



## Original article

# Validation of the English and Swahili Adaptation of the Patient Health Questionnaire–9 for Use Among Adolescents in Kenya



Albert Kimtai Tele, M.Sc.<sup>a,b,\*</sup>, Liliana Carvajal-Velez, M.Sc.<sup>c,d,1</sup>, Vincent Nyongesa<sup>a</sup>, Jill W. Ahs, M.Med.Sci.<sup>d,e</sup>, Shillah Mwaniga, M.Sc.<sup>f</sup>, Joseph Kathono, M.Sc.<sup>f</sup>, Obadia Yator, Ph.D.<sup>a</sup>, Simon Njuguna, M.Med.<sup>g</sup>, Ian Kanyanya, M.Med.<sup>g</sup>, Nabila Amin, M.Med.<sup>h</sup>, Brandon Kohrt, Ph.D.<sup>i,2</sup>, Grace Nduku Wambua, M.Sc.<sup>a,b,2</sup>, and Manasi Kumar, Ph.D.<sup>a,j,2</sup>

<sup>a</sup> Department of Psychiatry, University of Nairobi, Nairobi, Kenya

<sup>b</sup> Department of Clinical Psychology, Vrije Universiteit, Amsterdam, the Netherlands

<sup>c</sup> Division of Data, Analytics, Planning and Monitoring, Data and Analytics Section, UNICEF, New York, New York

<sup>d</sup> Department of Global Public Health, Karolinska Institutet, Stockholm, Sweden

<sup>e</sup> Swedish Red Cross University, Stockholm, Sweden

<sup>f</sup> Nairobi Metropolitan Services, Nairobi, Kenya

<sup>g</sup> Department of Mental Health, Kenyatta National Hospital, Nairobi, Kenya

<sup>h</sup> Mathare National Referral and Teaching Hospital, Nairobi, Kenya

<sup>i</sup> Department of Psychiatry, George Washington University, Washington, DC

<sup>j</sup> Brain and Mind Institute, Aga Khan University, Nairobi, Kenya

**Article history:** Received May 16, 2022; Accepted October 5, 2022

**Keywords:** Children and adolescents; Depression; Developing countries; Psychometrics; Screening; Validation

---

 A B S T R A C T

**Purpose:** Our study aimed to validate culturally adapted English and Swahili versions of the Patient Health Questionnaire–9 (PHQ-9) for use with adolescents in Kenya. Criterion validity was determined with clinician-administered diagnostic interviews using the Kiddie Schedule of Affective Disorders and Schizophrenia.

**Methods:** A total of 250 adolescents comprising 148 (59.2%) females and 102 (40.8%) males aged 10–19 years (mean = 14.76; standard deviation = 2.78) were recruited. The PHQ-9 was administered to all respondents concurrently in English and Swahili. Adolescents were later interviewed by clinicians using Kiddie Schedule of Affective Disorders and Schizophrenia to determine the presence or absence of current symptoms of major depressive disorder. Sensitivity specificity, positive predictive value (PPV) and negative predictive value (NPV), and likelihood ratios for various cut-off scores for PHQ-9 were analyzed using receiver operating characteristic curves.

**Results:** The internal consistency (Cronbach's  $\alpha$ ) for PHQ-9 was 0.862 for the English version and 0.834 for Swahili version. The area under the curve was 0.89 (95% confidence interval, 0.84–0.92) and 0.87 (95% confidence interval, 0.82–0.90) for English and Swahili version, respectively, on receiver operating characteristic analysis. A cut-off of  $\geq 9$  on the English-language version had a sensitivity of 95.0%, specificity of 73.0%, PPV of 0.23, and NPV of 0.99; a cut-off of  $\geq 9$  on the Swahili version yielded a sensitivity of 89.0%, specificity of 70.0%, PPV of 0.20, and NPV of 0.90.

**IMPLICATIONS AND CONTRIBUTION**

This study reports the psychometric properties of a culturally translated and adapted version of the PHQ-9 in Kenya for depression screening in adolescents aged 10–19 years and aged 15–19 years. The results provide information on a range of cut-off scores so that clinicians, policy-makers, and researchers can use thresholds that best fit their intended objectives.

**Conflicts of interest:** The authors have no conflicts of interest to declare.

**Disclaimer:** The article was published as part of supplement that was supported by the Bill & Melinda Gates Foundation [INV-001395] and UNICEF.

\* Address correspondence to: Albert Kimtai Tele, M.Sc., University of Nairobi, Nairobi, Kenya and Vrije Universiteit Amsterdam, The Netherlands, Department of Psychiatry, University of Nairobi, P.O. BOX 2743, Nairobi, Nairobi County, Kenya.

E-mail addresses: [telekim01@gmail.com](mailto:telekim01@gmail.com); [albert.tele@amhf.or.ke](mailto:albert.tele@amhf.or.ke) (A.K. Tele).

<sup>1</sup> Joint lead authors.

<sup>2</sup> Joint senior authors.

**Discussion:** Psychometric properties were comparable across both English-adapted and Swahili-adapted version of the PHQ-9, are reliable, and valid instrument to detect major depressive disorder among adolescents which can be used in resource-limited settings for early identification of adolescents in need of mental health support.

© 2022 Society for Adolescent Health and Medicine. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## Background

Anxiety and depressive disorders represent more than 40% of the burden of mental disorders for adolescents aged 10–19 years around the world [1]. Across low-income and middle-income countries (LMICs), a lack of country-level data on adolescent mental health is pervasive and a main bottleneck for prioritizing this important public health issue. For example, in Kenya, mental disorders top the list of causes of years lived with disabilities [2], and studies carried out locally have found the prevalence of adolescent depression and anxiety in Kenya to be an issue of significant concern [3,4]. However, estimates of the burden of mental disorders in Kenya have not used tools that have been culturally adapted and clinically validated locally; this means that current prevalence estimates are not cultural aspects of mental health. This risks potential overestimation or underestimation of the burden, which hinders the ability to work toward promoting and scaling up programs and services to support adolescent mental health [5]. Lack of valid data and the absence of related policy dedicated to child and adolescent mental health are limiting factors in implementing mental health strategies for adolescents in Kenya and many other LMICs [6,7].

Therefore, clinical validation of a depression screening tool for adolescents was undertaken in Kenya to establish the psychometric properties of both the local English dialect and Swahili versions. The goal was to develop data so that clinicians, public health researchers, and policymakers could make appropriate prevalence estimates and clinical or policy decisions to understand population level prevalence of depression. For example, validation outcomes will be useful to identify a range of cut-offs to detect depression while minimizing false positives, thus limiting additional burdens on health providers and potential stigmatization of respondents, while also minimizing false negatives to assure those in need of services can obtain them. Moreover, the simultaneous validation in English and Swahili would allow the use of both versions of the tool and adjustments for cut-offs based on language of completion so that data would be comparable. This research is a part of ongoing work led by United Nations Children's Fund (UNICEF) and in collaboration with key institutional and academic partners from Kenya to conduct and promote cultural adaptation and validation of adolescent mental health assessment tools. UNICEF's Measurement of Mental Health among Adolescents at the Population Level (MMAP) initiative [8] developed a protocol for cultural adaptation and criterion validation using structured diagnostic clinical assessments to harmonize efforts across countries.

The conceptual framework centers around UNICEF's impetus promote systematic data collection of adolescent mental health using tools that are culturally adapted for use among adolescents. In the UNICEF State of the World Children's report (2021) [9], MMAP initiative was identified as a robust and methodological

approach to collecting and managing mental health data for adolescents globally.

