HHS Public Access

Author manuscript

*Neurol Psychiatry Brain Res*. Author manuscript; available in PMC 2020 August 28.

Published in final edited form as:

*Neurol Psychiatry Brain Res*. 2019 February ; 31: 9–14. doi:10.1016/j.npbr.2018.11.002.

# Validating the Patient Health Questionnaire-9 (PHQ-9) for screening of depression in Tanzania

**Mary C. Smith Fawzi**[1,*], **Fileuka Ngakongwa**[2], **Yuanyuan Liu**[1], **Theonest Rutayuga**[2], **Hellen Siril**[3], **Magreat Somba**[4], **Sylvia F. Kaaya**[2]

[1]Harvard Medical School, Department of Global Health and Social Medicine, Boston MA, USA

[2]Muhimbili University of Health and Allied Sciences, Department of Psychiatry and Mental Health, Dar es Salaam, Tanzania

[3]Management and Development for Health, Dar es Salaam, Tanzania

[4]Africa Academy for Public Health, Dar es Salaam, Tanzania

## Abstract

**Background—**Major depression has a significant impact on years lived with disability (YLD) globally. In resource-limited countries, depression may accompany daily challenges of economic security that people face, hence there is a critical need to develop depression screening tools at primary levels of health care. The overall goal of the study is to validate the PHQ-9 in Tanzania.

**Methods—**A validation study was conducted from August to October 2014 among adults accessing primary health care at public clinics in Dar es Salaam. The Mini-International Neuropsychiatric Interview (MINI) was used as the gold standard for current major depressive episode.

**Results—**Among 180 patients recruited, six were not included in the analysis since the PHQ-9 and MINI assessments were conducted more than two weeks apart (n=174). The PHQ-9 demonstrated reasonable reliability in this setting ($\alpha$=0.83). Evidence for construct validity was observed through expected associations with female gender (r=0.16, p=0.04) and food insecurity (r=0.30, p<0.0001). Receiver Operating Characteristic analysis demonstrated good overall accuracy of the PHQ-9 (AOC=0.87, 95%CI: 0.77, 0.96). The optimal cut-off score in this population was 9, with a sensitivity of 78% and specificity of 87%.

[*]Address for correspondence: Mary C. Smith Fawzi, ScD, Harvard Medical School, Department of Global Health and Social Medicine, 641 Huntington Ave., Boston, MA 02115, Tel: 617 432 2574, Fax: 617 432 2565, mary_smith-fawzi@hms.harvard.edu.



Competing interests

None.

Ethical statement

The study was approved by the Institutional Review Board (IRB) at Harvard Medical School and the National Institute for Medical Research (NIMR) in Tanzania. All participants provided written informed consent.

**Limitations—**The study sample is from a primary health care setting, hence the findings may have some limited generalizability at the community level.

**Conclusions—**The PHQ-9 demonstrated reliability and validity among adults accessing primary health care in Dar es Salaam, indicating that it can serve as a useful tool in identifying patients with depression in primary care clinics in Tanzania and similar settings.

### Keywords

Depression; validity; primary health care; screening; Tanzania; PHQ-9

## Introduction

Worldwide, depression has demonstrated a significant burden on morbidity and mortality (Marcus, Yasamy, van Ommeren, & Chisholm, 2012). The Global Burden of Disease study reported that major depression was the second leading cause of years lived with disability (YLD) globally (Marcus et al., 2012; Vos et al., 2015). In particular, those living in low or middle-income countries can experience an increased risk of depression related to insufficient income and limited education (Kessler & Bromet, 2013), which can result in food insecurity and problems addressing other basic needs. In many cases poverty is coupled with political unrest or violence (Murthy & Lakshminarayana, 2006), and places a greater burden on vulnerable populations such as those living with HIV (Chishinga et al., 2011; Kaharuza et al., 2006; Myer et al., 2008) or other chronic conditions (Moussavi et al., 2007). In settings of persistent poverty, basic needs are often unmet such as adequate food, clothing, housing, clean water and sanitation. This can result in a higher burden of illness in resource-limited settings, often left untreated and leading to poor health outcomes compared to contexts with greater resources (Ferrari et al., 2013). Given that these factors typically overlap in resource-limited settings, individuals are at high risk for depression and this needs to be urgently addressed (Ferrari et al., 2013; Saraceno et al., 2007).

Despite the significant burden of depression, access to care for depression in resource-limited settings is often inadequate (Saraceno et al., 2007; Saxena, Thornicroft, Knapp, & Whiteford, 2007). In particular, Tanzania is one of the poorest countries worldwide with a Gross National Income (GNI) per capita of 920 USD (UNICEF, 2014). Depression, the most prevalent of a group of common mental disorders (CMDs), including major depressive episodes, anxiety and substance use disorders is also prevalent in the general population. In urban Tanzania, the six-month community level prevalence of CMDs has been reported to be 31 per 1,000 population, with rates in women and men being 36 and 25 per 1,000 population, respectively. In addition, lower income communities demonstrated a higher prevalence of CMDs when compared with more affluent localities (41 and 23 per 1,000 population, respectively) (Jenkins et al., 2010). In urban primary health care settings, rates of common mental disorders in Tanzania were reported to be high at 24% in primary public health care services, with a higher prevalence (48%) among clients of traditional healers (Ngoma et al., 2003). A significant proportion of patients (17.6% from primary care public services and 41.9% from traditional healers) had previously consulted at least four different allopathic health care providers with the same symptoms, raising concern as Ngoma (2003)

showed that 90.5% of persons with CMD accessing primary care services in Tanzania had symptoms of depression.

