild, moderate and severe psychological distress respectively. These findings are consistent with a prior study indicating a depression prevalence of 34.44% in the Southeast Asian context [67], as well as recent meta-analytic studies by Bueno-Notivol et al. [68] and Salari et al. [69]. Nevertheless, our results are higher than those reported in a German study, which suggested that 6.5% and 7.0% of participants had probable anxiety and depression, respectively [70].

### Clinical implications

Providing a reliable and valid Arabic version of the PHQ-4 could help gather precise epidemiological information regarding anxiety and depression symptoms in Arab nations. This initiative could enhance awareness surrounding mental health (anxiety and depression) screening and diagnosis in Arab contexts, and guide the creation of culturally appropriate interventions grounded in evidence.

### Limitations

Firstly, the snowball sampling method was used, which may have introduced sampling bias and restricted the representation of the general population. Hence, future studies should aim to employ more diverse and representative sampling techniques to enhance the external validity of the findings. Secondly, using cross-sectional data precludes the ability to assess the predictive validity and test–retest reliability of the PHQ-4. Additionally, despite the utilization of a substantial community sample of Lebanese participants in this study, access to the survey was restricted to individuals with internet connectivity, potentially leading to an incomplete representation of the entire adult general population. Finally, this study was conducted exclusively in Lebanon, thereby restricting the generalizability of our results to Arab-speaking individuals in other Arab and non-Arab countries.

### Conclusion

The PHQ-4 proves to be a reliable, valid, and cost-effective tool for assessing symptoms related to depression and anxiety. Using reliable mental health screening instruments lessens the load on participants in extensive data gathering, facilitates swift estimation by

researchers of the prevalence and intensity of mental health symptoms, assists in timely interventions and psychological support, and provides a sustainable method for monitoring and assessing mental health symptoms amidst economic crises and other humanitarian disasters. To evaluate the practical effectiveness of the Arabic PHQ-4 and to further enhance the data on its construct validity, future studies should assess the measure in diverse contexts and among specific populations.

### Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12888-024-05978-8>.

Supplementary Material 1.

### Acknowledgements

The authors would like to thank all participants.

### Authors' contributions

F.F.R, S.O and S.H designed the study; S.O and A.H drafted the manuscript; S.H carried out the analysis and interpreted the results; D.M, F.S, and M.D collected the data. J.X, Y.Y, C.J, and R.M authors reviewed the paper for intellectual content; all authors reviewed the final manuscript and gave their consent.

### Funding

None.

### Availability of data and materials

All data generated or analyzed during this study are not publicly available due the restrictions from the ethics committee, but are available upon a reasonable request from the corresponding author.

### Declarations

#### Ethics approval and consent to participate

Ethics approval for this study was obtained from the ethics committee of the School of Pharmacy at the Lebanese International University (2023RC-006-LIU-SOP). When filling out the online form, each participant provided written informed consent, as did their parents or legal guardian(s) if they were study participants under the age of sixteen. Every step of the process was carried out in compliance with all applicable laws and rules (such as the Declaration of Helsinki).

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare no competing interests.

#### Author details

<sup>1</sup>Social and Education Sciences Department, School of Arts and Sciences, Lebanese American University, Jbeil, Lebanon. <sup>2</sup>Faculty of Medicine, Lebanese University, Hadat, Lebanon. <sup>3</sup>College of Pharmacy, Gulf Medical University, Ajman, United Arab Emirates. <sup>4</sup>School of Pharmacy, Lebanese International University, Beirut, Lebanon. <sup>5</sup>School of Public Health, Hangzhou Normal University, Hangzhou, Zhejiang 311121, China. <sup>6</sup>School of Nursing, Hangzhou Normal University, Hangzhou, Zhejiang 311121, China. <sup>7</sup>Engineering Research Center of Mobile Health Management System, Ministry of Education, Hangzhou, Zhejiang 311121, China. <sup>8</sup>Department of Psychiatry "Ibn Omrane", The Tunisian Center of Early Intervention in Psychosis, Razi Hospital, Manouba 2010, Tunisia. <sup>9</sup>Faculty of Medicine of Tunis, Tunis El Manar University, Tunis, Tunisia. <sup>10</sup>School of Medicine and Medical Sciences, Holy Spirit University of Kaslik, P.O. Box 446, Jounieh, Lebanon. <sup>11</sup>Psychology Department, College of Humanities, Effat

University, Jeddah 21478, Saudi Arabia. <sup>12</sup>Applied Science Research Center, Applied Science Private University, Amman, Jordan.

