

Translation, cultural adaption, and validation of the PHQ-9 and GAD-7 in Kinyarwanda for primary care in the United States

Frank Müller, MD, PhD ^{a,b,c*}, Alexis Hansen^a, Megan Kube, MD ^{a,d}, Judith E. Arnetz, PhD^a, Omayma Alshaarawy, PhD^e, Eric D. Achtyes, MD, MS^{f*}, Harland T. Holman, MD^{* a,b}

^aDepartment of Family Medicine, Michigan State University, Grand Rapids, MI 49503, USA

^bCorewell Health Family Medicine Residency Clinic, Grand Rapids, MI 49503, USA

^cDepartment of General Practice, University Medical Center Göttingen, 37073 Göttingen, Germany

^dGraduate Medical Education in Psychiatry, Trinity Health Livonia Hospital, Livonia, MI 48154, USA

^eDepartment of Family Medicine, College of Human Medicine, Michigan State University, East Lansing, MI 48824, USA.

^fDepartment of Psychiatry, Western Michigan University Homer Stryker M.D. School of Medicine, Kalamazoo, MI 49048, USA

* Eric D. Achtyes and Harland T. Holman share senior authorship.

Word Count: 3,446

Corresponding author: Frank Müller, Department of Family Medicine, Michigan State University, 15 Michigan St NE, Grand Rapids MI 49503, muell313@msu.edu

Key words: Rwanda, mental health, communication barriers, screening, refugees, Kinyarwanda, PHQ-9, GAD-7

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

Abstract

Background: Depression and anxiety are significant health burdens that greatly impact the quality of life of refugees and migrants. In this study, we have translated and culturally adapted the Patient Health Questionnaire (PHQ-9) and Generalized Anxiety Disorder Screener (GAD-7) into Kinyarwanda and performed a validation study in a United States (US) primary care setting.

Methods: A committee of experts translated and culturally adapted the PHQ-9 and GAD-7, incorporating feedback from cognitive interviews with bilingual participants. The translated instruments were then tested in a cross-sectional validation study. Analyses include internal consistency, discriminative validity, principal component analyses, and confirmatory factor analyses.

Results: Analyses of 119 responses indicated overall good internal consistency with Cronbach's α of 0.85 (PHQ-9) and 0.92 (GAD-7). Both scales showed acceptable factor loadings between 0.44 and 0.90 in the principal component analyses and showed strong correlations with health-related quality of life and depression/anxiety symptoms measured with visual analog scales. Significantly higher scores for PHQ-9 and GAD-7 were shown among participants with known psychiatric conditions.

Discussion: PHQ-9 and GAD-7 demonstrated commendable applicability for Kinyarwanda-speaking patients in primary healthcare settings in the US. Culturally adapted translations of these tools may aid in mitigating health disparities among refugee and migrant populations. Future research should further validate these tools against gold-standard diagnostics in larger, geographically diverse samples.

Key words: Rwanda, mental health, communication barriers, screening, refugees, migrants, Kinyarwanda, PHQ-9, GAD-7

1. Introduction

Depression and anxiety are major health burdens and can lead to disability [1], mortality [2], reduced quality of life [3], and worsening of conditions like diabetes, hypertension, or heart failure [4–6]. The United States (US) Preventative Services Task Force (USPSTF) has recommended depression screening in primary care for all adults since 2016 [7] and recently expanded this guideline for anxiety screening [8].

Refugee populations, including refugees from Rwanda [9], are particularly affected by mental health problems such as depression, anxiety, post-traumatic stress disorder, and other mental conditions [10–12]. Western Michigan, including the metropolitan area of Grand Rapids, has become one of the main destinations for refugees and migrants from Rwanda, Uganda, and the Republic of Congo in the US forming a growing community of Kinyarwanda-speaking residents [13]. Research has demonstrated that these populations are often both highly traumatized and less likely to seek mental healthcare [14], further emphasizing the need for high-quality, culturally competent mental health screening.

While the Patient Health Questionnaire (PHQ-9), one of the most frequently used mental health screening instruments [15], has been translated in Kinyarwanda and evaluated among patients with epilepsy in Rwanda [16], a Kinyarwanda PHQ-9 has not been validated among refugee population in US primary care settings. Furthermore, a tool to screen for clinically relevant anxiety symptoms such as the Generalized Anxiety Disorder Screener (GAD-7) [17] has not been translated and validated in Kinyarwanda.

To address this gap in care, the objectives of this work were to translate and culturally adapt the PHQ-9 and GAD-7 into Kinyarwanda and validate both screeners in a cross-sectional study conducted in primary care clinics in Western Michigan. By publishing these validated instruments as a free resource for clinicians and researchers, we hope to contribute to the efforts to mitigate health disparities among refugee and migrant populations.

2. Methods

This study was comprised of two phases: In the first phase, the instruments were translated and culturally adapted using a committee framework approach including cognitive interviews with bilingual participants. In the second phase, a cross-sectional validation study was conducted in primary care clinics in Western Michigan. Results of this cross-sectional part are reported using STROBE guidelines [18].

Translation and Cultural Adaptation

The committee-based approach to translation and cultural adaptation uses a continuous feedback and review process within a team of language, culture, and subject matter experts [19]. All commonly used translation and adaption frameworks entail rigorous steps of planning, parallel translations, review, adjudication, pilot-testing, revision, and documentation, that can be iterated as necessary [20]. We extended this framework by conducting cognitive interviews instead of pilot-testing with bilingual participants and further added an independent back translation as a final quality control before conducting the validation study. The committee consisted of two clinicians (EDA, HTH) trained in family medicine and psychiatry, a clinician-researcher (FM) additionally trained in psychosomatic medicine, one psychiatry resident physician (MK), and one medical student (AH), and two trained medical Kinyarwanda interpreters who have been working for the healthcare system as on-site medical interpreters for several years. Interpreters were of male and female sex, middle-aged, and both grew up in Rwanda and the US and thus were proficient in both Kinyarwanda and the English language. The process of translation and cultural adaption was completed in six teleconference meetings using Zoom video conferencing software (Zoom Video Communications, Inc., San Jose, CA, USA), each lasting over 60 minutes between April and June 2022.