One of the ways that access to care for depression can improve is through increased identification (Kohrt et al., 2016). It has been shown that in primary health care clinics patients often present with depressive symptoms, which is reflected in the significant prevalence of depression in these settings, ranging from 10 – 20% (Abiodun, 1993; Gureje, Obikoya, & Ikuesan, 1992). When physicians are unable to identify a physical cause, people often leave without treatment. In addition, when depressive symptoms are exacerbated or recur patients often return for care. This cycle can persist if depression and other common mental disorders remain untreated (Kisely et al., 2006). Given the limited access to mental health professionals in resource-limited settings, and the frequent presentation of depression in primary care settings, screening tools are needed to identify those with elevated depressive symptoms for further evaluation and treatment.

The Patient Health Questionnaire-9 (PHQ-9) has been a tool that has been used widely to identify individuals with clinically relevant levels of depressive symptoms, particularly in resource-limited settings. It has demonstrated reasonable validity and reliability in some contexts, such as Uganda (Akena et al.,2013) and South Africa (Cholera et al., 2014), although sensitivity and/or specificity estimates have been lower in other settings, such as in Cameroon (Pence et al., 2012) and Ethiopia (Gelaye et al., 2013). However, to date the PHQ-9 has not been validated for use in Tanzania. Validating the PHQ-9 in Tanzania is useful from several vantage points. First, cut-off scores from other settings may not apply to Tanzania. Other studies have demonstrated that validating psychometric scales for different cultural contexts can result in varying cut-off scores (Betancourt et al., 2012; Fendrich et al., 1990). Second, it is important that the translation process adequately captures the local constructs and idioms for distress (Rasmussen et al., 2015). Third, having a locally validated cut-off score for the PHQ-9 will allow health care providers in primary care and other settings a simple tool that can be used to screen for depressive symptom levels comparable with major depression in Tanzania. Based on varied evidence from resource-limited settings and the need to identify an appropriate cut-off for recognition, treatment and/or referral of persons with depression in Tanzania in primary care settings, the goal of this paper is to evaluate the validity of the PHQ-9 among individuals accessing primary health care in Dar es Salaam, Tanzania in identifying a major depressive episode in this population.

## Methods

### Study design and sample selection

The design was a validation study for the PHQ-9 in Dar es Salaam, Tanzania. Adults (18 years of age and above) attending general out-patient clinics in eleven government-run primary health care facilities in Dar es Salaam who provided informed consent were eligible to participate in the study. From August to October 2014, adults waiting at the selected clinics were approached consecutively by research staff to inquire whether they may be interested in participating in the study. Prospective participants were informed that the intention of the study was to develop/validate a brief questionnaire for use in Tanzania to identify symptoms of depression, such as feeling sad for most of the time, feeling little or no

interest in things, feeling badly about yourself, feeling low in energy, among other symptoms; the overall goal was to identify those with symptoms of depression, so that they can be readily referred for treatment if necessary. All individuals approached decided to enroll in the study given the limited risk in participating.

An initial screening was conducted to recruit half of the study participants with a higher level of depressive symptoms and the other half to reflect a lower level of symptoms. The HSCL (Hopkins Symptom Checklist) was used to sample those with higher versus lower levels of depressive symptoms. The HSCL was validated previously in Dar es Salaam, with a shortened version (HSCL-8; a sub-set of items from the HSCL-25) demonstrating reasonable sensitivity and specificity estimates in this context. A cut-off score of greater than 1.06 on the HSCL-8 indicated a higher level of symptoms and those that scored 1.06 or less reflected a lower level of symptomatology (Kaaya et al., 2002). All of the study participants achieving above this cut-off score were included in the study until n=90 were recruited for this sub-group. For those who tested negative (cut-off score <=1.06), a sub-set of n=90 was randomly selected, since a larger percentage of participants achieved a score below this cut-off. The goal was to create groups of equal size to maximize statistical power for the study.

Study participants included in the validation were interviewed using the PHQ-9 by trained research assistants (RAs) different from those that performed the screening assessment; these RAs were blinded regarding the results of the screening assessment. Professionally trained mental health providers (a psychiatrist and psychiatric nurse), blinded to the results of both the screening and PHQ-9 assessments, administered the Mini-International Neuropsychiatric Interview (MINI; version 6.0, 2010) within two-weeks of the PHQ-9 interview, which served as the gold standard for the study. This measure was relatively brief and relied on Diagnostic and Statistical Manual of Mental Disorders-IV-TR (DSM-IV-TR) criteria for a major depressive episode (MDE) diagnosis. Given logistical constraints six out of the 180 participants had MINI assessments that were performed greater than two weeks from the administration of the PHQ-9 and were excluded from the analysis. Prospective participants found to be experiencing acute suicidal ideation during screening or administration of the PHQ-9 or MINI were referred for immediate assistance.

