



THE AGA KHAN UNIVERSITY

eCommons@AKU

Institute for Human Development

AKU in East Africa

12-2020

Validation of a Swahili version of the 9-item Patient Health Questionnaire (PHQ-9) among adults living with HIV compared to a community sample from Kilifi, Kenya

Paul Mwangi

Moses K. Nyongesa

Hans M. Koot

Pim Cuijpers

Charles R.J.C. Newton

See next page for additional authors

Follow this and additional works at: https://ecommons.aku.edu/eastafrica_ihd

 Part of the [Public Health Commons](#)

Authors

Paul Mwangi, Moses K. Nyongesa, Hans M. Koot, Pim Cuijpers, Charles R.J.C. Newton, and Abubakar Abdillah



Research Paper

Validation of a Swahili version of the 9-item Patient Health Questionnaire (PHQ-9) among adults living with HIV compared to a community sample from Kilifi, Kenya



Paul Mwangi^{a,1}, Moses K. Nyongesa^{b,c,1,*}, Hans M. Koot^c, Pim Cuijpers^c, Charles R.J.C. Newton^{b,d,e}, Amina Abubakar^{b,d,e,f}

^a Department of Biostatistics, University of Hasselt, Hasselt, Belgium

^b Neurosciences Group, KEMRI-Wellcome Trust Research Programme, Centre for Geographic Medicine Research (Coast), Box 230, Kilifi, Kenya

^c Department of Clinical, Neuro- and Developmental Psychology, Amsterdam Public Health Research Institute, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands

^d Department of Public Health, Pwani University, Kilifi, Kenya

^e Department of Psychiatry, University of Oxford, Oxford, UK

^f Institute for Human Development, Aga Khan University, Nairobi, Kenya

ARTICLE INFO

Keywords:

Psychometric properties
Factor analysis
Depressive symptoms
HIV infections
Adults
Kenya

ABSTRACT

Background: Depression remains under-investigated in people living with HIV in sub-Saharan Africa due to paucity of adequately validated measures. This study aimed to validate an adapted version of the 9-item Patient Health Questionnaire (PHQ-9) among adults living with HIV compared to those from the community in Kilifi, Kenya.

Methods: Analysis of data from 450 adults living with HIV and 337 adults from the community was conducted examining the reliability, factorial structure, measurement invariance and discriminant validity of interviewer-administered PHQ-9, Swahili version.

Results: Internal consistency of the Swahili PHQ-9 was good overall, in adults living with HIV and those from the community (Macdonald's omega > 0.80). The two-week test-retest reliability was acceptable among adults living with HIV ($ICC = 0.64$). A one-factor confirmatory factor analysis (CFA) model indicated the Swahili PHQ-9 was unidimensional in the overall sample, in adults living with HIV and those from the community. Multi-group CFA substantiated measurement invariance of this unidimensional scale across participant group (adults living with HIV vs. community), sex (females vs. males) and age category (young, middle-age and elderly adults). The Swahili PHQ-9 exhibited good discriminant validity between the two participant groups.

Limitations: No diagnostic interview for mental disorders was administered in the original studies limiting analysis of sensitivity and specificity of the Swahili PHQ-9.

Conclusion: The Swahili PHQ-9 is a reliable and valid unidimensional scale. It appears a valuable tool for assessing depressive symptoms that can be generalized across different demographic groups, in primary HIV clinics and the general community within this and similar settings.

Author statement

Contributors

PM, MKN, CN and AA conceptualized this study. MKN and AA supervised data collection for the original studies. PM managed the study data. PM and MKN analyzed the data. HK, PC, CN and AA contributed to the interpretation of the data. PM and MKN wrote

the first draft of the manuscript. HK, PC, CN and AA critically reviewed subsequent versions of the manuscript. All the authors have approved the final version of the manuscript for submission.

Role of funding source

The parent studies were supported through funding from the Medical Research Council (Grant number MR/M025454/1) to AA. This award is jointly funded by the UK Medical Research Council (MRC) and the UK Department for International Development.

* Corresponding author at: Neurosciences Group, KEMRI-Wellcome Trust Research Programme, Centre for Geographic Medicine Research (Coast), Box 230, Kilifi, Kenya.

E-mail address: Mkachama@kemri-wellcome.org (M.K. Nyongesa).

¹ These authors contributed equally to this work.

opment (DFID) under MRC/DFID concordant agreement and is also part of the EDCTP2 program supported by the European Union. During this work MKN was supported by funding from the Wellcome Trust International Master's Fellowship (Grant number 201310/Z/16/Z) and the DELTAS Africa Initiative [DEL-15-003]. The DELTAS Africa Initiative is an independent funding scheme of the African Academy of Sciences (AAS)'s Alliance for Accelerating Excellence in Science in Africa (AESA) and supported by the New Partnership for Africa's Development Planning and Coordinating Agency (NEPAD Agency) with funding from the Wellcome Trust [107769/Z/10/Z] and the UK government. The funders did not have a role in the design and conduct of the study or interpretation of study findings. The views expressed in this publication are those of the author(s) and not necessarily those of AAS, NEPAD Agency, Wellcome Trust or the UK government.

1. Introduction

Depression is projected to be the leading cause of disease burden worldwide by the year 2030 (Mathers and Loncar, 2006). As of 2018, the World Health Organization (WHO) estimated that over 300 million people of all ages across the globe suffered from depression (WHO, 2018). In Kenya, 1.9 million people are estimated to be suffering from depression, equivalent to a prevalence of 4.4% (WHO, 2017). People living with HIV/AIDS (PLWHA), compared to the general population, are 2–3 times more likely to suffer from this neuropsychiatric disorder (Bernard et al., 2017). Among PLWHA depression may arise from various factors including receiving a diagnosis of a chronic incurable disease, the challenges of adhering to lifetime antiretroviral medication, perceived stigma and discrimination associated with HIV or poor prognostic course of the disease such as advancement from HIV to AIDS and virological non-suppression (Global Initiative on Psychiatry, 2006; Stutterheim et al., 2009).