Figure 1 provides the theoretical framework for the validation study here, which is based on the MMAP protocol [8], which developed an approach for feasible and sustainable data collection on adolescent mental health at population level. Our validation study provides evidence on accurate identification of adolescents in need of a mental health assessment by using culturally adapted tools. The analytical approach includes sensitivity and predictive value of these measures. Through this approach we are generating evidence for recommendation of use of these tools in diverse country contexts where psychiatric morbidity is high and due to many factors including a lack of culturally adapted and feasible tools, an impact has not been widely assessed at country levels.

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 both the local English dialect and Swahili versions of Patient Health Questionnaire–9 (PHQ-9) for use among adolescents in Kenya (but also within the larger Swahili-speaking African countries). A previous study among head and neck cancer patients conducted in Kenya by Omoro et al. translated and validated a Swahili version of PHQ-9 [10], whereas Monahan et al. and Mwangi et al. administered and investigated the reliability and validity of the English version of PHQ-9; the latter used Swahili version of the PHQ-9 among adults living with HIV/AIDS [11,12].

In light of the aforementioned, we focus on reporting on (1) the internal consistency, (2) test-retest reliability, (3) receiver operator characteristics, and (4) discriminant validity of both English and Swahili PHQ-9 among adolescents in Kenya aged 10–19 years.

Age-specific data on health are required under the 2030 Sustainable Development Goals agenda and the World Health Organization and partners 'recommended standardized age groups when reporting on health data, especially when reporting on multiple diseases simultaneously, burden of disease, or a not previously recognized disease or condition. In this study, adolescents are grouped in two age groups: young adolescents aged 10–14 years and older adolescents aged 15–19 years [13].

## Methods

### Settings

Kenya is a lower-middle-income county, as per the World Bank Income classification in 2021 [14]. Located in Sub-Saharan Africa, Kenya is home to 47.5 million people as per the 2019 population census [15], of which one in four are adolescents aged 10–19 years. Kenya is a multilingual country, with English and Swahili, a Bantu language, as official languages [16]. The work

presented in this analysis was conducted in Nairobi, the capital, with an estimated population of 4.4 million people. The cultural adaptation and validation activities described in this article took place in Kariobangi and Kangemi health centers run under County Government's Nairobi Metropolitan Services.

#### Participants and recruitment

Recruitment of participants and administration of assessments took place between June and July 2021. The sampling for the validation was designed to be “enriched” for participants likely to have symptoms of depression for comparison with “controls” presumed not to have the condition. The desired ratio among the selected participants was 2:1 for likely cases of depression versus those likely without the condition. This approach was used to ensure a large enough sample size for participants with depression to allow for conducting a full range of psychometric analyses for validation against the clinical diagnosis [17].

Recruitment was conducted using a two-stage targeted approach (Figure A1): (1) The community case detection tool for children and adolescents [18,19] was used by community health workers to identify participants, participants identified by community health workers from households that had family issues, marital conflicts among parents were invited as likely to have depression and (2) pregnant and parenting adolescents subsample were identified from an ongoing study under which this work was embedded.

#### Instruments

**Patient health questionnaire—9.** The PHQ-9 is a nine-item self-report scale consisting of the nine criteria used for diagnosis of depression in the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) [20,21]. This scale is derived from the primary care evaluation of mental disorders, an instrument developed to assist primary care clinicians in making diagnoses of most common mental disorders [22].

The nine-item PHQ-9 scale is one of the tools included in the MMAP protocol for assessing symptoms of depression among adolescents. The tool was developed with the purpose of screening for and assessing the severity of depressive symptoms in both clinical and research contexts [20]. The initial validation research of PHQ-9 demonstrated strong reliability and criteria, concept, and external validity, in a large primary care sample [20]. Subsequently, the psychometric properties of PHQ-9 have been evaluated in different samples of patients, including other primary care samples [23–26], psychiatric patients [27,28], pregnant women [29], specific groups of medical patients [11,12,30,31], and also in diverse multicultural populations [32–34]. In an academic setting, validation of PHQ-9 scale was performed among university students of Korea [35], China [36], and Nigeria [37].

The PHQ-9 was translated into Swahili and adapted for English as spoken in Kenya based on the transcultural translation and adaptation process [38], previously adapted for use with children and adolescents [8,39] (adapted tool in [Supplemental Material Box 1](#) and [2](#)).

**Kiddie schedule for affective disorders and schizophrenia.** The Kiddie Schedule for Affective Disorders and Schizophrenia (K-SADS) is a semi-structured diagnostic clinical interview for children aged 6–18 years [40]. Clinicians receive formal training

to administer the K-SADS. The current version of the K-SADS provides criteria-based algorithms to assess presence of current mental health disorders consistent with DSM-5 criteria [40,41]. For this study, we used the depression section of the screening module and the depression supplementary module. K-SADS diagnosis of major depressive episodes was used as a criterion standard to determine validity of psychometrics of the Kenyan English and Swahili versions of the PHQ-9.

Participants first completed the adapted PHQ-9 alongside the other mental health modules in the MMAP tool [8]. Then, within the next 48 hours, trained clinicians assessed participants using the K-SADS. Clinicians administering the K-SADS were blind to the results for the participant's PHQ-9 score.

#### Training

Community Health Volunteers were trained to administer the MMAP tool, including the PHQ-9. Community Health Assistants were trained to provide supervision. Twenty clinicians (clinical officers, psychiatric nurses, psychologists, and psychiatrists) were trained to administer the K-SADS tool. Training included pair assessments and calculation of inter-rater reliability to ensure participants were being scored by the clinicians consistently and comparably. Of these, 12 clinicians who met the inter-rater reliability (IRR) requirements (IRR of 0.7 and more) were invited to conduct the diagnostic assessment using this tool. Data collection was conducted in English and Swahili, and all of our enumerators and clinicians and the respondents were fluent in both languages.

