



Validity and Utility of the Patient Health Questionnaire (PHQ)#2 and PHQ#9 for Screening and Diagnosis of Depression in Rural Chiapas, Mexico: A Cross#Sectional Study

Citation

Arrieta, J., M. Aguerrebere, G. Raviola, H. Flores, P. Elliott, A. Espinosa, A. Reyes, et al. 2017. "Validity and Utility of the Patient Health Questionnaire (PHQ)#2 and PHQ#9 for Screening and Diagnosis of Depression in Rural Chiapas, Mexico: A Cross#Sectional Study." *Journal of Clinical Psychology* 73 (9): 1076-1090. doi:10.1002/jclp.22390. <http://dx.doi.org/10.1002/jclp.22390>.

Published Version

doi:10.1002/jclp.22390

Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:34491853>

Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

Share Your Story

The Harvard community has made this article openly available. Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

Validity and Utility of the Patient Health Questionnaire (PHQ)-2 and PHQ-9 for Screening and Diagnosis of Depression in Rural Chiapas, Mexico: A Cross-Sectional Study

Jafet Arrieta,^{1,2,3} Mercedes Aguerrebere,³ Giuseppe Raviola,^{2,3} Hugo Flores,^{2,3,4} Patrick Elliott,^{2,3,4} Azucena Espinosa,³ Andrea Reyes,³ Eduardo Ortiz-Panozo,⁵ Elena G. Rodriguez-Gutierrez,⁶ Joia Mukherjee,^{2,3,4} Daniel Palazuelos,^{2,3,4} and Molly F. Franke²

¹Harvard T.H. Chan School of Public Health

²Harvard Medical School

³Partners In Health/ Compañeros En Salud

⁴Brigham and Women's Hospital

⁵National Institute of Public Health

⁶Tecnologico de Monterrey School of Medicine

Background: Depressive disorders are frequently under diagnosed in resource-limited settings because of lack of access to mental health care or the inability of healthcare providers to recognize them. The Patient Health Questionnaire (PHQ)-2 and the PHQ-9 have been widely used for screening and diagnosis of depression in primary care settings; however, the validity of their use in rural, Spanish-speaking populations is unknown. **Method:** We used a cross-sectional design to assess the psychometric properties of the PHQ-9 for depression diagnosis and estimated the sensitivity and specificity of the PHQ-2 for depression screening. Data were collected from 223 adults in a rural community of Chiapas, Mexico, using the PHQ-2, the PHQ-9, and the World Health Organization Quality of Life BREF Scale (WHOQOL- BREF). **Results:** Confirmatory factor analysis suggested that the 1-factor structure fit reasonably well. The internal consistency of the PHQ-9 was good (Cronbach's alpha > = 0.8) overall and for subgroups defined by gender, literacy, and age. The PHQ-9 demonstrated good predictive validity: Participants with a PHQ-9 diagnosis of depression had lower quality of life scores on the overall WHOQOL-BREF Scale and each of its domains. Using the PHQ-9 results as a gold standard, the optimal PHQ-2 cutoff score for screening of depression was 3 (sensitivity 80.00%, specificity 86.88%, area under receiver operating characteristic curve = 0.89; 95% confidence interval [0.84, 0.94]). **Conclusion:** The PHQ-2 and PHQ-9 demonstrated good psychometric properties, suggesting their potential benefit as tools for depression screening and diagnosis in rural, Spanish-speaking populations. © 2017 The Authors. *Journal of Clinical Psychology* published by Wiley Periodicals, Inc J. Clin. Psychol. 73:1076–1090, 2017.

We thank the following for their contributions to the manuscript. We are extraordinarily grateful to the physicians that work untiringly with Compañeros En Salud (CES) to provide primary care services in rural Chiapas. We are also grateful to the Tec de Monterrey medical students Alejandro Aguilar, David Aguirre, Victor Dominguez, Francisco Guadarrama, and Alejandra Rodriguez; CES physicians Eduardo Peters and Jesus Rocha; CES employees Jimena Maza and Enrique Valdespino; CES volunteers Ali Friedman, Andrea Echeverri, Araceli Gomez, Tonatiuh Lievano, and Patrick Salemme, who kindly assisted with the data collection. The assistance of Edgardo Lopez, Moises Mazariego, and Maria Eugenia Reyes from the CES logistics staff were essential to our work. We also thank Christina Lively for all her support throughout the design of the protocol and analysis of the data. We especially thank the people from the community in which the study was implemented and the participants of the study for warmly welcoming us during the implementation of the study.

Please address correspondence to: Jafet Arrieta, Department of Global Health and Social Medicine, Harvard Medical School, 641 Huntington Avenue, Boston, MA 02115. E-mail: jafet.arrieta@gmail.com

JOURNAL OF CLINICAL PSYCHOLOGY, Vol. 73(9), 1076–1090 (2017)

© 2017 The Authors. *Journal of Clinical Psychology* published by Wiley Periodicals, Inc

Published online in Wiley Online Library (wileyonlinelibrary.com/journal/jclp). DOI: 10.1002/jclp.22390

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

Keywords: mental health care; primary care; validation

Background

Mental disorders represent 14% of the global burden of disease (Alwan, 2011) and, along with substance disorders, were the leading cause of disability-adjusted life years (DALYs) and years lived with disability (YLD) in 2010. With a prevalence of 4% globally (Ferrari et al., 2013), depressive disorders accounted for 41% of DALYs caused by mental and substance use disorders (Whiteford et al., 2013). In Mexico, depression prevalence estimates range from 7% to 9% (Medina-Mora, Borges, Benjet, Lara, & Berglund, 2007; Spitzer, Williams, & Kroenke, 2014)

Depression is associated with severe physical and social impairment, lower quality of life (QOL; Andriopoulos, Lotti-Lykousa, Pappa, Papadopoulos, & Niakas, 2013; Papakostas et al., 2004; Pyne et al., 1997), and higher health care utilization in places where health services are available (Katon et al., 1990; Wu, Erickson, Piette, & Balkrishnan, 2012). Because patients living with depression often do not seek help for psychological problems but seek care for somatic symptoms instead, their depression often goes unrecognized (Katon & Ciechanowski, 2002; Roness, Mykletun, & Dahl, 2005), leading to treatment delays. Among patients presenting with depression and anxiety in the outpatient medical setting, it is estimated that > 50% patients present with physical rather than psychological complaints (Kroenke, 2003). In the World Health Organization's (WHO) Psychological Problems in General Health Care study, physicians correctly recognized depression in only 42% of patients attending a primary care consultation for depressive symptoms (Simon, Goldberg, Tiemens, & Ustun, 1999).

