Validation of brief screening instruments for internalizing and externalizing disorders in Mozambican adolescents

Kathryn L. Lovero1*, Salma Ebrahim Adam2, Carolina Ezequias Bila2, Elda D. Canda2, Maria Eduarda Fernandes2, Teresa I. Baltazar Rodrigues2, Mariel C. Tai Sander3, Claude A. Mellins3,4, Cristiane S. Duarte3, Palmira Fortunato dos Santos2 and Milton L. Wainberg3

Abstract

Background: Mental disorders are the leading cause of disability for youth worldwide. However, there is a dearth of validated, brief instruments to assess mental health in low- and middle-income countries (LMIC). We aimed to facilitate identification of mental disorders in LMIC contexts by adapting and validating measures of internalizing and externalizing disorders for adolescents in Mozambique, an LMIC in southeastern Africa.

Methods: We selected instruments with good support for validity in high-income and other LMIC settings: the Patient Health Questionnaire Adolescent (PHQ-A), Generalized Anxiety Disorders 7 (GAD-7), and Strengths and Difficulties Questionnaire (SDQ). Instruments were adapted by local and international mental health specialists followed by cognitive interviews (n = 48) with Mozambican adolescents. We administered the instruments along with the Miniature International Neuropsychiatric Interview for Children and Adolescents (MINI-KID) to 485 adolescents aged 12–19 years attending two secondary schools in Maputo City, Mozambique. One week later, we re-administered instruments to a randomly selected sample of 49 adolescents.

Results: Participants were 66.2% (n = 321) female and the average age was 15.9 (S.D = 1.7). Internal consistency (alpha = 0.80, PHQ-A; 0.84, GAD-7; 0.80, SDQ) and test–retest reliability (ICC = 0.74, PHQ-A; 0.70, GAD-7; 0.77, SDQ) were acceptable for the PHQ-A, GAD-7, and the full SDQ. The SDQ internalizing subscale showed poor test–retest reliability (ICC = 0.63) and the SDQ externalizing subscale showed poor internal consistency (alpha = 0.65). All instruments demonstrated good sensitivity and specificity (> 0.70). Youden’s index identified optimal cutoff scores of 8 for the PHQ-A, 5 for the GAD-7, 10 for the SDQ internalizing, and 9 for the SDQ externalizing subscales, though a range of scores provided acceptable sensitivity and specificity.

Conclusions: Our data supports reliability and validity of the PHQ-A, GAD-7, and SDQ instruments for rapidly assessing mental health problems in Mozambican adolescents. Use of these tools in other contexts with limited specialists may assist with expanding mental health assessment. Specific instrument and cutoff selection should be based on screening goals, treatment resources, and program objectives.

Keywords: Youth, LMIC, Psychometrics, PHQ, GAD, SDQ, Depression, Anxiety, Disruptive behavior disorders, ADHD

Background

Globally, mental disorders are the largest contributor to burden of disease in adolescents, accounting for 45% of years lived with disability [1]. It is estimated that 90% of...
adolescents live in low- and middle-income countries (LMIC) [2] and that 10–20% of these adolescents have one or more mental disorders [3]. Around 75% of mental disorders have their onset prior to or during adolescence [4], and mental illness in youth is associated with poor lifelong health, educational, social and economic outcomes [5–14]. Thus, there is an urgent need to increase early identification and treatment of adolescents with mental disorders, particularly those in LMIC settings.

Despite the high estimated burden of mental disorders in LMIC adolescents, there is an enormous shortage of mental health providers in these regions and almost no specialized child and adolescent mental health services [15]. Task-shifting of mental health services and integration of mental health care into primary and community care settings have been proposed as solutions to the vast treatment gap in LMIC [16]. However, in the absence of an adequate cadre of trained mental health specialists, locally validated brief assessment instruments are needed to help accurately a) estimate the prevalence of mental disorders and efficiently target resources, b) screen for these disorders in primary and community care settings and triage for treatment, and c) evaluate the impact of task-shifted mental health interventions. A recent systematic review found just 17 studies that validated mental health screening tools in LMIC adolescents [17], only four of which were in sub-Saharan Africa, where nearly 20% of the world’s adolescents reside [18].

In Mozambique, a low-income country located in southeastern Africa, adolescents account for one-third of the country’s over 31 million inhabitants and are the fastest growing population group [19]. While mental health services are offered free-of-charge through the national health system, there are less than 500 specialists (psychiatrists, psychologists, and psychiatric technicians) nationwide [20]. The Mozambique Ministry of Health has included child and adolescent mental health as a priority in its Strategy and Action Plan for Mental Health 2016–2026 [21] and is working to implement task-shifted adolescent mental health services. However, like in other LMIC [22], there has been almost no research on adolescent mental health locally and, though two studies have demonstrated the validity of brief mental health assessment tools adapted for Mozambican adults [23, 24], to our knowledge, none have been evaluated in Mozambican adolescents.

This study aimed to adapt commonly used brief screening tools to assess adolescent mental health problems and evaluate their measurement properties in Mozambican adolescents. We focused specifically on instruments previously validated in other LMIC adolescent populations to screen for internalizing disorders (depression and anxiety), as they are estimated to be the highest burden disorders in adolescents globally as well as in Mozambique [25], and externalizing disorders (conduct disorder, oppositional defiant disorder, and attention deficit hyperactivity disorder [ADHD]), as they are the strongly associated with poor educational and social outcomes [26]. Following a review of the literature and discussion with a team of international mental health experts, we selected the Patient Health Questionnaire Adolescent (PHQ-A) and the Generalized Anxiety Disorder 7 (GAD-7) because of their demonstrated performance identifying depression and anxiety, respectively, in adolescents of high-income countries and LMIC [27–32]. We chose to also adapt and evaluate the Strengths and Difficulties Questionnaire (SDQ), which assesses symptoms of both internalizing disorders and externalizing disorders, as it is one of the most widely used brief screening instruments in youth worldwide and the only open-access brief instrument for externalizing disorders, although it is not designed to discern specific disorders.

