



Validity, Reliability, and Diagnostic Cut-off of the Kinyarwandan Version of the Hamilton Depression Rating Scale in Rwanda

Peter Dedeken^{1,2,3*}, Joao Ricardo Nickenig Vissoci⁴, Fidele Sebera^{3,5,6}, Paul A. J. M. Boon³, Eugene Rutembesa⁶ and Dirk E. Teuwen²

OPEN ACCESS

Edited by:

Stéphane Bouchard,
University of Quebec in Outaouais,
Canada

Reviewed by:

Catherine M. Kirk,
Save the Children (United States),
United States
Giovanni Galeoto,
Sapienza University of Rome, Italy
Antonio Bulbena,
Autonomous University of Barcelona,
Spain

*Correspondence:

Peter Dedeken
peter.dedeken@heilighartlier.be;
info.dedekenneurologie@gmail.com
orcid.org/0000-0001-8493-6276

Specialty section:

This article was submitted to
Quantitative Psychology and
Measurement,
a section of the journal
Frontiers in Psychology

Received: 08 April 2020

Accepted: 20 May 2020

Published: 03 July 2020

Citation:

Dedeken P, Vissoci JRN, Sebera F,
Boon PAJM, Rutembesa E and
Tewen DE (2020) Validity, Reliability,
and Diagnostic Cut-off of the
Kinyarwandan Version of the
Hamilton Depression Rating
Scale in Rwanda.
Front. Psychol. 11:1343.
doi: 10.3389/fpsyg.2020.01343

¹Department of Neurology, Heilig Hart Hospital, Lier, Belgium, ²Corporate Societal Responsibility, UCB Pharma, Brussels, Belgium, ³Department of Neurology, Ghent University Hospital, Ghent, Belgium, ⁴Division of Emergency Medicine, Department of Surgery, Duke Global Health Institute, Duke University, Durham, NC, United States, ⁵Department of Neurology, Neuropsychiatric Hospital, CARAES Ndera, Kigali, Rwanda, ⁶Department of Neurology and Psychiatry, University of Rwanda, Kigali, Rwanda

Introduction: In Rwanda, major depressive disorder affects 11.9% of the population and up to 35% of genocide survivors. Mental health services remain underutilized due to stigma and lack of awareness. Increasing the ability and capacity to diagnose and treat mental disorders is considered important to close this gap. We describe the translation, validity, and reliability assessment of the Hamilton Depression Rating Scale (HDRS) as a diagnostic tool for moderate to severe depression in Rwanda.

Methods: The HDRS-21 was translated by a multi-group taskforce. We validated the translation against expert assessment in a comparative study on a sample of patients living with depression and of healthy volunteers. Psychometric properties, namely internal structure, reliability, and external validity were assessed using confirmatory factor analysis, three reliability calculations, and correlation analysis, respectively. Maximized Youden's index was used for determining diagnostic cut-off.

Results: The translated version demonstrated a kappa of 0.93. We enrolled 105 healthy volunteers and 105 patients with confirmed mild to severe depression. In the confirmatory factor analysis, HDRS had good factor loadings of 0.32–0.80. Reliability coefficients above 0.92 indicated strong internal consistency. External validity was shown by good sensitivity (0.95) and specificity (0.94) to differentiate depression from absence of depression. At a cut-off point of 17 for the diagnosis of depression, sensitivity and specificity were both 0.95 relative to gold standard.

Conclusion: The validated HDRS in Kinyarwanda with diagnostic cut-off provides mental healthcare staff with an accurate tool to diagnose moderate to severe depression, enabling closure of the diagnosis and treatment gap.

Keywords: Rwanda, Hamilton Depression Rating Scale, diagnosis, depression, cut-off, validation

INTRODUCTION

Depression is consistently ranked as one of the most important contributors to disability worldwide, even the single largest contributor to global disability in 2015 (World Health Organization, 2017). The global prevalence of depression is estimated to be 4.4%, with the highest at 5.9% reported among women in the African region (World Health Organization, 2017). In Rwanda, with nearly 13 million citizens (Worldometer, 2020), specifically, depressive disorders were the third cause of disability (Institute for Health Metrics and Evaluation, 2019). The burden of mental disorders is compounded by the continuing impact of the 1994 genocide (Munyandamutsa et al., 2012; Rugema et al., 2013, 2015; Ng and Harerimana, 2016), not surprisingly, as conflict causes extensive mental issues (Lafta, 2016).

Mental health care in Rwanda has a pyramidal structure, ensuring decentralized care (GOV.UK, 2017; Ministry of Health Republic of Rwanda, 2018) in all public health facilities, ranging from primary health centers under supervision of secondary district hospitals to tertiary reference hospitals. In 2016, more than 250,000 mental health consultations were performed, distributed over primary to tertiary centers. Depression accounted only for 4.7% of all visits (Ministry of Health Republic of Rwanda, 2018).