In sub-Saharan Africa (SSA) where the HIV/AIDS burden is greatest (UNAIDS, 2018), studies investigating depression or its severity among PLWHA are few, but research efforts are steadily increasing since 2010 (Bernard et al., 2017). Brief screening tools, developed in the West, are now commonly used to identify depressive symptoms in PLWHA, including those from SSA (Amare et al., 2018; Bernard et al., 2017). These include the 9-item Patient Health Questionnaire (PHQ-9) (Kroenke et al., 2001), Beck's Depression Inventory-II (BDI-II) (Beck et al., 1988), and the Center for Epidemiologic Studies Depression Scale (CES-D) (Radloff, 1977).

However, validation of local versions of these tools – a systematic and time-consuming process (Abubakar and Van De Vijver, 2017) – in settings such as SSA is rarely conducted (Pence et al., 2012; Tsai, 2014), and this may explain the dearth of literature on depression among PLWHA in this setting. Because of constraints including time and funding, or convenience, some investigators opt to use the original English versions of these depression screening measures or directly translate and use these measures developed in a different context, without checking their content, conceptual, semantic or idiomatic equivalence (Abubakar and Van De Vijver, 2017). This approach may be problematic because of contextual differences. In settings such as SSA, for example, distress is thought to be commonly expressed through somatic symptoms and local idioms (Patel and Wesseley, 1998). To contribute meaningfully to the epidemiological data on depression in the SSA context, it is paramount that researchers use contextually appropriate, and locally validated mental health measures adapted from elsewhere.

Accordingly, this study set out to assess the validity of a Swahili version of the PHQ-9 among adults living with HIV compared to adults sampled from the community in Kilifi, Coast of Kenya. We specifically chose to validate the PHQ-9 because it is a brief, easy to administer screening tool but also aligning with the 2016 guidelines from the Kenyan government recommending the use of the PHQ-9 as a screener for depressive symptoms in HIV primary care clinics (Ministry of Health, 2016). Addi-

tionally, many of the symptom items included in the PHQ-9 are relevant to the local conceptualization of depression in our setting (Bitta et al., 2019). In Kenya, Swahili is a national language and is used across different regions as the official communication language in addition to English. To the best of our knowledge, there is no study that has validated a Swahili version of the PHQ-9 among PLWHA in Kenya (but also within the larger Swahili speaking African countries). A previous study among PLWHA conducted in Western Kenya by Monahan et al. (2009), administered and investigated the reliability and validity of the English version of the PHQ-9. Omoro et al. (2006), translated and validated a Swahili version of the PHQ-9 but among only 48 head and neck cancer patients. Moreover, these two studies did not look at the factorial structure or measurement invariance of the PHQ-9.

Taking previous research forward, we specifically report on: i) the internal consistency; ii) test-retest reliability; iii) factorial structure; iv) measurement invariance; and v) discriminant validity of Swahili PHQ-9 among PLWHA compared to a community sample from Coastal Kenya.

2. Methods

2.1. Study setting and participants

Data used in this analysis were obtained from three studies conducted between February 2017 and October 2018 at the Center for Geographic Medicine Research (CGMR) in Kilifi County. Kilifi County is located at the Kenyan Coast and is largely a rural setting with a total population of 1.5 million people (Kenya National Bureau of Statistics, 2019). The HIV prevalence in Kilifi is estimated at 4.5% (1 in every 22) and this is slightly lower than the national average prevalence of 6% (National AIDS Control Council, 2016). The first study was cross-sectional in design looking at health-related quality of life and mental health outcomes among adults 18–60 years old living with HIV ($n = 450$). Participant recruitment procedure for this study is reported elsewhere (Nyongesa et al., 2020). The second study is an ongoing longitudinal study dubbed Adolescent Health Outcomes Study (AHOS). This study is investigating adolescent executive functioning and its association with various outcomes such as scholastic, mental health, health risk behavior, and medication adherence recruiting adolescent and primary caregiver dyads. From this longitudinal study, we used the baseline data of primary caregivers of adolescents. These caregivers were aged 20–64 years ($n = 199$). Participant recruitment into AHOS study is described in detail elsewhere (Sewanyana et al., 2020). The third study was a validation study of neurocognitive and neuropsychiatric measures among adolescents with sickle cell disease in Kilifi (main analysis and write up still ongoing), where adolescent and primary caregiver dyads were randomly recruited. From this study, we used data of primary caregivers of adolescents who were community controls. These caregivers were aged 25–80 years ($n = 138$). The total sample utilized from all the three studies was 787, including 450 adults living with HIV and 337 primary caregivers of adolescents identified from within the Kilifi Demographic and Health Surveillance System (Scott et al., 2012).

2.2. Measures relevant to the present analysis

Sociodemographic characteristics: In all three studies, a sociodemographic questionnaire capturing data on participant age, sex, level of education, employment, religion and marital status was administered by well-trained research assistants in a face-to-face interview. Additionally, an asset index form was used to collect information on disposable assets owned by participants (or their family) like television, radio, bicycle or motorbike as a proxy indicator of their socioeconomic status. HIV-related clinical data of adults living with HIV are reported elsewhere in detail (Nyongesa et al., 2020).

Depressive symptoms: PHQ-9 (Kroenke et al., 2001) was administered as a measure of depressive symptoms to all the study participants in the three aforementioned studies. It assesses depressive symptoms

experienced in the preceding two weeks. All the 9 PHQ items are rated on a Likert scale of “0” (not at all) to “3” (nearly every day). Item scores are summated to derive a total score that ranges from 0 to 27. In terms of severity, scores of 5–9, 10–14, 15–19, and 20–27 are indicative of mild, moderate, moderately severe, and severe levels of depressive symptoms, respectively. A cut-off score of ≥ 10 has been considered optimal for defining positive depression screen within East Africa (Akena et al., 2013; Gelaye et al., 2013). Conforming with international guidelines for translation of tools in health research (http://www.who.int/substance_abuse/research_tools/translation/en/), the English version of the PHQ-9 was independently translated to Swahili by two experienced translators fluent in both languages. Back-translations into English were done by another independent pair of translators. A harmonization meeting was then held including all the translators and researchers from the department (knowledgeable about the Kenyan culture, bilingual and fluent in both English and Swahili). Discrepancies in the translations were resolved by consensus. Where appropriate, additional locally relatable examples were added such as “reading the Bible or Quran” to item 7 of the PHQ-9 “trouble concentrating on things, such as reading the newspaper or watching television”. The final Swahili version of the PHQ-9 was interviewer-administered due to the low literacy level in the study setting.