Received: 1 April 2024 Accepted: 22 July 2024

Published online: 30 July 2024

## References

- Steel Z, Marnane C, Iranpour C, Chey T, Jackson JW, Patel V, Silove D. The global prevalence of common mental disorders: a systematic review and meta-analysis 1980–2013. *Int J Epidemiol*. 2014;43(2):476–93.
- Balta G, Paparrigopoulos T. Comorbid anxiety and depression: diagnostic issues and treatment management. *Psychiatriki*. 2010;21(2):107–14.
- Groen RN, Ryan O, Wigman JTW, Riese H, Penninx BWJH, Giltay EJ, Wichers M, Hartman CA. Comorbidity between depression and anxiety: assessing the role of bridge mental states in dynamic psychological networks. *BMC Med*. 2020;18:308.
- Lamers F, van Oppen P, Comijs HC, Smit JH, Spinhoven P, van Balkom AJ, Nolen WA, Zitman FG, Beekman AT, Penninx BW. Comorbidity patterns of anxiety and depressive disorders in a large cohort study: the Netherlands Study of Depression and Anxiety (NESDA). *J Clin Psychiatry*. 2011;72(3):341–8.
- Wittchen HU, Kessler RC, Pfister H, Höfler M, Lieb R. Why do people with anxiety disorders become depressed? A prospective-longitudinal community study: anxiety disorders predict development of depression. *Acta Psychiatr Scand*. 2000;102:14–23.
- Martin EIP, Ressler KJMDP, Binder EMDP, Nemeroff CBMDP. The neurobiology of anxiety disorders: brain imaging, genetics, and psychoneuroendocrinology. *Psychiatr Clin North Am*. 2009;32(3):549–75.
- Cummings CM, Caporino NE, Kendall PC. Comorbidity of anxiety and depression in children and adolescents: 20 years after. *Psychol Bull*. 2014;140(3):816–45.
- Chireh B, Essien SK, D'Arcy C, Gyan C, Ankrah M. Severe disability and self-reported depression and anxiety among persons living with Type 2 diabetes in Canada. *J Affect Disord Rep*. 2024;16:100720.
- Hohls JK, König H-H, Quirke E, Hajek A. Anxiety, depression and quality of life—a systematic review of evidence from longitudinal observational studies. *Int J Environ Res Public Health*. 2021;18(22):12022.
- Patterson SL, Marcus M, Goetz M, Vaccarino V, Gooding HC. Depression and anxiety are associated with cardiovascular health in young adults. *J Am Heart Assoc*. 2022;11(24):e027610.
- Zhang J, Liu X, Fang L. Combined effects of depression and anxiety on suicide: a case-control psychological autopsy study in rural China. *Psychiatry Res*. 2019;271:370–3.
- Garber J, Brunwasser SM, Zerr AA, Schwartz KTG, Sova K, Weersing VR. Treatment and prevention of depression and anxiety in youth: test of cross-over effects: 2015 ADA scientific research symposium: interventions for depression and anxiety. *Depress Anxiety*. 2016;33(10):939–59.
- Alonso J, Liu Z, Evans-Lacko S, Sadikova E, Sampson N, Chatterji S, Abdullmalik J, Aguilar-Gaxiola S, Al-Hamzawi A, Andrade LH, et al. Treatment gap for anxiety disorders is global: results of the world mental health surveys in 21 countries. *Depress Anxiety*. 2018;35(3):195–208.
- Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J. An inventory for measuring depression. *Arch Gen Psychiatry*. 1961;4(6):561–71.
- Hamilton M. A rating scale for depression. *J Neurol Neurosurg Psychiatry*. 1960;23(1):56–62.
- Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. *Br J Psychiatry*. 1979;134:382.
- Obeid S, Haddad C, Sacre H, Akel M, Chalhoub C, Rahme C, Abi Doumit C, Akiki M, Hallit S, Salameh P. The Lebanese depression scale: a scale to screen for depression in the Lebanese population. *Perspect Psychiatr Care*. 2021;57(2):620–6.
- Kroenke K, Spitzer RL, Williams JBW. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med : JGIM*. 2001;16(9):606–13.
- Kroenke K, Spitzer RL, Janet BWW. The patient health questionnaire-2: validity of a two-item depression screener. *Med Care*. 2003;41(11):1284–92.
- Beck AT, Epstein N, Brown G, Steer RA. An inventory for measuring clinical anxiety: psychometric properties. *J Consult Clin Psychol*. 1988;56(6):893–7.
- Hallit S, Obeid S, Haddad C, Hallit R, Akel M, Haddad G, Soufia M, Khansa W, Khoury R, Kheir N, et al. Construction of the Lebanese Anxiety Scale (LAS-10): a new scale to assess anxiety in adult patients. *Int J Psychiatry Clin Pract*. 2020;24(3):270–7.
- Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med* (1960). 2006;166(10):1092–7.
- Kroenke K, Spitzer RL, Williams JBW, Monahan PO, Löwe B. Anxiety disorders in primary care: prevalence, impairment, comorbidity, and detection. *Ann Intern Med*. 2007;146(5):317–25.
- Kroenke KMD, Spitzer RLMD, Williams JBWDSW, Löwe BMDPD. An ultra-brief screening scale for anxiety and depression: The PHQ-4. *Psychosomatics* (Washington, DC). 2009;50(6):613–21.
- Löwe B, Wahl I, Rose M, Spitzer C, Glaesmer H, Wingenfeld K, Schneider A, Brähler E. A 4-item measure of depression and anxiety: validation and standardization of the Patient Health Questionnaire-4 (PHQ-4) in the general population. *J Affect Disord*. 2010;122(1–2):86–95.
- Choi SK, Boyle E, Burchell AN, Gardner S, Collins E, Grootendorst P, Rourke SB; OHTN Cohort Study Group. Validation of Six Short and Ultra-short Screening Instruments for Depression for People Living with HIV in Ontario: Results from the Ontario HIV Treatment Network Cohort Study. *PLoS One*. 2015;10(11):e0142706. <https://doi.org/10.1371/journal.pone.0142706>.
- Trapp W, Röder S, Heid A, Billman P, Daiber S, Hajak G. Sensitivity and specificity of the Bamberg Dementia Screening Test's (BDST) full and short versions: brief screening instruments for geriatric patients that are suitable for infectious environments. *BMC Med*. 2021;19(1):65. <https://doi.org/10.1186/s12916-021-01927-4>.
- Mitchell AJ, Coyne JC. Do ultra-short screening instruments accurately detect depression in primary care? A pooled analysis and meta-analysis of 22 studies. *Br J Gen Pract*. 2007;57(535):144–51.
- Khubchandani J, Brey R, Kotecki J, Kleinfelder J, Anderson J. The psychometric properties of PHQ-4 depression and anxiety screening scale among college students. *Arch Psychiatr Nurs*. 2016;30(4):457–62.
- Rodríguez-Muñoz MdIF, Ruiz-Segovia N, Soto-Balbuena C, Le H-N, Olivares-Crespo ME, Izquierdo-Méndez N. The psychometric properties of the patient health questionnaire-4 for pregnant women. *Int J Environ Res Public Health*. 2020;17(20):7583.
- Havnen A, Lydersen S, Mandahl A, Lara-Cabrera ML. Factor structure of the patient health questionnaire-4 in adults with attention-deficit/hyperactivity disorder. *Front Psych*. 2023;14:1176298.
- Mills SD, Fox RS, Pan TM, Malcarne VL, Roesch SC, Sadler GR. Psychometric evaluation of the patient health questionnaire-4 in Hispanic Americans. *Hisp J Behav Sci*. 2015;37(4):560–71.
- Soares E, Leal G, Belchior C, Mendes S, Valente AR. Tradução, adaptação e validação do conteúdo do instrumento de avaliação preschooler awareness of stuttering survey para o português europeu. *Rev Portuguesa Educ*. 2023;36(1):e23025.
- Christodoulaki A, Baralou V, Konstantakopoulos G, Touloumi G. Validation of the Patient Health Questionnaire-4 (PHQ-4) to screen for depression and anxiety in the Greek general population. *J Psychosom Res*. 2022;160:110970.
- Kim H-W, Shin C, Lee S-H, Han C. Standardization of the Korean Version of the Patient Health Questionnaire-4 (PHQ-4). *Clinical psychopharmacology and neuroscience : the official scientific journal of the Korean College of Neuropsychopharmacology*. 2021;19(1):104–11.
- Kocalevent R-D, Finck C, Jimenez-Leal W, Sautier L, Hinz A. Standardization of the Colombian version of the PHQ-4 in the general population. *BMC Psychiatry*. 2014;14(1):205.
- Kazlauskas E, Gelezelyte O, Kvedaraite M, Ajdukovic D, Johannesson KB, Böttche M, Bondjers K, Dragan M, Figueiredo-Braga M, Grajewski P, et al. Psychometric properties of the Patient Health Questionnaire-4 (PHQ-4) in 9230 adults across seven European countries: findings from the ESTSS ADJUST study. *J Affect Disord*. 2023;335:18–23.
- Caro-Fuentes S, Sanabria-Mazo JP. A systematic review of the psychometric properties of the patient health questionnaire-4 in clinical and non-clinical populations. *J Acad Consult-Liaison Psychiatry*. 2023;65:178–94.
- Kliem S, Mößle T, Klatt T, Fleischer S, Kudlacek D, Kröger C, Brähler E, Beutel ME, Wiltink J. Psychometric evaluation of an Arabic version of the PHQ-4 based on a representative survey of Syrian refugees. *Psychother Psychosom Med Psychol*. 2016;66(9–10):385–92.