For the translation process, the English versions of the GAD7 and PHQ9 were used as sources and subsequently translated and adapted into Kinyarwanda. Kinyarwanda is a Bantu language with

different dialectal variants spoken in Rwanda and also in parts of Burundi, the Democratic Republic of the Congo, Uganda, and Tanzania. The aim for the translation and adaption process agreed upon among all authors was to generate a semantically similar version retaining format, scales, range of response options, and stimulus questions while using plain and familiar language allowing the instruments to be understood by people speaking different dialectal variants and from a diverse educational background.

In the preparation step we first conducted a literature review on existing healthcare survey instruments in the Kinyarwanda language to gain insights about potential obstacles that others faced during the translation process. Furthermore, these collected instruments served as a reference in the committee discussions. During the first meeting, the process of translation and cultural adaptation was discussed. The PHQ-9 and GAD-7 were introduced, and the clinicians described how these instruments were used and results interpreted in everyday practice. Furthermore, the interpreters provided insights into the cultural perception of mental health problems among the Kinyarwanda-speaking refugee community in the US and the relevant impact on the translation and adaptation process. They highlighted that mental health problems are often stigmatized among Kinyarwanda communities, and they expressed their fear that the use of stigmatized expressions would lead to biased responses and should thus be avoided. The interpreters furthermore shared that many in the community feared being labeled as “crazy” or “mad” and were afraid that a suspicious test result would implicate that their “children will be taken away”.

After this first meeting, both interpreters independently translated the instruments within two weeks (parallel forward translation). Additionally, clinicians provided paraphrases and elaborations independently on the meaning of each item, e.g. “This item is designed to address the idea of anhedonia, not enjoying things one previously used to enjoy”. The results of these first interpretations were compiled in a Microsoft Excel (Microsoft Corp., Redmond, WA) document together with the PHQ-9 version used in the study of Sebera et al. [16] as preparation for the following committee review.

In the following four review meetings each item was first introduced by the clinicians and explained what role it plays in patients with clinically relevant depression or anxiety symptoms. The interpreters then discussed matching phrasing to concepts and meanings as well as the format and distinction of questions and responses. Especially challenging was the translation of colloquial metaphors in the English version, such as “feeling on edge” and finding expressions that avoid stigma. Here it became apparent, that even basic wording like “mental health” itself would imply “madness” or “craziness” in a direct translation. In the following meetings, the items were further refined and compiled into a first draft, consisting of four pages.

This draft was then reviewed in the adjudication meeting, where interpreters checked also for contingency, grammar, and spelling. After final revisions, interpreters and clinicians prepared for pilot testing.

The pilot testing was conducted as cognitive interviews with n=5 bilingual English/Kinyarwanda participants (3 male, 2 female), that we recruited through the Michigan State University Department of African Language. All participants were faculty or students of various academic backgrounds located in the US and the Democratic Republic of the Congo. Cognitive interviews were carried out and video- and audio-recorded using Zoom teleconferencing software. The cognitive interviews were meant to assess the cognitive process in answering questions, e.g. if the items were understood (comprehension), and if necessary and relevant information and underlying concepts were retrieved (retrieval). Furthermore, we aimed to gain insights into the process of answer preparation of participants (judgment) and if challenges arose from formatting the answers (response) [21]. The pilot testing thus did not only focus on the survey items themselves but also the handling of the survey as a whole. Participants were first introduced to the purpose of the questionnaire and that it is administered as a standard screening instrument in a family medicine clinic during every patient visit. Then participants were asked to read out loud each item, answer it, then describe their thought process that led them to their final answer [22]. After completing all items, participants were asked to comment if any items were complex, hard to understand, perceived as odd or disturbing, or contained

grammatical or spelling errors. The recorded video of each interview was reviewed by the study team and concerns raised by interviewees were summarized. Furthermore, a certified external interpreter who was not involved in the research project provided back-translation of both survey instruments. The notes from the cognitive interviews and the results from the back-translation were then discussed in a final consensus meeting and the survey was finalized. The final version of the translated instruments are included with this manuscript as a supplemental file.

Participant Recruitment and Setting

Corewell Health is a large non-profit managed healthcare organization with 22 hospitals and more than 300 outpatient facilities in Michigan. Participants were enrolled in two ways: First, three family medicine clinics in Western Michigan that provided care to the Kinyarwanda community enrolled patients in their clinics during their routine clinic visits. Second, a Kinyarwanda interpreter at the healthcare system was briefed to enroll patients in all other primary care clinics in the health system. Results of the survey were reported to the provider with the notice that the information obtained was based on a not-yet-validated scale and therefore typical thresholds of the screening instruments may not apply. Recruitment took place from June 2022 to August 2023.

Patients were eligible to participate, if their preferred language for the encounter was (a) Kinyarwanda and a language interpreter was needed, (b) patients were aged 18 or older, and (c) patients agreed to participate. Exclusion criteria were (a) lack of Kinyarwanda proficiency or lack of need for language interpretation, (b) patients aged 17 or below, (c) patients refusing to participate, or not being able to provide informed consent.

Patients who met the sociodemographic eligibility criteria were asked by the staff of the participating clinics or the interpreter if they would like to participate. If they agreed, they were asked to read the

first part of the paper-and-pencil survey that introduces the objectives of the research, that participation was voluntary, and that only anonymized data were collected and processed. If they agreed with this statement, they were asked to proceed with the survey. Patients did not receive compensation for participating in this study.

Measures

Besides the two translated instruments PHQ-9 and GAD-7, the survey comprised questions on participants' sociodemographic characteristics (gender, age, education, timing of entrance into the US, self-perceived English language proficiency) as well as if the patient has ever been told by a health care provider that they have a mental health problem (such as depression, anxiety, schizophrenia, bipolar disorder, obsessive/compulsive disorder, a substance use disorder, or other mental health problems). These items were translated in the same process as described above. As illness concepts such as bipolar disorder do not have a direct translation in the Kinyarwanda language, we added a short description highlighting the main symptoms.