**Methods**

**Sample**

This cross-sectional study was conducted at two secondary schools in peripheral and central regions of the Mozambican capital, Maputo City. We selected these schools as typical of adolescents from lower and higher socioeconomic urban areas, respectively. During the school year our study was conducted, the peripheral urban school had enrolled a total of 3243 students (58.2% female); the central urban school had enrolled 2987 students (55.1% female).

The study population included adolescents aged 12–19 years in grades 8–12. We aimed to include 50 adolescents per grade at each school (n = 500 total participants) and randomly selected 2–3 classes (~100 students) per grade per school to be invited to participate. Two weeks prior to data collection, the research team presented an overview of the study to teachers of selected classrooms. Following standard procedures for research at Mozambican public schools, teachers were given copies of consent and assent forms. Teachers then instructed students to take the form home to review with their caregiver and return signed if they assent to participate and their caregiver consented to their participation. Consent and assent forms contained contact information for the research team to resolve any questions before consenting or assenting to participation. The research team collected consent and assent forms from teachers prior to data collection. Of the 1320 adolescents invited to participate, 1150 (87.1%) returned consent forms. Returned consent forms were order-randomized prior to data collection. From September 9th to October 2nd, 2019, students who had returned consent forms from two of the participating
classes were gathered in an outdoor area each day, where research assistants called on them to be interviewed from the randomly ordered consent forms. We were able to evaluate a total of 493 students for eligibility. One adolescent was excluded because they returned a consent form without a caregiver signature and three were excluded because they were over 19 years old. An additional four adolescents were excluded from present analyses because they did not complete the entire interview. The final sample of 485 adolescents included in this analysis were majority female (N=321, 66.2%), with a mean age of 15.9 years (SD = 1.7 years).

**Measures**

We first conducted a literature review to identify screening instruments for internalizing disorders and externalizing disorders that had been previously validated in other LMIC. We then met with a team of four American and three Mozambican mental health practitioners, policy makers, and researchers specializing in disorder measurement to review the evidence considering: 1) demonstrated performance in high-income countries; 2) demonstrated performance in LMIC, with special attention to those validated in Brazil, which is, like Mozambique, a lusophone LMIC; 3) number of items, with fewer items being preferable in low-resource settings; and 4) cost per use, as LMIC are very unlikely to have funds to pay for screen administration. From these meetings, the PHQ-A, GAD-7, and SDQ were selected for assessment in the present study. The PHQ-9 is one of the most commonly used depression screening instruments [33, 34], and has been validated for use in adults in numerous LMIC, including Brazil and Mozambique [17, 24]. The PHQ-A is a nine-item version of the PHQ-9 adapted for adolescents [35] and has been demonstrated to have good sensitivity and specificity for identifying depression using adolescent self-report in high-income countries and LMIC [28, 29, 31]. The GAD-7 is a commonly used instrument to screen for anxiety in high-income countries [36, 37] that has been validated for use in LMIC adults [38, 39] and adolescents from high-income countries [30, 32] and LMIC [27]. The SDQ is a widely-used, 25-item instrument for youth (ages 4–17 years) emotional and behavioral functioning [40]. The SDQ includes five subscales which can be grouped into two 10-item subscales: an internalizing subscale (i.e. emotional and peer problems subscales) and an externalizing subscale (i.e. conduct and hyperactivity/inattention subscales) [41]. The fifth subscale of the instrument, prosocial, is not included in either the internalizing nor externalizing subscales. While the SDQ includes both internalizing and externalizing subscales, it is not designed to distinguish between disorders within these diagnostic categories (i.e. depression or anxiety). The SDQ is the only open-access measure of externalizing disorders we identified and has been used in over 50 studies in Africa [42]; however, the performance of the SDQ has mixed results in both high-income country and LMIC settings [43–48]. All three instruments are available for use without charge.

To adapt these instruments, we followed World Health Organization procedures for translation and adaptation of psychiatric assessment instruments [49]. The adaptation team included five Mozambican psychologists (SA, CB, EC, MEF, TR) and one American global mental health researcher (KLL). We began with existing Portuguese versions from Brazil or Portugal and for all instruments, except the SDQ, made minor adjustments (e.g. verb tense, local slang for substances) to Mozambican Portuguese. All adaptations were collaboratively reviewed by a bilingual group of 9 Mozambican and 2 American mental health professionals (psychiatrists and psychologists), adjusted based on feedback, back-translated by a bilingual native English speaker not associated with the present study, and verified for accuracy to initial English versions of instruments. The translation team then conducted cognitive interviews with adolescents attending primary care clinics to assess the accuracy, appropriateness, and comprehensibility of measures [50]. We conducted cognitive interviewing using five focus groups [51] of Mozambican adolescents (n = 48) in which adolescents were asked a series of guide questions, followed by further probing, with regard to each item in the screening instruments: 1) Did you understand what this question was asking? 2) Was it difficult to respond to the question? 3) What do you think this question was asking? Can you repeat the question in your own words? 4) Did you understand the response options given? 5) Were you able to find a response option you would select? 6) Were there any words in the question or response options that you didn’t understand? 7) Were there any words or expressions in the question that you found inappropriate or offensive? The translation team iteratively updated instruments based on feedback between each focus group, such that the fifth focus group reviewed and approved the finalized instruments.