In a large-scale survey conducted in 2008 with the support of the Rwandan Ministry of Health, the prevalence of post-traumatic stress disorder and major depressive episodes were 26.1 and 22.7%, respectively (Munyandamutsa et al., 2012).

In a recent survey using the Kinyarwanda version of the Mini International Neuropsychiatric Interview, it was found that major depressive disorder affected 11.9% of the general population in Rwanda, with a higher prevalence in persons aged >46 years and with lower education. Major depressive disorder amounted to 35% among genocide survivors. Despite this high prevalence, significant proportions of the population either are not aware of the availability of mental health support or do not use the available services. While 62% of the population was aware of places offering mental health support, only 5.3% used mental health services, resulting in a large diagnosis and treatment gap. Among the recommended measures to improve mental health, increasing the ability and capacity of health personnel to assess and treat early symptoms of mental disorders was considered important (Rwandan Mental Health survey 2018, unpublished).

Enabling healthcare professionals (HCPs) to identify and diagnose depression is an important first step in mental health care. In a resource-limited environment with only few specialized HCPs, a broad use of scales to screen, diagnose, and follow-up depression can contribute to closure of the diagnosis and treatment gap (Anderson et al., 2002). The Hamilton Depression Rating Scale (HDRS) is a widely used tool to diagnose and rate the severity of depression (Williams, 2001; Ruhe et al., 2005). The HDRS-21 is composed of 17 diagnostic questions, of which eight are scored on a 5-point scale, ranging from 0 to 4 and nine are scored from 0 to 2. Questions 18–21 further qualify the depression (Medscape). The total HDRS score is calculated as the sum of the first 17 questions, ranging from 0 to 52. A score of 7 or less is considered as absence of depression, whereas it is generally agreed that a score of 17–23

represents moderate depression (Zimmerman et al., 2013). It has been used extensively for assessing severity of depression, changes over time, and treatment efficacy with high levels of internal consistency, inter-rater reliability, and test-retest reliability (Williams, 2001; Ruhe et al., 2005; Trajkovic et al., 2011).

We describe the translation process of the HDRS-21 for use in Rwanda. In many African societies, it is not culturally appropriate to ask directly about emotional states such as “Have you been depressed?” and careful wording is required. Also, words describing feeling states are sometimes totally absent in the indigenous language (Mbewe et al., 2013). Subsequently, we evaluated its psychometric properties in terms of internal structure, reliability, and cut-offs for diagnosis of moderate to severe depression in a Rwandan population.

MATERIALS AND METHODS

Translation of HDRS-21

The HDRS-21 was translated from French into Kinyarwanda by two groups of experts, who are fluent in French, English, and Kinyarwanda and are familiar with the HDRS questionnaire: one group of three clinical psychologists and one group of psychiatrists and psychiatric nurses. First, each group translated the HDRS from French to Kinyarwanda. After comparison of these two versions, an initial single translation after consensus was proposed. Second, both groups were joined by two linguists in a workshop to define a final version, which was back-translated to French by the linguists and compared to the original French version. Lastly, the translation groups scored each item of the final HDRS-21 Kinyarwanda version compared with the original French version, scoring for cultural and linguistic clarity. The quality of the translation was measured by the inter-rater agreement on the assessment of the linguistic and cultural clarity. Agreement was calculated for all 21 questions using Cohen's kappa coefficient (Cohen, 1960; McHugh, 2012).

Validation of the Translated HDRS-21

Study Set-up

Data collection was conducted in Rwanda. This was a comparative study on a sample of patients living with depression and of healthy volunteers. Data analysis was conducted at Duke University, Durham, United States in 2017. The study was approved by the Institutional Review Board of the Kigali University Hospital, Rwanda, and participants provided written informed consent.

Study Participants

Individuals with depression were recruited from six healthcare and support centers in Rwanda: Neuropsychiatric Hospital Ndera, CARAES Neuropsychiatric Center in Butare, TRAC (research and treatment center for HIV/AIDS), Association Rwandaise des Conseillers en Traumatisme (ARCT, counseling organization for trauma), Uyisenga N'Manzi center (community center for orphaned children to young adults after the genocide), and the Association des Veuves de la Genocide (counseling organization for widows following the genocide).

Individuals attending these centers were invited to participate in the study if they were aged ≥ 18 years and were able to communicate clearly. Patients with depression were selected based on their medical charts, and if diagnosed with depression according to the fourth revision of the American Psychiatric Association's Diagnostic and Statistical Manual (DSM-IV) or the tenth revision of the International Classification of Diseases and Related Health Problems (ICD-10) criteria by a structured HCP-conducted interview.