40. Karam EG, Mneimneh ZN, Karam AN, Fayyad JA, Nasser SC, Chatterji S, Kessler RC. Prevalence and treatment of mental disorders in Lebanon: a national epidemiological survey. *Lancet (British edition)*. 2006;367(9515):1000.
41. Farran N. Mental health in Lebanon: tomorrow's silent epidemic. *Mental Health Prev*. 2021;24:200218.
42. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine (Philadelphia, Pa 1976)*. 2000;25(24):3186–91.
43. Fenn J, Tan C-S, George S. Development, validation and translation of psychological tests. *BJPsych Adv*. 2020;26(5):306–15.
44. Ambuehl B, Inauen J. Contextualized measurement scale adaptation: a 4-step tutorial for health psychology research. *Int J Environ Res Public Health*. 2022;19:12775.
45. Ali AM, Hori H, Kim Y, Kunugi H. The depression anxiety stress scale 8-items expresses robust psychometric properties as an ideal shorter version of the depression anxiety stress scale 21 among healthy respondents from three continents. *Front Psychol*. 2022;13:799769.
46. Fekih-Romdhane F, Cherif W, Alhuwailah A, Fawaz M, Shuwiekh HAM, Helmy M, Hassan IHM, Naser AY, Zarrouq B, Chebly M. Cross-country validation of the Arabic version of the WHO-5 well-being index in non-clinical young adults from six Arab countries. 2023. <https://doi.org/10.21203/rs.3.rs-2988215/v1>.
47. Topp CW, Østergaard SD, Søndergaard S, Bech P. The WHO-5 well-being index: a systematic review of the literature. *Psychother Psychosom*. 2015;84(3):167–76.
48. Mundfrom DJ, Shaw DG, Ke TL. Minimum sample size recommendations for conducting factor analyses. *Int J Test*. 2005;5(2):159–68.
49. Hu Lt, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equ Model*. 1999;6(1):1–55.
50. Chen FF. Sensitivity of goodness of fit indexes to lack of measurement invariance. *Struct Equ Model*. 2007;14(3):464–504.
51. Vadenberg R, Lance C. A review and synthesis of the measurement invariance literature: suggestions, practices, and recommendations for organizational research. *Organ Res Methods*. 2000;3:4–70.
52. Dunn TJ, Baguley T, Brunsden V. From alpha to omega: a practical solution to the pervasive problem of internal consistency estimation. *Br J Psychol*. 2014;105(3):399–412.
53. Hair Jr JF, Sarstedt M, Ringle CM, Gudergan SP. *Advanced issues in partial least squares structural equation modeling*. SAGE publications; 2017.
54. Cohen J. Quantitative methods in psychology: A power primer. In: *Psychological bulletin*: 1992;112(1):155–9.
55. Olfson M, Shea S, Feder A, Fuentes M, Nomura Y, Gameroff M, Weissman MM. Prevalence of anxiety, depression, and substance use disorders in an urban general medicine practice. *Arch Fam Med*. 2000;9(9):876–83. <https://doi.org/10.1001/archfam.9.9.876>.
56. Schonfeld WH, Verboncoeur CJ, Fifer SK, Lipschutz RC, Lubeck DP, Bue-sching DP. The functioning and well-being of patients with unrecognized anxiety disorders and major depressive disorder. *J Affect Disord*. 1997;43(2):105–19.