2.3. Ethical statement

Ethical approval to conduct the three studies mentioned above was individually sought and granted by the local institutional review board, Scientific and Ethics Review Unit (SERU). The following are the approval reference numbers: KEMRI/SERU/CGMR-C/108/3594, KEMRI/SERU/CGMR-C/084/3454, and KEMRI/SERU/CGMR-C/031/3193. All participants in these studies provided informed written consent.

2.4. Statistical analysis

Study data were collected and managed using research electronic data capture (REDCap) hosted at the Center for Geographic Medicine Research-Coast. Descriptive analysis to summarize sample characteristics were performed using STATA version 15 (Stata Corp, College Station LP, USA). Categorical variables were compared using Pearson's chi-square test. Normally and non-normally distributed continuous variables were compared using either independent Student's *t*-test or Mann Whitney U test, as appropriate. All the psychometric analyses were performed using an open-source software for statistical computing and graphics, R version 3.5.0 (R Development Core Team: <https://www.r-project.org/>). The internal consistency of PHQ-9 scales was assessed using both Cronbach's alpha (α) and Macdonald's omega (ω), which were computed using “psych” package for R (Revelle, 2017). A threshold for Cronbach's α and Macdonald's $\omega > 0.70$ was considered satisfactory according to Nunnally and Bernstein criterion (Hair et al., 2010). The latter (ω) are considered the primary measures of reliability in this work. Intra-class correlation coefficient (ICC) assessed test-retest reliability of the Swahili PHQ-9 (using data of 60 randomly selected adult participants living with HIV rescreened 2 weeks following initial assessment). For factorial structure, we used confirmatory factor analysis (CFA) to test the known one-factor model of PHQ-9 (Gelaye et al., 2013; Yu et al., 2012). Multi-group CFA was used to investigate measurement invariance of Swahili PHQ-9 across participant group (adults living with HIV vs. those from the community), sex (female vs. male) and age category (young, middle-age, and elderly adults). We fitted these models using R lavaan package version 05–23 (Rosseel, 2012). We used diagonally weighted least squares (DWLS) estimator to account for the ordinal nature of the PHQ-9 Likert scale. We standardized the latent factor, allowing free estimation of all factor loadings. For all CFA analyses, three

goodness of fit indices; root mean square error of approximation (RMSEA), comparative fit index (CFI) and Tucker-Lewis index (TLI) were used to assess fit of the models. RMSEA values < 0.05 indicated good fit, and values as high as 0.08 represented reasonable errors of approximation (Chen et al., 2008) while CFI, and TLI values > 0.9 were considered representative of adequately fitting model (Jackson et al., 2009; Kline, 2005). In multi-group CFA, we tested configural, metric and scalar invariances. We did not test strict invariance since invariance of the item residuals is inconsequential to interpretation of latent mean differences as residuals are not part of the latent factor (Vandenberg and Lance, 2000). One incremental index for testing measurement of invariance; the comparative fit index (CFI), were obtained to compare constrained models in this sequence of invariance tests. If the change in CFI (Δ CFI) was ≤ 0.01 in magnitude, the two models were considered to have an equivalent fit (Chen, 2007). For discriminant validity, analysis of covariance (ANCOVA) was used to evaluate whether Swahili PHQ-9 is sensitive to participant group (adults living with HIV vs. those from the community). Less than 5% significance level was used as the cut-off for all tests of hypothesis.

3. Results

3.1. Sample characteristics

Table 1 depicts participant sociodemographic characteristics. Of the total sample ($N = 787$), 57.2% were adults living with HIV. The majority were females (79.7%), married (85.8%) and Christians (78.9%). The mean age of the study participants was 41.9 years (Standard Deviation [SD]=9.5). Over a quarter of the study participants had not attended formal education (34.1%) while only 7.8% were unemployed. There were significant sociodemographic differences between adults from the community and those living with HIV in terms of age, educational level, employment, religion and asset index score (p -value < 0.01 ; Table 1).

3.2. Distribution of PHQ-9 scores

As shown in Table 2, the overall median score value on the PHQ-9 was 3 (IQR=0 - 7). In terms of depressive symptom severity, overall scores were in the expected direction with majority of participants (62.9%) within the normal range score (0–4) and minority (0.5%) with scores in the severe range (between 20 and 27). Using the optimal cut-off score of ≥ 10 , 9.7% of the study participants had depressive symptoms. There were statistically significant differences between adults from the community and those living with HIV across PHQ-9 raw scores, severity levels, and positive screen for depression, with higher levels for the HIV group (Table 2; p -value < 0.01).

3.3. Item characteristics

Fig. 1 presents the distribution of responses across the ordinal Likert scale options for the 9 items of the Swahili PHQ-9. Responses were distributed across all the Likert options in the expected direction. The correlations of each item with the rest of the items ranged from 0.24 to 0.48, and all were statistically significant.

3.4. Reliability

The Swahili PHQ-9 showed an overall good internal consistency with Cronbach's $\alpha = 0.84$ (95% Confidence Interval [95% CI]: 0.82 - 0.86) and Macdonald's $\omega = 0.84$ (95% CI: 0.82 - 0.87). The internal consistency of this scale was slightly lower among adults living with HIV, female participants and elderly adults aged 50 to 80 years, but all within acceptable range (Table 3). Test-retest reliability of the Swahili PHQ-9 among adults living with HIV was acceptable, ICC=0.64 (95% CI: 0.46 - 0.77).

