Ann Ig 2021; 33(3): 254-267 doi:10.7416/ai.2021.2431

Translation, cross-cultural adaptation and psychometric evaluation of yoruba version of the short-form 12 health survey

C.E. Mbada[1], A.S. Awokoya[1], O.O. Oyewole[2], O.A. Idowu[3], A. Akinsulore[4], C. Fatoye[5], K.I. Oke[3], F. Fatoye[5]

Key words: Yoruba SF-12, Cross-Cultural Adaptation, Psychometric Evaluation, Health-related quality of life

Parole chiave: Yoruba DF-12, Adattamento transculturale, Valutazione psicrometrica, Vita di qualità legata alla salute

Abstract

Background. *Short Form 12 (SF-12) health survey has found its utility in clinical and research settings because of its short length that spares time. Though several translations into other languages do exist there is none available in Yoruba language. Hence, this study's objective was to culturally adapt and determine the reliability and validity of the Yoruba translated version of the SF-12.*

Methods. *Forward and backward translations of SF-12 into Yoruba version of SF-12 (Y-SF-12) were done using the International Quality of Life Assessment Project Guidelines. Healthy participants were assessed using both English and Yoruba versions of SF-12 for the validation phase, and two weeks later were re-assessed with the Y-SF-12 for the reliability phase.*

Results. *Participants were 225 males and 171 females. The mean scores for each scale range from 73.4 to 86.1, with no gender difference. All scale and domain scores evidenced a negative skew and ranges from -1.79 to -0.62. Concurrent validity (0.879 – 0.938) and convergent validity (0.786 – 0.907) appeared to be good as reflected by their correlation values. The internal consistency of Y-SF-12 was good as Cronbach's Alpha ranged between 0.899 and 0.968, while the Intraclass Correlation Coefficient (ICC) ranged between 0.775 and 0.949.*

Conclusion. *This is the first study to assess the psychometric properties of the Y-SF-12. It appears to be valid and may be an appropriate tool for assessing health-related quality of life among Yoruba population. The tool may help to improve the health outcomes of individuals, and redress health inequalities in low and middle-income countries.*

[1] *Department of Medical Rehabilitation, Obafemi Awolowo University, Ile-Ife, Nigeria*
[2] *Department of Physiotherapy, Olabisi Onabanjo University Teaching Hospital, Sagamu, Nigeria*
[3] *Department of Physiotherapy, University of Benin, Benin, Nigeria*
[4] *Department of Mental Health, Obafemi Awolowo University, Ile-Ife, Nigeria*
[5] *Department of Health Professions, Faculty of Health, Psychology and Social Care, Manchester Metropolitan University, UK*

Introduction

The need for brief but comprehensive health-related quality of life (HRQoL) led to the development of the 12-item health survey form (SF-12) (1). Especially for use in busy clinics and large surveys of general and specific populations. Since its development and validation, it has found its utility in clinics and research settings to assess HRQoL for specific clinical and general populations (2-9). Galenkamp et al (10) reported that SF-12 demonstrated no measurement invariance regarding gender, age and educational level. However, SF-12 was not measurement of invariant regarding ethnicity (10). This suggests the appropriateness of the SF-12 to assess physical and mental health differences across age groups, gender and educational levels but inappropriate to assess these or interpret differences between subpopulations with various ethnic backgrounds (10). Hence, this underscores the need to translate and culturally adapt SF-12 to various ethnic languages to assess general health.

Gandek et al. (11), in a cross-validation study, tested the standard 12-item selection suggested in the original United State study (12) for nine European countries (Denmark, France, Germany, Italy, the Netherlands, Norway, Spain, Sweden, and the United Kingdom). It was found that a high level of the variance of the PCS-36 and MCS-36 scores was achieved by the SF-12 as compared to the SF-36 across all nine countries. These results provided substantial evidence of concurrent validity for the SF-12. The Hausa SF-12 also established satisfactory construct validity, internal consistency, and test-retest reliability (13). The SF-12, since then, has been extensively used in health status studies involving the general population (7-9, 14) as well as in studies with disease groups (2–6).

The SF-12 has been translated into several languages by the International Quality of Life Assessment (IQOLA) project and concurrently tested in these languages (8, 15-17). However, depending on the population the instrument is applied, the validity and reliability can vary to a large extent, especially in situations where the individuals examined class under different sociodemographic characteristics and health status (18, 19). In this way it becomes a necessity to statistically examine the psychometric properties of a HRQoL assessment tool such as the SF-12 in population groups in different languages. Therefore, validating the applicability of such a tool for a given population, before it can be used to assess the HRQoL in the chosen cultural group where it has not been cross-culturally adapted, is warranted.