The control group consisted of healthy volunteers recruited at various sites (hospital, university, and health education institute).

Sample Size Calculation

Sample size calculation was estimated using Gorsuch's rule (MacCallum et al., 1999), which requires a sample size of five times the number of questions assessed, a total of 21 in this study, resulting in sample sizes of 105 per group. Given that the total score only concerns the first 17 questions, we assumed this to be an adequate sample size.

Procedure

The HDRS-21 in Kinyarwanda was administered through a semi-structured clinical interview conducted by trained nurses and psychologists, trained general practitioners, and trained psychiatrists. Severity of depression was based on expert assessment by the lead investigator, a trained psychiatrist.

Psychometric Parameters and Statistical Analysis

All analyses were conducted with R Language for Statistical Computing (R foundation, Vienna, Austria). Missing data was imputed for cases where individual questions were not answered, up to three missing item responses. Generated by the mice package in the R Language for Statistical Computing, we used 15 multiple imputation sets, with the other questions in the questionnaire and sociodemographic variables as the sources for imputation.

Internal Structure

Confirmatory factor analysis (CFA) was used to test the internal structure of the HDRS. Testing was based on a unidimensional model, indicating that the score based on all items of the HDRS referred to a single construct, depression. Given the categorical nature of the data, the weighted least square means and variance-adjusted (WLSMV) estimator was used to estimate model parameters. The relationship between each item and depression was determined by its loading, which corresponds to the correlation between the item and depression.

Through the CFA analysis, several indices that provide a measure of the goodness-of-fit of the models to the data were obtained. The fit indices, and the generally accepted reference levels for a good fit (in parentheses), were the following: chi-square (X^2 and p -value), root mean square error of approximation (RMSEA, ≤ 0.08), Tucker-Lewis index (TLI, > 0.90), comparative fit index (CFI, ≥ 0.90), and average variance extracted (AVE, > 0.50) (Hu and Bentler, 1999; Hair et al., 2005; Kline, 2011).

Reliability

To determine the reliability of the HDRS, its internal consistency was evaluated. Internal consistency, which indicates the degree to which all items in the instrument refer to the same construct, can be assessed by several coefficients, each with its own strengths and limitations (De Vellis, 2003; Padilla and Divers, 2016). For this study, Cronbach's alpha, composite reliability (CR), and McDonald's omega coefficient were calculated using CFA results.

Validity

The validity of the HDRS as a diagnostic tool was examined by evaluating its accuracy in distinguishing between levels of depression, in relation to the gold standard expert opinion. Sensitivity, specificity, positive predictive values, and negative predictive values of the tool were calculated. By plotting sensitivity versus specificity across the HDRS scores, receiver operator characteristic (ROC) curves were generated and the area under the ROC curve (AUC) was calculated. The AUC provides a summary measure of the sensitivity and specificity of the HDRS, relative to the expert opinion diagnostic, across the entire range of scores. The optimal cut-off score for maximum discrimination between depressed and non-depressed participants was determined from the ROC curves using the Youden index.

RESULTS

Translation

The mean Cohen's kappa coefficient of the composite score of all 21 items was 0.93, indicating a high degree of agreement between translators when comparing translated and original versions.

Study Participants

Overall, 210 adults participated in the study, divided into two groups of 105 participants (Table 1). Of the depressed group, 26 (24.8%) were recruited at the Neuropsychiatric Hospital Ndera, 5 (4.8%) at CARAES Neuropsychiatric Center in Butare, 14 (13.3%) at TRAC, 5 (4.8%) at ARCT, 10 (9.5%) at Uyisenga N'Manzi center, and 45 (42.9%) at the Association des Veuves de la Genocide.

Internal Structure

The CFA unidimensional model performed well, showing all items with factor loadings ranging from 0.32 to 0.81 for the HDRS 17 (Figure 1). Item 17 of the HDRS (patient insight into

TABLE 1 | Depression severity as per expert opinion.

| | Patient group (n = 105) | Control group (n = 105) |
|-----------------------------------|-------------------------|-------------------------|
| Depression severity, n (%) | | |
| No depression | 0 (0%) | 105 (100%) |
| Mild | 14 (13.3%) | 0 (0%) |
| Moderate | 0 (0%) | 0 (0%) |
| Severe | 91 (86.6%) | 0 (0%) |

his/her condition) showed the smallest factor loading, suggesting a problem with the item's performance. Another model that excluded item 17 was tested and was also shown to perform well, with factor loadings ranging from 0.65 to 0.81 (**Figure 1** and **Table 2**).

Further results of the CFA confirmed that the models show a good fit to the data; the goodness-of-fit indices all met the criteria standards for adequacy of fit (**Table 2**).

Reliability

Values obtained for all three reliability coefficients were above 0.80 for both HDRS 16 and 17 items, indicating strong internal consistency (**Table 2**).