Table 1
Participant sociodemographic characteristics by group.

Characteristic	Overall sample (n = 787)	Community sample (n = 337)	Adults living with HIV (n = 450)	Test statistic
<i>Age in years</i>				
Mean ± (SD)	41.89 ± 9.52	40.77 ± 9.10	42.74 ± 9.75	
Median (IQR)	42 (35 - 48)	41 (34 - 45)	43 (36 - 50)	$U = 87776.5, p < 0.001$
Min - Max	18 - 80	20 - 80	18 - 60	
<i>Sex, n (%)</i>				
Male	160 (20.3)	66 (19.6)	94 (20.9)	$\chi^2_{(1)} = 0.202, p = 0.653$
Female	627 (79.7)	271 (80.4)	356 (79.1)	
<i>Level of education, n (%)</i>				
No education	268 (34.1)	146 (43.5)	122 (27.2)	$\chi^2_{(3)} = 36.366, p < 0.001$
Primary	405 (51.6)	166 (49.4)	239 (53.2)	
Secondary	84 (10.7)	18 (5.4)	66 (14.7)	
Tertiary	28 (3.6)	6 (1.8)	22 (4.9)	
<i>Employment, n (%)</i>				
Formal employment	303 (38.9)	23 (6.9)	280 (62.2)	$\chi^2_{(2)} = 319.025, p < 0.001$
Non-formal employment	415 (53.3)	298 (90.6)	117 (26.0)	
Not employed	61 (7.8)	8 (2.4)	53 (11.8)	
<i>Religion, n (%)</i>				
Traditional/Non-believers	132 (16.8)	72 (21.4)	60 (13.4)	$\chi^2_{(2)} = 60.929, p < 0.001$
Christians	619 (78.9)	230 (68.5)	389 (86.6)	
Islam	34 (4.3)	35 (10.1)	0 (0.0)	
<i>Marital status, n (%)</i>				
Married	674 (85.8)	282 (83.6)	393 (87.3)	$\chi^2_{(1)} = 2.158, p = 0.142$
Not married	113 (14.3)	55 (16.4)	57 (12.7)	
<i>Total assets</i>				
Mean ± SD	0.97 ± 1.23	0.60 ± 0.89	1.24 ± 1.37	
Median (IQR)	1 (0 - 2)	1 (0 - 2)	0 (0 - 1)	$U = 95745.5, p < 0.001$

Notes. SD - Standard Deviation, IQR- Interquartile Range, χ^2 - Pearson Chi-Square, U- Mann Whitney U test.

Table 2
Distribution of Swahili PHQ-9 scores by participant group.

	Both groups	Adults from the community	Adults living with HIV	Test statistic
<i>PHQ-9 total score</i>				
Mean ± SD	4.02 ± 4.02	3.49 ± 3.82	4.41 ± 4.43	
Median (IQR)	3 (0 - 7)	2 (0 - 6)	3 (1 - 7)	$U = 86,103, p < 0.001$
<i>Severity scores, n (%)</i>				
Normal, 0 - 4	495 (62.9)	214 (63.5)	281 (62.4)	$\chi^2_{(4)} = 24.21, p < 0.001$
Mild, 5 - 9	216 (27.4)	109 (32.3)	107 (23.8)	
Moderate, 10 - 14	54 (6.9)	11 (3.3)	43 (9.6)	
Moderately severe, 15 - 19	18 (2.3)	2 (0.6)	16 (3.6)	
Severe, 20 - 27	4 (0.5)	1 (0.3)	3 (0.7)	
<i>Depressive symptoms, n (%)</i>				
No (PHQ-9 score ≤ 9)	711 (90.3)	323 (95.8)	388 (86.2)	$\chi^2_{(1)} = 19.37, p < 0.001$
Yes (PHQ-9 score ≥ 10)	76 (9.7)	14 (4.2)	62 (13.8)	

Notes. χ^2 - Pearson Chi-Square, U- Mann Whitney U test.

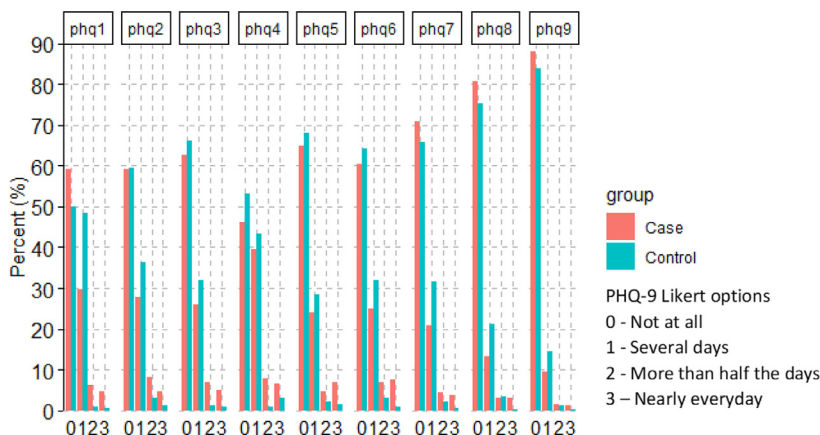


Fig. 1. Percent response across the Likert options for the 9 items of Swahili PHQ-9.

Table 3
Internal consistency of Swahili PHQ-9.

	Cronbach's alpha (95% CI)	Macdonald's Omega (95% CI)
Overall	0.84 (0.82 - 0.86)	0.84 (0.82 - 0.87)
Group		
Adults from the community	0.90 (0.87 - 0.92)	0.90 (0.87 - 0.93)
Adults living with HIV	0.81 (0.79 - 0.84)	0.81 (0.78 - 0.84)
Sex		
Females	0.83 (0.81 - 0.85)	0.84 (0.82 - 0.86)
Males	0.86 (0.78 - 0.90)	0.86 (0.77 - 0.91)
Age category		
Young adults (18 - 35 years)	0.85 (0.82 - 0.88)	0.85 (0.81 - 0.88)
Middle-age adults (36 - 49 years)	0.83 (0.80 - 0.86)	0.84 (0.81 - 0.87)
Elderly adults (50 - 80 years)	0.83 (0.78 - 0.86)	0.84 (0.77 - 0.88)

Notes. 95% CI - 95% Confidence Interval.