Lack of access to mental health care may further exacerbate treatment delays, particularly in resource-limited settings (Benjet, Borges, Medina-Mora, Fleiz-Bautista, & Zambrano-Ruiz, 2004; Familiar et al., 2013; Jesse, Dolbier, & Blanchard, 2008; Roness et al., 2005; WHO, 2011). In rural or underserved settings conditioned by a shortage of mental health professionals, medications, and infrastructure (Kitchen Andren et al., 2013) and a lack of available services in the health sector may lead people to seek care with traditional healers rather than with trained clinicians (Nigenda, Mora-Flores, Aldama-Lopez, & Orozco-Nunez, 2001; Pedersen & Baruffati, 1985; Roness et al., 2005; WHO, 2011). Furthermore, even when mental health services are available, stigma, exclusion, and discrimination may prevent care and treatment from reaching people with mental disorders (WHO, 2011). In Mexico, the delay to seek treatment has been reported at 10.6 years for early onset depression and 1.8 years for late onset depression, on average (Benjet et al., 2004).

Given that patients with depression often delay seeking care—and when they do, they seek care for somatic rather than psychological symptoms—it is a challenge for primary health care providers to identify depressive disorders early. Several screening questionnaires have been developed as tools to guide early detection of depression and clinical decision making. The Patient Health Questionnaire (PHQ)-2 (Kroenke, Spitzer, & Williams, 2003) and the PHQ-9 (Spitzer et al., 1999) were specifically designed as depression screening and diagnostic instruments for use in primary care settings, to facilitate the delivery of evidence-based mental health care interventions in places where there is a lack of specialized mental health providers (Spitzer, Kroenke, & Williams, 1999).

Spanish versions of the PHQ-9 have been reported to be reliable and valid measures of depression in urban clinical settings in Spain (Diez-Quevedo, Rangil, Sanchez-Planell, Kroenke, & Spitzer, 2001), Honduras (Wulsin, Somoza, & Heck, 2002), Chile (Baader et al., 2012), and Mexico (Familiar et al., 2015). However, the PHQ-9 has not yet been validated in rural and highly marginalized Spanish-speaking populations where the literacy rates are lower than urban settings (Stromquist, 2001) and the subjective experience of illnesses is shaped differently by the cultural background of individuals and the sociological characteristics of their context (Castro & Eroza, 1998).

The aim of this study was to assess the psychometric properties of the PHQ-9 for diagnosis of depression in a rural, Spanish-speaking population, and if the PHQ-9 proved to be valid, to estimate the sensitivity and specificity of the PHQ-2 for screening of depression.

Method

Setting and Study Population

We conducted a cross-sectional study between July and December 2014 in a rural and highly marginalized community of the Sierra Madre Mountains (the Sierra) of Chiapas, Mexico, in collaboration with Compañeros En Salud (CES), Partners In Health's sister organization in Mexico. Chiapas is the poorest state in Mexico and has a population of approximately 5 million (Instituto Nacional de Estadística y Geografía [INEGI], 2010). Approximately, 51% of the population lives in rural areas, 30% is indigenous, and 75% lives below the poverty line (INEGI, 2010). In the Sierra, living conditions are difficult, with approximately 30% of homes lacking water, 46% lacking sewage, and 15% lacking electricity (Instituto para el Federalismo y el Desarrollo Municipal, 2010).

Chiapas has the lowest level of effective health coverage in the country, which is affected by high levels of marginalization, geographic barriers to access care, poor communication infrastructure, and a shortage of human resources for health (Lozano et al., 2013). The prevalence of depression has been estimated at 6.3% for women and 2.6% for men in Chiapas (Belló, Puentes-Rosas, Medina-Mora, & Lozano, 2005), where there are only 26 psychiatrists for a rate of 0.54 per 100,000 population, compared to a rate of 2.7 per 100,000 population in Mexico (Heinze, del Carmen Chapa, Santiesteban, & Vargas, 2012).

In 2012, CES launched a community-based mental health program (Belkin et al., 2011; Raviola, Eustache, Oswald, & Belkin, 2012) that aimed to improve access to quality mental health care in the Sierra. This program includes active case-finding activities for depression, which comprises the use of the PHQ-2 for screening followed by the implementation of the PHQ-9 for diagnosis of depression. Patients with a PHQ-9 score greater than 9 are linked to the clinic for further diagnosis and treatment.

Ethics, Consent, and Permissions

The protocol received ethical approval from the Institutional Review Board (IRB) of the Harvard Medical School Office of Human Research Administration and the Tecnológico de Monterrey School of Medicine IRB. All participants provided verbal informed consent.

Measures

Sociodemographic characteristics. We assessed ability to read and write and financial situation, a proxy for socioeconomic status (SES), via self-report.

PHQ-9 (Spitzer et al., 1999). The PHQ was developed to make a criteria-based diagnosis of major depressive disorder. The PHQ-9 comprises nine items that evaluate the presence of the nine *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision* (American Psychiatric Association, 2000a) criteria for major depressive disorder in the previous 2 weeks. Each item of the PHQ-9 is rated on a 4-point scale, ranging from 0 (*not at all*) to 3 (*nearly every day*), for a total score ranging from 0 to 27. Higher scores indicated an increased severity of symptoms and an increased likelihood of major depressive disorder (Kroenke, Spitzer, & Williams, 2001). Questionnaires with up to two missing values are scored, replacing any missing values with the average score of the completed items (Kroenke, Spitzer, Williams, & Löwe, 2010).

Cutoffs of 5, 10, 15, and 20 represent mild, moderate, moderately severe, and severe levels of depressive symptoms, respectively (Kroenke et al., 2010). A cutoff of 10 or greater has been described as diagnostic in systematic reviews and meta-analysis of the PHQ-9 (Kroenke et al., 2010; Manea, Gilbody, & McMillan, 2015) and was used as the cutoff for a diagnosis of depression in the present study.

A final item assesses the perception of social, functional, and occupational impairment caused by the symptoms examined by the PHQ-9. Participants rate how difficult their depressive symptoms made it to do their work, take care of things at home, or get along with other people,

with four possible responses: not difficult at all, somewhat difficult, very difficult, and extremely difficult. We used the official Spanish translation of the PHQ-9 (Spitzer et al., 2014).

PHQ-2 (Kroenke, 2003). The PHQ-2 comprises the first two items of the PHQ-9 and evaluates the frequency of depressed mood and little interest or pleasure in doing things over the past 2 weeks. Items are rated on a 4-point scale, ranging from 0 (*not at all*) to 3 (*nearly every day*), for a total score ranging from zero to six (Kroenke et al., 2010). A cutoff of 3 or greater has been found to have the greatest sensitivity and specificity for screening of depression (Kroenke et al., 2003).