We used the Portuguese version [52] of the MINI International Neuropsychiatric Interview for Children and Adolescents (MINI-KID), a structured diagnostic interview for DSM-IV and ICD-10 disorders [53], as our gold standard mental health assessment tool. We assessed presence of a current disorder on the following modules: major depressive episode (past 2 weeks), generalized anxiety disorder (past 6 months), separation anxiety disorder (past month), social phobia (past month), agoraphobia (past 6 months), panic disorder (past month), ADHD (past 6 months), conduct disorder (past 12 months),
oppositional defiant disorder (past 6 months), PTSD (past month), alcohol use disorder (past 12 months), substance use disorder (past 12 months), suicidality (past month). In the present analyses, we focus only on results from modules corresponding to internalizing disorders (major depressive episode, generalized anxiety disorder, separation anxiety disorder, social phobia, agoraphobia, and panic disorder) and externalizing disorders (ADHD, conduct disorder, and oppositional defiant disorder). We also collected sociodemographic data including grade, age, gender, preferred language, religion, and household goods.

**Data collection**

Trained interviewers (13 local psychologists) administered the self-report sociodemographic questionnaire, MINI-KID, and screening instruments. All participants responded to the sociodemographics questionnaire first, and were randomized to respond to either the screening battery or the MINI-KID next. Immediately following completion, a different interviewer administered the remaining measure. We invited a random selection of 49 (10%) of participants to return one-week later and be re-tested with the screening battery. All data was collected on tablets using the REDCAP electronic data collection platform [54].

Following completion of interviews, on-site research team members (SA, CB, EC, MEF, TR) provided referrals to mental health services in the neighboring primary care clinic (<1 block from the schools) to all adolescents screened positive for any disorder on the MINI-KID. For adolescents with acute suicide risk, research team members immediately contacted caregivers and accompanied the adolescent to the psychologist at the neighboring primary care clinic.

**Data analyses**

Analyses were conducted in Stata IC 14. Summary statistics of participant demographics were calculated for the total sample and for those diagnosed with depression, anxiety, and externalizing disorders (based on MINI-KID). Cronbach’s alpha statistics were calculated for each instrument as well as the SDQ internalizing and externalizing subscales to evaluate internal consistency. Individual item correlations were also evaluated, with those whose exclusion would improve the Cronbach’s alpha of the scale by > 0.01 considered for removal. Intraclass correlation coefficients (ICC) and 95% confidence intervals (CI) were calculated to assess test–retest reliability. For the purpose of assessing the criterion validity of the PHQ-A, we compared scores to diagnosis of major depressive disorder in the MINI-KID; for the GAD-7, we compared scores to diagnoses of generalized anxiety disorder, agoraphobia, social phobia, separation anxiety, and/or panic disorder; SDQ internalizing subscale scores were compared to diagnoses of major depressive disorder, generalized anxiety disorder, agoraphobia, social phobia, separation anxiety, and/or panic disorder; SDQ externalizing subscale scores were compared to diagnoses of conduct disorder and oppositional defiant disorder (disruptive behavior disorders) and ADHD. We generated receiver operating characteristic curves (ROC) for each instrument and calculated the area under the curve (AUC). Youden’s index (J) was calculated to identify the optimal cutoff score that maximizes sensitivity and specificity for each scale [55]. Positive predictive value and negative predictive value for instruments were calculated using the prevalence in the study sample as well as a range of prevalences (1%, 5%, and 10%) identified in previous youth samples around the globe [22, 56].

**Results**

**Instrument adaptation**

Overall, minimal changes in content were required following cognitive interviews. Specifically, the grammatical phrasing had to be changed to “you” statements (e.g. “You felt…” instead of “Feeling…” on the PHQ-A and GAD-7) for adolescents to better understand the response options on the questionnaires. Other small changes included contextually relevant examples of possessions (e.g. pens instead of CDs).

**Participant characteristics**

Of the 485 included adolescents, 441 (90.9%) reported Portuguese as their preferred language and 423 (87.8%) reported being Christian or Catholic. In total, the interviews lasted on average 17.4 min (SD = 7.2) for the screening battery and 12.6 min (SD = 7.0) for the MINI-KID. Table 1 presents the rate of disorders identified in participants through MINI-KID interviews. Internalizing disorder, agoraphobia, social phobia, separation anxiety, and/or panic disorder; SDQ internalizing subscale scores were compared to diagnoses of major depressive disorder, generalized anxiety disorder, agoraphobia, social phobia, separation anxiety, and/or panic disorder; SDQ externalizing subscale scores were compared to diagnoses of conduct disorder and oppositional defiant disorder (disruptive behavior disorders) and ADHD. We generated receiver operating characteristic curves (ROC) for each instrument and calculated the area under the curve (AUC). Youden’s index (J) was calculated to identify the optimal cutoff score that maximizes sensitivity and specificity for each scale [55]. Positive predictive value and negative predictive value for instruments were calculated using the prevalence in the study sample as well as a range of prevalences (1%, 5%, and 10%) identified in previous youth samples around the globe [22, 56].

**Table 1** Participant psychiatric disorder diagnoses based on MINI-KID interviews

| Disorder                                      | Total N (%) |
|-----------------------------------------------|-------------|
| Internalizing Disorders                       | 95 (19.6%)  |
| Major Depressive Episode                      | 41 (8.5%)   |
| Anxietya                                      | 85 (17.5%)  |
| Externalizing Disorders                       | 36 (7.4%)   |
| ADHD                                          | 16 (3.3%)   |
| Disruptive Behavior Disorders                  | 32 (6.6%)   |

a Includes diagnoses of generalized anxiety disorder, separation anxiety disorder, social phobia, agoraphobia; and panic disorder
b Includes diagnoses of conduct and oppositional defiant disorders
disorders were over twice as common as externalizing disorders in the present sample.