## Measures

The initial screening measure for recruitment was the eight-item version of the HSCL previously validated in Tanzania. The validation study identified a sub-set of eight symptoms that performed as well as the HSCL-25 in detecting major depressive disorder as determined by psychiatrists administering the Structured Clinical Interview for the DSM-IV (SCID). Receiver Operating Characteristic (ROC) analysis was performed to compare the validity of HSCL-25 and the abbreviated eight-item version. The area under the ROC curve (AUC) was 0.86 and 0.88 for the HSCL-25 and revised HSCL-8, respectively. According to ROC analysis, the optimal cut-off score of 1.06 for the HSCL-8 resulted in a sensitivity of 88% and specificity of 89% (Kaaya et al., 2002).

The Patient Health Questionnaire-9 (PHQ-9) was included as the screening tool for detecting MDE. This measure, developed by Kroenke et al. (2001) has demonstrated good internal consistency (Cronbach’s alpha ranging from 0.86 to 0.89) and criterion validity (sensitivity

and specificity of 88%) (Kroenke et al., 2001). It has been used in a number of contexts worldwide, also demonstrating reasonable reliability and validity in those settings (Akena et al., 2013; Cholera et al., 2014), although with some variation (Gelaye et al., 2013; Pence et al., 2012).

The MINI, version 6.0, is a semi-structured interview that was used as the gold standard for diagnosis of MDE and was administered by mental health professionals within two weeks of the PHQ-9. Current MDE was defined as meeting MDE criteria according to the MINI within the past two weeks. The MINI is based on DSM-IV criteria and has been used in a wide range of settings as a gold standard measure for validation studies (Akena et al., 2013; Chishinga et al., 2011; Kessler & Bromet, 2013). Depression measures were initially translated from English to Swahili and then independently back-translated to English. Review of the English versions was performed by a translation team including general health care providers as well as mental health care professionals to identify discrepancies and to reach consensus on resolving these issues in the Swahili versions.

### Statistical analysis

Frequencies for categorical variables, as well as means and standard deviations for continuous variables, were reported for sociodemographic characteristics of the study population, including age, gender, and employment status, among other factors. For specific items of the PHQ-9, percentages of those endorsed were reported by gender and comparisons were made using the Fisher's exact test. Cronbach's alpha was used to estimate the internal consistency reliability of the PHQ-9. Receiving operating characteristic (ROC) analysis was applied to examine criterion validity of the PHQ-9, using the MINI assessment for MDE as the gold standard. The area under the ROC curve (AUC) estimates and corresponding 95% confidence intervals (CIs) were calculated. Sensitivity and specificity were calculated at different cut-off scores and the optimal cut-off score was determined based on results from the ROC analysis. Factor analysis with varimax rotation was also performed to examine construct validity of the PHQ-9 among adults accessing primary care in Dar es Salaam. Symptoms with factor loadings 0.40 or greater were included in Table 3, with the exception of redundant items. For these items, we only included those with the highest loading. In addition, correlational analyses were performed using Pearson's correlation coefficient to examine the association between the PHQ-9 and known sociodemographic risk factors for depression.

### Ethical considerations

The study was approved by the Institutional Review Board (IRB) at Harvard Medical School and the National Institute for Medical Research (NIMR) in Tanzania. All participants provided written informed consent.

## Results

The study population was relatively young, with 80% under the age of 35, and predominantly female (80%). Approximately 57% were married or living with a partner. Seven percent were unemployed or had less than a primary level of education. With respect

to their economic situation and living conditions, 82% reported having gas or electricity as an energy source for lighting and over 95% had vinyl, asphalt, or cement for flooring; however, over 75% indicated that they did not have a private flush toilet in their home (see Table 1). The most commonly reported symptoms of depression included poor appetite or over-eating (17%), trouble sleeping (16%), feeling tired or having little energy (15%) and reporting any level of difficulty in functioning as a result of endorsed depressive symptoms (15%) (see Table 2). In general, there was minimal variation by sex for the prevalence of specific depressive symptoms with the exception of "little interest or pleasure in doing things" (women reported 11.5% versus men indicating 0%; p=0.04) (see Table 2).

Regarding reliability, internal consistency of the PHQ-9 in this study population indicated a reasonable Cronbach's alpha of 0.83. The factor structure of the PHQ-9 was examined to assess construct validity of this measure in the study population. The results showed one factor having an eigenvalue greater than one (eigenvalue of 3.3). In the rotated factor pattern, four factors emerged; for the first factor, the items with the highest factor loadings were "moving or speaking slowly/ being fidgety and restless a lot more than usual" (0.62), suicidality (0.61), and "feeling down, depressed, or hopeless" (0.45). Factor two showed the highest loadings for "feeling tired or having little energy" (0.63) and trouble sleeping (0.52). The third factor had the highest loadings for "feeling bad about yourself" (0.57), trouble concentrating" (0.56), and "little interest or pleasure in doing things" (0.50). Factor four showed the highest loading for "poor appetite or over-eating" (0.58) (Table 3).