Furthermore, the survey contained the Kinyarwanda version of the EQ5D-VAS to measure patients self-rated overall health [23] on a scale from 0 (“worst health you can imagine”) to 100 (“best health you can imagine”). Although the Kinyarwanda EQ5D-VAS was thoroughly translated and provided as an official resource through the EuroQol group, we could not find any literature that these instruments have been previously validated in Kinyarwanda. To broadly assess depression and anxiety symptoms in our sample and to compare them to PHQ-9 and GAD-7 items, we added further visual analog scales with the questions “Overall in the last 30 days, how depressed have you felt?” and “Overall in the last 30 days, how anxious have you felt?”. Similar scales have been used and validated in other settings [24].

Besides these numerical measures, we also checked each completed survey for annotations made by the respondents.

Sample Size

For validation studies, typically an item-to-response ratio from 1:5 to 1:10 is favored [25], suggesting we needed at least 90 responses for the PHQ-9 and 70 responses for the PHQ-7.

Statistical Analysis

Descriptive statistics including frequencies, percentages, means and standard deviations (SD) were used to characterize our sample and the survey responses. Internal consistency of both scales was assessed with Cronbach's α and McDonald's Ω . Sensitivity analyses for consistency were conducted among various sociodemographic subgroups. For each item of both scales, inter-item and inter-scale correlations were calculated.

After assessing Kaiser-Meyer-Olkin-Criteria (KMO) and performing Bartlett's test of sphericity, principal component analyses (PCA) were used to extract factor loadings. Furthermore, confirmatory factor analyses (CFA) were used to test the known one-factor model of both instruments [26,27]. Standardized factor loadings were extracted for each item and are displayed as supplemental resources.

Items were tested whether they had a suitable explanatory power using item-total correlation thresholds of <0.3 [28] and factor loadings <0.32 [29]. Further, items were checked for inter-item collinearity using a threshold for correlation coefficient > 0.7 .

Discriminatory validity was assessed to compare the sum scores of participants with and without known psychiatric conditions using the Mann-Whitney- U test. Furthermore, Spearman's rho was used to calculate correlation coefficients between both psychometric scales and respective visual analog scales for depression, anxiety, and health-related quality of life (EQ-5D VAS score).

All descriptive and bivariate statistics were carried out using SPSS V29 (IBM Corp., Armonk NY) and the CFA was conducted using AMOS V29 (IBM Corp., Armonk NY).

Research Ethics

The study received approval from the Corewell Health Institutional Review Board (Approval #: 2022-075). Only anonymized data were collected. Participants received written information and provided written consent prior to enrollment.

3. Results

Sample characteristics

In total, we recorded 133 responses of which 14 were excluded due to missing values. Participants were predominantly female (68.9%) and had a mean age of 40 years (SD 15.5). Only 10 (8.9%) participants reported having a known psychiatric condition, of which anxiety disorder (n=6), PTSD, and Depression (each n=4) were the most common. Approximately 17.8% of the participants reported having no school education and 42.9% reported speaking English “not at all”. Further sociodemographic information is outlined in table 1.

Table 1: Participant characteristics

Characteristic*		Female N=82 n (%)	Male N=37 n (%)	Total N=119 n (%)
Sociodemographic characteristics				
Age	$\bar{x}$ (SD)	38.6 (13.6)	43.3 (19.1)	40.0 (15.5)
	18-29 years	21 (26.3)	9 (26.5)	30 (26.3)
	30-49 years	45 (56.3)	13 (38.2)	58 (50.9)
	50+ years	14 (17.5)	12 (35.3)	26 (22.8)
Education	0-4 school years	27 (33.3)	5 (13.5)	32 (27.1)
	5-9 school years	20 (24.7)	12 (32.4)	32 (27.1)
	10+ school years	34 (42)	20 (54.1)	54 (45.8)
Time in the US	<1 year	12 (14.6)	10 (27)	22 (18.5)
	1-2 years	10 (12.2)	8 (21.6)	18 (15.1)
	2-3 years	5 (6.1)	1 (2.7)	6 (5)

English proficiency	3-4 years	8 (9.8)	4 (10.8)	12 (10.1)
	4-5 years	19 (23.2)	4 (10.8)	23 (19.3)
	>5 years	28 (34.1)	10 (27)	38 (31.9)
	not at all	38 (46.3)	13 (35.1)	51 (42.9)
	Beginner	27 (32.9)	11 (29.7)	38 (31.9)
	intermediate	13 (15.9)	12 (32.4)	25 (21)
	Advanced	1 (1.2)	0 (0)	1 (0.8)
	Fluent	3 (3.7)	1 (2.7)	4 (3.4)
Psychiatric condition		5 (6.6)	5 (13.9)	10 (8.9)
Psychometrics [#]				
EQ-5D VAS score	$\bar{x}$ (SD)	81.1 (23.8)	75 (20.6)	79.3 (22.9)
Self-Perceived Anxiety	VAS 0-100 $\bar{x}$ (SD)	4.3 (13.9)	15.7 (27.8)	10.5 (21.0)
Self-Perceived Depression	VAS 0-100 $\bar{x}$ (SD)	7.8 (18.6)	16.8 (24.8)	7.8 (19.8)
PHQ-9 Sum Score	0-27 $\bar{x}$ (SD)	1.7 (2.7)	2.9 (4.5)	2.0 (3.4)
GAD-7 Sum Score	0-21, $\bar{x}$ (SD)	1.0 (2.6)	1.4 (3.6)	1.1 (2.9)

* Missing age n=5, Missing education n=1, psychiatric condition missing n=7, EQ-5D VAS missing n=5, Anxiety VAS missing n= 2, Depression VAS missing = 1, # EQ-5D: Higher scores indicate better self-perceived health, all other scores: Higher scores indicate higher symptom burden.

Psychometrics and item characteristics.

Most participants displayed normal depression (PHQ-9 sum score 0-4, 89.9%) and anxiety levels (GAD-7 sum score 0-4, 95%). Sum scores of both parameters exhibited a left-skewed distribution with a skewness of 3.57 (standard error 0.22) and kurtosis of 18.02 (standard error 0.44) for the PHQ-9 and a skewness of 4.37 (standard error 0.22) and kurtosis of 21.07 (standard error 0.44) for the GAD-7. The distribution of responses is shown in Figure 1.

