

Concurrent Validity of the PHQ-9 in English and Spanish

Robert S. Smith, DHSc, PA-C, DFAAPA; Helen Salisbury, MA, MBBS, MRCP, MRCGP

Corresponding Author

Robert S. Smith, DHSc, PA-C, DFAAPA; rsspac@aol.com

ABSTRACT:

Depression in the United States remains undertreated because of poor screening, resulting in half of the adult patients not diagnosed with major depression. There are barriers that lead to poor screening for depression, including language, availability of appropriate devices, and ease of use of a validated instrument for screening. The Patient Health Assessment Questionnaire – 9 (PHQ-9) is recommended for screening in primary care; it has been found to be effective and easy-to-use screening device and validated for this purpose, but only in English.

Purpose: A quantitative correlational study to demonstrate the concurrent validity of the Spanish PHQ-9 with the English PHQ-9 in screening for depression in Spanish-speaking patients.

Method: The study included a cohort of 10 primary care clinics in Dallas/Fort Worth, TX, which during a 6-week period collected data from 397 participants who were bilingual (read and write) in both English and Spanish. This was a quantitative correlational study using Intraclass Correlation Coefficient (ICC) analysis for this study; the researchers examined the concurrent validity of the Spanish version of the PHQ-9 for screening for depression as compared to the English version of the PHQ-9 item by item (paired items), comparing total scores and provisional diagnosis scores.

Results: The Spanish version of the PHQ-9 compared to the English version of the PHQ-9 had an ICC [$>.9$]. For each paired item an ICC [.970], for paired items total scores an ICC [.992], and for paired items provisional diagnosis scores of depression [.981].

Conclusions: The results demonstrated the translated psychometrics of the English PHQ-9 into Spanish are retained and the use of the Spanish version of the PHQ-9 is a valid instrument for assessing depressive symptoms in patients.

Keywords: Depression screening, PHQ-9 English and Spanish, PHQ-9 translations

1. Introduction

Depression in the United States remains undertreated in primary care. A recent study examining trends in U.S depression prevalence from 2015-2020 identified widespread increases in depression without treatment describing a need for a comprehensive plan for screening and treating the ongoing mental health crisis.¹ A literature review by Williams and Nieuwsma, supported the lack of care because of poor screening.² They found 50% of adult patients with major depression were not identified. Untreated depression leads to increased mortality and reduced quality of life, adversely affecting families, the workplace, and society as a whole.¹⁻² The United States Preventive Services Task Force (USPSTF) has evidence that depression screening can be effective in primary care for identifying and treating patients in all segments of society.² Additionally the USPSTF stated early detection and treatment improves clinical outcomes in at-risk populations with screening and providing a comprehensive treatment program that includes antidepressant therapy, psychotherapy, or a combination of both.¹⁻² In September 2013, the Institute for Clinical Systems Improvement (ICSI) published the 16th edition of the Adult Depression in Primary Care Guidelines, which incorporated an algorithm for screening patients in primary care adapted from the American Psychiatric Association (APA) *DSM-5* recommended a systematic screening for patients suspected of having depression.⁴ The recommended screening tools were the Patient Health Questionnaires (PHQ-2 and PHQ-9).⁵ This study examined the English PHQ-9, the use and validity of this instrument; appropriate translation into Spanish and the concurrent validity of the Spanish version of the PHQ-9 for screening for depression in Spanish.