In addition, correlations between PHQ-9 and known risk factors demonstrated expected associations, suggesting construct validity in this setting to some extent; although not all factors reached statistical significance. Factors such as female gender (r=0.160, p=0.04) and food insecurity (r=0.30, p<0.0001) were associated with a higher level of depressive symptoms. Other variables were not related to the PHQ-9, including age, and other socio-economic proxy measures such as education, employment status, and household characteristics such as type of floor, and source of energy for lighting.

To examine the criterion validity of the PHQ-9 to identify MDE in this population of primary clinic attendees in Dar es Salaam, we conducted ROC analysis. Findings from the PHQ-9 when compared with MDE as identified through mental health professionals administering the MINI diagnostic assessment indicated that this screening tool had a reasonable level of overall accuracy with an AOC estimate of 0.87 (95% CI: 0.77, 0.96) (see Figure 1). With respect to specific cut-off scores, Table 4 demonstrates the sensitivity and specificity estimates at different cut-off scores for the PHQ-9. Based on these findings, the highest sensitivity was 78% (for both cut-off scores of 8 and 9); however, specificity was higher for the cut-off score of 9 (87%), and was considered the optimal cut-off score for this population.

## Discussion

The findings suggest that the PHQ-9 is an appropriate tool for screening of depressive symptoms comparable with a major depressive episode in primary health care settings in Tanzania. Results support a cut-off score of greater than or equal to 9 for depressive

symptoms comparable with major depressive episode within this context, suggesting that a lower level of depressive symptoms are clinically relevant as compared to other study populations, in which the cut-off score was 10 (Gelaye et al., 2013; Hanwella, Ekanayake, & de Silva, 2014). In addition to demonstrating criterion validity, the measure showed good internal consistency reliability and construct validity through correlational and factor analysis.

A number of validation studies of the PHQ-9 have been conducted in other resource-limited settings. Interestingly, the cut-off scores for these studies range from 8 to 12 depending upon the context. For example, a validation study was conducted among adults in Uganda, and the optimal cut-off score was 10, which is the standard cut-off (Akena et al., 2013), with a sensitivity of 91% and specificity of 81%. Similar to the current study, the MINI administered by mental health professionals was used as the gold standard (Akena et al., 2013). In northwest South Africa, the PHQ-9 was validated among individuals with chronic conditions; although specificity was 94%, sensitivity was lower at 49% at a cut-off score of 9. However, reliability was reasonable in this study population (Cronbach's alpha of 0.76). Factors that may have influenced these findings included: 1) unreported length of time between PHQ-9 and SCID administration (gold standard); and 2) one psychologist performing the SCID interview was assisted by a translator, which may have affected the validity of the SCID diagnosis of major depressive disorder (Bhana et al., 2015). Similarly, sensitivity was low among HIV-positive patients in Cameroon (27%), although specificity was estimated as 94% at a cut-off score of 10 (Pence et al., 2012).

In contrast the specificity was lower in a population of adults in Ethiopia (67%), while the sensitivity was higher (86%), for a cut-off score of 10 (Gelaye et al., 2013). Similarly, in Zimbabwe among patients receiving primary care the sensitivity was estimated as 85% and specificity was lower (69%), although the optimal cut-off score in this context was higher (>=11) (Chibanda et al., 2016). A lower optimal cut-off score was reported for a validation study among adults in northern Johannesburg (Cholera et al., 2014); for a cut-off score of 8, sensitivity was estimated as 87% and specificity was 73%. In contrast, using a cut-off score of 12, demonstrated a much lower sensitivity (55%) and a higher specificity (89%) (Cholera et al., 2014). Similar to the current study, Gelaye et al. (2013) reported one factor having an eigenvalue greater than one among adults in Ethiopia. Although the order of factor loadings varied, two symptoms from the study in Ethiopia were included in the first factor for both analyses: feeling down, depressed or hopeless; and suicidal thoughts. Since a cardinal symptom of depression is depressed mood, this suggests some degree of construct validity for the PHQ-9 in the Tanzanian and Ethiopian contexts. In addition, although suicidality is not typically required to meet the diagnosis of depression, it is a critical symptom potentially demonstrating the importance of severity in resource-limited settings.

## Limitations

There are a number of limitations in the present study. One consideration is the fact that the study sample is from a primary health care setting, indicating that the findings may not be as generalizable if our study was conducted at the community level. In addition, 80% of the study population included women, which may also affect the generalizability of these

results. However, findings from prior research with a representative sample of patients accessing services at primary health clinics in Dar es Salaam also demonstrated that women comprised a majority of the study population (Ngoma et al., 2003). Therefore, it is feasible that findings can be generalized to these settings. The PHQ-9 showed a higher specificity (87%) compared to the sensitivity (78%) at a cut-off score of 9. This reduces the likelihood of false positives, which is often a concern in resource-limited settings. However, based on this sensitivity, over 20% of the true positives can be missed. Including additional local terms that reflect depression with the PHQ-9 items may enhance the sensitivity in this context. Statistical power was also limited, given that the number of individuals with MDE in the study population according to the MINI was 18, which was 10.3% of the study population. A strength of the study was the use of a structured clinical assessment administered by mental health professionals as the gold standard for MDE.