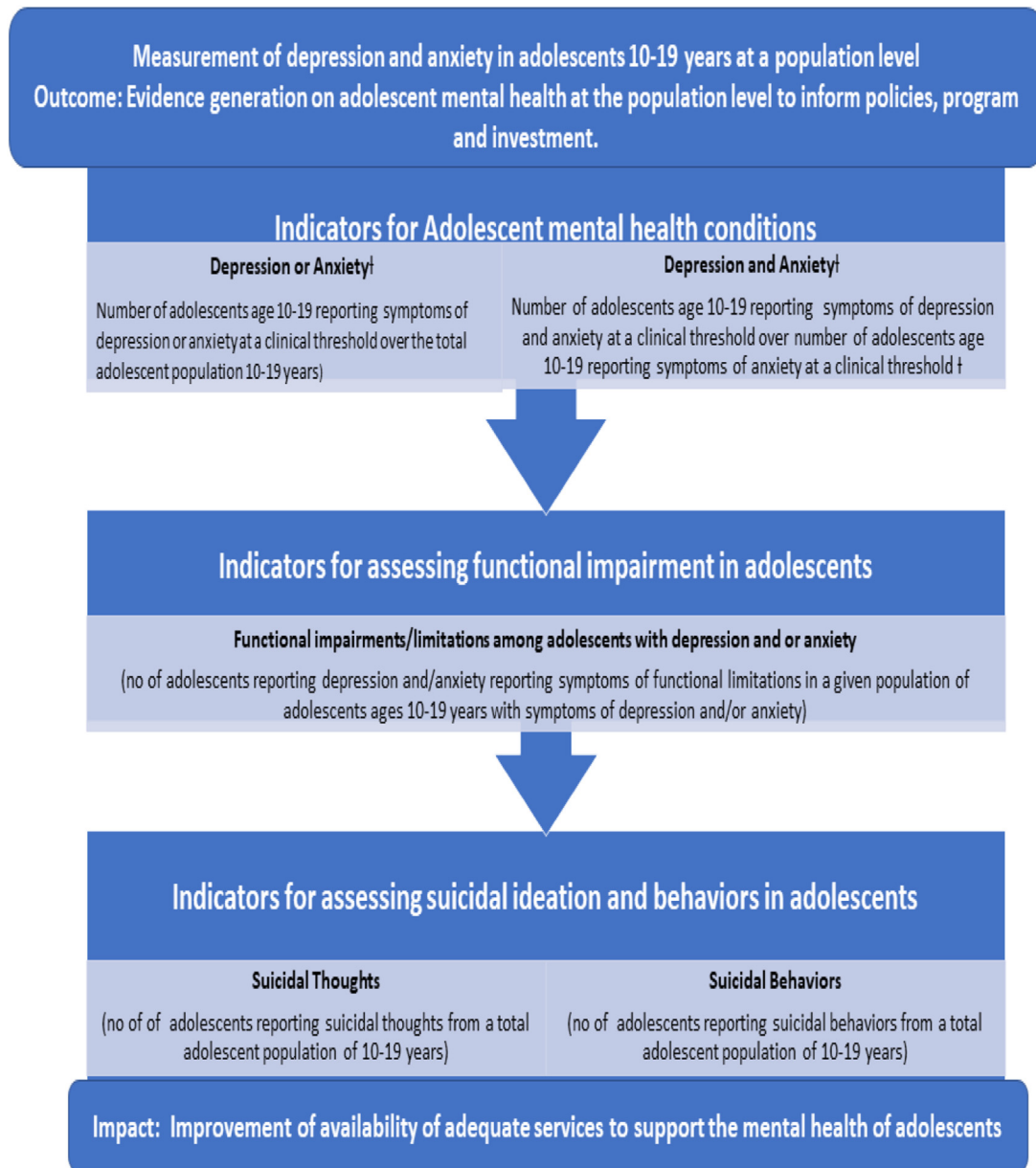
#### Statistical analysis

**Sample size.** The desired sample size was determined with the goal of obtaining a sample sufficient to estimate the receiver operating characteristic (ROC) curve and area under the curve (AUC) for the tool to be translated, adapted, and validated. Although a priori sample size determinations for AUC are highly susceptible to assumptions about the performance of the test [42], a sample size of 100 is generally sufficient to make a qualitative assessment of the utility of a test [43]. Over-recruitment of the target conditions is critical. That is, we require strategies to identify youth most likely to have the mental health concerns of interest. Given this backdrop, a sample of 150+ was considered for recruitment to test positive for depression/anxiety using community case detection tool (for caseness) and a sample of 100 participants negative for depressive or anxiety disorders. We added a sample of 50 pregnant adolescents to capture local contextual realities.

#### Data analysis

Item means, standard deviations (SDs), frequencies, and percentages were calculated for the sociodemographic variables. Independent samples *t*-tests were carried out to compare the mean PHQ-9 scores in the groups of depressed and not depressed participants as per the K-SADS.

**Reliability.** To investigate the reproducibility and consistency of PHQ-9, reliability coefficients as measured by Cronbach's alpha were calculated.



**Figure 1.** Theoretical framework of PHQ-9 validation.

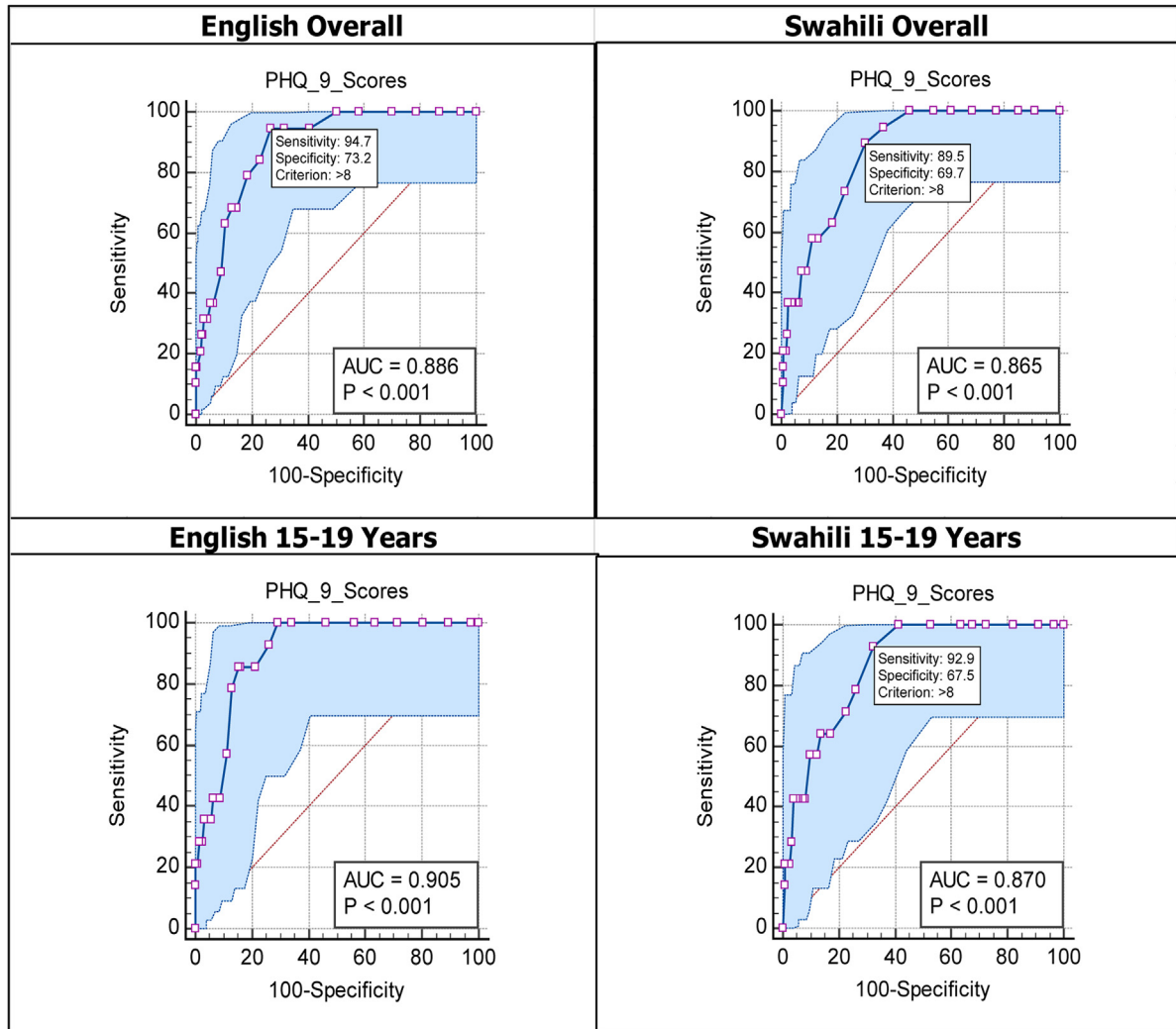
**Criterion validity.** The following measures of criterion validity were used in the analysis: sensitivity, specificity sensitivity, ROC curve, and AUC.