Internal consistency and test–retest reliability
The PHQ-A, GAD-7, and SDQ (25 items) showed good internal consistency (Cronbach's alpha ≥ 0.80; Table 2); whereas internal consistency for subscales was worse: moderate for the SDQ internalizing disorders (Cronbach's alpha = 0.73) and poor for the SDQ externalizing disorders (Cronbach's alpha = 0.65). Individual item covariance ranged from 0.19–0.20 on the PHQ-A, 0.29–0.31 on the GAD-7, 0.10–0.14 on the SDQ internalizing disorders subscale, and 0.06–0.09 on the SDQ externalizing disorders subscale (Table 3). Dropping any one item of the PHQ-A or GAD-7 did not improve internal consistency of the scale (alpha without item, Table 3). SDQ items 11 (“I have one good friend or more”), 16 (“I am nervous in new situations. I easily lose confidence”), and 23 (“I get along better with adults than with people my own age”) had low correlations with the internalizing disorders subscale (item test and item rest correlation, Table 3), and removal of each individually increased the Cronbach's alpha of the subscale from 0.73 to 0.75 (alpha without item, Table 3). Item 7 of the SDQ (“I usually do as I am told”) had poor correlation with the externalizing disorders subscale, and its removal increased the Cronbach's alpha of the subscale from 0.65 to 0.66.

The correlation between test and retest scores (Table 2) was moderate for the PHQ-A, GAD-7, full SDQ, and SDQ externalizing subscale (ICC ≥ 0.70) but poor for the SDQ internalizing subscale (ICC = 0.63, 95%CI = 0.43–0.77).

Criterion validity
A summary of the performance of the PHQ-A, GAD-7, and SDQ as compared to MINI disorder diagnoses is shown in Table 4. The area under the ROC curve (AUC) for all scales was > 0.70 (full ROC curves shown in Additional Fig. 1). Youdin's index identified an optimal cutoff score of 8 for the PHQ-A, 5 for the GAD-7, 10 for the SDQ internalizing and 9 for the SDQ externalizing subscales. At these cutoff scores, sensitivity and specificity of the PHQ-A for depression, GAD-7 for anxiety,
SDQ internalizing subscale for internalizing disorders, and SDQ externalizing subscale for externalizing disorders were all acceptable (> 0.70). The SDQ internalizing subscale showed slightly higher sensitivity but lower specificity for adolescents with depression (sens. = 0.83, spec. = 0.68) than adolescents with anxiety (sens. = 0.74, spec. = 0.72); the SDQ externalizing subscale had slightly lower specificity but higher sensitivity for adolescents with disruptive behavior disorders (sens. = 0.72, spec. = 0.70) than ADHD (sens. = 0.75, spec. = 0.57). Positive predictive values (PPV) and negative predictive values (NPV) for the scales at the identified optimal cutoff scores for different disorder prevalences can be found in Additional Table 1. Details on the performance of scales at all possible cutoff scores are in Additional Table 2.

Table 4 Performance of the PHQ-A, GAD-7, and SDQ Screening Tools in Mozambican Adolescents

|                          | AUC (95% CI) | Cutoff | Sensitivity | Specificity | % Correctly Classified |
|--------------------------|-------------|--------|-------------|-------------|------------------------|
| PHQ-A                    | 0.85 (0.76–0.90) | 8      | 0.78        | 0.80        | 80.2                   |
| GAD-7                    | 0.84 (0.79–0.89) | 5      | 0.81        | 0.72        | 73.6                   |
| SDQ Internalizing Subscale |             |        |             |             |                        |
| Any Internalizing Disorder | 0.80 (0.75–0.85) | 10     | 0.76        | 0.74        | 74.3                   |
| Depression               | 0.81 (0.74–0.87) | 10     | 0.83        | 0.68        | 69.7                   |
| Anxiety                  | 0.78 (0.73–0.84) | 10     | 0.74        | 0.72        | 72.6                   |
| SDQ Externalizing Subscale |             |        |             |             |                        |
| Any Externalizing Disorder | 0.72 (0.65–0.80) | 9      | 0.78        | 0.59        | 60.4                   |
| ADHD                     | 0.70 (0.60–0.81) | 9      | 0.75        | 0.57        | 57.9                   |
| Disruptive Behavior Disorders | 0.74 (0.67–0.82) | 10     | 0.72        | 0.70        | 70.1                   |

Abbreviations: AUC Area under the curve, CI Confidence interval, PHQ-A Patient Health Questionnaire for Adolescents, GAD-7 Generalized Anxiety Disorders 7, SDQ Strengths and Difficulties Questionnaire

Discussion

We sought to evaluate screening measures for internalizing and externalizing disorders that could be used easily and effectively by non-specialists in Mozambique, i.e., open access, few items, established validity in international settings. We selected the PHQ-A and GAD-7 owing to their strong performance identifying depression and anxiety symptoms, respectively, in adolescents of high-income countries and LMIC [27–32]. As in previous studies, the PHQ-A and GAD-7 in our sample showed acceptable internal consistency and test–retest reliability. Moreover, both instruments demonstrated good sensitivity and specificity for depression and anxiety, respectively.