Inter-item correlation coefficients for the PHQ-9 and GAD-7 ranged between $r=0.06$ to $r=0.65$, and $r=0.23$ to $r=0.65$, respectively. A heatmap with correlation coefficients can be found as supplemental material (Supplemental Table 1). PHQ-9 and GAD-7 sum scores were positively correlated ($r=0.59$, $p<0.001$).

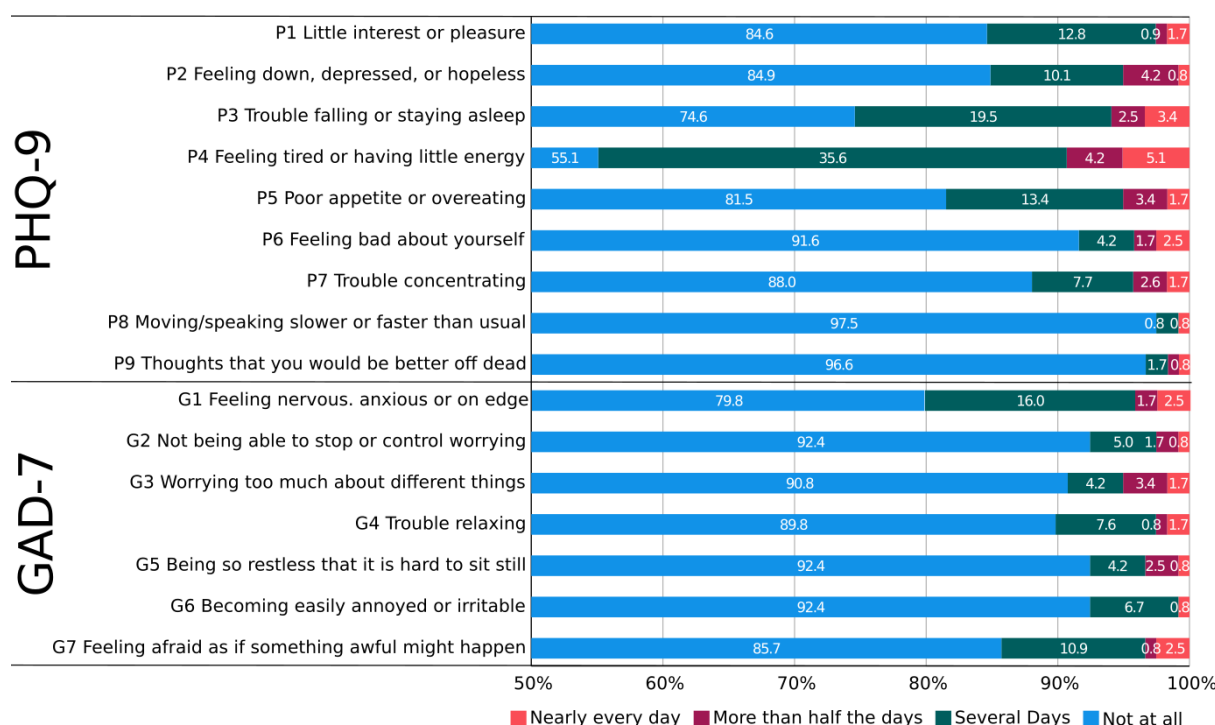


Figure 1: Participants' responses.

Reliability

PHQ-9 and GAD-7 displayed overall good internal consistency with a Cronbach's $\alpha = 0.85$ (McDonald's $\Omega = 0.87$) and $\alpha = 0.92$ (McDonald's $\Omega = 0.93$), respectively. Item-total correlations ranged between 0.40 and 0.69 for the PHQ-9 and 0.68 and 0.85 for the GAD-7. Cronbach's α did not significantly improve if items were removed from the scale.

In sensitivity analyses, lower α values for both PHQ-9 and GAD-7 were found for older respondents aged 50+ years, respondents who attempted 0-4 years in school, and respondents with lower self-perceived English proficiency and shorter stay in the US (table 2).

Table 2: Internal consistency among subgroups

Characteristic*		PHQ-9 α (95% CI)	GAD-7 α (95% CI)
Age group	18-29	0.79 (0.66 – 0.89)	0.94 (0.90 – 0.97)
	30-49	0.90 (0.86 – 0.94)	0.93 (0.90 – 0.95)
	50+	0.72 (0.52 – 0.86)	0.81 (0.68 – 0.91)
Education	0-4 school years	0.78 (0.64 – 0.88)	0.83 (0.72 – 0.91)

Time in the US	5-9 school years	0.75 (0.60 – 0.87)	0.86 (0.77 – 0.92)
	10+ school years	0.91 (0.86 – 0.94)	0.94 (0.91 – 0.96)
English proficiency	0-3 years	0.66 (0.49 – 0.79)	0.56 (0.33 – 0.73)
	> 3 years	0.88 (0.83 – 0.92)	0.94 (0.92 – 0.96)
Total	not at all	0.75 (0.62 – 0.84)	0.76 (0.64 – 0.85)
	beginner-fluent	0.91 (0.88 – 0.94)	0.95 (0.93 – 0.97)
Total		0.85 (0.81 – 0.89)	0.92 (0.89 – 0.94)

The Kaiser-Meyer-Olkin (KMO) value was 0.760 for PHQ-9 and Bartlett's test of sphericity was significant (χ^2 (36) = 552.53, $p < 0.001$). Similar results were shown for the GAD-7 (KMO value: 0.873, Bartlett's χ^2 (21) = 612.56, $p < 0.001$) suggesting both samples as appropriate for factor analyses. A principal component analysis for a single-factor design revealed factor loadings between 0.44 and 0.80 for the PHQ-9 and 0.76 to 0.90 for the GAD-7 (table 3).

Using CFA, factor loadings ranged between 0.31 and 0.82 for the PHQ-9 and 0.73 and 0.89 for the GAD-7, however, model fit indices for CFAs were poor (Supplement figures S2, S3, and supplement table S1).

Table 3: Item characteristics, item-total-correlations, and factor loadings for PHQ-9 (N=114) and GAD-7 (N=118).