**Sensitivity and specificity.** For each PHQ-9 cut-off point, sensitivity (or true positive rate) and specificity (or true negative rates) were calculated. In our case, sensitivity is defined by the proportion of individuals with current depressive episodes as per K-SADS criteria who were correctly identified by PHQ-9, whereas specificity is defined as the proportion of individuals without depression as per standard-SADS who were correctly identified as such by PHQ-9. Positive likelihood ratio (+LR) and negative likelihood ratio (–LR) and 95% confidence intervals (CIs) for each of these parameters were reported. We defined +LR as the

probability of an individual without the condition having a positive test, whereas –LR was defined as the probability of an individual without the condition having a negative test. Sensitivity and specificity were plotted on a curve, creating the ROC curve (Figure 2). The ROC curve is a plot of the sensitivity versus [1-specificity] overall possible threshold values of the test being validated.

The AUC, which is the area under the ROC curve, was calculated. AUC is a key psychometric that gives the probability that the instrument will yield a higher score for a randomly chosen individual with the target condition than for a randomly chosen individual without the condition [44].

Positive predictive value (+PPV; proportion of true positives on the K-SADS among all positives identified by the PHQ-9) and



**Figure 2.** Area under the curve for PHQ-9 by K-SADS depression in English and Swahili and age.

negative predictive value ( $-NPV$ ; proportion of true negatives on the K-SADS among all those who scored negative by PHQ-9) were calculated.

Diagnostic odds ratio (OR) (the odds of a positive test in those with disease relative to the odds of a positive test in those without disease [45]) was computed in terms of sensitivity and specificity and in terms of positive and negative likelihood ratios ( $DOR = LR+/LR-$ ). Thus, this single measure includes information about both sensitivity and specificity and tends to be reasonably constant despite diagnostic threshold.

Youden's index can be used as one type of criterion for choosing the "optimal" threshold value for the PHQ-9 test, the threshold value for which the value of [sensitivity + specificity - 1] is maximized. PHQ-9's accuracy (proportion of results, both positive and negative, correctly identified by the PHQ-9 at different cut-offs) was estimated by the area under the ROC curve.

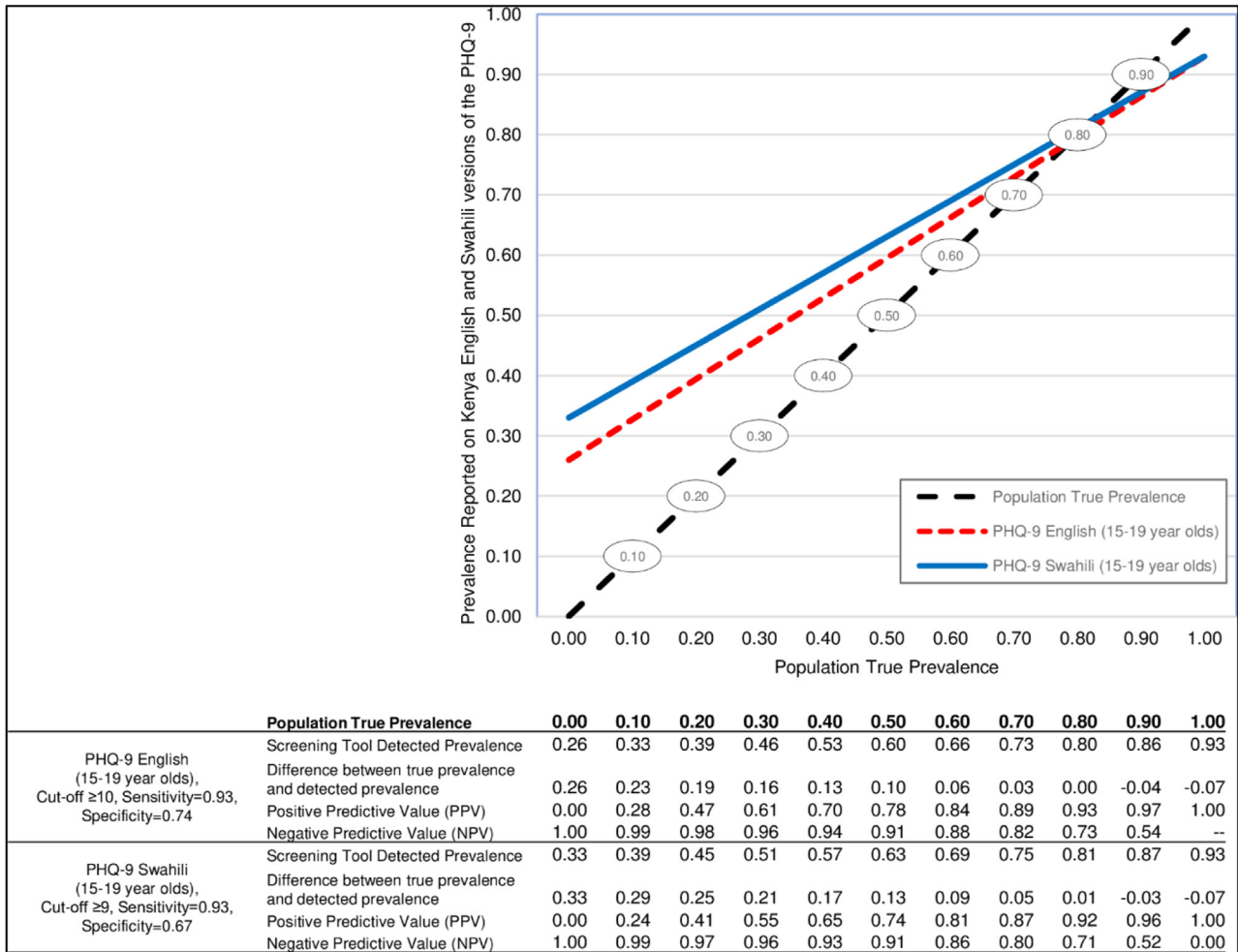
Discriminant ability of the adapted PHQ-9 items for adolescents with and without diagnosis as per K-SADS was also analyzed by calculating means and SDs and compared using

$t$ -tests, conducted separately for the overall sample and older adolescents aged 15–19 years with Bonferroni correction for multitesting.

Results were analyzed separately for English and Swahili responses for the overall sample in addition to respondents aged 15–19 years. All analyses were performed using Stata version 14.0 software and R-Version 4.1.2.

#### Ethical approval

The study was approved by the Kenyatta National Hospital/ University of Nairobi Ethical Review Committee (approval no. P694/09/2018). An approval was received from Nairobi County Health no. CMO/NRB/OPR/VOL1/2019/04 and a permit from the Kenyan National Commission for Science, Technology, and Innovation (NACOSTI/P/19/77,705/28,063) was subsequently obtained. Participants who were at a high risk of depressive disorder, defined as those who were experiencing high distress, or those who were suicidal (positive score on PHQ-9) and/or assessed by the clinician to be warranting further assessment or



**Figure 3.** Visualization for estimating true depression prevalence based on PHQ-9 obtained prevalence for the Kenyan English and Swahili versions using the psychometric properties obtained in this study.

treatment were referred to the Kenyatta National Hospital’s Department of Mental Health where three clinicians were requested to review their case and offer management. The community health workers embedded under the MMAP study made further family-based and community-based follow-ups as needed.

**Results**

A total of 250 adolescents comprising 148 (59.2%) females and 102 (40.8%) males, aged 10–19 years with mean of 14.76, SD = 2.78, of these the percentage of those aged 10–14 years was 45.2%, whereas those aged 15–19 years was 54.8%.

Among the 15–19-year-olds, 90 (65.7%) were girls and 47 (34.3%) were boys. All participants were fluent in English and Swahili. Regarding K-SADS results for clinical diagnoses, 19 adolescents fulfilled the criteria for K-SADS DSM-5 depression, with five adolescents among the 10–14 year-old group and 14 adolescents in the 15–19-year-old group (Table 1).