2. Background

Approximately 61 million people in the United States (21% of the U.S. population) speak a language other than English in their homes.⁶ Language barriers contribute to poor care, increased costs, and inherently, a potential for injury to patients.⁷⁻⁸ There are numerous barriers which lead to poor screening or no screening for depression, including (a) the effects of language, (b) the availability of appropriate screening tools, (c) the ease of use, and (d) provider expertise.⁴ Additional barriers include availability of screening tools in other languages, appropriate translations, and concurrent validity for use. The Patient Health Questionnaire (PHQ) was developed in 1999 as an easy-to-use self-assessment instrument to aid in the diagnosis of mental disorders in primary care.⁹ It differs from other instruments in that it was created to address the essentials of *DSM-5* criteria for screening for depression and was found to be valid for screening, monitoring treatment, and for establishing remission.⁵

Rationale

Language disparities exist in how patients interpret a written document, or how verbal statements are understood. The concern revolves around the subject of the questions asked and how they are answered by a patient.⁷⁻⁸ Patients must understand the context in which a question is being asked, and the responses made should be answered in that context.⁷⁻⁸ The PHQ-9 has a long history of use and validity, and most notably the most commonly used screening tools in primary care.³⁻⁴ The PHQ-9 has been translated into multiple languages by MAPI Research Trust, a leading linguistic organization using the protocol created by Beaton, Bombardier, Guillemin, and Ferraz.¹⁰ The presumption that a validated screening instrument will retain its psychometric properties when linguistically translated appropriately has not been demonstrated.¹¹

Conceptual framework

There are studies that have used the PHQ-9 in languages other than English to examine the reliability and validity of the PHQ-9; these include Chinese, Haitian Creole, Arabic, and others in Spanish.¹²⁻¹⁴ Examples of studies in Spanish include a cohort study in Peru using a Rasch Item Response Theory Analysis of the PHQ-9 with similar findings when compared with other similar groups.^{3,15} A correlational study completed in Chile with the specific objective of validating the PHQ-9 in Spanish, demonstrated validity in identifying depression when compared to those of its English counterpart.¹⁶ An Ecuadorian public health study, using a validity survey instrument that included sociodemographic data, the PHQ-9, and other standards of measurement such as the Beck Depression Inventory-II (BDI-II), the Generalized Anxiety Disorder Scale (GAD-7), the Positive and Negative Affect Scale (PANAS), and the Quality-of-Life Index (QLI) for evaluating for validity of the psychometric properties and concluded that the PHQ-9 could aid screening for depression.¹⁷ In Argentina a cross-sectional study aimed to validate and corroborate the PHQ-9 to determine appropriate cut-off points for varying degrees of the severity of depression using the PHQ-9, additional correlation with interviews with trained clinicians using the Mini International Neuropsychiatric Interview (MINI) and the Beck Depression Inventory-II (BDI-II) to validate the scores of the PHQ-9. The outcomes demonstrated acceptable validity and reliability for both screening and assessing the severity of depressive symptoms.¹⁸ There are no previous studies that have examined the concurrent validity of the PHQ-9 in English to Spanish with participants who are bilingual (able to read and write) and who have completed both versions of the PHQ-9 sequentially.

Significance

The significance of the proposed study is to bridge the gap in the literature by validating the use of the PHQ-9 in Spanish for screening for depression. A validated instrument in Spanish benefits both clinicians and patients in identifying depression and improves the morbidity and mortality of depression in a Spanish speaking population. Currently, there are limited validity studies for the use of the PHQ-9 in Spanish for screening for depression and the studies completed have demonstrated low evidence, because of limitations.¹⁻⁴ There is good evidence for proper translation of the PHQ-9, however questions remain as to the retention of its psychometric properties and if validity is maintained when compared to the English version of the PHQ-9. The purpose of

this quantitative correlational study was to further demonstrate the concurrent validity of the Spanish version of the PHQ-9 with the English version in screening for depression in Spanish-speaking patients. The outcome of this research is to have a validated instrument for screening for depression in the Spanish-speaking community and to create a methodology of validation that can be replicated for use other languages.

3. Research Question(s) and Hypotheses

RQ₁: What is the concurrent validity of the English and Spanish language versions of the PHQ-9 depression screening tool?

$H_0: < .8$

$H_1: > .8$

4. Methods

Research Design

The study was quantitative, applying a correlational design, examining the concurrent validity of the diagnostic cut-offs both in the English and Spanish PHQ-9 questionnaires.