For effective documentation of the HRQoL items and scales, the need to assess patients in the language they can read and understand perfectly is extremely important and Nigeria, a multi-ethnic and multi-lingual country that it is, has not had Yoruba cross-cultural adaptation and psychometric testing of this SF-12 tool. Yoruba, one of the most widely spoken languages in Nigeria and other West African countries, has had the psychometric testing of its translated version of the SF-36 done (20). Yoruba language has been estimated to be spoken by about 40 million people including secondary speakers in Nigeria (21). However, the psychometric testing of the Yoruba indigenous language's version of SF-12 is yet to be determined and since it has been reported that a significant number of patients in Nigeria neither speaks nor writes English (22). Therefore, in order to support the global health and Sustainable Development Goal (SDG) initiatives of improving the health outcomes of individuals and redressing health inequalities in low and middle-income countries (LIMCs), it becomes vital to carry out the cross-cultural adaptation and its psychometric properties of Y-SF-12. Hence, the objective of this

study was to culturally adapt and determine the reliability and validity of the Y-SF-12.

Methods

Participants and study site

Healthy individuals and residents of Ile-Ife, Osun state, Nigeria volunteered for this study. Eligible participants were literate in both English and Yoruba languages and have no cognitive or mental impairments. Student participants were recruited from four selected halls of residence of Obafemi Awolowo University (OAU), Ile-Ife (two each for male and *female* students' halls). Every odd number room in all the blocks were sampled. Staff participants were randomly selected from ten departments using a fishbowl technique. A multistage cluster sampling technique were used to recruit participants from the community (23). In the community, participants were randomly selected from five out of eleven political wards into which Ile-Ife central local government is divided. Houses with odd numbers were selected for survey. Prior to data collection sample size was determined at 5% precision and 95% confidence interval (24). Minimum sample of 396 participants were found to power the study. In order for all groups making up the sample size be adequately represented; 198 slots were allocated to Ile-Ife residents and the other 198 slots were allocated to students and lecturers in OAU.

Research design

The design for this study was a cross-sectional survey.

Instruments

The tools employed in this study were: the English and Yoruba translated version of Short Form 12 (SF-12) questionnaires.

The English version of Short Form 12 (SF-12) questionnaire.

The SF-12, a subset of 12 items and 8 scales from the SF-36. It includes six items from the physical summary measure, namely one item each from the bodily pain (BP), general health (GH) and two items each from the physical functioning (PF) and role-physical (RP), and six items from the mental summary measure, namely one item each from the vitality (VT) and social functioning (SF) scales and two items each from the role-emotional (RE) and mental health (MH) scales. In both, the score ranges on a scale of 0 (the worst) to 100 (the best), with higher scores associated with the highest levels of quality of life (25). All 12 items are used to calculate the Physical Component Summary (PCS-12) and the Mental Component Summary (MCS-12) by applying empirically derived scoring algorithms (The Quality Metric Health Outcome scoring software 2) (26).

Translation procedure

The SF-12 was translated from English to Yoruba based on the IQOLA translation procedure (27). The methodological process included forward and backward translation. Two bilingual translators with significant clinical experience translated the original questionnaire into the Yoruba version. Each translator prepared a separate translation and encountered difficulties with obtaining conceptually equivalent expressions in Yoruba were assessed. Both versions were discussed between the translators in order to reach consensus. Subsequently, the two translators had a meeting to produce the first adapted version. To assess the conceptual equivalence, this adapted version was translated back into English by two other bilingual translators who were unfamiliar with the original SF-12. Their translations back into the language were compared with the original version in order to identify items or words that were precisely equivalent.

A pilot study was carried out to gather opinions from 10 participants with diverse

health status and were recruited in the study to test the face validity and administrative feasibility, and to assess comprehension of the translated questionnaire and the terminology used. It should be noted that, because of no funding and resources, the response choices were forward and backward-translated without passing through the process of selecting the response choices that will not be equally distanced. However, the IQOLA states that these steps are optional. For the same reason, the rating of the difficulty and quality of the forward translations were skipped. Again, this is an optional step as per the IQOLA protocol.

The English and Yoruba translated questionnaires with questions on socio-economic and demographics characteristics were given simultaneously to respondents to answer on two different days with interval of two weeks. On day 1, participants were assessed using both English and Yoruba versions, while on day 2 (two weeks later), participants were re-assessed on only the Yoruba translated version of the SF-12 (Appendix). Calculation of raw scores for the eight scales were performed using the transformed scores (range: 0-100). The psychometric testing of the Yoruba translated questionnaire was then carried out by testing its validity and reliability. Ethical approval was obtained from the Ethics and Research Committee, Institute of Public Health, College of Health Sciences, Obafemi Awolowo University, Ile-Ife, Nigeria (Approval No: IPHOAU/12/915).

Data analysis

Data were summarized using descriptive statistics of mean, standard deviation and percentages. Spearman correlation was carried out to determine the relationship between the English and Yoruba Translated version of SF-12 questionnaire. Known-groups validity was tested by comparing scales and domains scores, adjusted for gender using student t-test across groups known to differ. Multi trait scaling analysis was used to confirm item convergent and discriminant validity. The reliability of the SF-12 scales was calculated with intra class correlation and Cronbach's coefficient alpha to confirm item's test-retest reliability and internal consistency. The level of significance was accepted as $p \leq 0.05$. The 23.0 version of the Statistical Package for Social Sciences (SPSS) was used for data analysis.