Item		Score		Item-Total correlation	α if item deleted	Factor Loading PCA
		Mean	SD			
P1	Little interest or pleasure in doing things.	0.19	0.53	0.60	0.83	0.70
P2	Feeling down, depressed, or hopeless.	0.22	0.56	0.69	0.82	0.79
P3	Trouble falling or staying asleep.	0.34	0.68	0.59	0.84	0.66
P4	Feeling tired or having little energy.	0.58	0.80	0.55	0.85	0.64
P5	Poor appetite or overeating.	0.25	0.59	0.40	0.85	0.44
P6	Feeling bad about yourself.	0.16	0.57	0.69	0.82	0.80
P7	Trouble concentrating.	0.18	0.56	0.64	0.83	0.74
P8	Moving/speaking slower or faster than usual.	0.04	0.31	0.56	0.84	0.72
P9	Thoughts that you would be better off dead.	0.06	0.36	0.66	0.84	0.80
G1	Feeling nervous, anxious or on edge.	0.27	0.62	0.68	0.92	0.76
G2	Not being able to stop or control worrying.	0.11	0.43	0.80	0.90	0.85
G3	Worrying too much about different things.	0.16	0.55	0.76	0.91	0.83
G4	Trouble relaxing.	0.14	0.49	0.73	0.91	0.81
G5	Being so restless that it is hard to sit still.	0.12	0.46	0.85	0.90	0.90
G6	Becoming easily annoyed or irritable.	0.09	0.37	0.77	0.91	0.84
G7	Feeling something awful might happen.	0.20	0.58	0.77	0.90	0.84

Discriminant Validity

Participants with psychiatric comorbidities had significantly higher scores on the PHQ-9 (mean 9.1 [SD 6.7]) compared to those without comorbidity (mean 1.3 [SD 1.8], $p < 0.001$). The same pattern was observed for the GAD-7 scores, with those with comorbidities scoring higher (mean 5.2 [SD 6.3]) than those without (mean 0.6 [SD 1.8], $p < 0.001$) (see Figure 2).

Self-perceived depression, as measured on a Visual Analog Scale (VAS), exhibited a strong positive correlation with the PHQ-9 sum score ($r = 0.53$, $p < 0.001$). Similarly, self-perceived anxiety VAS scores showed a strong positive correlation with the GAD-7 sum score ($r = 0.65$, $p < 0.001$).

Health-related quality of life using the EQ-5D VAS score showed a strong negative correlation with the PHQ-9 sum score ($r = -0.59$, $p < 0.001$) and a moderate negative correlation with the GAD-7 sum score ($r = -0.42$, $p < 0.001$).

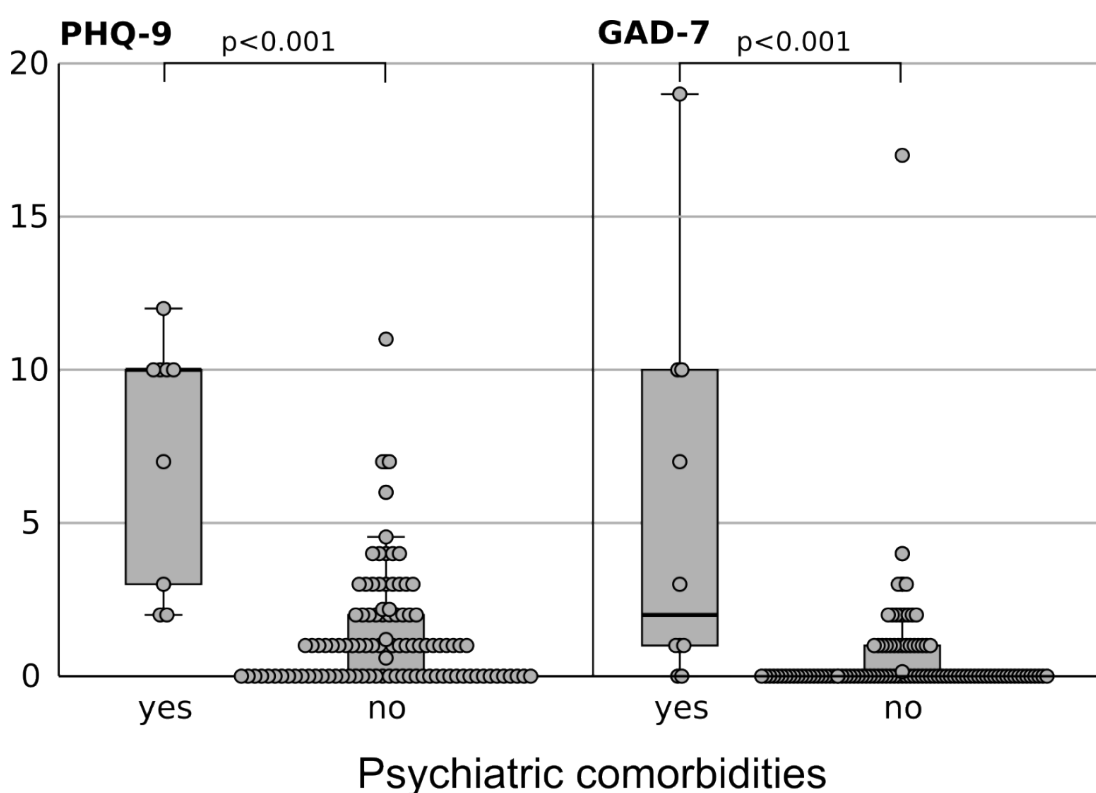


Figure 2: PHQ-9 and GAD-7 values among participants with preexisting psychiatric conditions (Mann-Whitney-U test)

Comments

Only few participants made written comments or annotations to the survey, mainly elaborating on why participants faced trouble falling asleep (P3), referring to environmental noise due to housing situation or busy roads.

4. Discussion

Studies validating psychometric instruments in a primary care situation for migrants and refugees are limited, especially for smaller language communities. The lack of availability of these instruments, however, is a key factor in insufficient mental health screening among refugee and migrant populations that are particularly affected by the trauma of relocation and mental health problems [30]. This work aimed to translate and culturally adapt the commonly used mental health screening instruments PHQ-9 and GAD-7 into the Kinyarwanda language and conduct a validation study by determining internal consistency, factorial structure, and discriminant validity in a cross-sectional sample of migrant and refugee patients seeking care in primary care clinics in the Midwest.