*Psychometric analysis*

The AUC for the English language version of the scale for sample was 0.886 (95% CI, 0.840–0.923). The AUC for the Swahili version of the scale was 0.865 (95% CI, 0.817–0.905; Figure 2).

The AUC for the English language version of the scale for adolescents aged 15–19 years was 0.905 (95% CI, 0.843–0.949) and 0.870 (95% CI, 0.801–0.921) for the Swahili version.

The PHQ-9 had high internal consistency (Cronbach alpha [α] = 0.834 for the English version and α 0.862 Swahili version and α = 0.835 for 15–19-year-olds for both English and Swahili version). The test-retest reliability for English version was r = 0.530, P < .001 and Swahili was r = 0.773, P < .001.

Results in English and Swahili did not vary significantly. Table 2 presents the psychometric properties of the adapted PHQ-9 in English and Swahili compared with K-SADS depression.

The appropriate cut-off point was ≥9 for both the English and Swahili versions. At this point, the sensitivity was 95.0% and the specificity was 73.0%. The PPV and NPV were 23.0% and 99.0%, respectively. The positive and negative likelihood ratios at this point were 3.53 and 0.77, respectively. The diagnostic OR was

**Table 1**  
Demographics of participants

Variable	Category	10–14 years (n = 114)	15–19 Years (n = 136)	Total sample (n = 250)
Age	Mean (SD)	12.07 (1.45)	16.97 (1.24)	14.76 (2.78)
Gender	Male	58 (51.3%)	90 (65.7%)	102 (40.8%)
	Female	55 (48.7%)	47 (34.3%)	148 (59.2%)
Highest level of education attained	Primary	110 (97.3%)	31 (22.6%)	141 (56.4%)
	Secondary	3 (2.7%)	97 (70.8%)	100 (40%)
	Higher	0 (0%)	9 (6.6%)	9 (3.6%)
Community Case Detection Tool (CCDT) for recruitment	Negative (no condition)	58 (51.3%)	49 (35.8%)	107 (42.8%)
	Depression CCDT positive	55 (48.7%)	88 (64.2%)	143 (57.2%)
Kiddie Schedule of Affective Disorders and Schizophrenia (K-SADS)	No Depression	108 (95.6%)	123 (89.8%)	231 (92.4%)
	Depression diagnosis	5 (4.4%)	14 (10.2%)	19 (7.6%)

49.1 and the accuracy was 74.8% for the English version. The sensitivity was 89.0% for the Swahili version and the specificity was 70.0%. The PPV and NPV were 20.0% and 99.0%, respectively. The positive and negative likelihood ratios at this point were 2.95 and 0.15, respectively. The diagnostic OR was 19.6 and the accuracy was 71.2%.

The appropriate cut-off point for the adolescents aged 15–19 years was  $\geq 10$  for English version and  $\geq 9$  for Swahili version. At this point, the sensitivity was 93.0% for both English and Swahili versions and specificity was 74.0% for English version and 67.0% for Swahili version. The PPV and NPV were 29.0% and 99.0%, respectively, for English version and 25.0% and 99.0%, respectively, for Swahili version. The accuracy was 75.9% for the English version and 70.1% for Swahili version. The diagnostic OR was 37.0 for English and 27.0 for Swahili version.

#### Item analyses

Item means were calculated for adolescents diagnosed with depression versus adolescents who did not receive any diagnosis. The discriminant ability of the items is presented in Table 3. Most of the PHQ-9 items in both Swahili performed well; this is made evident by significant differences between respondents with and without diagnosis in all the items. However, some items performed poorly in discriminating between those with and without K-SADS depression diagnoses. When using the entire sample, the items performing poorly included PHQ Item #1 (*Little pleasure/interest*), Item #3 (*Sleep*) in English, and Item #7 (*Shida ya kuzingatia au kuwa makini/Trouble concentrating*) in Swahili; there were no significant differences in these items based on depression status. There were significant differences among the older adolescents in all the items except Item #1 (*little interest or pleasure*) in English. However, the total score for both language versions demonstrated a significant difference in their discriminant ability overall. For the total sample PHQ-9 score in English, the mean among the adolescents with depression (mean = 15.84, SD = 6.34  $p < .001$ ) was more than the mean among the adolescents without depression (mean = 6.55, SD = 5.05  $p < .001$ ).

The means for each of the nine PHQ items by K-SADS diagnostic status comparing depression diagnosis with no diagnosis are presented in Figure A2. In English, item number 6, “Feeling bad about yourself,” was the most commonly endorsed symptom among diagnosed adolescents. Item number 2 “Feeling down/depressed” was the most commonly endorsed symptom among nondiagnosed participants. This was also the most commonly endorsed item in the Swahili scale among all (diagnosed and

nondiagnosed) participants. All items had good discriminatory ability between the groups.

#### Application of the validation results: adjustments for population prevalence

The validation results can aid policymakers in improving the accuracy of population estimates of adolescent depression. Using cut-offs that balance sensitivity and specificity, Figure 3 presents visualization for estimating true depression prevalence based on PHQ-9—obtained prevalence for the Kenyan English and Swahili versions using the psychometric properties obtained in this study. When adjusting for false positive and false negatives, the PHQ-9—identified prevalence rates for depression and anxiety can be adjusted to approximate what the true prevalence may be in the population. For example, if a prevalence of 39% is identified for depression among 15–19-year-old adolescents using the English language version, the estimated true prevalence is likely closer to 20%. A true prevalence of 20% would correspond with a detected prevalence rate of 45%. The degree of adjustment differs based on prevalence because of the contribution of false positive versus false negatives to the estimates made. The PPV and NPV for individual adolescents also vary by prevalence rates (also included in Figure 3). Policy makers can use algorithms or figures such as this to make adjusted prevalence estimates when allocating resources and designing programs. This is important because at common true prevalence rates of 10%–20%, the PHQ-9 overestimates population prevalence by two-fold to three-fold.

#### Discussion

Lack of culturally adapted and validated tools has been a barrier to national-level data on adolescent mental health, which can inform policies and investment in implementing or scaling up programs. This work contributes to the evidence of availability of suitable and validated tools in LMICs such as Kenya. The culturally adapted PHQ-9 in Kenyan English and Swahili languages for use among adolescents uses a terminology that is clearly understood by adolescents, thus allowing for easy administration and communication with respondents, and has been developed through a stakeholder-led translation and adaptation process. The internal consistency of PHQ-9 in this study was similar to the values found in other studies, which ranged from 0.824 to 0.869 [37,46]. For the adolescents aged 15–19 years, the sensitivity at the cut-off value of  $\geq 10$  for English version and  $\geq 9$  for Swahili version or more was 93.0%, and the specificity was between 67.0% and 74.0%. These values, in

particular, the sensitivity, are higher than those reported in two meta-analyses using PHQ-9, although those studies reported findings of adults aged more than 18 years [47,48]. These cut-offs among Kenyan adolescents are with the range of meta-analyses identified the optimal cut-off points for diagnosing major depressive episodes, ranging from 8 to 11 [49].

The validation results among adolescents demonstrated psychometric properties comparable to results for adults and high-income settings [34]. A cut-off score of > 8 produced psychometric properties comparable to findings in the literature across other LMICs like Argentina [50], Thailand [51], Malaysia [52], and South Africa [53].

### Limitations

This analysis is among the first to include adolescents as young as 10 years old in clinical validation approaches using population-level mental health measures in Kenya. Due to the small number of adolescents aged 10–14 years who were found to have depression on screening (N = 5), were not able to report results for this group. Despite challenges posed by restrictions

related to the COVID-19 pandemic, it was possible to conduct the work, although during a longer period than initially planned.