Additionally, we selected the SDQ because it has demonstrated satisfactory performance and is one of the most commonly used instruments globally for assessing both internalizing disorders and externalizing disorders [48], and to our knowledge the only brief screen for externalizing disorders without a cost-per-use. In African youth populations, the internal consistency of the SDQ has been demonstrated to vary widely (Cronbach's alpha = 0.18–0.89) [42]. In our sample, internal consistency of the full SDQ (Cronbach's alpha = 0.80) and internalizing subscale (Cronbach's alpha = 0.73) were good, while that of the externalizing subscale was weaker (Cronbach's alpha = 0.65). These findings are similar to data from self-report, non-clinical adolescent samples in high-income countries and other LMIC, in which acceptable internal consistency of the full instrument but not the subscales has been found [57–63]. Moreover, previous research has demonstrated that the factor structure of the SDQ differs cross-culturally, suggesting that individual SDQ items may be more culturally-dependent than others and that further testing of the measure in diverse populations is required [64]. Our assessment of the performance of individual items of the SDQ subscales identified four questions (items 7, 11, 16, and 23) that had poor correlation with other items. However, removal of the items individually only minimally improved internal consistency of the subscales (Cronbach's alpha = 0.73 to 0.75 for internalizing subscale and Cronbach's alpha = 0.65 to 0.66 for externalizing subscale). Studies of adolescents from diverse cultural settings, including South Africa, Norway, Israel, Germany, and China, similarly found that item 7 (“I usually do as I am told”) and items 11 and 23 (“I have one good friend or more” and “I get along better with adults than with people my own age”) had poor factor loadings for conduct problems and peer relationships, respectively [47, 57, 58, 61, 65, 66], suggesting these items may be generally less
predictive of problem areas. Previous research in Zambian youth described challenges in determining a local linguistic equivalent to translate item 16 (“I am nervous in new situations. I easily lose confidence”) [67]. Because our cognitive interviews with adolescents did not reveal any difficulties in interpretation of this item, we do not believe its inconsistency is due to a translation issue. However, further research is needed to determine if this item is not culturally relevant or if its interpretation is disparate among Mozambican youth.

Finally, the optimal cutoff score for each instrument in our sample was lower than recommended cutoff scores from high-income countries [30, 31, 68], a finding that has been demonstrated in multiple LMIC for both children [29] and adults [24, 69, 70]. It is important to note, though, that while the optimal cutoff scores in the present study provide the best balance of sensitivity and specificity for identifying a disorder in this population, they do not account for severity of symptoms or functioning. As such, these optimal cutoff scores are more appropriate for certain screening purposes, e.g., estimating disorder prevalence in the general adolescent population, than others, e.g. identifying adolescents who should receive treatment in a resource-limited setting where inclusion of adolescents with mild symptoms could overwhelm the system. Still, we recommend proceeding with caution when considering increasing the cutoff score, owing to the fairly steep decline in sensitivity using higher cutoff scores. For example, at the recommended cutoff score of 10 for moderate depression, the sensitivity of the PHQ-A in our sample drops to 59%; at the recommended cutoff score of 10 for moderate anxiety, the sensitivity of the GAD-7 in our sample drops to 46%. Therefore, screening purpose, expected disorder prevalence, and available treatment resources should be carefully considered before selecting a particular cutoff score for instrument implementation.

Our study had a few important limitations. First, adolescents were recruited at public secondary schools in urban and peri-urban areas, and thus may not be generalizable to Mozambican adolescents who do not attend school or who live in rural regions of the country. The cognitive interviews conducted to establish comprehensibility of the questions was done with adolescents who were attending primary care clinics in Maputo City, who represent both adolescents in and out of school. Still, though the instruments may be understood by these adolescents, this does not indicate they would perform similarly in indentifying mental disorders in this population. Additionally, all instruments were validated in Portuguese, which was the predominant language of the study sample but is not the primary language of many Mozambicans. We therefore recommend that future studies aim to assess performance of these instruments in rural, non-school attending, and non-Portuguese speaking Mozambican adolescents. Finally, the timeframe of symptom assessment in the screening instruments differed from the MINI-KID for some disorders (e.g. MINI-KID conduct disorder = past 12 months, SDQ externalizing disorders = past 6 months). Still, our data demonstrated acceptable to good criterion validity of all screening instruments.

Despite these limitations, the present study had a number of strengths. For one, this is the first study to validate any type of mental health screening instrument in Mozambican adolescents. Moreover, our instrument adaptation was driven by input from local mental health experts and adolescents. We also used a gold standard (the MINI-KID) to establish criterion validity. Finally, we included a random selection of adolescents from schools that spanned a wide range of adolescents ages (12–19).

The PHQ-A, GAD-7, and full SDQ demonstrated good internal and test–retest reliability. All scales showed adequate to good criterion validity based on diagnoses made using the MINI-KID interview. These findings indicate these instruments can serve as a useful tool for rapidly assessing adolescent mental health problems in Mozambique without the need of specialized providers and may be of interest to other LMIC looking to expand task-shifted mental health services for adolescents. In terms of selecting when to use the PHQ-A, GAD-7 or SDQ, we recommend that the specific needs of the population being assessed be used to inform decision-making. For example, in a setting where it is important to determine whether a child has depression and/or anxiety, such as when a provider is selecting their specific treatment approach, we recommend the use of the PHQ-A and GAD-7. However, if the need is to identify youth with general internalizing and/or externalizing problems, such as in a community-based screen to provide referrals to mental health services, it may be preferable to use the SDQ.

Conclusions
Here, we for the first time adapt and evaluate the performance of mental health screening measures in Mozambican adolescents and add to the very limited literature on screening measures in LMIC youth, who account for the vast majority of the world’s young people. Our results suggest that the PHQ-A, GAD-7, and SDQ have satisfactory psychometric properties for screening for symptoms of depressive anxiety, internalizing, and externalizing disorders, respectively. Future research should focus on the performance of these measures in rural and non-Portuguese-speaking Mozambican adolescents, as well as their performance in other LMIC.
Supplementary Information
The online version contains supplementary material available at https://doi.org/10.1186/s12888-022-04189-3.