The WHO Quality of Life BREF (WHOQOL-BREF) assessment instrument. We used the WHOQOL-BREF to assess QOL (Harper, 1996). This scale is an abbreviated 26-item version of the WHOQOL-100 and comprises items that were extracted from the WHOQOL-100 field trial data (Harper, 1996). The WHOQOL-BREF has one item from each of the 24 facets of QOL included in the WHOQOL-100, plus two items from the general facet on overall QOL and general health. The facets are subsumed in four domains: physical health, psychological health, social relationships, and environment. The scores for each of the domains are transformed on a scale from 0 to 20, to enable comparisons to be made between domains composed of unequal numbers of items. A higher score for the WHOQOL-BREF scale reflects a higher level of functioning and QOL (Harper, 1996).

The WHOQOL-BREF instrument has proved to be reliable and valid for assessing QOL among diverse Spanish-speaking populations, and its psychometric properties have been shown to be similar to those of the English version (Colbourn, Masache, & Skordis-Worrall, 2012; Espinoza, Osorio, Torrejon, Lucas-Carrasco, & Bunout, 2011; Lucas-Carrasco, Laidlaw, & Power, 2011; Lucas-Carrasco, 2012).

Data Collection

Seven medical students visited 152 households as part of a programmatic census and active case-finding activity, and they invited all eligible adult residents to participate in the study (N = 250). Eligible participants were those who were 18 years or older, resided within the study catchment area, and were native-Spanish speakers. The students informed the potential participants about the study and asked for verbal informed consent. After obtaining informed consent, the medical students conducted a face-to-face interview using the PHQ-2, the PHQ-9, and the WHOQOL-BREF assessment questionnaires. A total of 223 (89%) people completed the interview. Those items that were not answered or understood by the participants were computed as missing data.

Participants with a PHQ-9 score greater than 9 and those who reported self-harm ideation or attempts were referred to the clinic for further evaluation by the local physician. Additionally, in a random subset (10%) of participants with a PHQ-9 score greater than 9, a local psychiatrist, blind to the PHQ-9 results, conducted a second depression evaluation using the Hamilton Depression Rating Scale (Hamilton, 1960) and clinical criteria.

Data Analysis

We excluded eight participants with more than two missing values in the PHQ-9. We conducted confirmatory factor analyses using weighted least squares and compared goodness of fit statistics between a one-factor solution and two two-factor solutions proposed for Mexican women and a United States-based population, including Latina women, respectively (Familiar et al., 2015; Granillo, 2012). Model goodness of fit was assessed using the comparative fit index (Bentler, 1990) and the Bentler-Bonnet non-normed index (Bentler & Bonett, 1980). Values > 0.90 for these indices indicate acceptable fit. We also examined root mean square error of approximation RMSEA, for which a value < 0.08 indicates an excellent model fit. Finally, we examined factor loadings under each solution and their statistical significance.

Once we identified the optimal model, we assessed internal consistency of the PHQ-9 in the study population and across subgroups defined by gender, literacy, and age, using the Cronbach's

alpha coefficient. We assessed predictive validity in two ways. First, we compared WHOQOL-BREF domain scores among participants with PHQ-9 scores greater than 9 and less, or equal to 9 and tested for differences using Wilcoxon rank-sum test. Based on previous research studies (Andriopoulos et al., 2013; Papakostas et al., 2004; Pyne et al., 1997), we anticipated that the median WHOQOL-BREF scores for the overall scale and each of the domains would be lower for participants with a PHQ-9 score greater than 9.

Second, we conducted univariable logistic regression analysis to evaluate the association between sociodemographic characteristics and a PHQ-9 diagnosis of depression. Based on prior studies, we hypothesized that the presence of depression symptoms would be more common among women than men, and among those with lower SES compared to those with higher SES (Andrade et al., 2003; Andriopoulos et al., 2013; Belló et al., 2005; Bromet et al., 2011; Popoola & Adewuya, 2012; Rancans, Vrublevska, Snikere, Koroleva, & Trapencieris, 2014; Slone et al., 2006). After establishing the validity of the PHQ-9, we assessed the sensitivity, specificity, and positive and negative likelihood ratios of the PHQ-2 as a screening instrument, using the PHQ-9 as a gold standard. We used receiver operating characteristic (ROC) curve analysis to find the optimal PHQ-2 cutoff for depression screening. Statistical significance for all tests was determined at a p -value < 0.05 . We computed descriptive and analytic statistics of the quantitative data obtained in this study using STATA (version 13.1).

Results

Sociodemographic Characteristics

Table 1 shows the sociodemographic characteristics of the study population. Of the 215 participants included for analysis, 152 (71%) were women. The mean age of the participants was 38 (standard deviation [*SD*] 16) years, and 21% of the participants were unable to read or write. Agriculture was the main economic activity for 90% of the participants, and 48% of the participants reported not having enough money for basic expenses. Ninety-one percent had access to at least one social program, including Seguro Popular, a health insurance program, and Oportunidades, a conditional cash transfer program.

PHQ-9 Scores, Scale Comprehension, and Follow-Up

For the 215 participants, the overall mean PHQ-9 score was 6.29 (*SD* 5.47). A total of 65 (30.2%) participants had a PHQ-2 score equal to or greater than 3. A total of 55 (25.6%) participants had a PHQ-9 score greater than 9: 35 (63.6%) participants had a PHQ-9 score between 10 and 14, 13 (23.6%) between 15 and 19, and seven (12.7%) greater than 20, corresponding to moderate, moderately severe, and severe depression, respectively. Of the 55 participants with a PHQ-9 score greater than 9, eight (14.5%) had a previous diagnosis of depression and were already receiving treatment at the local clinic. The remaining participants were referred to their local clinic for follow-up, 35 (63.6%) of whom attended the appointment; 28 of the 35 (80.0%) received a confirmatory diagnosis of depression by the local physician. When the local psychiatrist conducted an additional evaluation in a random subset of seven (13%) participants with a PHQ-9 score greater than 9, she confirmed the depression diagnosis in all of them.

Table 2 shows the distribution of responses to each of the PHQ-9 items and the proportion of participants who did not understand each of the items. Feeling tired or having little energy and feeling down, depressed, or hopeless were the most common symptoms, reported in 68% and 56%, respectively. Furthermore, 26% of the study population reported thoughts that they would be better off dead or of hurting themselves in some way in the previous 2 weeks, 9% of which reported having these thoughts on more than half the days. Six percent of participants reported not understanding item 1 (“Having little interest or pleasure in doing things”) and/or item 8 (“Moving or speaking so slowly that other people could have noticed, or feeling so fidgety or restless that you have been moving around a lot more than usual”). Inability to read or write was positively and statistically significantly associated with not understanding item 1 (p -value = 0.03) and borderline positively associated with not understanding item number 8 (p -value = 0.05).