Results

A total of 396 participants (males, 225) participated in this study. The participants' ages ranged between 18 and 70 years, with mean of 25.4 ±9.3 years. The socio-demographic characteristics of the respondents are presented in Table 1.

Concurrent validity and descriptive statistics of the Yoruba version of the SF-12

Table 2 shows the results of concurrent validity by correlation of scores of English and Yoruba versions of the SF-12 using Spearman correlation. The Spearman correlation (r) of scales were within the range of 0.879 and 0.938. The highest was scored by Physical Functioning (PF) scale (r = 0.938, p = 0.001) and the lowest scored by Role Limitation-Physical Health (RP) (r = 0.879, p = 0.001). The scores for the Physical Components and Mental Components were (r = 0.928, p = 0.001) and (r = 0.916, p = 0.001), respectively.

The mean score, confidence interval, skewness and Kurtosis of each scale and domain of the Yoruba version of the SF-12 are also presented in table 2. The average mean scores for each scale range from 73.4 to 86.1. Physical Functioning (PF) had the highest mean score (86.1), while the lowest mean score was obtained for Role Emotional (RE) (73.4). The domain scores for Physical Health Component (PHC) and the Mental Health Component (MHC) were 83.4 ±

Appendix 1

SF-12 Health Survey

This survey asks for your views about your health. This information will help keep track of how you feel and how well you are able to do your usual activities. **Answer each question by choosing just one answer**. If you are unsure how to answer a question, please give the best answer you can.

1. **In general, would you say your health is**:

□ Excellent □ Very good □ Good □ Fair □ Poor

The following questions are about activities you might do during a typical day. Does your health now limit you in these activities? If so, how much?

	YES limited a lot	YES limited a little	NO, not limited at all
2. **Moderate activities** such as moving a table, pushing a vacuum cleaner, bowling, or playing golf.	□	□	□
3. Climbing **several** flights of stairs.	□	□	□

During the past 4 weeks, have you had any of the following problems with your work or other regular daily activities as a result of your physical health?

	YES	NO
4. **Accomplished less** than you would like.	□	□
5. Were limited in the **kind** of work or other activities.	□	□

During the past 4 weeks, have you had any of the following problems with your work or other regular daily activities as a result of any emotional problems (such as feeling depressed or anxious)?

	YES	NO
6. **Accomplished less** than you would like.	□	□
7. Did work or activities **less carefully than usual**.	□	□

8. **During the past 4 weeks, how much did pain interfere with your normal work (including work outside the home and housework)?**

□ Not at all □ A little bit □ Moderately □ Quite a bit □ Extremely

These questions are about how you have been feeling during the past 4 weeks.
For each question, please give the one answer that comes closest to the way you have been feeling.
How much of the time during the past 4 weeks…

	All of the time	Most of the time	A good bit of the time	Some of the time	A little of the time	None of the time
9. Have you felt calm & peaceful?	□	□	□	□	□	□
10. Did you have a lot of energy?	□	□	□	□	□	□
11. Have you felt down-hearted and blue?	□	□	□	□	□	□

12. **During the past 4 weeks, how much of the time has your physical health or emotional problems interfered with your social activities (like visiting friends, relatives, etc.)?**

□ All of the time □Most of the time □Some of the time □A little of the time □None of the time

Appendix 1

SF-12 Àgbéyẹ̀wò Ìlera

Àgbéyẹ̀wò yìí bèèrè fún àwọn èrò rẹ lórí ìlera rẹ. Àkójọpọ̀ èrò yìí yóò fi ìmọ̀lára rẹ àti bí o ṣe le ṣe àwọn iṣẹ́ òòjọ́ rẹ dáradára sí hàn. Dáhùn àwọn ìbéèrè yìí nípa mímú ìdáhùn kan ṣoṣo. Tí o bá ń ṣe iyèméjì lórí àtidáhùn ìbéèrè jọ̀wọ́ sa ipá rẹ láti le pèsè ìdáhùn tí ó tọ̀nà jùlọ.

1. Ní àkótán, ipò wo ni o lè sọ pé ìlera rẹ wà:

❑ Ó dára jùlọ ❑Ó dára púpọ̀ ❑Ó dára ❑Ó dára díẹ̀ ❑ Kò dára rárá

Àwọn ìbéèrè wọ̀nyí dálé orí àwọn àmúṣe rẹ ní òòjọ́. Njẹ́ ìlera rẹ báyìí ń mú ìdíwọ́ bá awọn iṣẹ́ òòjọ́? Tó bá rí bẹ́ẹ̀, dé gbèdéke wo ni?