Validity

The translated HDRS was associated with a high sensitivity (0.95) and specificity (0.94) in differentiating participants in the control group who did not have depression from those in the patient group who had moderate to severe depression (**Figure 2**).

At a cut-off point of 17 for diagnosis of depression, sensitivity and specificity were both 0.95 relative to gold standard. Positive and negative predictive values were 96.5 and 93% respectively, and the AUC was 0.98.

DISCUSSION

This study evaluated the psychometric properties of the HDRS-21 in Kinyarwanda and established the optimal cut-off for depression at a score of ≥ 17 .

The translation process, using a mixed model including specialized HCPs and bilingual experts with multiple iterations, allowed us to address cultural sensitivities on depression and words describing feeling states. High consistency between the French and Kinyarwanda versions was reflected by Cohen's kappa >0.8 and demonstrated excellent inter-group consistency of all translated items.

Subsequently, the psychometric properties of the translated HDRS were evaluated. Results of the various tests conducted to assess these properties support the use of the HDRS as a reliable and valid instrument for evaluating depression. Using CFA, the internal structure of the HDRS was found to be adequate in relation to a single construct, depression, as reported in other studies (Williams, 2001; Ruhe et al., 2005; Trajkovic et al., 2011). For a unidimensional model, all measuring items should have acceptable factor loadings. In our analysis, item 17 was associated with a smaller factor loading relative to the other 16 items, suggesting a poorer performance.

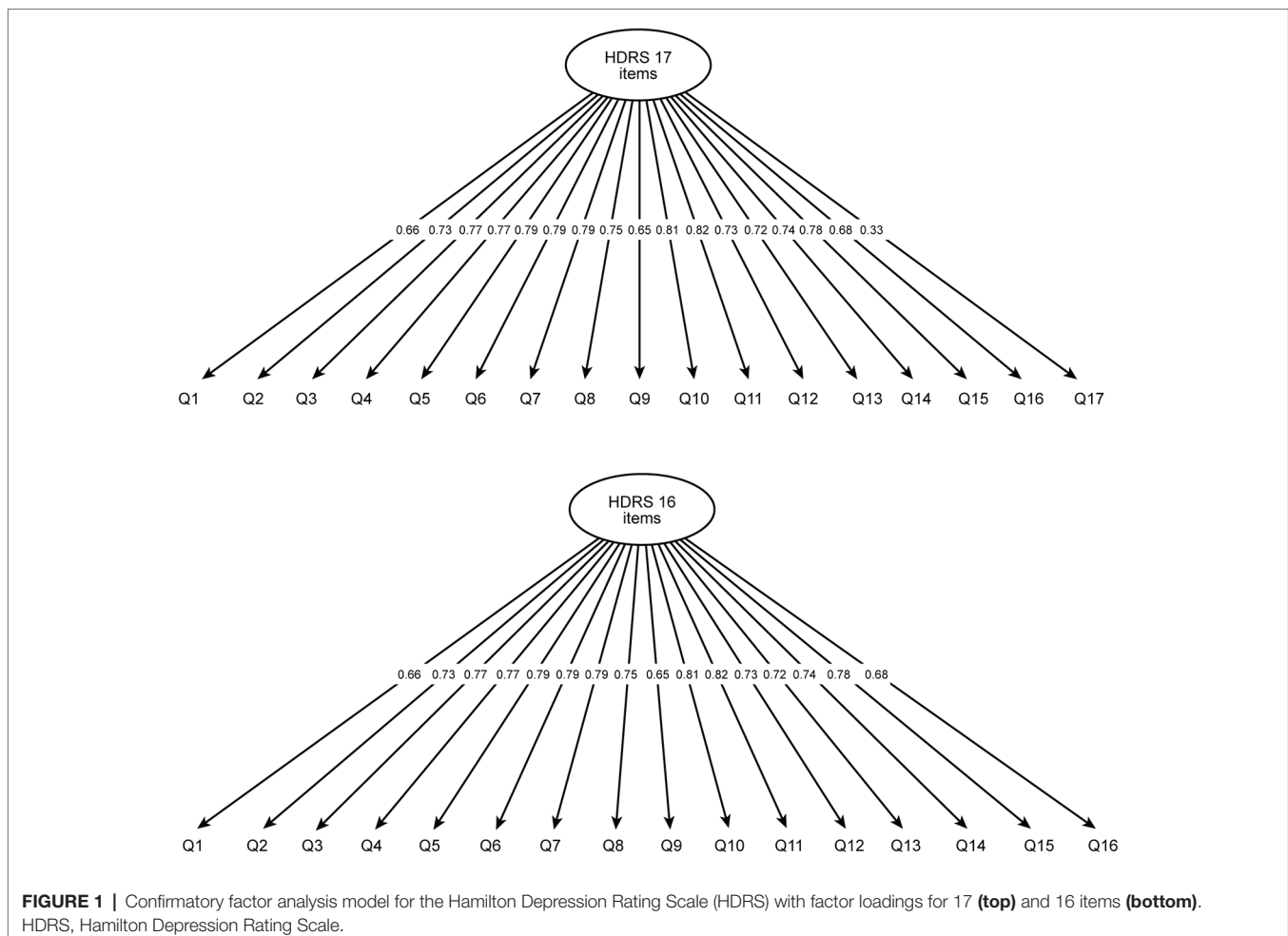
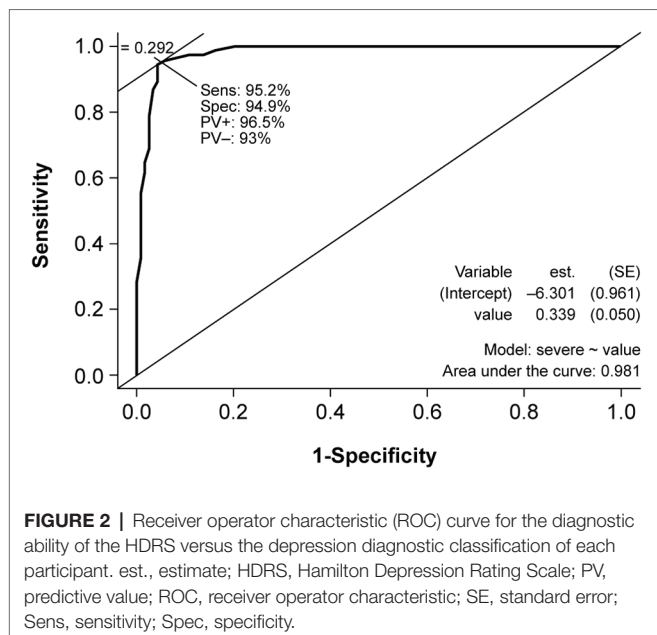