Study Participants

A cohort of ten clinics located in the Dallas/Fort Worth, TX, were recruited to collect the data sample. The Medical Director of the organization approved permission to perform the research project. The participant pool came from patients 18 years or older who presented to the clinic for medical services were asked to participate and based on participant response and bilingual status resulted in either inclusion or exclusion from the research study. There were 411 participants who completed the PHQ-9 in both English and Spanish, of those 397 participants met the inclusion criteria.

Inclusion criteria. The inclusion criteria included bilingual patients who were (a) 18 years or older, (b) able to read and write in Spanish, and (c) were able to complete both versions of the PHQ-9 sequentially.

Exclusion criteria. The exclusion criteria: if the patient was (a) *not* 18 years or older or is (b) *unable* to read and write in both English and Spanish, or (c) declined to participate.

Sampling methodology.

The collection of data was accomplished using a consecutive sampling methodology. Each patient who presented for clinical services and was 18 years or older and bilingual Spanish/English (read & write) was asked to participate in the research project. Those who met the inclusion criteria were asked to complete a demographics questionnaire and the paper versions of the PHQ-9 in both English and Spanish, sequentially – random order English – Spanish or Spanish – English (see Appendices A - D).

Institutional Review Board approval.

Approval for this study was granted by the Institutional Review Board of A. T. Still University. No data was collected prior to approval. A waiver of consent was given; no changes to the study have been made.

Survey Instrument

The criterion was the English version of the Patient Health Questionnaire -9 (PHQ-9) which is considered the gold standard of measurement for screening for depression and has been validated and confirmed for use in English.^{5,8,19} The PHQ-9 has been translated in over 55 languages, but has not been validated in any language other than English.²⁰ The Spanish version of the PHQ-9 is formatted in the same manner as the English. For this study, the translated Spanish version by MAPI has been linguistically validated was used and correlated to the

English version.^{10,21} Both versions of the PHQ-9 incorporate the Likert scale. There are 9 questions and optional 10th question with corresponding columns labeled: 0 - *not all days*, 1- *several days*, 2- *more than half the days* and 3 - *nearly every day*. Each column scores were summed individually, and the subtotals are summed for a total score that provided a provisional diagnosis of depression that was scored: 0-4 none – *minimal depression*, 5-9 *mild depression*, 10-14 *moderate depression*, 15-19 *moderately severe depression*, and 20-27 *severe depression*.

Data Collection

The collection of data took place over 6 weeks at 10 clinics. The front office medical office assistant(s) (MOA) were trained in the study protocol (Appendix I). The MOA at a front desk / reception area screened for participants (18 years and older; bilingual able to read and write in both English and Spanish), upon entering the clinic for an appointment. If a patient met the inclusion criteria and agreed to participate, the person was given a survey to complete; demographics form, and both versions of PHQ-9 (randomly ordered) sequentially. Once completed, the patient returned (no identifiers) the forms in a provided envelope and placed into a secured drop box at the front desk. This method of data collection ensured a more diverse population completing the forms to help with validation and ensured anonymity which removed the need for informed consent. The completed forms were collected from the clinics at regular intervals by the principal investigator.

Statistical Analysis

Data analysis was performed using IBM SPSS Statistic Version 24. Demographic variables collected were primarily using a questionnaire that was completed by the participants which included sex, country of origin, and generation of arrival to United States, highest education level, income status, current work status, insurance status, and marital status, apart from the level variable age. Frequencies were completed on the variables with valid percent's reported to demonstrate the number of occurrences and to identify any missing data. Normality testing was conducted using Shapiro-Wilk analysis, ($p>.05$) for the interval of age. Each question individually yielded data which is scored from 0-3 and scored accordingly. The columns are summed, and a total score achieved to establish a provisional diagnosis score. Intraclass correlations between both versions of the PHQ-9 were completed (English Item 1 to Spanish Item1, [paired item]) and total scores. A variable was created to identify the provisional depression diagnostic category ranges based on the total scores for each PHQ-9 (English and Spanish) that was completed by each participant. The two variables labeled English provisional diagnosis and Spanish provisional diagnosis were analyzed using frequencies to illustrate the valid percentage of the provisional diagnosis of depression levels. Additionally, Intraclass correlation was completed on the paired items to demonstrate concurrent validity. The data are displayed in tables.