The validation only addressed Kenyan English and Swahili, Kenya's two official working languages. We did not validate the tool for other languages spoken in Kenya like other Bantu languages widely spoken, including Kikuyu, Kamba, and Luhya. Also, the work was conducted around two periurban informal settlements (Nairobi) with diverse ethnicities, although this does not include rural settings. Therefore, the extrapolation of the tool to rural settings should be taken with caution. Despite these limitations, this work contributes to an improved measurement of adolescent mental health in Kenya using a tool only used among adults and is now culturally adapted and validated for use among adolescents in the country.

### Conclusion

Validation results of the culturally adapted PHQ-9 in Kenya showed satisfactory performance for detecting depression among adolescents. Through this effort, we are contributing to strengthening the evidence and availability of culturally and

**Table 2**  
Validation psychometrics of patient health questionnaire–9 (PHQ-9)

Cut-off	Sensitivity	Specificity	PLR	NPR	Diagnostic OR	Youden's J	TP	FN	TN	FP	PPV	NPV	% Accuracy
<b>English Overall</b>													
≥ 6	1.00	0.50	1.99	0.00	-	0.498	19	0	115	116	0.14	1.00	53.6
≥ 7	0.95	0.59	2.33	0.09	26.2	0.540	18	1	137	94	0.16	0.99	62
≥ 8	0.95	0.68	3.00	0.08	39.0	0.631	18	1	158	73	0.20	0.99	70.4
≥ 9	<b>0.95</b>	<b>0.73</b>	<b>3.53</b>	<b>0.07</b>	<b>49.1</b>	<b>0.679</b>	<b>18</b>	<b>1</b>	<b>169</b>	<b>62</b>	<b>0.23</b>	<b>0.99</b>	<b>74.8</b>
≥ 10	0.84	0.77	3.67	0.20	17.9	0.613	16	3	178	53	0.23	0.98	77.6
≥ 11	0.79	0.81	4.24	0.26	16.4	0.603	15	4	188	43	0.26	0.98	81.2
≥ 12	0.68	0.85	4.65	0.37	12.6	0.537	13	6	197	34	0.28	0.97	84
<b>Swahili Overall</b>													
≥ 6	1.00	0.45	1.82	0.00	-	0.450	19	0	104	127	0.13	1.00	49.2
≥ 7	1.00	0.54	2.16	0.00	-	0.537	19	0	124	107	0.15	1.00	57.2
≥ 8	0.95	0.63	2.57	0.08	30.9	0.579	18	1	146	85	0.17	0.99	65.6
≥ 9	<b>0.89</b>	<b>0.70</b>	<b>2.95</b>	<b>0.15</b>	<b>19.6</b>	<b>0.592</b>	<b>17</b>	<b>2</b>	<b>161</b>	<b>70</b>	<b>0.20</b>	<b>0.99</b>	<b>71.2</b>
≥ 10	0.74	0.77	3.21	0.34	9.4	0.507	14	5	178	53	0.21	0.97	76.8
≥ 11	0.63	0.81	3.39	0.45	7.5	0.445	12	7	188	43	0.22	0.96	80
≥ 12	0.58	0.87	4.31	0.49	8.9	0.445	11	8	200	31	0.26	0.96	84.4
≥ 13	0.58	0.89	5.14	0.47	10.8	0.466	11	8	205	26	0.30	0.96	86.4
Cut-Off	Sensitivity	Specificity	PLR	NPR	Diagnostic OR	Youden's J	TP	FN	TN	FP	PPV	NPV	% Accuracy
<b>15–19 Years English</b>													
≥ 7	1.00	0.54	2.16	0.00	-	0.537	14	0	66	57	0.20	1.00	58.4
≥ 8	1.00	0.66	2.93	0.00	-	0.659	14	0	81	42	0.25	1.00	69.3
≥ 9	1.00	0.71	3.42	0.00	-	0.707	14	0	87	36	0.28	1.00	73.7
≥ 10	<b>0.93</b>	<b>0.74</b>	<b>3.57</b>	<b>0.10</b>	<b>37.0</b>	<b>0.668</b>	<b>13</b>	<b>1</b>	<b>91</b>	<b>32</b>	<b>0.29</b>	<b>0.99</b>	<b>75.9</b>
≥ 11	0.86	0.79	4.05	0.18	22.4	0.646	12	2	97	26	0.32	0.98	79.6
≥ 12	0.86	0.84	5.27	0.17	30.9	0.695	12	2	103	20	0.38	0.98	83.9
≥ 13	0.86	0.85	5.55	0.17	32.8	0.703	12	2	104	19	0.39	0.98	84.7
≥ 14	0.79	0.87	6.04	0.25	24.5	0.656	11	3	107	16	0.41	0.97	86.1
<b>15–19 Years Swahili</b>													
≥ 7	1.00	0.47	1.89	0.00	-	0.472	14	0	58	65	0.18	1.00	52.6
≥ 8	1.00	0.59	2.41	0.00	-	0.585	14	0	72	51	0.22	1.00	62.8
≥ 9	<b>0.93</b>	<b>0.67</b>	<b>2.86</b>	<b>0.11</b>	<b>27.0</b>	<b>0.603</b>	<b>13</b>	<b>1</b>	<b>83</b>	<b>40</b>	<b>0.25</b>	<b>0.99</b>	<b>70.1</b>
≥ 10	0.79	0.74	3.02	0.29	10.4	0.526	11	3	91	32	0.26	0.97	74.5
≥ 11	0.71	0.77	3.14	0.37	8.5	0.487	10	4	95	28	0.26	0.96	76.6
≥ 12	0.64	0.83	3.77	0.43	8.7	0.472	9	5	102	21	0.30	0.95	81.0
≥ 13	0.64	0.86	4.65	0.41	11.2	0.505	9	5	106	17	0.35	0.95	83.9
≥ 14	0.57	0.88	4.69	0.49	9.6	0.449	8	6	108	15	0.35	0.95	84.7

Validation psychometrics of Patient Health Questionnaire (PHQ-9, Kenyan English, and Swahili versions) from comparison with the Kiddie Schedule of Affective Disorders and Schizophrenia (K-SADS DSM-5 depression). Appropriate cut-off point for screening of depression are indicated in bold. PHQ-9, Patient Health Questionnaire; PPV, Positive predictive value; NPV, Negative predictive value; PLR, Positive likelihood ratio; NLR, Negative likelihood ratio; OR, Odds ratio; TP, True positive; TN, True negative; FP, False positive; FN, False negative.