TABLE 2 | Goodness-of-fit indices from the confirmatory factor analysis and reliability coefficients.

| | HDRS 16 items | HDRS 17 items |
|--------------------------------------|-------------------------------|-------------------------------|
| Confirmatory factor analysis | | |
| X ² (degrees of freedom); | 147.10 (104); <i>p</i> = 0.01 | 159.03 (119); <i>p</i> = 0.01 |
| <i>p</i> -value | | |
| RMSEA (95% CI) | 0.05 (0.03, 0.06) | 0.04 (0.02, 0.06) |
| Tucker-Lewis index | 0.99 | 0.99 |
| Comparative fit index | 0.99 | 0.99 |
| Average variance extracted | 0.56 | 0.53 |
| Factor loading range | 0.65–0.81 | 0.32–0.81 |
| Reliability | | |
| Cronbach's alpha (95% CI) | 0.92 (0.91, 0.93) | 0.92 (0.90, 0.93) |
| Composite reliability | 0.95 | 0.95 |
| McDonald's omega | 0.93 | 0.93 |

CI, confidence interval; HDRS, Hamilton Depression Rating Scale; RMSEA, root mean square error of approximation.

**FIGURE 2** | Receiver operator characteristic (ROC) curve for the diagnostic ability of the HDRS versus the depression diagnostic classification of each participant. est., estimate; HDRS, Hamilton Depression Rating Scale; PV, predictive value; ROC, receiver operator characteristic; SE, standard error; Sens, sensitivity; Spec, specificity.

Reliability is the ability of an instrument to measure a given construct consistently. The most widely used method for evaluating the reliability of an instrument is Cronbach's alpha (Tavakol and Dennick, 2011). As a sensitivity analysis, composite reliability and McDonald's omega were also calculated. The strong internal consistency of the translated HDRS, with reliability coefficients >0.80, suggests that it is a highly reliable tool for assessing depression.

In addition to reliability, another fundamental element in the evaluation of a measurement instrument is its validity (Tavakol and Dennick, 2011). The validity of the translated HDRS was demonstrated by the AUC of 0.98, close to a value of 1.0 that indicates a perfectly accurate instrument.

The present study aimed to establish a diagnostic cut-off for the HDRS in Kinyarwanda. Based on the Youden index, a cut-off point of 17 was deemed optimal with sensitivity and specificity >95%. This cut-off for diagnosis is in line with similar cut-offs of ≥ 16 and 18/19 observed in a mixed community/

outpatient sample (Mottram et al., 2000) and in a Parkinson's disease outpatient setting in Ecuador, respectively (Serrano-Duenas and Soledad Serrano, 2008). In disease-specific populations, lower cut-offs have been observed for Parkinson's disease (Leentjens et al., 2000) (≥ 12 and 15/16) (Naarding et al., 2002), stroke (10/11) (Naarding et al., 2002), and Alzheimer's disease (13/14) (Naarding et al., 2002). In our study, only 11 patients had mild depression, skewing diagnosis to moderate to severe depression, which may have resulted in a high HDRS cut-off. This could be an advantage in settings with limited availability of drugs and where priorities for medical treatment need to be set toward moderate to severe depression. For example, in the treatment of epilepsy, the treatment of depression may not be suitable for patients with minor or mild depression (Rathore et al., 2014).