5. Results

Characteristics of the Sample

There were 411 participants who completed the survey instrument. Of these, 14 were excluded because of age, leaving 397 participants who met the inclusion criteria of bilingual (able to read and write) both in English and Spanish, and consented to participate in the research. Normality testing was conducted on the variable, *Age*, via the Shapiro-Wilk test, $p=.0001$. The median age was 30 years (IRQ =17) patients ranged in age from 18 to 86 years old. The majority of participants were female ($n=238$; 64.2%). Most of the participants were from the United States ($n=183$; 50%), and Mexico ($n=133$; 36%) (see Table 1).

The largest groups of participants were first generation ($n=138$; 40.6%) as compared to participants who have been in the United States three generations or longer ($n=89$; 27.5%). The educational level of the participants varied from 6 years of education or less ($n=36$; 9.7%), with the largest group having a secondary education (7-12 years) ($n=139$; 37.1%). A majority of the participants reported working either full-time or part-time ($n=275$; 76.6%). Most participants reported an average income of \$10,000 - \$49,000 ($n=174$; 47.5%). The greatest

number of participants had either commercial health insurance ($n=145$; 40%) or were self-pay ($n=138$; 38.3%). Most of the respondents reported been married ($n=181$; 47.9%).

6. Findings

Research question.

The question of concurrent validity of the English and Spanish language versions of the PHQ-9 depression screening tools when compared using Intraclass correlation (ICC) item by item (item pair) for both versions of the PHQ-9, demonstrated that the English to Spanish versions of the PHQ-9 does correlate, with an ICC [$>.9$] in every item pair (see Table 2). Additionally, ICC has demonstrated a high degree of confidence that the Spanish version of the PHQ-9 is equally valid to the English version for use in screening for depression in Spanish. The results reject the null Hypothesis ICC [$<.8$], there is insufficient evidence to illustrate no concurrent validity between the English version of the PHQ-9 and the Spanish version of the PHQ-9 for screening for depression (see Table 2). Additionally, the results have demonstrated ICC [$>.8$], supporting the alternative hypothesis, that there is concurrent validity of the English version of the PHQ-9 and the Spanish version of the PHQ-9 for screening for depression (see Table 2). Further analysis using ICC for examining the total scores of depression in the English and Spanish versions were also shown to highly correlate with an ICC of [$.992$] and CI $>.991$.

An additional analysis was performed to compare the frequencies of the provisional diagnoses of depression for both the English and Spanish versions. Both versions of the PHQ-9 demonstrated a provisional diagnosis score of minimal depression (0-4) of 55.7%. Those with a severe depression score (20-27) slightly varied English 2.5% and Spanish 2.3%. Additional analysis of the provisional diagnosis paired items demonstrated a high degree of correlation with an ICC of [$.981$] and a CI $>.977$ (see Table 3). The findings demonstrate a retention of the translated psychometrics and concurrent validity of the Spanish PHQ-9 to the English PHQ-9 and therefore valid for use in screening for depression in Spanish.