	Bẹ́ẹ̀ni, ó ṣe àkùdé	**Bẹ́ẹ̀ni, ó ṣe àkùdé díẹ̀**	**Rárá, Bẹ́ẹ̀kọ́ kò sí àkùdé**
2. **Àwọn àmúṣe** pẹ̀pẹ̀pẹ́ bíi: sísún tábìlì, nínú ilẹ̀, ṣíṣe eré ìdárayá	❑	❑	❑
3. Gígun àtẹ̀gùn ọ̀lọ́pọ̀ọ̀pọ̀ ìpele	❑	❑	❑

Níwọ̀n ọ̀sẹ̀ mẹ́rin séyìn, látàrí ìlera rẹ, ṣé o ti dojúkọ àwọn ìṣòro wọ̀nyí nínú àwọn iṣẹ́ òòjọ́ rẹ?

	Bẹ́ẹ̀ni	**Bẹ́ẹ̀kọ́**
4. Ṣíṣe kéré ju bí o ti fẹ́	❑	❑
5. Àdínkù dé bá irúfẹ́ iṣẹ́ tàbí àmúṣe iṣẹ́ òòjọ́ rẹ	❑	❑

Níwọ̀n ọ̀sẹ̀ mẹ́rin séyìn, ǹjẹ́ o ti ní àwọn ìdojúkọ wọ̀nyí pẹ̀lú àwọn iṣẹ́ òòjọ́ rẹ látàrí ìṣòro ìmọ̀lára (bíi ìrẹ̀wẹ̀sì tàbí ílàílo)?

	Bẹ́ẹ̀ni	**Bẹ́ẹ̀kọ́**
6. Ṣíṣe kéré ju bí o ti fẹ́	❑	❑
7. Ṣíṣe iṣẹ́ òòjọ́ pẹ̀lú ìjáfara ju ti tẹ́lẹ̀	❑	❑

8. **Níwọ̀n ọ̀sẹ̀ mẹ́rin séyìn, gbèdéke àkùdé wo ni ìrora mú bá iṣẹ́ rẹ (pẹ̀lúpẹ̀lú, iṣẹ́ òde àti iṣẹ́-ilé)?**

❑Kò nípa rárá ❑Ó nípa fẹ́ẹ́rẹ́fẹ́ ❑Ó mọ níwọ̀nba ❑Ó nípa lọ́pọ̀ ❑Ó kọjá àfaradà

Àwọn ìbéèrè yìí dá lórí bí o ṣe ń ṣe ọ́ láti bí ọ̀sẹ̀ mẹ́rin séyìn. Fún ìbéèrè kọ̀ọ̀kan, jọ̀wọ́ pèsè ìdáhùn kan tí ò súnmọ́ bí ó ṣe ń ṣe ọ́ gírígírí.

Ní gbèdéke ìgbà mélòó ni láàrin, ọ̀sẹ̀ mẹ́rin séyìn…

	Ìgbà gbogbo	lọ́pọ̀ ìgbà	níwọ̀n ìgbà díẹ̀	ní ẹ̀kọ̀ọ̀kan	ní ìgbà kéréje	kò ìgbà kan
9. Ǹjẹ́ o ní ìbàlẹ̀ ọkàn àti àlàáfíà?	❑	❑	❑	❑	❑	❑
10. Ǹjẹ́ o ní okun tí ó péye?	❑	❑	❑	❑	❑	❑
11. Ǹjẹ́ o ti ní ìbànújẹ́ àti ọgbẹ́ ọkàn?	❑	❑	❑	❑	❑	❑

12. **Níwọ̀n bíi ọ̀sẹ̀ mẹ́rin séyìn, bí ìgbà mélòó ni àìlera tàbí ìṣòro ìmọ̀lára rẹ ti mú àkùdé bá ìbáṣepọ̀ rẹ pẹ̀lú àwọn tí ó wà láyìíká rẹ (bíi lílọ kí àwọn ọ̀rẹ́, ojúlùmọ̀, abbl)?**

❑ Ìgbà gbogbo ❑lọ́pọ̀ọ̀pọ̀ ìgbà ❑níwọ̀n ìgbà díẹ̀ ❑ní ẹ̀kọ̀ọ̀kan ❑ní ìgbà kéréje ❑kò sí ìgbà kan

Table 1 - Socio-demographic characteristics of the respondents (n = 396)

Variables	Frequency	Percentage (%)
Age group		
< 20 years	28	7.1
20-30 years	291	73.5
30-40 years	30	7.6
>40 years	47	11.9
Gender		
Male	255	64.4
Female	141	35.6
Marital status		
Single	342	86.4
Married	48	12.1
Widowed	5	1.3
Divorced	1	0.3
Ethnic group		
Hausa	1	0.3
Igbo	14	3.5
Yoruba	371	93.7
Others	10	2.5
Religion		
Christianity	312	78.8
Islam	80	20.2
Traditional	2	0.5
Others	2	0.5
Educational level		
No Formal	1	0.3
Primary	0	0.0
Secondary	31	7.8
Tertiary	364	91.9

18.1 and 76.0 ± 21.9, respectively. There were no significant gender differences in domain scores ($p > 0.05$). However, results showed that males had significant higher mean scores in VT ($p = 0.025$) scale. Mean scores of male range from 70.6 to 86.6 while female mean scores range from 74.0 to 85.2. The highest mean scale score for male and female was PH (86.6 for male and 85.2 for female) and the lowest scale score for male was RE (70.6) and for female VT (74.0). The scale and domain scores show a negative skew, ranges from -1.79 to -0.62.