The main limitation of the study was the patient sample enrolled. First, the sample size according to Gorsuch's rule allowed validation and assessment of its psychometric properties, but larger sample sizes have been proposed as well (MacCallum et al., 1999). This may explain the decreased factor loading of item 17, with a supplementary analysis using a 16-item model excluding item 17 also showing to be a good fit to the data. Until larger studies confirm our findings and validate a 16-item HDRS total score, we recommend to maintain the 17-item HDRS total score. Second, we recruited at six different sites, yet patients with a background of war trauma represented more than half of our sample, which may not have been representative of the general population. HDRS cut-offs also vary by concomitant disease or background and require confirmation if used in specific populations. Lastly, our study design did not allow for testing of inter- and intra-rater reliability. These limitations warrant future studies in a larger sample of the general population, including intraclass correlation as well as confirmation of the observed factor loading for item 17.

Evaluating psychometric properties of screening and assessment tools is a necessary process, since even minor issues in wording, connotation, and question structure can reduce the cross-cultural relevance of the tools (Sweetland et al., 2014). This is particularly important since use of these scales by trained nurses and other non-specialist staff is pivotal for scaling up mental health care in Rwanda and to a broader extent, in low-income countries, it increases the number of staff that can provide such care (Belkin et al., 2011; Sweetland et al., 2014). In 2012, the Rwandan Ministry of Health, in collaboration with Partners in Health, adapted a program for primary care nurses providing training to participate in the clinical care of people with mental disorders (Smith et al., 2017). In this stepped-up pathway, trained staff can participate more and more in screening and triage, engagement, follow-up, and monitoring, as well as referral when and where required (Sweetland et al., 2014). Consequently, availability of HDRS in Kinyarwanda is an important step forward given the high prevalence of depression.

The HDRS is a valuable addition to the armamentarium of validated tools in Kinyarwanda such as the Patient Health Questionnaire (PHQ)-9, validated for screening for depression, and the Center for Epidemiological Studies Depression Scale for Children (CES-DC) (Betancourt et al., 2012), addressing a vulnerable population. Although HDRS has also been used for

screening for depression (Aben et al., 2002; Costa et al., 2016), we believe that the high number of questions and duration of administration impair its user-friendliness as a screening tool, for which PHQ-9 is available as a validated alternative.

Addressing the recommendation in the Rwandan Mental Health survey, the validation of the HDRS in Kinyarwanda (the Kinyarwanda version of HDRS can be accessed online here: <http://caraesnderahospital.rw/research-2/>) provides primary and secondary healthcare staff with an accurate tool for diagnosis of moderate to severe depression, enabling closure of the diagnosis and treatment gap.

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available due to patient confidentiality. Requests to access the datasets should be directed to peter.dedeken@heilighartlier.be.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Institutional Review Board of the Kigali University Hospital, Rwanda. The patients/participants provided their written informed consent to participate in this study.

REFERENCES

- Aben, I., Verhey, F., Lousberg, R., Lodder, J., and Honig, A. (2002). Validity of the beck depression inventory, hospital anxiety and depression scale, SCL-90, and hamilton depression rating scale as screening instruments for depression in stroke patients. *Psychosomatics* 43, 386–393. doi: 10.1176/appi.psy.43.5.386
- Anderson, J. E., Michalak, E. E., and Lam, R. W. (2002). Depression in primary care: tools for screening, diagnosis, and measuring response to treatment. *B. C. Med. J.* 44, 415–419.
- Belkin, G. S., Unutzer, J., Kessler, R. C., Verdelli, H., Raviola, G. J., Sachs, K., et al. (2011). Scaling up for the bottom billion: 5 x 5 implementation of community mental health care in low-income regions. *Psychiatr. Serv.* 62, 1494–1502. doi: 10.1176/appi.ps.000012011
- Betancourt, T., Scorza, P., Meyers-Ohki, S., Mushashi, C., Kayiteshonga, Y., Binagwaho, A., et al. (2012). Validating the center for epidemiological studies depression scale for children in Rwanda. *J. Am. Acad. Child Adolesc. Psychiatry* 51, 1284–1292. doi: 10.1016/j.jaac.2012.09.003
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educ. Psychol. Meas.* 20, 37–46. doi: 10.1177/001316446002000104
- Costa, M. V., Diniz, M. F., Nascimento, K. K., Pereira, K. S., Dias, N. S., Malloy-Diniz, L. F., et al. (2016). Accuracy of three depression screening scales to diagnose major depressive episodes in older adults without neurocognitive disorders. *Braz. J. Psychiatry* 38, 154–156. doi: 10.1590/1516-4446-2015-1818
- De Vellis, R. F. (2003). *Scale development: Theory and applications*. Vol. 26. 2nd Edn. Thousand Oaks, CA, USA: SAGE Publications Inc.
- GOV.UK (2017). *Closing the gap: decentralising mental health care to primary care centres in one rural district of Rwanda* [Online]. GOV.UK. Available at: <https://www.gov.uk/dfid-research-outputs/closing-the-gap-decentralising-mental-health-care-to-primary-care-centres-in-one-rural-district-of-rwanda> (Accessed May 13, 2020).
- Hair, J. F., Black, W. C., Anderson, R. E., and Tatham, R. L. (2005). *Análise Multivariada de Dados*. Porto Alegre, Brazil: Bookman.
- Hu, L. T., and Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct. Equ. Model.* 6, 1–55.