Table 4
Item factor loadings under one-factor confirmatory factor analysis model.

Item code	Factor loadings Item description*	Overall sample	Community sample	Adults living with HIV
phq1	Little interest or pleasure (anhedonia)	0.604	0.631	0.611
phq2	Feeling down, depressed, or hopeless	0.688	0.728	0.673
phq3	Trouble falling/staying asleep/sleeping too much	0.632	0.748	0.578
phq4	Feeling tired or having little energy (fatigue)	0.651	0.756	0.602
phq5	Poor appetite or overeating	0.550	0.685	0.476
phq6	Feeling bad about self	0.637	0.775	0.577
phq7	Trouble concentrating	0.618	0.732	0.561
phq8	Moving or speaking very slowly	0.578	0.690	0.509
phq9	Suicidal thoughts	0.447	0.516	0.392

Notes.

* item requires respondents to consider the past 2 weeks.

3.5. Factorial structure

Prior to CFA, we examined sampling adequacy using the Kaiser-Meyer-Olkin (KMO) value and Bartlett's test of sphericity. The KMO value was 0.89 while Bartlett's test of sphericity was highly significant ($\chi^2 = 2045.50$, p -value < 0.001) suggesting that there were correlations within the dataset appropriate for factor analysis. Under CFA, all the Swahili PHQ-9 items loaded well on one-factor solution (item factor loadings > 0.3 ; Table 4). All goodness of fit indices for the one-factor structure of the Swahili PHQ-9 were not only excellent for the overall sample, but also among adults living with HIV and those from the community (Table 5). These patterns of results suggest the Swahili PHQ-9 is an unidimensional scale.

3.6. Measurement invariance across participant group, sex and age category

Overall, as shown in Table 5, multi-group CFA confirmed measurement invariance of the unidimensional Swahili PHQ-9 across participant group (adults living with HIV vs. those from the community), sex (females vs. males) and age category (young, middle-age, and elderly adults).

In a model assuming the same item-factor assignment (configural invariance), the one-factor solution of Swahili PHQ-9 fitted the data well across participant group, sex and age category (Table 5). Assuming same item-factor assignment and additionally constraining factor loadings to equivalence across participant group, sex and age category (metric invariance), all the goodness of fit indices for the one-factor solution of the Swahili PHQ-9 were excellent (Table 5). Comparing this metric invariance model to the least stringent configural invariance model, ΔCFI were 0.002, < 0.001 and 0.003 for participant group, sex and age category, respectively.

Additionally, constraining item intercepts to equivalence across participant group, sex and age category to the model above (scalar invariance), all goodness of fit indices for the one-factor solution of Swahili

PHQ-9 were excellent (Table 5). Comparing this scalar invariance model to the less restrictive metric invariance model, ΔCFI were 0.004, < 0.001 and 0.007 for participant group, sex and age category, respectively.

3.7. Discriminant validity

Raw total scores on the Swahili version of PHQ-9 were significantly higher among adults living with HIV compared to those from the community ($U = 86,103$, p -value < 0.001 ; Table 2). Controlling for the effect of age, education, employment, and socioeconomic status, the observed differences remained statistically significant, $F(1777) = 8.53$, p -value = 0.004. This initial finding suggests good discriminant validity of Swahili PHQ-9 total score.

4. Discussion

4.1. Summary of study findings

Studies validating measures of depression or its severity for use with PLWHA in SSA are limited. To fill this gap, this study set out to validate a local version of the PHQ-9 in Kenya. The specific objectives were to determine the internal consistency, test-retest reliability, factorial structure, measurement invariance and discriminant validity of an interviewer-administered Swahili PHQ-9 among PLWHA compared to a community sample from Kilifi, Kenya. We found that the Swahili PHQ-9 had good internal consistency overall, in both adults living with HIV and those in the community sample. The test-retest reliability was acceptable among adults living with HIV. In terms of factorial structure and measurement invariance, we found the Swahili PHQ-9 to be a unidimensional scale that is invariant across participant group, sex and age categories. Swahili PHQ-9 was very sensitive and could discriminate adults living with HIV and those in the community sample.

Table 5

Goodness of fit indices for one-factor structure and measurement invariance of Swahili PHQ-9 using confirmatory factor analysis.

	Chi-square statistic	RMSEA ^a	TLI ^b	CFI ^b	ΔCFI ^c
<i>One-factor structure of Swahili PHQ-9</i>					
Overall sample	$\chi^2_{(27, n = 787)} = 34.08, p = 0.164$	0.018	0.995	0.996	N/A
Community sample	$\chi^2_{(27, n = 337)} = 19.21, p = 0.862$	0.000	1.006	1.000	N/A
Adults living with HIV	$\chi^2_{(27, n = 450)} = 27.32, p = 0.447$	0.005	1.000	1.000	N/A
<i>Measurement invariance</i>					
<i>Participant group (adults living with HIV vs. community sample)</i>					
Configural	$\chi^2_{(54, n = 787)} = 46.52, p = 0.755$	0.000	1.009	1.000	-
Metric	$\chi^2_{(62, n = 787)} = 66.43, p = 0.327$	0.014	0.998	0.998	0.002
Scalar invariance	$\chi^2_{(70, n = 787)} = 85.28, p = 0.103$	0.023	0.994	0.994	0.004
<i>Sex (Females vs. Males)</i>					
Configural	$\chi^2_{(54, n = 787)} = 40.95, p = 0.905$	0.000	1.001	1.000	-
Metric	$\chi^2_{(62, n = 787)} = 50.83, p = 0.844$	0.000	1.007	1.000	0.000
Scalar invariance	$\chi^2_{(70, n = 787)} = 56.40, p = 0.886$	0.000	1.008	1.000	0.000
<i>Age category (young, middle-age and elderly adults)</i>					
Configural	$\chi^2_{(81, n = 787)} = 57.07, p = 0.976$	0.000	1.018	1.000	-
Metric	$\chi^2_{(97, n = 787)} = 103.40, p = 0.309$	0.016	0.996	0.997	0.003
Scalar invariance	$\chi^2_{(113, n = 787)} = 131.71, p = 0.110$	0.025	0.991	0.990	0.007

Notes. RMSEA- root mean squared error of approximation, CFI- Comparative Fit Index, TLI- Tucker-Lewis Index, ΔCFI - change in Comparative Fit Index, N/A – not applicable.