Tests of Yoruba SF-12 scaling assumptions (convergent validity) and discriminant validity

Four items had a correlation coefficient high of 1.000 (i.e. item 1, 8, 10, 12) because their corresponding scale (i.e. GH, BP, VT, SF) had only them as the scale's only item, however, scales PF, RP, RE and MH have 2 items each as their scale's item. Correlations between each item and its hypothesised scale were all above 0.50. The highest score was 0.907 (item 3) "Climbing several flights of stairs" for the Physical Functioning (PF)

Table 2 - Pearson correlation (r) analysis for the Concurrent validity and descriptive statistics of the Yoruba version SF-12

Scale/domain	r	Male mean±SD	Female mean±SD	Both gender mean±SD	95%CI	Skewness	Kurto-sis
PF	0.938*	86.6 ± 21.7	85.2 ± 20.9	86.1±21.41	84.0 – 88.2	-1.42	1.14
RP	0.879*	85.1 ± 35.7	80.1 ± 40.0	83.2±37.4	79.5 – 87.0	-1.79	1.22
BP	0.926*	81.4 ± 22.1	80.0 ± 21.2	80.8±21.8	78.7 – 83.0	-0.96	0.14
GH	0.906*	84.1 ± 17.3	81.5 ± 17.1	83.2±17.2	81.5 – 84.9	-1.39	2.80
VT	0.890*	78.5 ± 18.1†	74.0 ± 20.4	77.1±18.8	75.2 – 79.6	-1.04	1.79
SF	0.903*	74.5 ± 43.6	78.7 ± 41.0	76.0±42.8	71.7 – 80.2	-1.22	-0.052
RE	0.923*	70.6 ± 45.7	78.0 ± 41.6	73.4±44.2	69.0 – 77.8	-1.06	-0.87
MH	0.909*	77.3 ± 15.4	78.0 ± 15.9	77.6±15.6	76.0 – 79.1	-0.62	0.08
PCS	0.928*	84.3 ± 17.8	81.7 ± 18.7	83.4±18.1	81.6 – 85.1	-1.45	1.446
MCS	0.916*	75.3 ± 22.3	77.7 ± 21.2	76.0±21.9	73.8 – 78.2	-0.90	0.351

Key: **Scales**: physical functioning (PF), role limitations due to physical problems (RP), bodily pain (BP), general health (GH), vitality (VT), social functioning (SF), role limitations due to emotional problems (RE), and mental health (MH). **Domains**: physical health component (PC) and mental health components (MC).
*r significant at *P*=0.001; † significant at *P*=0.025

Table 3 - Item-scale correlations (discriminant validity) of the Yoruba SF-12 (n = 396)

Items/scales	PF	RP	BP	GH	VT	SF	RE	MH
1	0.186	0.367	0.331	1.000	0.477	0.180	0.146	0.409
2	0.873	0.226	0.228	0.098	0.232	0.183	0.117	0.168
3	0.907	0.263	0.270	0.224	0.197	0.199	0.106	0.176
4	0.205	0.890	0.285	0.336	0.252	0.329	0.271	0.280
5	0.286	0.884	0.331	0.314	0.291	0.192	0.232	0.271
6	0.091	0.201	0.196	0.140	0.148	0.138	0.861	0.307
7	0.123	0.286	0.329	0.108	0.170	0.174	0.849	0.238
8	0.281	0.347	1.000	0.331	0.310	0.275	0.306	0.421
9	0.150	0.217	0.412	0.421	0.530	0.162	0.234	0.786
10	0.239	0.306	0.310	0.447	1.000	0.219	0.186	0.443
11	0.163	0.283	0.281	0.254	0.209	0.325	0.283	0.836
12	0.215	0.295	0.275	0.180	0.219	1.000	0.182	0.306

Key: **Scales**: physical functioning (PF), role limitations due to physical problems (RP), bodily pain (BP), general health (GH), vitality (VT), social functioning (SF), role limitations due to emotional problems (RE), and mental health (MH).

scale and the lowest score was 0.786 (item 9) “Have you felt calm & peaceful?”. The details for the results of tests of scaling assumption for the general population data are summarized in table 3.

For item discriminant validity, correlation of items in PF, RP, RE and MH were less than 0.3 except item 4 “Accomplished less than you would like” with GH (0.336) and SF (0.329); item 5 “Were limited in the kind

of work or other activities" with BP (0.331) and GH (0.314); item 6 "Accomplished less than you would like" with MH (0.307); item 7 "Did work or activities less carefully than usual" with BP (0.329); item 9 "Have you felt calm & peaceful" with BP (0.412), GH (0.421) and VT (0.530); and item 11 "Have you felt down-hearted and sad" with SF (0.306). The details of values in discriminant validity are provided in Table 3.