**Table 3**

Discriminant ability of PHQ-9 items for adolescents with and without depressive disorder diagnoses on the kiddie schedule of affective disorders and schizophrenia

PHQ-9 items English	Overall (N = 250)			15–19 Years (n = 137)		
	Not depressed	Depressed	p value	Not depressed	Depressed	p value
	Mean ± SD	Mean ± SD		Mean ± SD	Mean ± SD	
Item 1: little interest/pleasure	0.90 ± 0.89	1.58 ± 1.07	.123	0.94 ± 0.91	1.64 ± 1.15	.080
item 2: feeling down/depressed	0.93 ± 0.91	1.84 ± 1.07	.000	0.93 ± 0.93	2.00 ± 1.18	.001
Item 3: sleep	0.70 ± 0.89	1.53 ± 1.22	.083	0.76 ± 0.94	1.71 ± 1.20	.005
Item 4: tired	0.84 ± 0.95	2.00 ± 0.94	< .001	0.86 ± 0.95	2.00 ± 0.96	< .001
Item 5: appetite/overeating	0.69 ± 0.98	2.00 ± 1.20	< .001	0.80 ± 1.08	2.14 ± 1.17	< .001
Item 6: bad about yourself	0.77 ± 0.98	2.42 ± 0.96	< .001	0.97 ± 1.06	2.64 ± 0.74	< .001
Item 7: trouble concentrating	0.77 ± 0.95	1.42 ± 1.17	.049	0.78 ± 0.98	1.71 ± 1.20	.011
Item 8: moving/speaking	0.58 ± 0.83	1.47 ± 1.02	< .001	0.62 ± 0.89	1.57 ± 1.02	.002
Item 9: thoughts dead	0.39 ± 0.80	1.58 ± 1.17	.003	0.37 ± 0.77	1.79 ± 1.12	< .001
PHQ-9 Total Scores	6.55 ± 5.05	15.84 ± 6.34	< .001	7.03 ± 5.14	17.21 ± 5.92	< .001
Swahili	Overall (N = 250)			15–19 Years (n = 137)		
Item 1: kupoteza hamu/ furaha	0.80 ± 0.80	1.79 ± 0.98	.003	0.87 ± 0.82	1.93 ± 1.00	< .001
item 2: Kukosa furaha au kuboeka/mawazo mengi	0.94 ± 0.82	2.21 ± 0.98	< .001	1.12 ± 0.83	2.36 ± 0.93	< .001
Item 3: usingizi	0.73 ± 0.98	1.58 ± 1.12	.004	0.76 ± 1.00	1.57 ± 1.09	.044
Item 4: kuchoka	0.77 ± 0.83	1.63 ± 1.01	.015	0.89 ± 0.90	1.64 ± 1.08	.042
Item 5: hamu ya kula/kula sana	0.79 ± 0.96	1.68 ± 1.16	.034	0.87 ± 0.97	1.86 ± 1.17	.005
Item 6: vibaya kujihusu	0.90 ± 1.01	1.95 ± 0.97	< .001	1.08 ± 1.08	2.00 ± 1.04	.026
Item 7: shida ya kuzingatia au kuwa makini	0.77 ± 0.89	1.32 ± 1.00	.102	0.77 ± 0.84	1.50 ± 1.09	.031
Item 8: kutembea/kuongea	0.59 ± 0.84	1.74 ± 1.15	.004	0.65 ± 0.90	1.93 ± 1.14	< .001
Item 9: mawazo kufa	0.37 ± 0.70	1.53 ± 1.22	.006	0.40 ± 0.72	1.71 ± 1.20	< .001
PHQ-9 Total Scores	6.66 ± 5.15	15.42 ± 6.87	< .001	7.41 ± 5.26	16.50 ± 6.96	< .001

clinically valid measurement tools to determine the prevalence of depression at the population level, increasing the feasibility of data collection by using tools that trained interviewers can administer in the context of population-based data collection. In addition, the range of cut-off scores provided with this analysis can inform practitioners, including clinicians, policymakers, and researchers, to identify the best fit for the intended use in Kenyan settings.

### Acknowledgments

We acknowledge and thank members of the MMAP Technical Advisory Group for their contributions in shaping this work: Abiodun Adewuya, Claudia Cappa, Regina Guthold, Vikram Patel, George Patton, James Scott, Chiara Servili, and Danuta Wasserman. We are also grateful to Jill Ahs in her role as UNICEF mental health MMAP consultant for providing technical guidance and support with documentation that informed validation field work. Kasra Zarei and Georgia Eleftheriou for editing the document.

We are also grateful to Priscilla Idele from the UNICEF Office of Research for her support in facilitating the Kenya field pilot work and technical guidance in the process.

### Funding

UNICEF participation in this work was supported, in part, by the Bill & Melinda Gates Foundation [INV-001395]. Under the grant conditions of the Foundation, a Creative Commons Attribution 4.0 Generic License has already been assigned to the Author Accepted Manuscript version that might arise from this submission. MMAP validation field work in Kenya presented in

this manuscript was funded by the Government of Italy, through a grant to the UNICEF Office of Research—Innocenti. Research reported in this publication was also supported by the Fogarty International Center of the National Institutes of Health under Award Number K43TW010716, which also supported the contributions of M.K. to this work. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

### Supplementary Data

Supplementary data related to this article can be found at <https://doi.org/10.1016/j.jadohealth.2022.10.003>.

### References

- [1] Institute for Health Metrics and Evaluation (IHME). Global burden of disease collaborative Network. Global burden of disease study 2019 (GBD 2019) results. Seattle, United States. 2020. Available at: <http://ghdx.healthdata.org/gbd-results-tool>. Accessed March 15, 2022.
- [2] World Health Organization. Global health estimates for 2019. World Health Organization, Geneva, Switzerland. Available at: <https://www.who.int/data/global-health-estimates>. Accessed November 25, 2022.
- [3] Ndeti DM, Khasakhala L, Nyabola L, et al. The prevalence of anxiety and depression symptoms and syndromes in Kenyan children and adolescents. *J Child Adolesc Ment Health* 2008;20:33–51.
- [4] Khasakhala LI, Ndeti DM, Mutiso V, et al. The prevalence of depressive symptoms among adolescents in Nairobi public secondary schools: Association with perceived maladaptive parental behaviour. *Afr J Psychiatry* 2012;15:106–13.
- [5] APHRC. Adolescent mental health in Kenya: Where is the data? - APHRC. Available at: <https://aphrc.org/blogarticle/adolescent-mental-health-in-kenya-where-is-the-data/>. Accessed November 25, 2022.