7. Discussion

The results of this study demonstrated concurrent validity of the PHQ-9 in English and Spanish using ICC. Every item pair was correlated in both versions of the PHQ-9 that demonstrated that the translated psychometrics was retained from English to Spanish with high ICC correlation rates between both versions of the PHQ-9, with minimal divergence in the ICC scores. These findings support previous research which examined the validity of the Spanish version of the PHQ-9, using a one language methodology.^{15,21,22}

All of these studies were considered low evidence because of study limitations.²⁻⁴ There have been several multi-lingual studies completed - Chinese, Haitian - Creole, Arabic and Spanish and there still remains a gap in the validity of PHQ-9 in any other language but English.¹²⁻¹⁵ A meta-analysis has been completed examining the use of the PHQ-9 with analysis and review of 27 validation studies, in numerous languages including Spanish.²⁴⁻²⁵ The two major studies in Spanish were both shown to be appropriately translated, however, there was only one study which had reported translated psychometric properties although neither studies was considered as strong evidence of validity.^{11,24-25}

A presumption has been made that the translation process provides retention of the psychometric attributes analogous to validity and reliability has been shown not to be true.¹¹ Although a translated statement or word may be correct linguistically, it may not be interpreted by the examinee in the same context as when asked, thereby making the assessment tool invalid.²¹ A standardized process has been created for how a translation of a self-assessment questionnaire is completed from its original version to a new language while achieving linguistic validity, maintaining cultural relevance, and assuring its concurrent validity for use.^{10,21} The Spanish version of the PHQ-9 translated by MAPI meeting the standards established by Beaton et al. and Ramada-Rodilla was used for this study.^{10,21} The participants in this research were diverse in country of origin providing strength

to the translation of the PHQ-9 in the retention of the translated psychometrics as the same in both English and Spanish for validation and use.

Implications of Findings

The implications of this research demonstrated validity for using the Spanish version of the PHQ-9 in screening for depression. The clinical benefits include having an additional validated instrument, removes a language disparity and primary care providers can have confidence in using the Spanish language PHQ-9 as a screening device for depression knowing the psychometric properties of the original English version are retained. Additionally, the methodology used in this research can be replicated for use in validating other language translations of the PHQ-9 to improve screening of patients for depression in other languages.

Limitations of the Study

There were a few limitations noted in the research process. There was a larger sample size anticipated with the 10 clinic cohort, the smaller size could be related to the initial numbers projected because of the original definition of bilingual; speaking both languages, rather than reading and writing in both languages, this would account for the smaller sample size. A limitation may have been not defining a champion at each cohort clinic for collecting data, at the reception desk rather than as a group effort. Additionally, complacency may have played a role in collecting data; over time there tended to be less enthusiasm for collecting data. A strategy to mitigate this from occurring might be to create a competition or incentive to increase the motivation for collecting data. Additional emphasis during the training highlighting the importance of data collection over an extended period might make the task a higher priority for the individuals involved. A more aggressive style in the management of the collection process such as calling, texting, and encouragement to those assisting in the collection process may have increased the number of participants. Finally, completing the research using a paper process was cumbersome considering the age of technology present; using electronic tablets or similar devices would have expedited the collection process, as well as speeding up the scoring and data entry portions of the research.

Recommendations for Future Research

The next step in the process would be to validate the PHQ-9 in Spanish for monitoring of treatment of depression and then for remission of depression. The English version PHQ-9 has been shown to be valid in both monitoring the treatment and remission of depression.^{5,26-27} Currently, no studies have been validated in any other languages. Further research should include validating the PHQ-9 for screening for depression in adolescents (12-18 years) in Spanish and other languages as well.²⁸

8. Conclusions

This research was to determine if the translated PHQ-9 in Spanish was concurrently valid to the PHQ-9 in English for screening depression, and to demonstrate that the translated psychometrics were equal linguistically and understood by the patient to be the same as the English version of PHQ-9, in screening for depression in Spanish. The findings demonstrated a high Intraclass Correlation Coefficient between each item pair for both versions of the PHQ-9 English and Spanish, bridging the gap in the literature and supporting the findings of previous research. This research was the first of its kind using bilingual (read and write both languages) participants completing both versions of the PHQ-9 sequentially allowing for direct analysis using ICC item by item leading to these findings. The results demonstrated that the translated psychometrics of the English PHQ-9 into Spanish is retained and the use of the Spanish version of the PHQ-9 is a valid instrument for assessing depressive symptoms in patients.