Table 1
Sociodemographic Characteristics of the Study Population (N = 215)^a

Variable	N (%) or mean (SD)
Woman	152 (71)
Age	38 (16)
Literate	169 (79)
Level of education	
Secondary or more	69 (32)
Primary or less	122 (57)
None	24 (11)
Has a partner	171 (80)
Has children	185 (86)
Access to services (electricity, water and sanitation)	181 (84)
Water source	
Tap (inside house)	7 (3)
Tap (in yard)	180 (84)
Communal tap	7 (3)
River/stream/pond	21 (10)
Toilet type	
Flush toilet	6 (3)
Bucket toilet	202 (94)
Pit latrine	1 (0)
None	6 (3)
Energy source	
Electricity	209 (97)
Wood	5 (2)
None	1 (0)
Economic activity	
Agriculture	194 (90)
Small business	2 (1)
Both	10 (5)
Other	9 (4)
Religion (N=213)	
Catholic	198 (93)
Presbyterian	6 (3)
Other	5 (2)
None	4 (2)
Access to social programs	196 (91)
<i>Seguro Popular</i> (health insurance program)	55 (26)
<i>Oportunidades</i> (conditional cash transfer program)	10 (5)
More than one social program	131 (61)
Financial situation	
Enough money for food/clothes and other things	16 (7)
Enough money for food/clothes but not other things	95 (44)
Not enough money for basic expenses	103 (48)
Refused to answer	1 (0)

Note. SD = standard deviation.

^aUnless otherwise specified.

Confirmatory Factor Analyses

Results of confirmatory factor analyses are shown in Table 3. For the one-factor solution, the comparative fit index indicated acceptable goodness of fit, with a value of 0.91; however, the value for the Bentler-Bonnett non-normed index was just below the threshold for acceptability (0.88). All factor loadings were statistically significant at a level of < 0.001, and all but one factor loading (item 1, “Having little interest or pleasure in doing things”) indicated moderate (> 0.40)

Table 2
PHQ-9 Questions and Answer Frequency (N = 215)

Questions	Not at all	Several days	More than half the days	Nearly every day	Did not understand
	%	%	%	%	%
1. Little interest or pleasure in doing things	51.6	23.7	6.0	12.6	6.0
2. Feeling down, depressed, or hopeless	44.2	32.1	5.6	17.7	0.5
3. Trouble falling or staying asleep, or sleeping too much	60.0	28.8	1.4	9.3	0.5
4. Feeling tired or having little energy	31.2	38.1	9.3	20.5	0.9
5. Poor appetite or overeating	57.7	26.5	2.8	13.0	0.0
6. Feeling bad about yourself, or that you are a failure or have let yourself or your family down	55.3	28.8	5.1	8.8	1.9
7. Trouble concentrating on things, such as watching television	61.4	22.3	4.7	8.4	3.3
8. Moving or speaking so slowly that other people could have noticed or being so fidgety or restless that you have been moving around a lot more than usual	63.7	20.0	2.8	7.0	6.5
9. Thoughts that you would be better off dead or of hurting yourself in some way	74.0	16.3	2.8	6.5	0.5

association with the factor. With regard to the two-factor solutions, results for each were nearly identical to the one-factor solution, and the factors within the two-factor solutions were highly correlated (0.92) for both models. Given that neither of the two factor solutions appeared to improve fit, relative to the single factor solution, all subsequent analyses were conducted based on the single-factor solution.

The two items with the highest factor loadings in the confirmatory factor analysis were item 2 (“Feeling down, depressed, or hopeless”) and item 9 (“Having thoughts that you would be better off dead or of hurting yourself in some way”). The two items with the lowest factor loadings were item 1 (“Having little interest or pleasure in doing things”) and item 8 (“Moving or speaking so slowly that other people could have noticed, or feeling so fidgety or restless that you have been moving around a lot more than usual”).

Reliability

The internal consistency of the PHQ-9 was good for both the overall PHQ-9 scores and each of the subgroups evaluated. The Cronbach’s alpha coefficient was 0.81 for the overall PHQ-9; 0.85 for men and 0.80 for women; 0.81 for literate participants and 0.83 for illiterate participants; 0.80 for participants younger than 60 years of age and 0.90 for participants 60 years of age and older.

Table 3
Confirmatory Factor Analysis of the PHQ-9 Among Adults in Rural Mexico: Standardized Factor Loadings and Goodness of Fit Statistics

Depression	Model III			
	Model I	Model II		Two-factor solution identified by Familiar et al.
	Affect	Somatic	F1	F2
1. Little interest or pleasure in doing things	36	39		39
2. Feeling down, depressed, or hopeless	82	83		83
3. Trouble falling or staying asleep, or sleeping too much	61		62	59
4. Feeling tired or having little energy	61		63	62
5. Poor appetite or overeating	73		73	73
6. Feeling bad about yourself, or that you are a failure or have let yourself or your family down	54	55		58
7. Trouble concentrating on things, such as watching television	74	73		77
8. Moving or speaking so slowly that other people could have noticed or being so fidgety or restless that you have been moving around a lot more than usual	44		47	48
9. Thoughts that you would be better off dead or of hurting yourself in some way	78	79		82
Correlation between factors	–	0.92		0.92
NNFI	0.88	0.88		0.88
CFI	0.91	0.91		0.92
RMSEA	0.09	0.09		0.09

Note. NNFI = Bentler-Bonnet non-normed fit index; CFI = comparative fit index.

Predictive Validity

Table 4 shows the association between the PHQ-9 scores and the WHOQOL-BREF scores. The median WHOQOL-BREF scores for the overall scale and each of the domains were statistically significantly lower for the participants with a PHQ-9 score greater than 9, except for the social relationships domain, which was borderline significant ($p = 0.05$). PHQ-2 and PHQ-9 scores were positively correlated, with a Pearson correlation coefficient of 0.75.