Correlations between most scales and hypothesised Physical Components and Mental Components in Yoruba version were all 0.40, except for scale General Health (0.293), Physical Functioning scale (0.351), Role Limitation-Physical Health (0.315), all to Mental Component (MC) and Social Functioning (0.321), Role Limitation-Mental Health (0.319), all to Physical Component (PC). The highest score was for Role Limitation-Physical Health (RP) (0.848) against Physical Component (PC). Only two scales were strongly associated to Physical Component; Physical Functioning (PF) and Role Limitation-Physical Health (RP): (0.760) and (0.848), respectively. Only two scales correlated, were strongly associated with Mental Component: Social

Table 4 - Intraclass correlation coefficient (ICC), Cronbach alpha (α) and Correlation of domains with the 8 scales in test-retest of the Yoruba version of SF-12

Scales/domains	Value of correlation			
	PCS	MCS	ICC	Cronbach alpha (α)
PF	0.760^{+}	0.351*	0.906	0.968
RP	0.848^{+}	0.315*	0.775	0.935
BP	0.657*	0.409*	0.933	0.961
GH	0.596*	0.293^{-}	0.775	0.949
VT	0.449*	0.449*	0.949	0.942
SF	0.321*	0.808^{+}	0.878	0.949
RE	0.319*	0.815^{+}	0.929	0.960
MH	0.425*	0.526*	0.899	0.952
PCS			0.883	0.963
MCS			0.942	0.956
Item by item				
1			0.775	0.949
2			0.808	0.943
3			0.885	0.963
4			0.896	0.950
5			0.775	0.935
6			0.929	0.960
7			0.878	0.949
8			0.933	0.961
9			0.911	0.936
10			0.949	0.942
11			0.801	0.899
12			0.829	0.928

Key: $^{+}$strong association ($r > 0.70$); *moderate association ($0.30 < r < 0.70$); $^{-}$weak association ($r < 0.30$).
Scales: physical functioning (PF), role limitations due to physical problems (RP), bodily pain (BP), general health (GH), vitality (VT), social functioning (SF), role limitations due to emotional problems (RE), and mental health (MH).
Domains: physical health component (PCS) and mental health components (MCS).

Functioning (SF) and Role Limitation-Mental Health (RE): (0.808) and (0.815). The correlation of General Health (GH) with Mental Component (MC) was the weakest association (0.293). The full results of the correlations between scales and components (domains) are summarized in Table 4.

Reliability (test-retest)

The internal consistency of the Yoruba version of the SF-12 using Cronbach's Alpha and Intraclass Correlation Coefficient (ICC) at the level of twelve items are shown in table 4 while for the eight scales are also shown in Table 4. All the items had Cronbach's Alpha ranged between 0.899 - 0.963. Item 3 "Climbing several flights of stairs" had the highest value (0.963) while item 11 "Have you felt down-hearted and sad" had the lowest value (0.899). The Intraclass Correlation Coefficient (ICC) ranges 0.775 to 0.949. Item 10 "Did you have a lot of energy?" had the highest value (0.949), while item 1 "In general, would you say your health is" and item 5 "Were limited in the kind of work or other activities" had the lowest score (0.775). All the scales had Cronbach's Alpha ranged between 0.935-0.968. Physical Functioning scale (PF) had the highest value (0.968) while Role Limitation-Physical Health (RP) had the lowest (0.935). The two Components (Domains) had Cronbach's Alpha of 0.956 and 0.963 for MC and PC respectively. The Intraclass Correlation Coefficient (ICC) ranges 0.775 to 0.949 for scales. Vitality scale (VT) had the highest value (0.949) while General Health scale (GH) and Role Limitation-Physical Health (RP) had the lowest value (0.775). For the components, values were 0.883 and 0.942 for PC and MC, respectively.

Discussion

To our knowledge this is the first study to translate, culturally adapt and validate the Yoruba version of the SF12 (Y-SF-12) according to the International Quality of Life Assessment (IQOLA) Project Guidelines. The Y-SF-12 had a high rate of data completion, with good data obtained not only in institutional settings but also in the context of self-administration in the general population. A high response rate was also achieved, suggesting that the Y-SF-12 was an acceptable tool for measuring health perception in the general Yoruba population. The concurrent validity of the Y-SF-12 questionnaire, when correlated with the English version, was high with scale correlations (r) ranging from 0.879 to 0.938 and domain correlations of 0.928 and 0.916 for Physical and Mental Health components, respectively. The average mean scores for scales ranged from 73.4 to 86.1 and domains of 83.4 and 76.0 for Physical Health (PH) and Mental Health (MC) respectively. The negative skewness distribution observed with the Y-SF-12 scale is also similar to the pattern observed in previous studies (28, 29). The scales and domains score showed a negative skew in the Yoruba population samples, suggesting that participants scored towards the positive end of the health spectrum (30).