### Conclusions

In conclusion, the PHQ-9 is a useful tool for screening for major depressive episode in this context. To a significant extent, reliability and validity of this scale were reasonable. However, future studies may consider enhancing the sensitivity of this measure by including locally derived symptoms of depression. Although it is important to note that the study population was based on a primary health care sample and included predominantly women, we anticipate that it is more reflective of the general population compared to more specialized care settings. Given the high burden of depression in Tanzania and the limited access to care, having a validated screening tool, such as the PHQ-9 can assist professional and lay primary health care workers in identifying patients with symptoms comparable to major depression and promoting necessary referrals for much-needed mental health services.

## Acknowledgements

We would like to express our sincere gratitude to the study participants and to staff related to this study that made this research possible. We also wish to thank the City of Dar es Salaam Health Department for their collaboration through this study.

Funding source

This research was funded with support from the National Institute of Mental Health (Grant No. R01-MH100338).

## Abbreviations

| | |
|---|---|
| **AUC** | Area under the ROC curve |
| **CMDs** | Common mental disorders |
| **DSM-IV-TR** | Diagnostic and Statistical Manual of Mental Disorders-IV-TR |
| **GNI** | Gross National Income |
| **HSCL** | Hopkins Symptom Checklist |
| **MDE** | Major Depressive Episode |
| **MINI** | Mini-International Neuropsychiatric Interview |

| | |
|---|---|
| **PHQ-9** | Patient Health Questionnaire-9 |
| **ROC** | Receiver Operating Characteristic |
| **YLD** | Years Lived with Disability |

## References


Abiodun OA (1993). A study of mental morbidity among primary care patients in Nigeria. Comprehensive Psychiatry, 34(1), 10–13. [PubMed: 8425386]

Akena D, Joska J, Obuku E, & Stein DJ (2013). Sensitivity and specificity of clinician administered screening instruments in detecting depression among HIV-positive individuals in Uganda. AIDS Care, 25(10), 1245–52. [PubMed: 23398282]

Betancourt TS, Scorza P, Meyers-Ohki S, Mushashi C, Kayiteshonga Y, Binagwaho A, Stulac S, Beardslee W (2012). Validating the Center for Epidemiological Studies Depression Scale for Children in Rwanda. J Am Acad Child Adolesc Psychiatry, 51, 1284–1292. [PubMed: 23200285]

Bhana A, Rathod SD, Selohilwe O, Kathree T, Petersen I (2015). The validity of the Patient Health Questionnaire for screening depression in chronic care patients in primary health care in South Africa. BMC Psychiatry, 15, 118. [PubMed: 26001915]

Chibanda D, Verhey R, Gibson LJ, Munetsi E., Machando D, Rusakaniko S, … Abas M (2016). Validation of screening tools for depression and anxiety disorders in a primary care population with high HIV prevalence in Zimbabwe. Journal of Affective Disorders, 198, 50–55. [PubMed: 27011359]

Chishinga N, Kinyanda E, Weiss HA, Patel V, Ayles H, & Seedat S (2011). Validation of brief screening tools for depressive and alcohol use disorders among TB and HIV patients in primary care in Zambia. BMC Psychiatry, 11(1), 75. [PubMed: 21542929]

Cholera R, Gaynes BN, Pence BW, Bassett J, Qangule N, Macphail C, … Miller WC (2014). Validity of the Patient Health Questionnaire-9 to screen for depression in a high-HIV burden primary healthcare clinic in Johannesburg, South Africa. Journal of Affective Disorders, 167, 160–166. [PubMed: 24972364]

Fendrich M, Weissman MM, Warner V (1990). Screening for depressive disorder in children and adolescents: validating the Center for Epidemiologic Studies Depression Scale for Children. Am J Epidemiol, 131, 538–551. [PubMed: 2301363]

Ferrari AJ, Charlson FJ, Norman RE, Patten SB, Freedman G, Murray C, … Whiteford HA (2013). Burden of depressive disorders by country, sex, age, and year: Findings from the Global Burden of Disease Study 2010. PLoS Medicine, 10(11), e1001547. [PubMed: 24223526]

Galea S, Ahern J, Nandi A, Tracy M, Beard J, Vlahov D (2007). Urban neighborhood poverty and the incidence of depression in a population-based cohort study. Annals of Epidemiology, 17(3), 171–179. [PubMed: 17320784]

Gelaye B, Williams MA, Lemma S, Deyessa N, Bahretibeb Y, Shibre T, … Andrew Zhou XH (2013). Validity of the Patient Health Questionnaire-9 for depression screening and diagnosis in East Africa. Psychiatry Research, 210(2), 653–61. [PubMed: 23972787]

Gureje O, Obikoya B, Ikuesan BA (1992). Prevalence of specific psychiatric disorders in an urban primary care setting. East African Medical Journal, 69(5), 282–7. [PubMed: 1644049]

Hanwella R, Ekanayake S, de Silva VA (2014). The validity and reliability of the Sinhala translation of the Patient Health Questionnaire (PHQ-9) and PHQ-2 Screener. Depression Research and Treatment, 2014, 768978. [PubMed: 24795822]