Future research on use of the PHQ-9 in Spanish for monitoring treatment and for remission along with validating the PHQ-9 other languages is warranted.

Disclaimer: The views expressed in this paper are those of the authors and not necessarily of the authors' organizations or the National Hispanic Medical Association (NHMA). The paper is intended to help inform and stimulate discussion. It is not a report of NHMA. Copyright by the National Hispanic Medical Association. All rights reserved.

References

1. Goodwin RD, Dierker LC, Wu M, Galea S, Hoven CW, Weinberger AH. Trends in U.S. Depression Prevalence from 2015 to 2020: The Widening Treatment Gap. *AM J Prev Med.* 2022;63(5):726-733, doi:10.1016/j.amepre.2022.05.014
2. Williams J, Nieuwsma J. Screening for depression. Up to date. Published 2015. <http://www.uptodate.com/contents/screening-for-depression-in-adults>
3. Siu AL, Bibbins-Domingo K, Grossman DC, et al. Screening for Depression in Adults. *JAMA.* 2016;315(4):380. doi:https://doi.org/10.1001/jama.2015.18392
4. Mitchell J, Trangle M, Degnan B, et al. Adult Depression in Primary Care. Published 2013. https://www.icsi.org/_asset/fnhdm3/Depr-Interactive0512b.pdf
5. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. 5th ed. *American Psychiatric Publishing*; 2013.
6. Ryan, C. Language use in the United States: 2011. American Community Survey Reports, US Census Bureau, August 16. <https://www.census.gov/prod/2013pubs/acs-22.pdf>
7. Nápoles AM, Santoyo-Olsson J, Karliner LS, Gregorich SE, Pérez-Stable EJ. Inaccurate Language Interpretation and its Clinical Significance in the Medical Encounters of Spanish-speaking Latinos. *Medical care.* 2015;53(11):940-947. doi:https://doi.org/10.1097/MLR.0000000000000422
8. Parés-Avila JA, Sobralske MC, Katz JR. No Comprendo: Practice Considerations When Caring for Latinos with Limited English Proficiency in the United States Health Care System. *Hispanic Health Care International.* 2011;9(4):159-167. doi:https://doi.org/10.1891/1540-4153.9.4.159
9. Spitzer RL. Validation and utility of a self-report version of PRIME-MD: The PHQ primary care study. *JAMA.* 1999;282(18):1737. doi:https://doi.org/10.1001/jama.282.18.1737
10. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine.* 2000;25(24):3186-3191. doi:https://doi.org/10.1097/00007632-200012150-00014
11. Manea L, Gilbody S, McMillan D. A diagnostic meta-analysis of the Patient Health Questionnaire-9 (PHQ-9) algorithm scoring method as a screen for depression. *General Hospital Psychiatry.* 2015;37(1):67-75. doi:https://doi.org/10.1016/j.genhosppsych.2014.09.009
12. Wang W, Bian Q, Zhao Y, et al. Reliability and validity of the Chinese version of the Patient Health Questionnaire (PHQ-9) in the general population. *General Hospital Psychiatry.* 2014;36(5):539-544. doi:https://doi.org/10.1016/j.genhosppsych.2014.05.021
13. Marc LG, Henderson WR, Desrosiers A, Testa MA, Jean SE, Akom EE. Reliability and Validity of the Haitian Creole PHQ-9. *Journal of General Internal Medicine.* 2014;29(12):1679-1686. doi:https://doi.org/10.1007/s11606-014-2951-5
14. Al-Ghafri G, Al-Sinawi H, Al-Muniri A, et al. Prevalence of depressive symptoms as elicited by Patient Health Questionnaire (PHQ-9) among medical trainees in Oman. *Asian Journal of Psychiatry.* 2014;8(8):59-62. doi:https://doi.org/10.1016/j.ajp.2013.10.014
15. Zhong Q, Gelaye B, Fann J R, et al. Cross-cultural validity of the Spanish version of PHQ-9 among pregnant Peruvian women: a Rasch item response theory analysis. *Journal of Affective Disorders,* 2014;158:148–153. <https://doi.org/10.1016/j.jad.2014.02.012>
16. Baader M T, Molina F JL, Venezian B S, et al. 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.* 2012;50(1):10-22. doi:https://doi.org/10.4067/S0717-92272012000100002
17. Quiñonez-Freire C, Vara D, Tomás JM, Baños R. Psychometric properties of the Spanish version of the Patient Health Questionnaire-9 in users of the Ecuadorian public health care system. *Revista Latinoamericana de Psicología,* (2021);53: 210-217. <https://doi.org/10.14349/rlp.2021.v53.23>
18. Urtasun M, Daray FM, Teti GL, et al. Validation and calibration of the patient health questionnaire (PHQ-9) in Argentina. *BMC Psychiatry.* 2019;19(1). doi:https://doi.org/10.1186/s12888-019-2262-9
19. Deneke DE, Schultz H, Fluent TE. Screening for Depression in the Primary Care Population. *Primary Care: Clinics in Office Practice.* 2014;41(2):399-420. doi:https://doi.org/10.1016/j.pop.2014.02.011
20. Pfizer. Patient Health Questionnaire (PHQ) Screeners. Free Download. www.phqscreeners.com. Accessed March 3, 2024. <http://www.phqscreeners.com/select-screener>