^a acceptable fit ≤ 0.08 , excellent fit ≤ 0.05 .

^b acceptable fit ≥ 0.90 , excellent fit ≥ 0.95 .

^c acceptable cut-off to detect invariance ≤ 0.01 .

4.2. Comparison of study findings with findings from previous PHQ-9 validation studies

4.2.1. Internal consistency and test-retest reliability

We used two statistical approaches to examine the internal consistency of Swahili PHQ-9, the conventional Cronbach's alpha and Macdonald's omega. Macdonald's omega is thought to provide a more accurate estimate of reliability of a measure than Cronbach's alpha, more so in the ideal world where measurement must consider tau-equivalence, scale unidimensionality, and sample characteristics such as sample size and whether samples are heterogeneous (Dunn et al., 2014; Revelle and Zinbarg, 2009). Using both approaches, results were very similar indicating that the Swahili PHQ-9 is internally consistent. This finding corroborates findings from other validation studies of the PHQ-9 conducted in SSA and elsewhere among participants living with HIV (Monahan et al., 2009), those with other chronic illnesses (Omoro et al., 2006; Udedi et al., 2019), but also studies recruiting from primary health care settings (Gelaye et al., 2013) and the general population (Yu et al., 2012).

In this study, we report an acceptable test-retest reliability of the Swahili PHQ-9 among adults living with HIV based on a random sample of 60 re-assessed participants (two weeks apart). Other validation studies also report stability of PHQ-9 scores (Gelaye et al., 2013; Monahan et al., 2009; Omoro et al., 2006; Yu et al., 2012). Test-retest data from the community sample was not collected in the primary study. An investigation of the test-retest reliability of the Swahili PHQ-9 may be worthwhile when intended for use with the general population in this setting. Other studies investigating test-retest reliability of PHQ-9 using samples from the general population report acceptable stability (Yu et al., 2012).

4.2.2. Factorial structure and measurement invariance

Among the few validation studies of the PHQ-9 involving PLWHA from SSA of which we are aware of (Akena et al., 2013; Bhana et al., 2015; Cholera et al., 2014; Monahan et al., 2009; Pence et al., 2012), this is the first study to report on factorial structure and measurement invariance of the PHQ-9, Swahili version. Most of these studies have focused on reporting the performance characteristics of the PHQ-9 (sensitivity, specificity, and receiver operating characteristic curves) in addition to its reliability and various validity aspects and suggest a cut-off score of

≥ 10 as optimal for positive probable depression screen (Akena et al., 2013; Cholera et al., 2014; Pence et al., 2012).

We observed that the Swahili PHQ-9 is a unidimensional scale evidenced by good item factor loadings and excellent goodness of fit indices to a pre-specified one-factor structure. Similar findings have been previously reported (Gelaye et al., 2013; González-Blanch et al., 2018; Yu et al., 2012). In contrast, a study from Japan (Doi et al., 2018) and Germany (Petersen et al., 2015) reported that a bi-factor model and a two-factor model, respectively, provided a better fit to the data than the one-factor structure of the PHQ-9. In terms of measurement invariance, we observed that the unidimensional Swahili PHQ-9 can be generalized across participant group (adults living with HIV and those from the community), sex (females and males) and age (young, middle-age, and elderly adults) categories. Age and sex invariance of PHQ-9 has previously been reported (González-Blanch et al., 2018; Yu et al., 2012), but we are unaware of studies reporting invariance between PLWHA and a control group to compare or contrast our findings. We call for more studies of this investigative nature to substantiate our findings.

4.3. Discriminant validity

In this study, the Swahili PHQ-9 exhibited good discriminant validity. It could discriminate between adults living with HIV and adults from community sample after adjusting for the observed between-group sociodemographic differences. Past studies also report good discriminant validity of the PHQ-9 using different approaches (Amtmann et al., 2015; Cameron et al., 2011).

4.4. Study strengths and limitations

The strength of this study is comparing the psychometric properties of the Swahili PHQ-9 among HIV adult cases and a sample of adults from the community. We also specifically focus on an under-investigated area – factorial structure and measurement invariance. However, the study has some limitations that need to be considered when interpreting the findings. First, the study was conducted in one geographical setting in Kenya, therefore caution should be taken in generalizing the study findings especially to regions where Swahili is not commonly spoken or has considerable variation from that of Coastal region of Kenya. Further, the controls were entirely primary caregivers and may be atypical of the

general adult population in the community. Second, we cannot claim that all the recruited participants in the community sample were HIV negative as HIV status was not definitively determined in the parent studies. Based on the reported HIV prevalence in Kilifi County, we acknowledge that around 15 participants in the control group may have been HIV-infected at the time of the study but this number is unlikely to invalidate our study findings. Third, we did not have data from the community sample to investigate test-retest reliability. Lastly, we could not analyze the diagnostic accuracy of Swahili PHQ-9 as no 'gold standard' diagnostic interview was administered alongside Swahili PHQ-9 in the original studies.

5. Conclusion

Our findings show that the Swahili PHQ-9 is a reliable and valid unidimensional scale. It appears a valuable tool for assessing depressive symptoms that can be applied across different demographic groups but also in primary HIV care and in the general community within this and similar settings. However, for the community set-up, further research is needed to investigate its test re-test stability. Since a standard cut-off score of ≥ 10 on the PHQ-9 is considered optimal for positive probable depression screen among PLWHA in SSA (Akena et al., 2013; Cholera et al., 2014; Pence et al., 2012), it is important for future studies to examine the diagnostic accuracy of Swahili PHQ-9, at different cut-off points, against a standard diagnostic interview to see if the same cut-off will apply.