Honikman S, van Heyningen T, Field S, Baron E, Tomlinson M (2012). Stepped care for maternal mental health: a case study of the perinatal mental health project in South Africa. PLoS Medicine, 9(5), e1001222. [PubMed: 22666181]

Jenkins R, Mbatia J, Singleton N, White B (2010). Common mental disorders and risk factors in urban Tanzania. International Journal of Environmental Research and Public Health, 7(6), 2543–2558. [PubMed: 20644689]

Kaaya SF, Smith Fawzi MC, Mbwambo JK, Lee B, Msamanga GI, Fawzi W (2002). Validity of the Hopkins Symptom Checklist-25 amongst HIV-positive pregnant women in Tanzania. Acta Psychiatr Scand, 106(1), 9–19. [PubMed: 12100343]

Kaharuza FM, Bunnell R, Moss S, Purcell DW, Bikaako-Kajura W, Wamai N, … Mermin J (2006). Depression and CD4 cell count among persons with HIV infection in Uganda. AIDS and Behavior, 10(SUPPL. 7), 105–111. [PubMed: 16477512]

Kessler RC, Bromet EJ (2013). The epidemiology of depression across cultures. Annual Review of Public Health, 34, 119–138.

Kisely S, Scott A, Denney J, Simon G (2006). Duration of untreated symptoms in common mental disorders: association with outcomes: international study. The British Journal of Psychiatry: The Journal of Mental Science, 189, 79–80. [PubMed: 16816310]

Kohrt BA, Luitel NP, Acharya P, Jordans MJ (2016). Detection of depression in low resource settings: validation of the Patient Health Questionnaire (PHQ-9) and cultural concepts of distress in Nepal. BMC Psychiatry, 16, 58. [PubMed: 26951403]

Kroenke K, Spitzer RL, Williams JBW (2001). The PHQ-9: Validity of a brief depression severity measure. Journal of General Internal Medicine, 16(9), 606–613. [PubMed: 11556941]

Marcus M, Yasamy MT, van Ommeren M, Chisholm D (2012). Depression, a global public health concern. Retrieved from http://www.who.int/mental_health/management/depression/who_paper_depression_wfmh_2012.pdf

Moussavi S, Chatterji S, Verdes E, Tandon A, Patel V, Ustun B (2007). Depression, chronic diseases, and decrements in health: results from the World Health Surveys. The Lancet, 370(9590), 851–858.

Murthy RS, Lakshminarayana R (2006). Mental health consequences of war: a brief review of research findings. World Psychiatry, 5(9), 25–30. [PubMed: 16757987]

Myer L, Smit J, Roux L, Parker S, Stein DJ, Seedat S (2008). Common mental disorders among HIV-infected individuals in South Africa: prevalence, predictors, and validation of Brief Psychiatric Rating Scales. AIDS Patient Care and STDs, 22(2), 147–158. [PubMed: 18260806]

Ngoma MC, Prince M, & Mann A (2003). Common mental disorders among those attending primary health clinics and traditional healers in urban Tanzania. The British Journal of Psychiatry, 183(4), 349–355. [PubMed: 14519614]

Pence BW, Gaynes BN, Atashili J, O'Donnell JK, Tayong G, Kats D, … Ndumbe PM (2012). Validity of an interviewer-administered Patient Health Questionnaire-9 to screen for depression in HIV-infected patients in Cameroon. Journal of Affective Disorders, 143(1–3), 208–13. [PubMed: 22840467]

Rasmussen A, Eustache E, Raviola G, Kaiser B, Grelotti DJ, Belkin GS (2015). Development and validation of a Haitian Creole screening instrument for depression. Transcultural Psychiatry 2015, 52(1), 33–57. [PubMed: 25080426]

Saraceno B, van Ommeren M, Batniji R, Cohen A, Gureje O, Mahoney J, … Underhill C (2007). Barriers to improvement of mental health services in low-income and middle-income countries. The Lancet, 370(9593), 1164–1174.

Saxena S, Thornicroft G, Knapp M, Whiteford H (2007). Resources for mental health: scarcity, inequity, and inefficiency. The Lancet, 370(9590), 878–889.

Simmons LA, Braun B, Charnigo R, Havens JR, Wright DW (2008). Depression and poverty among rural women: A relationship of social causation or social selection? Journal of Rural Health, 24(3), 292–298. [PubMed: 18643807]

UNICEF (2014). The state of the world's children in Numbers. Retrieved December 3, 2014 from http://www.unicef.org/sowc2014/numbers/.

Vos T, Barber RM, Bell B, Bertozzi-Villa A, Biryukov S, Bolliger I, … Murray C (2015). Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: A systematic analysis for the Global Burden of Disease Study 2013. The Lancet, 386(9995), 743–800.