Additional file 1.

Acknowledgements
The authors wish to thank the research participants and the research assistants (Lourena Chemane, Lino Conjo, Ana Cristina Garnito, Olência Gulube, Marcia Mathe, Carolina Matsolo, Meriba Mazove, Aldo Saute, Leonor Sebastião, Maria Helena Simbine, Nilza Simbine, Lourenço Zilria) for their support in data collection. We also thank the mental health professionals who provided care for adolescents diagnosed with mental disorders in the present study and Ministry of Health Department of Mental Health staff (Wilza Fumo, Joaquim Maravel and Lidia Gouveia) for facilitating contacts with Ministry of Education, school administration and health facilities.

Authors' contributions
KLL and MLW conceived of the study. KLL, SEA, CEB, EDC, MEF, and TIBR trained interviewers and supervised data collection. KLL conducted the analysis and drafted the manuscript. All authors contributed to the study design, interpretation of data and revisions of the manuscript, and have reviewed and approved of the final version.

Funding
This study was support by grant numbers K01MH120258 and U19MH113203 from the National Institute of Mental Health and grant number D43TW009675 from the Fogarty International Center. The funders had no role in the study design, data collection and analysis, interpretation of data, or writing of the manuscript.

Availability of data and materials
The dataset generated during the current study are not publicly available due to Mozambican ethics committee regulations, but are available from the corresponding author (Kathryn Lovero, kll2153@cumc.columbia.edu) on reasonable request.

Declarations
Ethics approval and consent to participate
All research was conducted in accordance with the Declaration of Helsinki. All study materials and procedures were approved by the New York State Psychiatric Institute Institutional Review Board (#47/9) and the Eduardo Mondlane University Institutional Health Bioethics Council (ICBS FM & HCM/54/2017). All participants provided written assent and their caregivers provided written informed consent to participate.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
1Department of Sociomedical Sciences, Columbia University Mailman School of Public Health, 722 W. 168th St. 9th Floor, New York, NY 10032, USA.
2Department of Mental Health, Ministry of Health, Maputo, Mozambique.
3Department of Psychiatry, New York State Psychiatric Institute and Columbia University Vagelos College of Physicians and Surgeons, New York, NY, USA.
4HIV Center for Clinical and Behavioral Health, Department of Psychiatry, New York State Psychiatric Institute and Columbia University, New York, NY, USA.

Received: 5 December 2021 Accepted: 18 July 2022
Published online: 12 August 2022

References
1. Gove FM, Bloem PJ, Patton GC, Ferguson J, Joseph V, Coffey C, Sawyer SM, Matthers CD. Global burden of disease in young people aged 10–24 years: a systematic analysis. Lancet. 2011;377(9783):2093–102.
2. The World Bank: World Bank Open Data. 2022. https://data.worldbank.org/. Accessed 31 Mar 2022.
3. Erskine HE, Moffitt TE, Copeland WE, Costello EI, Ferrari AJ, Patton G, Degenhardt L, Vos T, Whiteford HA, Scott JG. A heavy burden on young minds: the global burden of mental and substance use disorders in children and youth. Psychol Med. 2021;51(7):1551–63.
4. Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. Arch Gen Psychiatry. 2005;62(6):593–602.
5. Brown LK, Danovsky MB, Lourie KJ, DiClemente RJ, Ponton LE. Adolescents with psychopathological disorders and the risk of HIV. J Am Acad Child Adolesc Psychiatry. 1997;36(1):1609–17.
6. Chen H, Cohen P, Kasen S, Johnson JG, Berenson GS, Gumpert S. Impact of adolescent mental disorders and physical illnesses on quality of life 17 years later. Arch Pediatr Adolesc Med. 2006;160(1):93–9.
7. Colman I, Murray J, Abbott RA, Maughan B, Kuh D, Croudace TJ, Jones PB. Outcomes of conduct problems in adolescence: 40 year follow-up of national cohort. BMJ (Clinical research ed). 2009;338:a2981.
8. Oommen GR, Pado M. Youths and HIV/AIDS: psychiatry’s role in a changing epidemic. J Am Acad Child Adolesc Psychiatry. 2005;44(8):728–47.
9. Esch P, Bocquet V, Pull C, Couffignal S, Lheurtin T, Graas M, Fend-Harmant L, Anseau M. The downward spiral of mental disorders and educational attainment: a systematic review on early school leaving. BMC Psychiatry. 2014;14(1):237.
10. Fergusson DM, Horwood LJ, Lynskey MT. The effects of conduct disorder and attention deficit in middle childhood on offending and scholastic ability at age 13. J Child Psychol Psychiatry. 1993;34(6):899–916.
11. Fergusson DM, Woodward LJ. Mental health, educational, and social role outcomes of adolescents with depression. Arch Gen Psychiatry. 2002;59(3):225–31.
12. Kjellberg E. DSM-IV conduct disorder symptoms in adolescents as markers of registered criminality. Eur Child Adolesc Psychiatry. 2002;11(1):2–9.
13. Sayal K, Washbrook E, Propper C. Childhood behavior problems and academic outcomes in adolescence: longitudinal population-based study. J Am Acad Child Adolesc Psychiatry. 2015;54(5):360-368.e362.
14. Wickersham A, Sugg HV, Epstein S, Stewart R, Ford T, Downs J. Systematic review and meta-analysis: the association between child and adolescent depression and later educational attainment. J Am Acad Child Adolesc Psychiatry. 2021;60(1):105–18.
15. World Health Organization: Mental Health ATLAS 2017. Geneva: WHO, 2018. Available from http://www.who.int/mental_health/evidence/atlas/mental_health_atlas_2017/en/. (Accessed 3 Oct 2018).
16. Patel V, Kieling C, Makulik PK, Divan G. Improving access to care for children with mental disorders: a global perspective. Arch Dis Child. 2013;98(5):323–7.
17. Ali GC, Ryan G, De Silva M. Validated screening tools for common mental disorders in low and middle income countries: a systematic review. PLoS One. 2016;11(6):e0156939.
18. Adolescents demographic [https://data.unicef.org/topic/adolescents/ demographics/]
19. World Bank: Open Data [https://data.worldbank.org/]
20. Dos Santos PF, Wainberg ML, Caldas-de-Almeida JM, Saraceno B, Mari JD. Overview of the mental health system in Mozambique: addressing the treatment gap with a task-shifting strategy in primary care. Int J Ment Health Syst. 2016;10(1):1–9.
21. Ministry of Health. Estratégia e Plano de Acção Para a Saúde Mental 2016 – 2026. Maputo: MISAU, 2015.
22. Kieling C, Baker-Henningham H, Belfer M, Conti G, Erturm I, Omigbodun O, Rohde LA, Sinath S, Ullker N, Rahaman A. Child and adolescent mental health worldwide: evidence for action. Lancet. 2011;378(9801):1515–25.
23. Lovero KL, Basaraba C, Khan S, Suleman A, Mabunda D, Mabunda A, Falciano P, Dos Santos P, Fumo W, Mandate F, Greene MC, Fiks Salem A. Brief screening...
tool for Stepped-Care management of mental and substance use disorders. Psychiatr Serv. 2021;72(8):891–7.