Declaration of Competing Interest

None. No author of this work has any conflict(s) of interest to declare.

Acknowledgments

This work was made possible by our study participants, we are grateful for their voluntary participation. We appreciate the effort and dedication of our different data collection teams involved in the parent studies. We also acknowledge the permission from the Director, Kenya Medical Research Institute (KEMRI) to publish this work.

References

- Abubakar, A., Van De Vijver, F., 2017. How to adapt tests for sub-Saharan Africa. In: *Handbook of Applied Developmental Science in Sub-Saharan Africa*. Springer, New York, NY, pp. 197–212.
- Akena, D., Joska, J., Obuku, E.A., Stein, D.J., 2013. Sensitivity and specificity of clinician administered screening instruments in detecting depression among HIV-positive individuals in Uganda. *AIDS Care* 25, 1245–1252.
- Amare, T., Getinet, W., Shumet, S., Asrat, B., 2018. Prevalence and associated factors of depression among PLHIV in Ethiopia: systematic review and meta-analysis, 2017. *AIDS Res. Treat.* e5462959. doi:10.1155/2018/5462959.
- Amtmann, D., Bamer, A.M., Johnson, K.L., Ehde, D.M., Beier, M.L., Elzea, J.L., Bombardier, C.H., 2015. A comparison of multiple patient reported outcome measures in identifying major depressive disorder in people with multiple sclerosis. *J. Psychosom. Res.* 79, 550–557.
- Beck, A.T., Steer, R.A., Carbin, M.G., 1988. Psychometric properties of the Beck depression inventory: twenty-five years of evaluation. *Clin. Psychol. Rev.* 8, 77–100.
- Bernard, C., Dabis, F., de Rekeneire, N., 2017. Prevalence and factors associated with depression in people living with HIV in sub-Saharan Africa: a systematic review and meta-analysis. *PLoS ONE* 12, e0181960.
- Bhana, A., Rathod, S.D., 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.
- Bitta, M.A., Kariuki, S.M., Gona, J., Abubakar, A., Newton, C.R., 2019. Priority mental, neurological and substance use disorders in rural Kenya: traditional health practitioners' and primary health care workers' perspectives. *PLoS ONE* 14, e0220034. doi:10.1371/journal.pone.0220034.
- Cameron, I., Cardy, A., Crawford, J., 2011. Assessing the Validity of the PHQ-9, HADS, BDI-II and QIDS-SR16 in Measuring of Depression in a UK Sample of Primary Care Patients With a Diagnosis of Depression. *Healthcare Improvement Scotland, Edinburgh*.
- Chen, F., Curran, P.J., Bollen, K.A., Kirby, J., Paxton, P., 2008. An empirical evaluation of the use of fixed cutoff points in RMSEA test statistic in structural equation models. *Sociol. Methods Res.* 36, 462–494.
- Chen, F.F., 2007. Sensitivity of goodness of fit indexes to lack of measurement invariance. *Struct. Equ. Model. Multidiscip. J.* 14, 464–504.
- Cholera, R., Gaynes, B., Pence, B., Bassett, J., Qangule, N., Macphail, C., Bernhardt, S., Pettifor, A., Miller, W., 2014. Validity of the patient health questionnaire-9 to screen for depression in a high-HIV burden primary healthcare clinic in Johannesburg, South Africa. *J. Affect. Disord.* 167, 160–166.
- Doi, S., Ito, M., Takebayashi, Y., Muramatsu, K., Horikoshi, M., 2018. Factorial validity and invariance of the patient health questionnaire (PHQ)-9 among clinical and non-clinical populations. *PLoS ONE* 13, e0199235. doi:10.1371/journal.pone.0199235.
- Dunn, T.J., Baguley, T., Brunsden, V., 2014. From alpha to omega: a practical solution to the pervasive problem of internal consistency estimation. *Br. J. Psychol.* 105, 399–412.
- Gelaye, B., Williams, M.A., Lemma, S., Deyessa, N., Bahretibeb, Y., Shibre, T., Wondim-agegn, D., Lemenhe, A., Fann, J.R., Vander Stoep, A., 2013. Validity of the patient health questionnaire-9 for depression screening and diagnosis in East Africa. *Psychiatry Res.* 210, 653–661.
- Global Initiative on Psychiatry, Double stigma, double challenge: mental health and HIV/AIDS in central and eastern Europe and the newly independent states, Global Initiative on Psychiatry, Hilversum, the Netherlands. <http://www.bibalex.org/Search4Dev/files/417303/362887.pdf>.
- González-Blanch, C., Medrano, L.A., Muñoz-Navarro, R., Ruiz-Rodríguez, P., Moriana, J.A., Limonero, J.T., Schmitz, F., Cano-Vindel, A., Group, P.R., 2018. Factor structure and measurement invariance across various demographic groups and over time for the PHQ-9 in primary care patients in Spain. *PLoS ONE* 13, e0193356. doi:10.1371/journal.pone.0193356.
- Hair, J.F., Black, W.C., Babin, B.J., 2010. *RE Anderson Multivariate Data Analysis: A Global Perspective*. Pearson Prentice Hall, New Jersey, pp. 661–699.
- Jackson, D.L., Gillaspay Jr, J.A., Purc-Stephenson, R., 2009. Reporting practices in confirmatory factor analysis: an overview and some recommendations. *Psychol. Methods* 14, 6.
- Kenya National Bureau of Statistics, 2019. Kenya Population and Housing Census Volume I: Population by County and Sub-county. Kenya National Bureau of Statistics (KNBS), Nairobi <https://www.knbs.or.ke/?wpdmpromo=2019-kenya-population-and-housing-census-volume-i-population-by-county-and-sub-county>.
- Kline, R.B., 2005. *Principles and Practice of Structural Equation Modeling*, 2nd edition Guilford Press, New York.
- Kroenke, K., Spitzer, R.L., Williams, J.B., 2001. The PHQ-9: validity of a brief depression severity measure. *J. Gen. Intern. Med.* 16, 606–613.
- Mathers, C.D., Loncar, D., 2006. Projections of global mortality and burden of disease from 2002 to 2030. *PLoS Med.* 3, e442.
- Ministry of Health, 2016. Guidelines on Use of Antiretroviral Drugs for Treating and Preventing HIV Infection in Kenya. Ministry of Health (MOH), Kenya, Nairobi https://aidsfree.usaid.gov/sites/default/files/kenya_art_2016.pdf.
- Monahan, P.O., Shacham, E., Reece, M., Kroenke, K., Ong'or, W.O., Omollo, O., Yebei, V.N., Ojwang, C., 2009. Validity/reliability of PHQ-9 and PHQ-2 depression scales among adults living with HIV/AIDS in western Kenya. *J. Gen. Intern. Med.* 24, 189.
- National AIDS Control Council, 2016. Kenya HIV County Profiles. National AIDS Control Council (NACC), Nairobi <http://nacc.or.ke/wp-content/uploads/2016/12/Kenya-HIV-County-Profiles-2016.pdf>.
- Nyongesa, M.K., Mwangi, P., Wanjala, S.W., Mutua, A.M., Koot, H.M., Cuijpers, P., Newton, C.R., Abubakar, A., 2020. Correlates of health-related quality of life among adults receiving combination antiretroviral therapy in coastal Kenya. *Health Qual. Life Outcomes* 18, 1–14.
- Omoro, S.A., Fann, J.R., Weymuller, E.A., Macharia, I.M., Yueh, B., 2006. Swahili translation and validation of the patient health questionnaire-9 depression scale in the Kenyan head and neck cancer patient population. *Int. J. Psychiatry Med.* 36, 367–381.
- Patel, V., Wessely, S., 1998. *Culture and Common Mental Disorders in Sub-Saharan Africa*. Psychology Press, East Sussex.
- Pence, B.W., Gaynes, B.N., Atashili, J., O'Donnell, J.K., Tayong, G., Kats, D., Whetten, R., Whetten, K., Njamnshi, A.K., Ndumbe, P.M., 2012. Validity of an interviewer-administered patient health questionnaire-9 to screen for depression in HIV-infected patients in Cameroon. *J. Affect. Disord.* 143, 208–213.
- Petersen, J.J., Paulitsch, M.A., Hartig, J., Mergenthal, K., Gerlach, F.M., Gensichen, J., 2015. Factor structure and measurement invariance of the patient health questionnaire-9 for female and male primary care patients with major depression in Germany. *J. Affect. Disord.* 170, 138–142.
- Radloff, L.S., 1977. The CES-D scale: a self-report depression scale for research in the general population. *Appl. Psychol. Meas.* 1, 385–401.
- Revelle, W., 2017. *Psych: Procedures for Personality and Psychological Research*. Software. Northwestern University, Illinois <https://www.scholars.northwestern.edu/en/publications/psych-procedures-for-personality-and-psychological-research>.
- Revelle, W., Zinbarg, R.E., 2009. Coefficients alpha, beta, omega, and the glb: comments on Sijtsma. *Psychometrika* 74, 145.
- Rosseel, Y., 2012. Lavaan: an R package for structural equation modeling. *J. Stat. Softw.* 48, 1–36.
- Scott, J.A., Bauni, E., Moisi, J.C., Ojal, J., Gatakaa, H., Nyundo, C., Molyneux, C.S., Kombe, F., Tsofa, B., Marsh, K., Peshu, N., Williams, T.N., 2012. Profile: the Kilifi health and demographic surveillance system (KHDSS). *Int. J. Epidemiol.* 41, 650–657.
- Ssewanyana, D., Newton, C.R., van Baar, A., Hassan, A.S., Stein, A., Taylor, H.G., Van De Vijver, F., Scerif, G., Abubakar, A., 2020. Beyond their HIV status: the occurrence of multiple health risk behavior among adolescents from a rural setting of sub-Saharan Africa. *Int. J. Behav. Med.* 27, 426–443.
- Stutterheim, S.E., Pryor, J.B., Bos, A.E., Hoogendijk, R., Muris, P., Schaalma, H.P., 2009. HIV-related stigma and psychological distress: the harmful effects of specific stigma manifestations in various social settings. *AIDS* 23, 2353–2357.