**Highlights:**

- Major depression exacts a significant burden in resource-limited settings
- The PHQ-9 is a screening tool for depression that can increase access to care
- Validity of the PHQ-9 was demonstrated through 78% sensitivity and 87% specificity
- The PHQ-9 can be used as a screening tool in primary care settings in Tanzania
- Identification of depression in Tanzania can significantly increase access to care

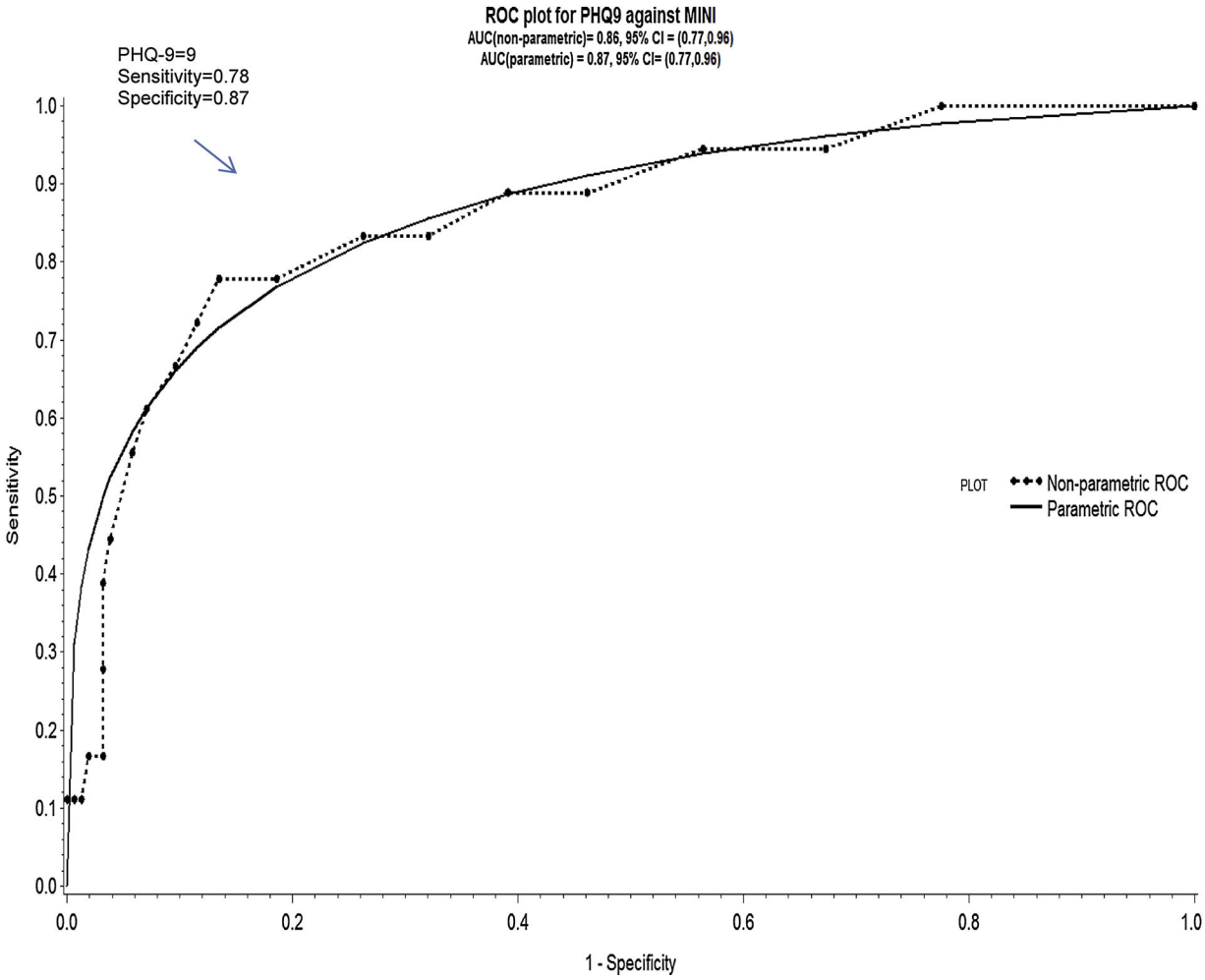


**Figure 1.**
Receiver Operating Characteristic curve for the PHQ-9 in Tanzania

**Table 1.**

Socio-demographic characteristics of study participants, Dar es Salaam, Tanzania

| | n or mean ± SD[1] (n=174) | % |
|---|---|---|
| ***Age*** *(range 18–63 years)* | 28.9 ± 8.9 | |
| *18–24* | 64 | 36.8 |
| *25–34* | 75 | 43.1 |
| *35+* | 35 | 20.1 |
| ***Gender*** | | |
| *Male* | 35 | 20.1 |
| *Female* | 139 | 79.9 |
| ***Marital Status*** | | |
| *Legally or traditionally married* | 87 | 50.0 |
| *Cohabiting* | 12 | 6.9 |
| *Single/Divorced/Separated/Widowed* | 75 | 43.1 |
| ***Did not complete primary education*** | 12 | 6.9 |
| ***Employment Status*** | | |
| *Employed/Self-employed* | 107 | 61.5 |
| *Housewife* | 43 | 24.7 |
| *Unemployed* | 12 | 6.9 |
| *Student* | 12 | 6.9 |
| ***Energy sources for lighting***[2] | | |
| *Gas/Electricity* | 131 | 81.9 |
| *Solar/Paraffin/Other* | 29 | 18.1 |
| ***Main material of the floor in household*** | | |
| *Earth and/or sand* | 6 | 3.4 |
| *Rudimentary floor* | 2 | 1.1 |
| *Vinyl or asphalt strips* | 138 | 79.3 |
| *Cement screed* | 28 | 16.1 |
| ***Significant problems securing food*** | 4 | 2.3 |
| ***Lack of a private flush toilet*** | 131 | 75.3 |

1. SD: Standard Deviation

2. 14 participants did not provide this information.