24. Cumbe VF, Muando A, Manca MN, Fumo H, Chiruca F; Hicks L, de Jesus MJ, Wagenaar BH. Validity and item response theory properties of the Patient Health Questionnaire-9 for primary care depression screening in Mozambique (PHQ-9-MZ). BMC Psychiatry. 2020;20(1):382.

25. Institute for Health Metrics and Evaluation (IHME). GBD Compare Data Visualization. Seattle, WA, University of Washington, 2017. Available from http://vizhub.healthdata.org/gbd-compare. (Accessed 03.27.20).

26. Erskine HE, Norman RE, Ferrari AJ, Chan GC, Copeland WE, Whiteford HA, Scott JG. Long-term outcomes of attention-deficit/hyperactivity disorder and conduct disorder: a systematic review and meta-analysis. J Am Acad Child Adolesc Psychiatry. 2016;55(10):841–50.

27. Adjorlolo S. Generalised anxiety disorder in adolescents in Ghana: Examination of the psychometric properties of the Generalised Anxiety Disorder-7 scale: Afr J Psychol Assessment. 2019;1(1).

28. Anum A, Adjorlolo S, Kugbey N. Depressive symptomatology in adolescents in Ghana: examination of psychometric properties of the Patient Health Questionnaire-9. Afr J Psychol Assessment. 2019;25:261–8.

29. Ganguly S, Samanta M, Roy P, Chatterjee S, Kaplan DW, Basu B. Patient health questionnaire-9 as an effective tool for screening of depression among Indian adolescents. J Adolesc Health. 2013;52(5):546–51.

30. Mossman SA, Luft MJ, Schroeder HK, Varney ST, Fleck DE, Barzman DH, Gilman R, Davidoff MP, Strawn JR. The Generalized Anxiety Disorder-7 item (GAD-7) scale in adolescents with generalized anxiety disorder: signal detection and validation. Ann Clin Psychiatry. 2017;29(4):227.

31. Richardson LP, McCauley E, Grossman DC, McCarty CA, Richards J, Russo JE, Rockhill C, Katon W. Evaluation of the Patient Health Questionnaire-9 item for detecting major depression among adolescents. Pediatrics. 2010;126(6):1117–23.

32. Tiilikainen K, Haravuori H, Ranta K, Kaltiala-Heino R, Marttunen M. Psychometric properties of the 7-item Generalized Anxiety Disorder Scale (GAD-7) in a large representative sample of Finnish adolescents. Psychiatry Res. 2019;272:30–5.

33. Kroenke K, Spitzer RL. The PHQ-9: a new depression diagnostic and severity measure. Psychiatr Ann. 2002;32(9):599–15.

34. Moriarty AS, Gilbody S, McMillan D, Manea L. Screening and case finding for major depressive disorder using the Patient Health Questionnaire (PHQ-9): a meta-analysis. Gen Hosp Psychiatry. 2015;37(6):567–76.

35. Johnson JG, Harris ES, Spitzer RL, Williams JB. The patient health questionnaire for adolescents: validation of an instrument for the assessment of mental disorders among adolescent primary care patients. J Adolesc Health. 2002;30(3):196–204.

36. Plummer F, Manea L, Trepel D, McMillan D. Screening for anxiety disorders with the GAD-7 and GAD-2: a systematic review and diagnostic metaanalysis. Gen Hosp Psychiatry. 2016;39:24–31.

37. Spitzer RL, Kroenke K, Williams JB, Lowe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med. 2002;162(1):1092–7.

38. Chibanda D, Verhey R, Gibson LJ, Munetsi E, Machando D, Rusakaniko S, Chiruca P, Jaensch D, Frost DJ, Kandemir H, Knez R, Franic T, Petrov P. The Strengths and Difficulties Questionnaire (SDQ) with inpatient adolescents. Psychiatry Res. 2014;219(3):651–7.