AUTHOR CONTRIBUTIONS

FS and ER performed data collection. JV performed statistical analysis. JV, PD, DT, and PB developed concept of analysis. JV and PD co-authored the manuscript. FS, ER, JV, DT, PB, and PD reviewed and approved the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

The authors declare that this study received funding from UCB Biopharma. The funder was not involved in the study design, collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication.

ACKNOWLEDGMENTS

We acknowledge and thank the participants of this study and the patients for rising above the stigma associated with depression. The authors acknowledge Fabien Debailleul, PhD (UCB Pharma, Brussels, Belgium) and Envision Pharma Group (Horsham, United Kingdom), funded by UCB Pharma, for managing the development of the manuscript.

- Institute for Health Metrics and Evaluation (2019). *Measuring what matters: Rwanda* [Online]. Available at: <http://www.healthdata.org/rwanda> (Accessed December 12, 2019).
- Kline, R. B. (2011). “Convergence of structural equation modeling and multilevel modeling” in *The SAGE handbook of innovation in social research methods*. eds. M. Williams and W. P. Vogt (London, UK: SAGE Publications Ltd).
- Lafta, R. K. (2016). Health in times of uncertainty. *Lancet Glob. Health* 4, e666–e667. doi: 10.1016/S2214-109X(16)30184-X
- Leentjens, A. F., Verhey, F. R., Lousberg, R., Spitsbergen, H., and Wilms, F. W. (2000). The validity of the Hamilton and Montgomery-Asberg depression rating scales as screening and diagnostic tools for depression in Parkinson's disease. *Int. J. Geriatr. Psychiatry* 15, 644–649. doi: 10.1002/1099-1166(200007)15:7<644::AID-GPS167>3.0.CO;2-L
- MacCallum, R. C., Widaman, K. F., Zhang, S., and Hong, S. (1999). Sample size in factor analysis. *Psychol. Methods* 4, 84–99. doi: 10.1037/1082-989X.4.1.84
- Mbewe, E. K., Uys, L. R., Nkwanyana, N. M., and Birbeck, G. L. (2013). A primary healthcare screening tool to identify depression and anxiety disorders among people with epilepsy in Zambia. *Epilepsy Behav.* 27, 296–300. doi: 10.1016/j.yebeh.2013.01.025
- McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochem. Med.* 22, 276–282.
- Ministry of Health Republic of Rwanda (2018). *Annual Health Statistics Booklet 2016* [Online]. Available at: http://www.moh.gov.rw/fileadmin/user_upload/HMIS/2016_Annual_Statistical_booklets_V9_08_03_2018.pdf (Accessed May 13, 2020).
- Mottram, P., Wilson, K., and Copeland, J. (2000). Validation of the Hamilton Depression Rating Scale and Montgomery and Asberg Rating Scales in terms of AGE-CAT depression cases. *Int. J. Geriatr. Psychiatry* 15, 1113–1119. doi: 10.1002/1099-1166(200012)15:12<1113::AID-GPS253>3.0.CO;2-Y
- Munyandamutsa, N., Mahoro Nkubamugisha, P., Gex-Fabry, M., and Eytan, A. (2012). Mental and physical health in Rwanda 14 years after the genocide. *Soc. Psychiatry Psychiatr. Epidemiol.* 47, 1753–1761. doi: 10.1007/s00127-012-0494-9
- Naarding, P., Leentjens, A. F., van Kooten, F., and Verhey, F. R. (2002). Disease-specific properties of the rating scale for depression in patients with stroke, Alzheimer's dementia, and Parkinson's disease. *J. Neuropsychiatr. Clin. Neurosci.* 14, 329–334. doi: 10.1176/jnp.14.3.329