The test of the known group validity of the Y-SF-12 indicated that gender does not significantly influence it. From the result, male respondents had higher mean scores in most scales except SF, RE and MH and MHC on the domain. Interestingly, this finding has not been reported anywhere in literature. While, reasons for higher HRQoL scores in men is still a subject of debate, however, Hopman et al. (31) believed that women report higher incidences of psychological symptoms and greater psychological distress than men; and that female patients with physical illnesses have worse scores than men in terms of symptoms and well-being. The higher HRQoL scores in male in this study may be due to these factors reasons.

Multi trait scaling assumption was employed in confirming the hypothesised scale structure of the Y-SF-12. In item scaling assumption, item internal consistency was supported because all item-scale correlations (i.e., correlations of an item with its own scale) were far greater than the minimum value of 0.4 recommended by Ware and Gandek (25), which showed high level of item internal consistency for all items. As expected, all items correlated better with its hypothesised scale than scales measuring other concepts. Item-scale correlations were more or less similar within each scale. As recommended and expected; all the correlation between each scale and its related summary measure were greater than its correlation with the other summary measure (Components or Domains) this is similar to result obtained in previous study (8). The internal consistency using Cronbach's alpha at the level of items, scales and domains have Cronbach alpha greater than 0.7, which is the standard minimum reliability coefficient for group-level analyses (27, 32). The Intraclass correlation coefficient (ICC) was also acceptable, because they scored greater than 0.7 which is the standard minimum measure (33). The recent published Hausa SF-12 among low back pain population was similar to our observation of strong internal consistency and convergent validity of the tool among Nigeria ethnic groups (13).

The results of this study indicated that the Y-SF-12 is a tool that can be used to measure HRQoL among the Yoruba populace. However, item 11 "Have you felt down-hearted and sad" was the least satisfactory in the discriminant validity testing. It is suggested that future research should test this and other known group validity by age, educational level and marital status for better understanding. Generally, the Y-SF-12 produced similar psychometric properties on item scale assumption, concurrent validity of scales and domains compared to the original American and other translated versions. In summary, this study revealed that the culturally adapted Y-SF-12 have reasonable psychometrics properties that are satisfactory with benchmarks. However, the potential limitation of this study is that the psychometric testing of the Y-SF-12 was carried out among residents of Ile-Ife region only, and as such be limit the generalizability of the finding. Nonetheless, Ile-Ife is far becoming a metropolitan city with a heterogeneous mix of people from different parts of the Yoruba tribe, and Nigeria at large. We do not have equal sample of males and females and as such equal variance may not be assumed among gender. It is also worthy of note that participants were healthy individuals and such could be a limitation. Future research should examine the psychometric properties of Y-SF-12 on clinical populations. Despite these limitations, this is the first translation and cultural adaptation of Y-SF-12 that showed good satisfactory psychometric benchmark as the original and can be used among Yoruba population.

Conclusion

The psychometric properties of the Y-SF-12 health survey appear to be valid. This may be appropriate for clinicians and researchers for measuring HRQoL among Yoruba population. It may also help to support the global health and SDG initiatives of improving the health outcomes of individuals and redressing health inequalities in LMICs.

Acknowledgements
Funding: None
Competing interests: None declared
Ethical approval: Approval was obtained from the Ethics and Research Committee, Institute of Public Health, College of Health Sciences, Obafemi Awolowo University, Ile-Ife, Nigeria (Approval No: IPHOAU/12/915). The procedures used in this study adhere to the tenets of the Declaration of Helsinki.

Riassunto

Traduzione in Lingua Yoruba del Modello breve 12 per le Indagini sulla Salute, suo Adattamento interculturale e suo Utilizzo per la Valutazione Psicometrica

Premessa. L'indagine sulla salute con modello breve 12 (SF-12) ha trovato il suo impiego negli ambienti clinici e di ricerca proprio perché la sua brevità fa risparmiare tempo. Benchè ne esistano traduzioni in molte altre lingue, manca quella nella lingua Yoruba. Pertanto il presente studio si è posto l'obiettivo di adattare culturalmente alla lingua Yoruba il modello S-12, e di valutarne la validità e l'affidabilità.

Metodi. Traduzione in lingua Yoruba del modello SF-12 e contro-traduzione in inglese sono state effettute utilizzando le Linee Guida dello "International Quality of Life Assessment Project". Partecipanti sani sono stati valutati usando l'una o l'altra versione del modello SF-12 per la fase dei validazione, e due settimane dopo sono stati ri-valutati con le versioni alternative per la valutazione di attendibilità.