21. Ramada-Rodilla JM, Serra-Pujadas C, Delclós-Clanchet GL. Adaptación cultural y validación de cuestionarios de salud: revisión y recomendaciones metodológicas. *Salud Pública de México*. 2013;55(1):57-66. doi:<https://doi.org/10.1590/s0036-36342013000100009>
22. Granillo MT. 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*. 2012;3(2):80-93. doi:<https://doi.org/10.5243/jsswr.2012.6>
23. Merz EL, Malcarne VL, Roesch SC, Riley N, Sadler GR. A multigroup confirmatory factor analysis of the Patient Health Questionnaire-9 among English- and Spanish-speaking Latinas. *Cultural Diversity and Ethnic Minority Psychology*. 2011;17(3):309-316. doi:<https://doi.org/10.1037/a0023883>
24. Díez-Quevedo C, Rangil T, Sanchez-Planell L, Kroenke K, Spitzer RL. Validation and Utility of the Patient Health Questionnaire in Diagnosing Mental Disorders in 1003 General Hospital Spanish Inpatients. *Psychosomatic Medicine*. 2001;63(4):679-686. doi:<https://doi.org/10.1097/00006842-200107000-00021>
25. Navinés R, Castellví P, Moreno-España J, et al. Depressive and anxiety disorders in chronic hepatitis C patients: Reliability and validity of the Patient Health Questionnaire. *Journal of Affective Disorders*. 2012;138(3):343-351. doi:<https://doi.org/10.1016/j.jad.2012.01.018>
26. Moore M, Ali S, Stuart B, et al. Depression management in primary care: an observational study of management changes related to PHQ-9 score for depression monitoring. *British Journal of General Practice*. 2012;62(599):e451-e457. doi:<https://doi.org/10.3399/bjgp12x649151>
27. Schueller SM, Kwasny MJ, Dear BF, Titov N, Mohr DC. Cut points on the Patient Health Questionnaire (PHQ-9) that predict response to cognitive-behavioral treatments for depression. *General Hospital Psychiatry*. 2015;37(5):470-475. doi:<https://doi.org/10.1016/j.genhosppsych.2015.05.009>
28. Zenlea IS, Milliren CE, Mednick L, Rhodes ET. Depression Screening in Adolescents in the United States: A National Study of Ambulatory Office-Based Practice. *Academic Pediatrics*. 2014;14(2):186-191. doi:<https://doi.org/10.1016/j.acap.2013.11.006>