- Ng, L. C., and Harerimana, B. (2016). Mental health care in post-genocide Rwanda: evaluation of a program specializing in posttraumatic stress disorder and substance abuse. *Glob. Ment. Health* 3:e18. doi: 10.1017/gmh.2016.12
- Padilla, M. A., and Divers, J. (2016). A comparison of composite reliability estimators: coefficient omega confidence intervals in the current literature. *Educ. Psychol. Meas.* 76, 436–453. doi: 10.1177/0013164415593776
- Rathore, J. S., Jehi, L. E., Fan, Y., Patel, S. I., Foldvary-Schaefer, N., Ramirez, M. J., et al. (2014). Validation of the Patient Health Questionnaire-9 (PHQ-9) for depression screening in adults with epilepsy. *Epilepsy Behav.* 37, 215–220. doi: 10.1016/j.yebeh.2014.06.030
- Rugema, L., Mogren, I., Ntaganira, J., and Gunilla, K. (2013). Traumatic episodes experienced during the genocide period in Rwanda influence life circumstances in young men and women 17 years later. *BMC Public Health* 13:1235. doi: 10.1186/1471-2458-13-1235
- Rugema, L., Mogren, I., Ntaganira, J., and Krantz, G. (2015). Traumatic episodes and mental health effects in young men and women in Rwanda, 17 years after the genocide. *BMJ Open* 5:e006778. doi: 10.1136/bmjopen-2014-006778
- Ruhe, H. G., Dekker, J. J., Peen, J., Holman, R., and de Jonghe, F. (2005). Clinical use of the Hamilton Depression Rating Scale: is increased efficiency possible? A post hoc comparison of Hamilton Depression Rating Scale, Maier and Bech subscales, clinical global impression, and symptom checklist-90 scores. *Compr. Psychiatry* 46, 417–427. doi: 10.1016/j.comppsy.2005.03.001
- Serrano-Duenas, M., and Soledad Serrano, M. (2008). Concurrent validation of the 21-item and 6-item Hamilton Depression Rating Scale versus the DSM-IV diagnostic criteria to assess depression in patients with Parkinson's disease: an exploratory analysis. *Parkinsonism Relat. Disord.* 14, 233–238. doi: 10.1016/j.parkreldis.2007.08.013
- Smith, S. L., Misago, C. N., Osrow, R. A., Franke, M. F., Iyamuremye, J. D., Dusabeyezu, J. D., et al. (2017). Evaluating process and clinical outcomes of a primary care mental health integration project in rural Rwanda: a prospective mixed-methods protocol. *BMJ Open* 7:e014067. doi: 10.1136/bmjopen-2016-014067
- Sweetland, A. C., Belkin, G. S., and Verdelli, H. (2014). Measuring depression and anxiety in sub-saharan Africa. *Depress. Anxiety* 31, 223–232. doi: 10.1002/da.22142
- Tavakol, M., and Dennick, R. (2011). Making sense of Cronbach's alpha. *Int. J. Med. Educ.* 2, 53–55. doi: 10.5116/ijme.4dfb.8dfd
- Trajkovic, G., Starcevic, V., Latas, M., Lestarevic, M., Ille, T., Bukumiric, Z., et al. (2011). Reliability of the Hamilton Rating Scale for Depression: a meta-analysis over a period of 49 years. *Psychiatry Res.* 189, 1–9. doi: 10.1016/j.psychres.2010.12.007
- Williams, J. B. (2001). Standardizing the Hamilton Depression Rating Scale: past, present, and future. *Eur. Arch. Psychiatry Clin. Neurosci.* 251(Suppl 2), Ii6–Ii12. doi: 10.1007/bf03035120
- World Health Organization (2017). *Depression and other common mental disorders: Global health estimates*. Available at: <https://apps.who.int/iris/bitstream/handle/10665/254610/WHO-MSD-MER-2017.2-eng.pdf> (Accessed January 8, 2020).
- Worldometer (2020). *Rwanda demographics* [Online]. Available at: <https://www.worldometers.info/demographics/rwanda-demographics/> (Accessed May 13, 2020).
- Zimmerman, M., Martinez, J. H., Young, D., Chelminski, I., and Dalrymple, K. (2013). Severity classification on the Hamilton Depression Rating Scale. *J. Affect. Disord.* 150, 384–388. doi: 10.1016/j.jad.2013.04.028

Conflict of Interest: PD has received consultancy fees from UCB Pharma and Novartis. DT is an employee of UCB Pharma. PB has received speaker fees and grants from several pharmaceutical companies, outside the scope of this study.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Dedeken, Vissoci, Sebera, Boon, Rutembesa and Teuwen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.