We found for both instruments an overall good consistency; however, lower consistency levels were observed among patients who were new to the US, had lower English proficiency levels, or were older (50+ years). Particularly the PHQ-9 items on poor appetite/overeating and sleeping problems showed lower factor loadings. However, five items had factor loadings >0.6, suggesting the overall measured factor as reliable [31].

Some participants made annotations to the survey, indicating that their impaired sleep conditions were due to environmental noise exposure that was associated with their housing situation. Research has indicated that newly arrived refugees and migrants are often accommodated in shelters, boarding houses, or substandard apartment buildings that lack privacy and are often exposed to noise that may impair their sleep [32,33]. Similarly, newly arriving refugees often face considerable challenges in retaining their food-related practices [34,35], with particular hardships in acquiring goods commonly consumed in their country of origin while facing a limited budget [36]. As a result newcomers to the US often face food insecurity [37], which is often further aggravated by language barriers [38]. These

aspects likely contribute to a feeling of decreased appetite, especially among newly arrived migrants without being a clear sign that would indicate an impairment of mental health. This may help explain the lower consistency among newcomers and lower factor loadings of the respective items in this cohort.

The rather small number of incomplete surveys suggests that the survey questions were understood well by various patient groups, including those with limited education. The quality of the translation was confirmed in cognitive interviews.

While we have performed a confirmatory factor analysis, the indicated model of fit indices were poor. It is likely that this is due to a combination of lower factor loadings among some of the previously discussed items and the vulnerability of goodness of fit indices when analyzing smaller samples [39].

Although we were not able to validate our instruments against an established gold standard, we found that respective sum scores correlated well with visual analog scales for health-related quality of life, depression, and anxiety. Furthermore, patients with a known history of mental health problems scored significantly higher for both instruments.

This study comes with further limitations that need to be considered when interpreting the results. Although our sample size met the suggested item-to-response ratios, a considerably larger sample size would have led to more robust findings, especially among subgroups of the sample, and might have yielded better goodness-of-fit indices in the confirmatory factor analyses. The majority of enrolled patients in our sample indicated no or few mental health problems thus reducing variability within our sample. This could be addressed with either a larger sample or a sample that includes more people with such problems, e.g. recruited through mental health providers.

Recently other psychometric instruments in the Kinyarwanda language have been introduced, such as the Beck Depression Inventory [40] or the Mini International Neuropsychiatric Interview [41], however, these were not used in our study as they have not been validated in the respective setting of primary care clinics in the US.

348 Further research should investigate validity against a standard diagnostic interview in a larger and
 349 more heterogenic sample, e.g. including patients from mental health clinics and psychiatric hospitals.
 350 Such a study could also test diagnostic accuracy and determine cut-off scores for depression and
 351 anxiety screening. With our study, we provide the necessary groundwork including a translated and
 352 culturally adopted version of the PHQ-9 and GAD-7 and the first results of a validation study.
 353

354

355 References

- 356 1. Weinberger AH, Gbedemah M, Martinez AM, Nash D, Galea S, Goodwin RD. Trends in
357 depression prevalence in the USA from 2005 to 2015: widening disparities in vulnerable groups.
358 Psychol Med. 2018;48: 1308–1315. doi:10.1017/S0033291717002781
- 359 2. Ösby U, Brandt L, Correia N, Ekblom A, Sparén P. Excess Mortality in Bipolar and Unipolar
360 Disorder in Sweden. Arch Gen Psychiatry. 2001;58: 844. doi:10.1001/archpsyc.58.9.844
- 361 3. Lépine J-P, Briley M. The increasing burden of depression. NDT. 2011; 3.
362 doi:10.2147/NDT.S19617
- 363 4. Barth J, Schumacher M, Herrmann-Lingen C. Depression as a Risk Factor for Mortality in
364 Patients With Coronary Heart Disease: A Meta-analysis. Psychosomatic Medicine. 2004;66:
365 802–813. doi:10.1097/01.psy.0000146332.53619.b2
- 366 5. Birk JL, Kronish IM, Moise N, Falzon L, Yoon S, Davidson KW. Depression and multimorbidity:
367 Considering temporal characteristics of the associations between depression and multiple
368 chronic diseases. Health Psychology. 2019;38: 802–811. doi:10.1037/hea0000737
- 369 6. Egede LE, Zheng D, Simpson K. Comorbid Depression is Associated With Increased Health Care
370 Use and Expenditures in Individuals With Diabetes. Diabetes Care. 2002;25: 464–470.
371 doi:10.2337/diacare.25.3.464
- 372 7. Siu AL, and the US Preventive Services Task Force (USPSTF), Bibbins-Domingo K, Grossman DC,
373 Baumann LC, Davidson KW, et al. Screening for Depression in Adults: US Preventive Services
374 Task Force Recommendation Statement. JAMA. 2016;315: 380. doi:10.1001/jama.2015.18392
- 375 8. O'Connor E, Henninger M, Perdue LA, Coppola EL, Thomas R, Gaynes BN. Screening for
376 Depression, Anxiety, and Suicide Risk in Adults: A Systematic Evidence Review for the U.S.
377 Preventive Services Task Force. Agency for Healthcare Research and Quality; 2022. Report No.:
378 AHRQ Publication No. 22-05295-EF-1.
- 379 9. Tonui BC. The cultural perceptions on mental health and post-resettlement challenges among
380 Rwandan refugees in the U.S. Psychiatry Research. 2022;313: 114642.
381 doi:10.1016/j.psychres.2022.114642
- 382 10. Mesa-Vieira C, Haas AD, Buitrago-Garcia D, Roa-Diaz ZM, Minder B, Gamba M, et al. Mental
383 health of migrants with pre-migration exposure to armed conflict: a systematic review and
384 meta-analysis. The Lancet Public Health. 2022;7: e469–e481. doi:10.1016/S2468-
385 2667(22)00061-5
- 386 11. Close C, Kouvonen A, Bosqui T, Patel K, O'Reilly D, Donnelly M. The mental health and wellbeing
387 of first generation migrants: a systematic-narrative review of reviews. Global Health. 2016;12:
388 47. doi:10.1186/s12992-016-0187-3
- 389 12. Abubakar I, Aldridge RW, Devakumar D, Orcutt M, Burns R, Barreto ML, et al. The UCL–Lancet
390 Commission on Migration and Health: the health of a world on the move. The Lancet.
391 2018;392: 2606–2654. doi:10.1016/S0140-6736(18)32114-7
- 392 13. Rahal S. Michigan becomes haven for Congo refugees. In: The Detroit News [Internet]. 12 Apr
393 2019 [cited 25 Mar 2024]. Available:

- 394 [https://eu.detroitnews.com/story/news/local/michigan/2019/04/12/michigan-becomes-](https://eu.detroitnews.com/story/news/local/michigan/2019/04/12/michigan-becomes-haven-congo-refugees/3287895002/)
395 [haven-congo-refugees/3287895002/](https://eu.detroitnews.com/story/news/local/michigan/2019/04/12/michigan-becomes-haven-congo-refugees/3287895002/)
- 396 14. Byrow Y, Pajak R, Specker P, Nickerson A. Perceptions of mental health and perceived barriers
397 to mental health help-seeking amongst refugees: A systematic review. *Clinical Psychology*
398 *Review*. 2020;75: 101812. doi:10.1016/j.cpr.2019.101812
- 399 15. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9: Validity of a brief depression severity measure.
400 *J Gen Intern Med*. 2001;16: 606–613. doi:10.1046/j.1525-1497.2001.016009606.x
- 401 16. Sebera F, Vissoci JRN, Umwiringirwa J, Teuwen DE, Boon PE, Dedeken P. Validity, reliability and
402 cut-offs of the Patient Health Questionnaire-9 as a screening tool for depression among
403 patients living with epilepsy in Rwanda. Tran TD, editor. *PLoS ONE*. 2020;15: e0234095.
404 doi:10.1371/journal.pone.0234095
- 405 17. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A Brief Measure for Assessing Generalized Anxiety
406 Disorder: The GAD-7. *Arch Intern Med*. 2006;166: 1092. doi:10.1001/archinte.166.10.1092
- 407 18. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, et al. The
408 Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement:
409 guidelines for reporting observational studies. *J Clin Epidemiol*. 2008;61: 344–349.
410 doi:10.1016/j.jclinepi.2007.11.008
- 411 19. Behr D. Translating questionnaires for cross-national surveys: A description of a genre and its
412 particularities based on the ISO 17100 categorization of translator competences. *Trans-Int*.
413 2018;10: 5–20. doi:10.12807/ti.110202.2018.a02
- 414 20. Valdez D, Montenegro MS, Crawford BL, Turner RC, Lo W-J, Jozkowski KN. Translation
415 frameworks and questionnaire design approaches as a component of health research and
416 practice: A discussion and taxonomy of popular translation frameworks and questionnaire
417 design approaches. *Social Science & Medicine*. 2021;278: 113931.
418 doi:10.1016/j.socscimed.2021.113931
- 419 21. Schwarz N. Cognitive aspects of survey methodology. *Appl Cognit Psychol*. 2007;21: 277–287.
420 doi:10.1002/acp.1340
- 421 22. Tourangeau R, Rips LJ, Rasinski K. *The Psychology of Survey Response*. 1st ed. Cambridge
422 University Press; 2000. doi:10.1017/CBO9780511819322
- 423 23. Rabin R, Charro FD. EQ-SD: a measure of health status from the EuroQol Group. *Annals of*
424 *Medicine*. 2001;33: 337–343. doi:10.3109/07853890109002087
- 425 24. Huang Z, Kohler IV, Kämpfen F. A Single-Item Visual Analogue Scale (VAS) Measure for
426 Assessing Depression Among College Students. *Community Ment Health J*. 2020;56: 355–367.
427 doi:10.1007/s10597-019-00469-7
- 428 25. MacCallum RC, Widaman KF, Zhang S, Hong S. Sample size in factor analysis. *Psychological*
429 *Methods*. 1999;4: 84–99. doi:10.1037/1082-989X.4.1.84
- 430 26. Gelaye B, Williams MA, Lemma S, Deyessa N, Bahretibeb Y, Shibre T, et al. Validity of the
431 Patient Health Questionnaire-9 for depression screening and diagnosis in East Africa. *Psychiatry*
432 *research*. 2013;210: 653–661. doi:10.1016/j.psychres.2013.07.015