- Tsai, A.C., 2014. Reliability and validity of depression assessment among persons with HIV in sub-Saharan Africa: systematic review and meta-analysis. *J. Acquir. Immune Defic. Syndr.* 66, 503.
- Udedi, M., Muula, A.S., Stewart, R.C., Pence, B.W., 2019. The validity of the patient health questionnaire-9 to screen for depression in patients with type-2 diabetes mellitus in non-communicable diseases clinics in Malawi. *BMC Psychiatry* 19, 81.
- UNAIDS, 2018. Global HIV & AIDS Statistics - 2018 Fact Sheet. Joint United Nations Programme on HIV/AIDS (UNAIDS) http://www.unaids.org/sites/default/files/media_asset/UNAIDS_FactSheet_en.pdf.
- Vandenberg, R.J., Lance, C.E., 2000. A review and synthesis of the measurement invariance literature: suggestions, practices, and recommendations for organizational research. *Organ. Res. Methods* 3, 4–70.
- WHO, 2017. Depression and Other Common Mental Disorders: Global Health Estimates. World Health Organization (WHO), Geneva <http://apps.who.int/iris/bitstream/10665/254610/1/WHO-MSD-MER-2017.2-eng.pdf?ua=1&ua=1>.
- WHO, 2018. Depression: Key Facts. World Health Organization (WHO), Geneva <https://www.who.int/news-room/fact-sheets/detail/depression>.
- Yu, X., Tam, W.W., Wong, P.T., Lam, T.H., Stewart, S.M., 2012. The patient health questionnaire-9 for measuring depressive symptoms among the general population in Hong Kong. *Compr. Psychiatry* 53, 95–102.