The univariable regression analysis (Table 5) showed that having a partner was significantly associated with a 64% reduction in the odds of having a PHQ-9 greater than 9 ($p\text{-value} = 0.003$). No other statistically significant associations between sociodemographic characteristics and a PHQ-9 score greater than 9 were found. However, although not statistically significant, not

Table 4
Association Between WHOQOL-BREF Domains and PHQ-9 Scores

	PHQ-9 \leq 9		PHQ-9 $>$ 9		p-value ^a
	N, median (25th, 75th)		N, median (25th, 75th)		
WHOQOL-Physical	152, 16.00 (14.29, 17.14)		51, 12.57 (10.86, 14.87)		<0.0001
WHOQOL-Psychological	151, 16.00 (14.00, 17.33)		51, 12.80 (11.20, 15.33)		<0.0001
WHOQOL-Social Relationships	156, 16.00 (14.67, 17.33)		54, 16.00 (13.33, 16.00)		0.05
WHOQOL-Environment	147, 14.00 (12.50, 15.50)		51, 13.00 (11.50, 14.29)		0.0007
Total WHOQOL Score	141, 61.24 (57.62, 64.67)		48, 53.77 (46.99, 59.29)		<0.0001

Note. WHOQOL = World Health Organization Quality of Life BREF Scale; PHQ = Patient Health Questionnaire.

^aWilcoxon rank sum p-value.

Table 5
Association Between Sociodemographic Characteristics and PHQ-9 Score > 9

Population (N = 215)		PHQ-9 Score ^a		
Variable	OR	95% CI	p-value	
Age (per increase in 10 years)	1.01	[0.83, 1.23]	0.92	
Female gender	1.14	[0.58, 2.26]	0.70	
Has a partner	0.36	[0.17, 0.70]	0.003	
Has children	0.54	[0.24, 1.22]	0.14	
Literate	1.54	[0.69, 3.43]	0.29	
Has access to electricity, water and sanitation	0.95	[0.41, 2.18]	0.90	
Not enough money for food/clothes	1.74	[0.94, 3.24]	0.08	
Social programs	0.72	[0.26, 2.00]	0.53	

Note. PHQ = Patient Health Questionnaire; OR = odds ratio; CI = confidence interval.

^aUnivariable regression analysis.

Table 6
Sensitivity and Specificity of the PHQ-2 for Screening of Depression

PHQ-2 score	Sensitivity (%)	Specificity (%)	LR+	LR-
\geq 1	98.2 (54/55)	45.0 (72/160)	1.79	0.04
\geq 2	89.1 (49/55)	71.3 (114/160)	3.10	0.15
\geq 3	80.0 (44/55)	86.9 (139/160)	6.10	0.23
\geq 4	45.5 (25/55)	95.6 (153/160)	10.39	0.57
\geq 5	27.3 (15/55)	98.1 (157/160)	14.55	0.74
6	20.0 (11/55)	98.8 (158/160)	16.00	0.81

Note. PHQ = Patient Health Questionnaire; LR+ = positive likelihood ratio; LR- = negative likelihood ratio.

having enough money for basic expenses was associated with an increase in the odds of having a PHQ-9 greater than 9; while having children as well as access to services and social programs were associated with a reduction in the odds of having a PHQ-9 greater than 9.

Sensitivity and Specificity of the PHQ-2

Table 6 shows the sensitivity, specificity, and positive and negative likelihood ratios of the PHQ-2 for depression screening. Using a cutoff of 3 and greater, the PHQ-2 had a sensitivity of 80.0%;

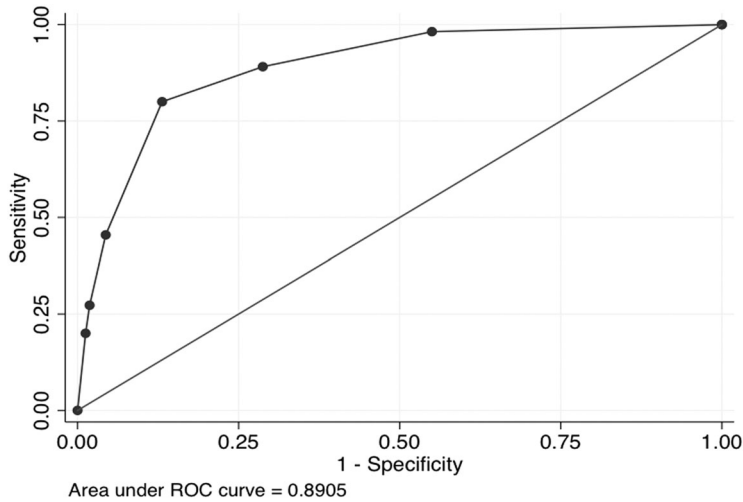


Figure 1. The receiver operating characteristic curve of the PHQ-2 for screening of depression.

specificity of 86.9%; and positive and negative likelihood ratios of 6.10 and 0.23, respectively. In ROC analysis (Figure 1), the PHQ-2 performed well in detecting participants with a PHQ-9 diagnosis of depression (area under the ROC curve was 0.89 (95% confidence interval [0.84, 0.94]). Given that with a score of 3 there was a balance between sensitivity and specificity, a PHQ-2 score of 3 was identified as the optimal cutoff for screening of depression. Using the sensitivity results of the PHQ-2, at a cutoff of 3, the estimated prevalence of depression was 20% for this study population.

Discussion

This study is the first report of the validity and reliability of the PHQ-2 and PHQ-9 as screening and diagnostic instruments for depression in rural Spanish-speaking populations in Mexico and elsewhere. Familiar and colleagues (2015) previously validated the PHQ-9 among educated women teachers in Mexico; we expand on their work to fill a knowledge gap regarding the generalizability of their findings and offer strong evidence for the validity and reliability of the PHQ-9 among men and less educated adults in rural Mexico.

Specifically, our data support the notion of a one-factor structure for the PHQ-9 in this population. The internal consistency was adequate for the overall PHQ-9 and by subgroups. Predictive validity was supported by the statistically significant inverse association between PHQ-9 scores and the overall WHOQOL-BREF scores, and between the PHQ-9 scores and the scores of each of the physical, psychological and environment domains. These associations have been previously observed in primary care settings (Andriopoulos et al., 2013; Brenes, 2007; Papakostas et al., 2004; Pyne et al., 1997). Predictive validity was also suggested by a positive association between a PHQ-9 score greater than 9 and a low SES (not having enough money for basic expenses), although this relationship was not statistically significant ($p = 0.08$).

In this study, the mean PHQ-9 score was 6.29 and the prevalence of a PHQ-9 diagnosis of depression among the participants was 25.6%. The Mexican teachers' cohort study, which used the PHQ-9 to assess female teachers for depression, reported a mean PHQ-9 score of 4.5 and a prevalence of a PHQ-9 diagnosis of 12.6% (Familiar et al., 2015). This difference in depression prevalence may be explained by several factors. First, clinic physicians confirmed the PHQ-9 depression diagnosis in only 80% of people with a score greater than 9 who attended a follow-up visit. This underscores the critical role of clinical evaluation in conjunction with deployment of the PHQ-9. Second, participants in the teachers' cohort study differed from participants in the present study in that they were highly educated (73% with a university degree and 16%

with a postgraduate degree) and tended to be well off (45% reported a medium SES and 20% high SES), suggesting the possibility of better access to mental health care. Of participants in the present study, 21% and 48% were unable to read and write and reported a low SES, respectively.