57. Cano-Vindel A, Muñoz-Navarro R, Medrano LA, Ruiz-Rodríguez P, González-Blanch C, Gómez-Castillo MD, Capafons A, Chacón F, Santolaya F, Psic APRG. A computerized version of the patient health questionnaire-4 as an ultra-brief screening tool to detect emotional disorders in primary care. *J Affect Disord*. 2018;234:247–55.
58. Ghaheri A, Omani-Samani R, Sepidarkish M, Hosseini M, Maroufizadeh S. The four-item patient health questionnaire for anxiety and depression: a validation study in infertile patients. *Int J Fertil Steril*. 2020;14(3):234–9.
59. Carranza Esteban RF, Caycho Rodríguez T, Mamani Benito OJ, Cjuno J, Vilca Quiro LW, Barrios I, Tito Betancur M, Torales J. Adaptation and validation of the Patient Health Questionnaire for depression and anxiety (PHQ-4) in a sample of Quechua-speaking Peruvians. *Med Clin Soc*. 2024;8(1):63–74.
60. Malone C, Wachholtz A. The relationship of anxiety and depression to subjective well-being in a Mainland Chinese sample. *J Relig Health*. 2018;57(1):266–78.
61. Stein MB, Heimberg RG. Well-being and life satisfaction in generalized anxiety disorder: comparison to major depressive disorder in a community sample. *J Affect Disord*. 2004;79(1):161–6.
62. Postler KB, Helms HM, Anastopoulos AD. Examining the linkages between marital quality and anxiety: a meta-analytic review. *Fam Process*. 2022;61(4):1456–72.
63. Huh HJ, Kim S-Y, Yu JJ, Chae J-H. Childhood trauma and adult interpersonal relationship problems in patients with depression and anxiety disorders. *Ann Gen Psychiatry*. 2014;13(1):26.
64. Deady M, Collins D, Johnston D, Glozier N, Calvo R, Christensen H, Harvey S. The impact of depression, anxiety and comorbidity on occupational outcomes. *Occup Med*. 2022;72(1):17–24.
65. Gabarrell-Pascuet A, García-Mieres H, Giné-Vázquez I, Moneta MV, Koy-anagi A, Haro JM, Domènech-Abella J. The association of social support and loneliness with symptoms of depression, anxiety, and posttraumatic stress during the COVID-19 pandemic: a meta-analysis. *Int J Environ Res Public Health*. 2023;20(4):2765.
66. Tiller JW. Depression and anxiety. *Med J Aust*. 2013;199(6):S28–31.
67. Mendoza NB, Frondoza CE, Dizon JIWT, Buenconsejo JU. 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. *Curr Psychol*. 2022:1–10. <https://doi.org/10.1007/s12144-022-02833-5>.
68. Bueno-Notivol J, Gracia-García P, Olaya B, Lasheras I, López-Antón R, Santabárbara J. Prevalence of depression during the COVID-19 outbreak: a meta-analysis of community-based studies. *Int J Clin Health Psychol*. 2021;21(1):100196.
69. Salari N, Hosseini-Far A, Jalali R, Vaisi-Raygani A, Rasoulpoor S, Mohammadi M, Rasoulpoor S, Khaledi-Paveh B. Prevalence of stress, anxiety, depression among the general population during the COVID-19 pandemic: a systematic review and meta-analysis. *Global Health*. 2020;16:1–11.
70. Wicke FS, Krakau L, Löwe B, Beutel ME, Brähler E. Update of the standardization of the Patient Health Questionnaire-4 (PHQ-4) in the general population. *J Affective Disord*. 2022;312:310–4.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.