- [6] Mathai M, Mwayo AW, Mutavi T, et al. Child and adolescent mental health in Kenya: Do we need a child and adolescent mental health policy? *Child Behav Heal Sub-saharan Africa* 2022;1:125–43.
- [7] Richardson R. Exploration of the state of mental Healthcare services for youth in Nairobi. University of North Carolina at Chapel Hill: Independent Study Project (ISP); 2016:5–32.
- [8] Carvajal L, Ahs JW, Requejo JH, et al. Measurement of mental health among adolescents at the population level: A Multicountry protocol for adaptation and validation of mental health measures. *J Adolesc Health* 2022;1–7.
- [9] UNICEF. The State of the World's Children 2021: On My Mind – promoting, protecting and caring for children's mental health. New York. October 2021-UNICEF.
- [10] Omoro SAO, Fann JR, Weymuller EA, et al. 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* 2006;36:367–81.
- [11] Mwangi P, Nyongesa MK, Koot HM, et al. 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. *J Affect Disord Rep* 2020;1:100013.
- [12] Monahan PO, Shacham E, Reece M, et al. Validity/reliability of PHQ-9 and PHQ-2 depression scales among adults living with HIV/AIDS in western Kenya. *J Gen Intern Med* 2009;24:189–97.
- [13] Diaz T, Strong KL, Cao B, et al. A call for standardized age-disaggregated health data. *Lancet Heal Longev* 2021;2:e436–43.
- [14] World Bank. Countries and Economies. Available at: <https://data.worldbank.org/country/kenya?view=chart>. Accessed August 18, 2022.
- [15] Kenya National Bureau of Statistics. Kenya population and housing census. 2019. Available at: [www.knbs.or.ke](http://www.knbs.or.ke). Accessed March 16, 2022.
- [16] University of Pennsylvania. African studies center, East Africa living Encyclopedia. Available at: <https://www.africa.upenn.edu/NEH/klanguages.htm>. Accessed March 16, 2022.
- [17] Burkey MD, Adhikari RP, Ghimire L, et al. Validation of a cross-cultural instrument for child behavior problems: The disruptive behavior international scale – Nepal version. *BMC Psychol* 2018;6:51.
- [18] Jordans MJD, Kohrt BA, Luitel NP, et al. Accuracy of proactive case finding for mental disorders by community informants in Nepal. *Br J Psychiatry* 2015;207:501–6.
- [19] van den Broek M, Hegazi L, Ghazal N, et al. Accuracy of a proactive case detection tool for internalizing and externalizing problems among children and adolescents. *J Adolesc Health* 2021;1–8. <https://doi.org/10.1016/j.jadohealth.2021.03.011>.
- [20] Kroenke K, Spitzer RL, Williams JBW. The PHQ-9: Validity of a brief depression severity measure. *J Gen Intern Med* 2001;16:606–13.
- [21] American Psychiatric Association. Diagnostic and statistical mental disorders manual of fifth edition dsm-5, vol. 17; 2013.
- [22] Kroenke K, Spitzer RL. The PHQ-9: A new depression diagnostic and severity measure. *Psychiatr Ann* 2002;32:509–15.
- [23] Liu S-I, Yeh Z-T, Huang H-C, et al. Validation of Patient Health Questionnaire for depression screening among primary care patients in Taiwan. *Compr Psychiatry* 2011;52:96–101.
- [24] Arroll B, Goodyear-Smith F, Crengle S, et al. Validation of PHQ-2 and PHQ-9 to screen for major depression in the primary care population. *Ann Fam Med* 2010;8:348–53.
- [25] González-Blanch C, Medrano LA, Muñoz-Navarro R, et al. Factor structure and measurement invariance across various demographic groups and over time for the PHQ-9 in primary care patients in Spain. *PLoS One* 2018;13:e0193356.
- [26] Hanlon C, Medhin G, Selamu M, et al. Validity of brief screening questionnaires to detect depression in primary care in Ethiopia. *J Affect Disord* 2015;186:32–9.
- [27] Sun Y, Fu Z, Bo Q, et al. The reliability and validity of PHQ-9 in patients with major depressive disorder in psychiatric hospital. *BMC Psychiatry* 2020;20:1–7.
- [28] Beard C, Hsu KJ, Rifkin LS, et al. Validation of the PHQ-9 in a psychiatric sample. *J Affect Disord* 2016;193:267–73.
- [29] Woldetsensay YK, Belachew T, Tesfaye M, et al. Validation of the Patient Health Questionnaire (PHQ-9) as a screening tool for depression in pregnant women: Afaan Oromo version. *PLoS One* 2018;13:e0191782.
- [30] Seo J-G, Park S-P. Validation of the patient health questionnaire-9 (PHQ-9) and PHQ-2 in patients with migraine. *J Headache Pain* 2015;16:1–7.
- [31] van Steenberg-Weijnenburg KM, de Vroeghe L, Ploeger RR, et al. Validation of the PHQ-9 as a screening instrument for depression in diabetes patients in specialized outpatient clinics. *BMC Health Serv Res* 2010;10:1–6.
- [32] Wang W, Bian Q, Zhao Y, et al. Reliability and validity of the Chinese version of the Patient Health Questionnaire (PHQ-9) in the general population. *Gen Hosp Psychiatry* 2014;36:539–44.
- [33] Lamela D, Soreira C, Matos P, et al. Systematic review of the factor structure and measurement invariance of the patient health questionnaire-9 (PHQ-9) and validation of the Portuguese version in community settings. *J Affect Disord* 2020;276:220–33.
- [34] Kohrt BA, Luitel NP, Acharya P, et al. Detection of depression in low resource settings: Validation of the patient health questionnaire (PHQ-9) and cultural concepts of distress in Nepal. *BMC Psychiatry* 2016;16:58.
- [35] Kim YE, Lee B. The psychometric properties of the patient health questionnaire-9 in a sample of Korean university students. *Psychiatry Investig* 2019;16:904.
- [36] Du N, Yu K, Ye Y, et al. Validity study of patient health questionnaire-9 items for internet screening in depression among Chinese university students. *Asia Pac Psychiatry* 2017;9:10.
- [37] Adewuya AO, Ola BA, Afolabi OO. Validity of the patient health questionnaire (PHQ-9) as a screening tool for depression amongst Nigerian university students. *J Affect Disord* 2006;96:89–93.
- [38] van Ommeren M, Sharma B, Thapa S, et al. Preparing instruments for transcultural research: Use of the translation monitoring form with Nepali-speaking Bhutanese refugees. *Transcult Psychiatry* 1999;36:285–301.
- [39] Kohrt BA, Jordans MJD, Tol WA, et al. Validation of cross-cultural child mental health and psychosocial research instruments: Adapting the depression self-rating scale and child PTSD symptom scale in Nepal. *BMC Psychiatry* 2011;11:127.
- [40] Kaufman J, Birmaher B, Brent D, et al. Schedule for Affective disorders and Schizophrenia for School-age children-Present and Lifetime version (K-SADS-PL): Initial reliability and validity data. *J Am Acad Child Adolesc Psychiatry* 1997;36:980–8.
- [41] DSM-IV. Diagnostic and statistical manual of mental disorders. 4th ed. Washington, DC: American Psychiatric Association; 1994.
- [42] Bradley AP, Longstaff ID. Sample size estimation using the receiver operating characteristic curve. In: *Proc 17th Int Conf Pattern Recognition*, 2004. ICPR 2004, vol. 4. The University of Queensland, St Lucia, QLD 4072, Australia: IEEE; 2004:428–31.
- [43] Metz CE. Basic principles of ROC analysis. *Semin Nucl Med* 1978;8:283–98.
- [44] Hosmer DW Jr, Lemeshow S, Sturdivant RX. *Applied logistic regression*, vol. 398. New York: John Wiley & Sons; 2013.
- [45] Glas AS, Lijmer JG, Prins MH, et al. The diagnostic odds ratio: A single indicator of test performance. *J Clin Epidemiol* 2003;56:1129–35.
- [46] Rogers WH, Adler DA, Bungay KM, et al. Depression screening instruments made good severity measures in a cross-sectional analysis. *J Clin Epidemiol* 2005;58:370–7.
- [47] Levis B, Benedetti A, Thombs BD. Accuracy of patient health questionnaire-9 (PHQ-9) for screening to detect major depression: Individual participant data meta-analysis. *BMJ* 2019;365:11476.
- [48] Negeri ZF, Levis B, Sun Y, et al. Accuracy of the patient health questionnaire-9 for screening to detect major depression: Updated systematic review and individual participant data meta-analysis. *BMJ* 2021;375:n2183.
- [49] Manea L, Gilbody S, McMillan D. Optimal cut-off score for diagnosing depression with the patient health questionnaire (PHQ-9): A meta-analysis. *Cmaj* 2012;184:E191–6.
- [50] Urtasun M, Daray FM, Teti GL, et al. Validation and calibration of the patient health questionnaire (PHQ-9) in Argentina. *BMC Psychiatry* 2019;19:291.
- [51] Lotrakul M, Sumrithe S, Saipanish R. Reliability and validity of the Thai version of the PHQ-9. *BMC Psychiatry* 2008;8:46.
- [52] Sherina MS, Arroll B, Goodyear-Smith F. Criterion validity of the PHQ-9 (Malay version) in a primary care clinic in Malaysia. *Med J Malaysia* 2012;67:309–15.
- [53] Bhana A, Rathod SD, Selohilwe O, et al. The validity of the Patient Health Questionnaire for screening depression in chronic care patients in primary health care in South Africa. *BMC Psychiatry* 2015;15:118.