The lack of understanding of some of the items asked in the PHQ-9 was significantly associated with inability to read and write. Also, the data collectors reported that some participants did not understand some of the words (i.e. pleasure) used in the questionnaire. Interestingly, the lowest factor loadings were observed for the two scale items (1 and 8) with the highest percent of individuals who did not understand the items. The influence of low literacy on the accuracy of depression screening instruments has been documented elsewhere (Dickens Akena et al., 2012) and should be taken into account through careful piloting and adaptation when using them in settings like ours where illiteracy rates are high.

Previous studies have highlighted the use of the PHQ-9 in primary care to identify individuals at increased risk of suicide attempt or death (Simon et al., 2013; Uebelacker, German, Gaudiano, & Miller, 2011). In our study, thoughts of being better off dead or hurting oneself (item 9) were present in 26% of the participants, and those thoughts were present more than half the days in 9% of the participants. This item had the second highest factor loading in our confirmatory factor analysis. Although the PHQ-8, which excludes the self-harm item, has been recommended for community surveys and has been successfully used in primary care settings, given the high prevalence of thoughts of self-harm and its high factor loading in the confirmatory factor analysis, we chose not to exclude this item. Further studies should be conducted to assess the presence, frequency, and intensity of these thoughts, and their causes and potential implications to develop strategies aimed at preventing suicide in these settings.

Brief tools are needed to effectively screen and diagnose patients for depression in primary care settings where demand for services is high and health professionals, including highly trained mental health professionals, are scarce (Kroenke, 1997; Williams, 1998). In this study, the PHQ-2 and PHQ-9 questionnaires were implemented in the context of an active case-finding activity for depression. Both instruments have shown good case-finding properties compared to patient-rated major depression in general medical practice (Olsson, Mykletun, & Dahl, 2005). Once the psychometric properties of the PHQ-9 were assessed and its validity for this population was established, we compared the PHQ-2 with the PHQ-9 to determine its validity for screening of depression.

Our study provides strong evidence to support the use of the PHQ-2 to identify individuals who may be at high risk for depression. The PHQ-2 showed high sensitivity and specificity as a screening instrument for a diagnosis of major depression when compared with the PHQ-9. A cutoff of 3 and greater was optimal for the study population maximizing sensitivity (80.0%) and specificity (86.9%). However, selecting an optimal PHQ-2 cutoff should be done cautiously: Selecting a high sensitivity and lower specificity would increase false positive screens, which in turn would potentially lead to increased burden on clinicians and/or cost to the patient, while selecting a lower sensitivity in favor of a higher specificity would contribute to miss depression diagnoses. The sensitivity and specificity of the PHQ-2 in this study population were similar to previous studies in the literature from primary care populations (Arroll et al., 2010; Kroenke et al., 2003; Richardson et al., 2010; Zuthoff et al., 2010).

The positive likelihood ratio increased as the cutoff increased and was 6.10 for a cutoff of 3 or greater, which means that given a PHQ-2 score of 3 or greater, the individual is 6.10 times more likely to have a PHQ-9 score greater than 9. The positive and negative likelihood ratios are a combination of sensitivity and specificity and are useful for clinicians especially in settings like ours where the true prevalence of depression is unknown (Arroll et al., 2010).

The PHQ-2 showed good psychometric properties for community-based screening of depression, is easy for care providers to remember the core symptoms it assesses, and is a brief depression screen; therefore, we recommend the use of the PHQ-2 as a screening tool in rural Spanish-speaking settings to improve early detection of depression. Given its good reliability and predictive validity, we recommend using a contextually adapted version of PHQ-9 for further assessment of depression for those patients with a PHQ-2 of 3 or greater, followed by a further diagnostic work-up for those patients with a PHQ-9 greater than 9.

Limitations

Our study has several limitations. First, we conducted the study in only one rural community; however, we believe that our results are likely generalizable to other rural, low literacy Spanish-speaking communities in Mexico. These findings are also likely generalizable to other populations with similar sociodemographic characteristics (i.e., rural, low literate) in Latin American countries; however, it is unknown how these findings would apply to non-Spanish-speaking populations, including those who might speak an indigenous dialect as their first or primary language. Second, we used the PHQ-9 as a gold standard to assess the sensitivity and specificity of the PHQ-2 as a screening instrument for depression. Therefore, if the sensitivity or the specificity of the PHQ-9 is suboptimal, then our analyses related to the sensitivity and specificity of the PHQ-2 may be biased, and a study using a different diagnostic tool from the PHQ-9 would be recommended. Given that some the participants did not understand some of the items in the PHQ-9, we would recommend adapting it to relate to the local context and the literacy level of the population.

Conclusion

Given the large health and social burden of depression and considering the need for brief, structured, reliable, and valid tools to aid primary care providers in assessing patients for depression, both the PHQ-2 and PHQ-9 are useful and valuable instruments for screening and diagnosis of depression in rural, Spanish-speaking populations. However, to reduce the global burden of mental health disorders and improve QOL of patients, the use of these tools should be coupled with expanded access to appropriate mental health and both pharmacological and nonpharmacological interventions that have proven useful for these settings.

References

- Alwan, A. (2011). Global status report on non-communicable diseases 2010. Geneva: World Health Organization.
- American Psychiatric Association. (2000a). Diagnostic and statistical manual of mental disorders, fourth edition, text revision. Washington, DC: American Psychiatric Publishing.
- Andrade, L., Caraveo-Anduaga, J. J., Berglund, P., Bijl, R. V., De Graaf, R., Vollebergh, W., . . . Wittchen, H. U. (2003). The epidemiology of major depressive episodes: Results from the International Consortium of Psychiatric Epidemiology (ICPE) surveys. *International Journal of Methods in Psychiatric Research*, 12(1), 3–21.
- Andriopoulos, P., Lotti-Lykousa, M., Pappa, E., Papadopoulos, A. A., & Niakas, D. (2013). Depression, quality of life and primary care: A cross-sectional study. *Journal of Epidemiology and Global Health*, 3, 245–252.
- Arroll, B., Goodyear-Smith, F., Crengle, S., Gunn, J., Kerse, N., Fishman, T., . . . Hatcher, S. (2010). Validation of PHQ-2 and PHQ-9 to screen for major depression in the primary care population. *Annals of Family Medicine*, 8(4), 348–353.
- Baader M. T., Molina F, J. L., Venezian B. S., Rojas C. C., Farias S. R., Fierro-Freixenet, C., . . . Mundt, C. (2012). Validity and utility of PHQ9 (Patient Health Questionnaire) in the diagnosis of depression in user patients of primary care in Chile. *Rev. Chil. Neuro-Psiquiat.*, 50(1), 10.
- Belkin, G. S., Unützer, J., Kessler, R. C., Verdelli, H., Raviola, G. J., Sachs, K., . . . Eustache, E. (2011). Scaling up for the “bottom billion”: “5×5” implementation of community mental health care in low-income regions. *Psychiatric Services*, 62(12), 1494–1502.
- Belló, M., Puentes-Rosas, E., Medina-Mora, M. E., & Lozano, R. (2005). Prevalencia y diagnóstico de depresión en población adulta en México [Prevalence and diagnosis of depression in Mexico]. *Salud Pública De México*, 47(1), 4.
- Benjet, C., Borges, G., Medina-Mora, M. E., Fleiz-Bautista, C., & Zambrano-Ruiz, J. (2004). La depresión con inicio temprano: Prevalencia, curso natural y latencia para buscar tratamiento [Early onset depression: Prevalence, course, and treatment latency]. *Salud Pública De México*, 46(5), 417–424.
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107(2), 237–246.

- Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88(3), 588–606.
- Brenes, G. A. (2007). Anxiety, depression, and quality of life in primary care Patients. *Primary Care Companion to the Journal of Clinical Psychiatry*, 9(6), 437–443.
- Bromet, E., Andrade, L. H., Hwang, I., Sampson, N. A., Alonso, J., de Girolamo, G., . . . Kessler, R. C. (2011). Cross-national epidemiology of DSM-IV major depressive episode. *BMC Medicine*, 9(90).
- Castro, R., & Eroza, E. (1998). Research notes on social order and subjectivity: Individuals' experience of susto and fallen fonatnelle in a rural community in central Mexico. *Culture, Medicine and Psychiatry*, 22(2), 203–230.
- Colbourn, T., Masache, G., & Skordis-Worrall, J. (2012). Development, reliability and validity of the Chichewa WHOQOL-BREF in adults in Lilongwe, Malawi. *BMC Research Notes*, 5(346).
- Dickens Akena, D., Joska, J., Okubu, E. A., Amos, T., Musisi, S., & Stein, D. J. (2012). Comparing the accuracy of brief versus long depression screening instruments which have been validated in low and middle income countries: A systematic review. *BMC Psychiatry*, 12(1), 187–193.
- Diez-Quevedo, C., Rangil, T., Sanchez-Planell, L., Kroenke, K., & Spitzer, R. 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.
- Espinoza, I., Osorio, P., Torrejon, M. J., Lucas-Carrasco, R., & Bunout, D. (2011). Validation of the WHOQOL-BREF quality of life questionnaire among Chilean older people. *Revista Medica De Chile*, 139(5), 579–586.
- Familiar, I., Ortiz-Panozo, E., Hall, B., Vieitez, I., Romieu, I., Lopez-Ridaura, R., & Lajous, M. (2015). Factor structure of the Spanish version of the Patient Health Questionnaire-9 in Mexican women. *International Journal of Methods in Psychiatric Research*, 24(1), 74–82.
- Familiar, I., Sharma, S., Ndayisaba, H., Munyentwari, N., Sibomana, S., & Bass, J. K. (2013). Community perceptions of mental distress in a post-conflict setting: A qualitative study in Burundi. *Global Public Health*, 8(8), 943–957.
- Ferrari, A. J., Charlson, F. J., Norman, R. E., Patten, S. B., Freedman, G., Murray, C. J., . . . Whiteford, H. A. (2013). Burden of depressive disorders by country, sex, age, and year: Findings from the Global Burden of Disease study 2010. *PLoS Medicine*, 10(11).
- 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.
- Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery and Psychiatry*, 23(56), 56–62.
- Harper, A. (1996). WHOQOL-BREF introduction, administration, scoring and generic version of the assessment. Geneva: World Health Organization.
- Heinze, G., del Carmen Chapa, G., Santiesteban, J., & Vargas, I. (2012). Los especialistas en psiquiatría en México: Su distribución, ejercicio profesional y certificación. *Salud Mental*, 35, 279–285.
- Instituto Nacional de Estadística y Geografía. (2010). *Conteo de Población y Vivienda*. Mexico City. Retrieved from <http://www.inegi.org.mx/est/contenidos/proyectos/ccpv/cpv2010/>
- Instituto para el Federalismo y el Desarrollo Municipal. (2010). Estado de Chiapas [Chiapas State]. Mexico City: Secretaria de Gobernación. Retrieved from <https://www.gob.mx/inafed>
- Jesse, D. E., Dolbier, C. L., & Blanchard, A. (2008). Barriers to seeking help and treatment suggestions for prenatal depressive symptoms: Focus groups with rural low-income women. *Issues in Mental Health Nursing*, 29(1), 3–19.
- Katon, W., & Ciechanowski, P. (2002). Impact of major depression on chronic medical illness. *Journal of Psychosomatic Research*, 53(4), 859–863.
- Katon, W., Von Korff, M., Lin, E., Lipscomb, P., Russo, J., Wagner, E., & Polk, E. (1990). Distressed high utilizers of medical care. DSM-III-R diagnoses and treatment needs. *General Hospital Psychiatry*, 12(6), 355–362.
- Kitchen Andren, K. A., Mckibbin, C. L., Wykes, T. L., Lee, A. A., Carrico, C. P., & Bourassa, K. A. (2013). Depression treatment among rural older adults: Preferences and factors influencing future service use. *Clinical Gerontologist*, 36(3), 241–259.
- Kroenke, K. (1997). Discovering depression in medical patients: Reasonable expectations. *Annals of Internal Medicine*, 126(6), 463–465.