39. Myers LA, Xu Q, Nestadt DF, Knox J, Kauchali S, Arapdi S, Kvalsvig J, Shroot PE, Davidson LL. Screening for mental health among young South African children: the use of the Strengths and Difficulties Questionnaire (SDQ). Glob Soc Welf. 2018;5(1):29–38.

40. Stone LL, Otten R, Engels RCM, Vermulst AA, Janssens JMM. Psychometric properties of the parent and teacher versions of the strengths and difficulties questionnaire for 4- to 12-year-olds: a review. Clin Child Fam Psychol Rev. 2010;13(3):254–74.

41. Sartorius N, Janca A. Psychiatric assessment instruments developed by the World Health Organization. Soc Psychiatry Psychiatr Epidemiol. 1996;31(2):55–69.

42. Willis GB. 2004. Cognitve interviewing: A tool for improving questionnaire design. United States of America: SAGE publications.

43. Norris AE, Torres-Thomas S, Williams ET. Adapting cognitive interviewing for early adolescent Hispanic girls and sensitive topics. Hisp Health Care Int. 2014;12(3):111–9.

44. Rio D, Brazaño R, Baroso R, Silva DR, Vagas P, Veira A, Lavado A, Macedo AM. Mental health problems in male young offenders in custodial versus community-based-programs: implications for juvenile justice interventions. Child Adolesc Psychiatry Ment Health. 2016;10:40.

45. Sheehan DV, Sheehan KH, Shytle RD, Janavs J, Bannon Y, Rogers JE, Milo KM, Stock SL, Wilkinson B. Reliability and validity of the mini international neuropsychiatric interview for children and adolescents (MINI-KID). J Clin Psychol. 2010;71(13):1739.

46. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDcap)—a metadata-driven methodology and workflow work for providing translational research informatics support. J Biomed Inform. 2009;42(2):377–81.

47. Youden WI. Index for rating diagnostic tests. Cancer. 1950;3(1):32–5.

48. Polanczyk GV, Salum GA, Sugaya LS, Caye A, Rohde LA. Annual research review: a meta-analysis of the worldwide prevalence of mental disorders in children and adolescents. J Child Psychol Psychiatry. 2015;56(3):345–65.

49. Becker A, Wang B, Kunze B, Otto C, Schlack R, Holling H, Ravens-Sieberer U, Klasen F, Rogge J, Iseesee C, et al. Normative data of the self-report version of the German strengths and difficulties questionnaire in an epidemiological setting. Zeitschrift fur Kinder- und Jugendpsychiatrie und Psychotherapie. 2018;46(6):523–33.

50. de Vries PJ, Davids EL, Mathews C, Aare LE. Measuring adolescent mental health around the globe: psychometric properties of the self-report Strengths and Difficulties Questionnaire in South Africa, and comparison with UK, Australian and Chinese data. Epidemiol Psychiatric Sci. 2018;27(4):369–80.

51. Giannakopoulos G, Tzavara C, Dimitrakiakis C, Kolaitis G, Rotiski V, Touantas Y. The factor structure of the Strengths and Difficulties Questionnaire (SDQ) in Greek adolescents. Ann Gen Psychiatry. 2009;8(1):20.

52. Goodman R. Psychometric properties of the strengths and difficulties questionnaire. J Am Acad Child Adolesc Psychiatry. 2001;40(11):1337–45.

53. Van Roy B, Veenstra M, Clench-Asa J. Construct validity of the five-factor Strengths and Difficulties Questionnaire (SDQ) in pre-, early, and late adolescence. J Child Psychol Psychiatry. 2008;49(12):1304–12.

54. Vuigueveer J, de Beldt A, Theunissen M, Reinjveld SA, Timmerman M. Validity Aspects of the Strengths and Difficulties Questionnaire (SDQ) Adolescent Self-Report and Parent-Report Versions Among Dutch Adolescents. Assessment. 2019;28(2):2601–16.

55. Yao S, Zhang C, Zhu X, Jing X, McWhinnie CM, Abela JRZ. Does the Strengths and Difficulties Questionnaire (SDQ) scale perform better among Chinese adolescents? Child Adolesc Health. 2009;45(1):55–62.

56. Stevanovic D, Urbanc R, Atitola O, Vostanis P, Singh Balhara YP, Avicenna M, Kandemir H, Knez R, Franic T, Petrov P. Does the Strengths and Difficulties Questionnaire - self report yield invariant measurements across different
Data from the International Child Mental Health Study Group. Epidemiol Psychiatric Sci. 2015;24(4):323–34.

Du Y, Kou J, Coghill D. The validity, reliability and normative scores of the parent, teacher and self report versions of the Strengths and Difficulties Questionnaire in China. Child Adolesc Psychiatry Ment Health. 2009;2(1):8.

Mansbach-Kleinfeld I, Apter A, Farbstein I, Levine SZ, Ponizovsky AM. A population-based psychometric validation study of the strengths and difficulties questionnaire - hebrew version. Front Psych. 2010;1:151.

Menon A, Glazebrook C, Ngoma M. Mental health of HIV positive adolescents in Zambia. Med J Zambia. 2009;36(4):151–6.

SDQ Normative Data [https://www.sdqinfo.org/gt.html]

Bhana A, Rathod SD, Selohilwe O, Kathree T, Petersen I. 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.

Smith Fawzi MC, Ngakongwa F, Liu Y, Rutayuga T, Siril H, Somba M, Kaaya SF. Validating the Patient Health Questionnaire-9 (PHQ-9) for screening of depression in Tanzania. Neurol Psychiatry Brain Res. 2019;31:9–14.

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.