27. Johnson SU, Ulvenes PG, Øktedalen T, Hoffart A. Psychometric Properties of the General Anxiety Disorder 7-Item (GAD-7) Scale in a Heterogeneous Psychiatric Sample. *Frontiers in psychology*. 2019;10: 1713. doi:10.3389/fpsyg.2019.01713
28. Cristobal E, Flavián C, Guinalú M. Perceived e-service quality (PeSQ). *Managing Service Quality: An International Journal*. 2007;17: 317–340. doi:10.1108/09604520710744326
29. Comrey AL, Lee HB. *A first course in factor analysis*. Second edition. New York: Psychology Press; 2016.
30. Müller F, Abdelnour AM, Rutaremara DN, Arnetz JE, Achtyes ED, Alshaarawy O, et al. Mental Health Screening Differences in Non-English Speaking Patients: Results From a Retrospective Cohort Study. *J Prim Care Community Health*. 2023;14: 21501319231200304. doi:10.1177/21501319231200304
31. Field A. *Discovering statistics using SPSS for Windows: Advanced techniques for the beginner*. Repr. London: Sage Publ; 2004.
32. Bakker L, Cheung SY, Phillimore J. The Asylum–Integration Paradox: Comparing Asylum Support Systems and Refugee Integration in The Netherlands and the UK. *International Migration*. 2016;54: 118–132. doi:10.1111/imig.12251
33. Gewalt SC, Berger S, Ziegler S, Szecsenyi J, Bozorgmehr K. Psychosocial health of asylum seeking women living in state-provided accommodation in Germany during pregnancy and early motherhood: A case study exploring the role of social determinants of health. *PloS one*. 2018;13: e0208007. doi:10.1371/journal.pone.0208007
34. Hersey J, Anliker J, Miller C, Mullis RM, Daugherty S, Das S, et al. Food shopping practices are associated with dietary quality in low-income households. *Journal of nutrition education*. 2001;33 Suppl 1: S16–26. doi:10.1016/s1499-4046(06)60066-3
35. Sellen D, Tedstone A, Kerhsen J. *Food in the Migrant Experience*. 2002.
36. Drewnowski A, Specter SE. Poverty and obesity: the role of energy density and energy costs. *The American journal of clinical nutrition*. 2004;79: 6–16. doi:10.1093/ajcn/79.1.6
37. Hadley C, Sellen D. Food security and child hunger among recently resettled Liberian refugees and asylum seekers: a pilot study. *Journal of immigrant and minority health*. 2006;8: 369–375. doi:10.1007/s10903-006-9007-9
38. Himmelgreen D, Pérez-Escamilla R, Segura-Millán S, Peng Y-K, Gonzalez A, Singer M, et al. Food Insecurity Among Low-Income Hispanics in Hartford, Connecticut: Implications for Public Health Policy. *Human Organization*. 2000;59: 334–342. doi:10.17730/humo.59.3.76557m317748l414
39. Sun J. Assessing Goodness of Fit in Confirmatory Factor Analysis. *Measurement and Evaluation in Counseling and Development*. 2005;37: 240–256. doi:10.1080/07481756.2005.11909764
40. Dedeken P, Nickenig Vissoci JR, Sebera F, Boon PAJM, Rutembesa E, Teuwen DE. Validity, Reliability, and Diagnostic Cut-off of the Kinyarwanda Version of the Hamilton Depression Rating Scale in Rwanda. *Frontiers in psychology*. 2020;11: 1343. doi:10.3389/fpsyg.2020.01343
41. Kayiteshonga Y, Sezibera V, Mugabo L, Iyamuremye JD. Prevalence of mental disorders, associated co-morbidities, health care knowledge and service utilization in Rwanda - towards a

472 blueprint for promoting mental health care services in low- and middle-income countries? BMC
473 public health. 2022;22: 1858. doi:10.1186/s12889-022-14165-x
474

PHQ-9

GAD-7

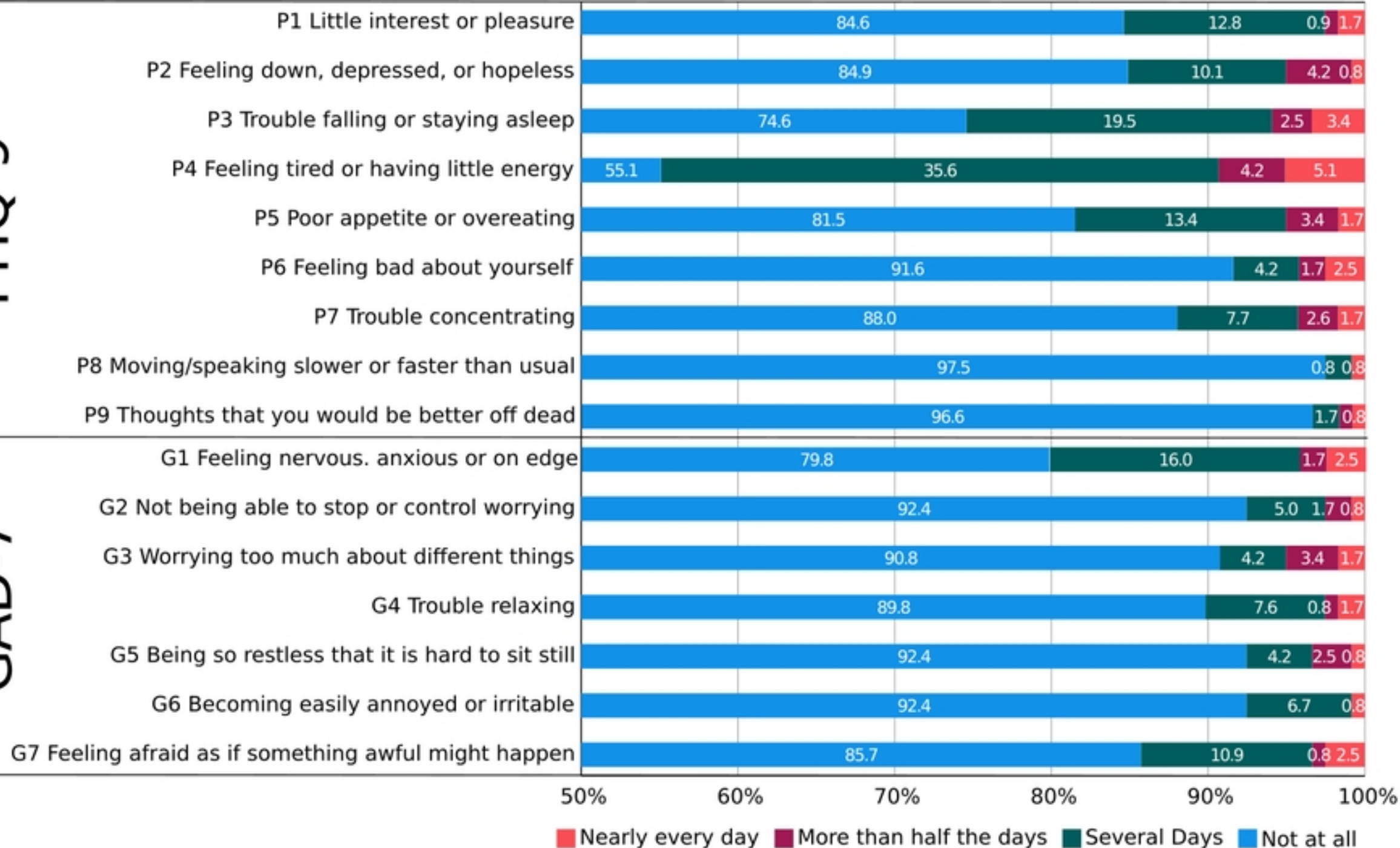


Figure 1