- Kroenke, K. (2003). The interface between physical and psychological symptoms. *The Primary Care Companion to The Journal of Clinical Psychiatry*, 5(7), 11.
- Kroenke, K., Spitzer, R., & Williams, J. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine*, 16(9), 606–613.
- Kroenke, K., Spitzer, R., & Williams, J. (2003). The Patient Health Questionnaire-2: Validity of a two-item depression screener. *Medical Care*, 41(11), 1284–1292.
- 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.
- Lozano, R., Gómez-Dantés, H., Garrido-Latorre, F., Jiménez-Corona, A., Campuzano-Rincón, J., Franco-Marina, F., . . . Murray, C. (2013). La carga de enfermedad, lesiones, factores de riesgo y desafíos para el sistema de salud en México [Burden of disease, injuries, risk factors and challenges for the health system in Mexico]. *Salud Pública De México*, 55(6), 580.
- Lucas-Carrasco, R. (2012). The WHO Quality of Life (WHOQOL) questionnaire: Spanish development and validation studies. *Quality of life research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, 21(1), 161–165.
- Lucas-Carrasco, R., Laidlaw, K., & Power, M. J. (2011). Suitability of the WHOQOL-BREF and WHOQOL-OLD for Spanish older adults. *Aging & Mental Health*, 15(5), 595–604.
- 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.
- Medina-Mora, M., Borges, G., Benjet, C., Lara, C., & Berglund, P. (2007). Psychiatric disorders in Mexico: Lifetime prevalence in a nationally representative sample. *The British Journal of Psychiatry*, 190(6), 521–528.
- Nigenda, G., Mora-Flores, G., Aldama-Lopez, S., & Orozco-Nunez, E. (2001). La práctica de la medicina tradicional en América Latina y el Caribe: El dilema entre regulación y tolerancia [The practice of traditional medicine in Latin America and the Caribbean: The dilemma between regulation and tolerance]. *Salud Pública De México*, 43(1), 41–51.
- Olsson, I., Mykletun, A., & Dahl, A. (2005). The hospital anxiety and depression rating scale: A cross-sectional study of psychometrics and case finding abilities in general practice. *BMC Psychiatry*, 5(46).
- Papakostas, G., Petersen, T., Mahal, Y., Mischoulon, D., Nierenberg, A., & Fava, M. (2004). Quality of life assessments in major depressive disorder: A review of the literature. *General Hospital Psychiatry*, 26(1), 13–17.
- Pedersen, D., & Baruffati, V. (1985). Health and traditional medicine cultures in Latin America and the Caribbean. *Social Science & Medicine*, 21(1), 5–12.
- Popoola, A. O., & Adewuya, A. O. (2012). Prevalence and correlates of depressive disorders in outpatients with breast cancer in Lagos, Nigeria. *Psycho-Oncology*, 21(6), 675–679.
- Pyne, J., Patterson, T., Kaplan, R., Gillin, J., Koch, W., & Grant, I. (1997). Assessment of the quality of life of patients with major depression. *Psychiatric Services*, 48(2), 224–230.
- Rancans, E., Vrublevska, J., Snikere, S., Koroleva, I., & Trapencieris, M. (2014). The point prevalence of depression and associated sociodemographic correlates in the general population of Latvia. *Journal of Affective Disorders*, 156, 104–110.
- Raviola, G., Eustache, E., Oswald, C., & Belkin, G. S. (2012). Mental health response in Haiti in the aftermath of the 2010 earthquake: A case study for building long-term solutions. *Harvard Review of Psychiatry*, 20(1), 68–77.
- Richardson, L., Rockhill, C., Russo, J., Grossman, D., Richards, J., McCarty, C., . . . Katon, W. (2010). Evaluation of the PHQ-2 as a brief screen for detecting major depression among adolescents. *Pediatrics*, 125, e1–e7.
- Rones, A., Mykletun, A., & Dahl, A. A. (2005). Help-seeking behaviour in patients with anxiety disorder and depression. *Acta Psychiatrica Scandinavica*, 111(1), 51–58.
- Simon, G. E., Goldberg, S. D., Tiemens, B. G., & Ustun, T. B. (1999). Outcomes of recognized and unrecognized depression in an international primary care study. *General Hospital Psychiatry*, 21(2), 97–105.
- Simon, G. E., Rutter, C. M., Peterson, D., Oliver, M., Whiteside, U., Operskalski, B., & Ludman, E. J. (2013). Does response on the PHQ-9 depression questionnaire predict subsequent suicide attempt or suicide death? *Psychiatry Services*, 64(12), 1195–1202.

- Slone, L. B., Norris, F. H., Murphy, A. D., Baker, C. K., Perilla, J. L., Diaz, D., . . . Gutierrez Rodriguez J. J. (2006). Epidemiology of major depression in four cities in Mexico. *Depression and Anxiety, 23*(3), 158–167.
- Spitzer, R. L., Kroenke, K., & Williams, J. B. (1999). Validation and utility of a self-report version of PRIME-MD: The PHQ primary care study. *Primary Care Evaluation of Mental Disorders. Patient Health Questionnaire. JAMA: The Journal of the American Medical Association, 282*(18), 1737–1744.
- Spitzer, R. L., Williams, J. B. W. & Kroenke, K. (2014). Patient health questionnaire (PHQ) screeners. Retrieved from http://www.phqscreeners.com/sites/g/files/g10016261/f/201412/PHQ9_Spanish%20for%20the%20USA.pdf
- Stromquist, N. P. (2001). What poverty does to girls' education: The intersection of class, gender and policy in Latin America. *Compare: A Journal of Comparative Education, 31*(1), 39–56.
- Uebelacker, L. A., German, N. M., Gaudiano, B. A., & Miller, I. W. (2011). Patient Health Questionnaire depression scale as a suicide screening instrument in depressed primary care patients: A cross-sectional study. *The Primary Care Companion for CNS Disorders, 13*(1).
- Whiteford, H. A., Degenhardt, L., Rehm, J., Baxter, A. J., Ferrari, A. J., Erskine, H. E., . . . Vos, T. (2013). Global burden of disease attributable to mental and substance use disorders: Findings from the Global Burden of Disease study 2010. *Lancet, 382*(9904), 1575–1586.
- Williams, J. W. (1998). Competing demands: Does care for depression fit in primary care? *Journal of General Internal Medicine, 13*(2), 137–139.
- World Health Organization. (2011). *Mental health atlas 2011* (no. 1). Italy: Author.
- Wu, C. H., Erickson, S. R., Piette, J. D., & Balkrishnan, R. (2012). Mental health resource utilization and health care costs associated with race and comorbid anxiety among Medicaid enrollees with major depressive disorder. *Journal of the National Medical Association, 104*(1-2), 78–88.
- Wulsin, L., Somoza, E., & Heck, J. (2002). The feasibility of using the Spanish PHQ-9 to screen for depression in primary care in Honduras. *Primary Care Companion to the Journal of Clinical Psychiatry, 4*(5), 191–195.
- Zuithoff, N., Vergouwe, Y., King, M., Nazareth, I., Van Wezep, M., Moons, K., & Geerlings, M. (2010). The Patient Health Questionnaire-9 for detection of major depressive disorder in primary care: Consequences of current thresholds in a cross-sectional study. *BMC Family Practice, 11*(98).
- Janzen, G., & Hawlik, M. (2005). Orientierung im Raum: Befunde zu Entscheidungspunkten [Orientation in space: Findings about decision points]. *Zeitschrift für Psychologie, 213*(4), 179–186.