**Table 2.**

Endorsement frequencies of depressive symptoms and functioning from the PHQ-9

| PHQ-9[1] items | Endorsement Frequencies | | | | | | |
|---|---|---|---|---|---|---|---|
| | *All (n=174)* | | *Men (n=35)* | | *Women (n=139)* | | *p-value*[2] |
| | *N* | *(%)* | *N* | *(%)* | *N* | *(%)* | |
| Poor appetite or over-eating | 29 | (16.7) | 5 | (14.3) | 24 | (17.3) | 0.80 |
| Trouble falling asleep or staying asleep, or sleeping too much | 27 | (15.5) | 7 | (20.0) | 20 | (14.4) | 0.44 |
| If any problems were checked off, have these problems made it difficult for you to do your work, take care of things at home, or get along with other people[3] | 23 | (15.1) | 3 | (8.6) | 20 | (14.4) | 0.57 |
| Feeling tired or having little energy | 26 | (14.9) | 2 | (5.7) | 24 | (17.3) | 0.11 |
| Feeling down, depressed, or hopeless | 26 | (14.9) | 4 | (11.4) | 22 | (15.8) | 0.61 |
| Feeling bad about yourself, or that you are a failure or have let yourself or your family down | 16 | (9.2) | 4 | (11.4) | 12 | (8.6) | 0.74 |
| Little interest or pleasure in doing things | 16 | (9.2) | 0 | (0.00) | 16 | (11.5) | 0.04 |
| Moving or speaking so slowly / being so fidgety and restless a lot more than usual | 11 | (6.3) | 1 | 2.9 | 10 | (7.2) | 0.70 |
| Thoughts that you would be better off dead or of hurting yourself in some way | 8 | (4.6) | 0 | (0.00) | 8 | (5.8) | 0.36 |
| Trouble concentrating on things | 6 | (3.5) | 1 | (2.9) | 5 | (3.6) | 1.00 |

*1.* Patient Health Questionnaire-9

*2.* p-values were calculated using the Fisher's exact test.

*3.* 22 participants (5 men and 17 women) have missing values for this item.

**Table 3.**

Factor structure of the PHQ-9 among primary health clinic attendees in Dar es Salaam, Tanzania (n=174)

| **PHQ-9**[1] **items** | ***Factor Loadings*** |
|---|---|
| *Factor 1: Depressed mood/ self-harm (explained 76% of variance)* | |
| Moving or speaking so slowly / being so fidgety and restless a lot more than usual | 0.62 |
| Thoughts that you would be better off dead or of hurting yourself in some way | 0.61 |
| Feeling down, depressed, or hopeless | 0.45 |
| *Factor 2: Physiological symptoms (explained 13% of variance)* | |
| Feeling tired or having little energy | 0.63 |
| Trouble falling asleep or staying asleep, or sleeping too much | 0.52 |
| *Factor 3: Negative/impaired cognitions & anhedonia (explained 6% of variance)* | |
| Feeling bad about yourself, or that you are a failure or have let yourself or your family down | 0.57 |
| Trouble concentrating on things | 0.56 |
| Little interest or pleasure in doing things | 0.50 |
| *Factor 4: Appetite (explained 5% of variance)* | |
| Poor appetite or over-eating | 0.58 |

[1] PHQ-9: Patient Health Questionnaire-9

**Table 4.**

Sensitivity, specificity, positive predictive value, and negative predictive value of depression symptoms (95% CIs[1]) for various PHQ-9 cut-off scores

| PHQ-9 cut-off score | *Sensitivity* | *Specificity* | *Positive predictive value* | *Negative predictive value* |
|---|---|---|---|---|
| ≥ 7 | 0.83 (0.59, 0.96) | 0.74 (0.66, 0.80) | 0.27 (0.16, 0.40) | 0.97 (0.93, 0.99) |
| ≥ 8 | 0.78 (0.52, 0.94) | 0.81 (0.74, 0.87) | 0.33 (0.19, 0.49) | 0.97 (0.92, 0.99) |
| ≥ 9 | 0.78 (0.52, 0.94) | 0.87 (0.80, 0.92) | 0.42 (0.24, 0.58) | 0.97 (0.93, 0.99) |
| ≥ 10 | 0.72 (0.47, 0.90) | 0.88 (0.82, 0.93) | 0.40 (0.25, 0.61) | 0.97 (0.92, 0.99) |
| ≥ 11 | 0.67 (0.41, 0.87) | 0.90 (0.85, 0.95) | 0.44 (0.26, 0.65) | 0.96 (0.91, 0.99) |

1. Confidence Intervals