Risultati. Hanno partecipato 396 soggetti, 225 maschi e 171 femmine. Il punteggio medio per ogni scala andava da73,4 a 86,1, senza differeza di genere. Tutti i punteggi di scala e di dominio hanno evidenziato un'inclinazione negative, con un intervallo da -1,79 a -0,62. Buone sono risultate la validità concorrente (0,879 - 0,938) e la validità convergente (0,786 - 0,907), come emerge dai loro valori di correlazione. La coerenza interna dell'Y-SF-12 è risultata buona, dato che l'Alfa di Cronbach era compresa tra 0,899 e 0,968, mentre il coefficiente di correlazione intraclasse (ICC) era compreso tra 0,775 e 0,949.

Conclusioni. Questo è il primo studio per valutare le proprietà psicometriche della versione Yoruba del modello SF-12. Sembra essere valido e può essere uno strumento appropriato per valutare la qualità della vita correlata alla salute tra la popolazione Yoruba. Lo strumento può contribuire a migliorare gli esiti sanitari degli individui e correggere le disuguaglianze sanitarie nei paesi a basso e medio reddito.

References

1. Ware J, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. Med Care 1996; **3**4(3): 220-33. doi: 10.1097/00005650-199603000-00003.
2. Patel AR, Lester RT, Marra CA, et al. The validity of the SF-12 and SF-6D instruments in people living with HIV/AIDS in Kenya. Health Qual Life Outcomes 2017; **15**(1): 143. doi: 10.1186/s12955-017-0708-7.
3. Tawiah AK, Al Sayah F, Ohinmaa A, Johnson JA. Discriminative validity of the EQ-5D-5 L and SF-12 in older adults with arthritis. Health Qual Life Outcomes 2019; ***17***(1): 68. doi: 10.1186/s12955-019-1129-6.
4. Wang X, Guo G, Zhou L, et al. Health-related quality of life in pregnant women living with HIV: a comparison of EQ-5D and SF-12. Health Qual Life Outcomes 2017; **15**(1):158. doi: 10.1186/s12955-017-0731-8.
5. Okonkwo OC, Roth DL, Pulley L, Howard G. Confirmatory factor analysis of the validity of the SF-12 for persons with and without a history of stroke. Qual Life Res Int J Qual Life Asp Treat Care Rehabil 2010; **19**(9): 1323-31. doi: 10.1007/s11136-010-9691-8.
6. Luo X, George ML, Kakouras I, et al. Reliability, validity, and responsiveness of the short form 12-item survey (SF-12) in patients with back pain. Spine 2003; **28**(15): 1739-45. doi: 10.1097/01.BRS.0000083169.58671.96.
7. Campolina AG, Lopez RVM, Nardi EP, Ferraz MB. Quality of life in a sample of Brazilian adults using the generic SF-12 questionnaire. Rev Assoc Medica Bras (1992) 2018; **64**(3): 234-42. doi: 10.1590/1806-9282.64.03.234.
8. Kontodimopoulos N, Pappa E, Niakas D, Tountas Y. Validity of SF-12 summary scores in a Greek general population. Health Qual Life Outcomes 2007; **5**: 55. doi: 10.1186/1477-7525-5-55.
9. Larson CO, Schlundt D, Patel K, Beard K, Hargreaves M. Validity of the SF-12 for use in a low-income African American community-based research initiative (REACH 2010). Prev Chronic Dis 2008; **5**(2): A44.
10. Galenkamp H, Stronks K, Mokkink LB, Derks EM. Measurement invariance of the SF-12 among different demographic groups: The HELIUS study. PloS One 2018; **13**(9): e0203483. doi: 10.1371/journal.pone.0203483.
11. Gandek B, Ware JE, Aaronson NK, et al. Cross-validation of item selection and scoring for the SF-12 Health Survey in nine countries: results from the IQOLA Project. International Quality of Life Assessment. J Clin Epidemiol 1998; **51**(11): 1171-8. doi: 10.1016/s0895-4356-(98)00109-7.
12. Ware JE, Snow KK, Kosinski M, Gandek B; New England Medical Center Hospital Health

Institute. SF-36 Health Survey: Manual and Interpretation Guide. Boston MA: Health Institute, New England Medical Center, 1993.
13. Ibrahim AA, Akindele MO, Ganiyu SO, et al. The Hausa 12-item short-form health survey (SF-12): Translation, cross-cultural adaptation and validation in mixed urban and rural Nigerian populations with chronic low back pain. PloS One. 2020;15(5): e0232223. doi: 10.1371/journal.pone.0232223.
14. Ware JE, Keller SD, Kosinski M. SF-12: How to Score the SF-12 Physical and Mental Health Summary Scales. Boston, MA: Health Institute, New England Medical Center, 1995.
15. Failde I, Medina P, Ramirez C, Arana R. Construct and criterion validity of the SF-12 health questionnaire in patients with acute myocardial infarction and unstable angina. J Eval Clin Pract 2010; **16**(3): 569-73. doi: 10.1111/j.1365-2753.2009.01161.x.
16. Montazeri A, Vahdaninia M, Mousavi SJ, Omidvari S. The Iranian version of 12-item Short Form Health Survey (SF-12): factor structure, internal consistency and construct validity. BMC Public Health 2009; **9**: 3 41. doi: 10.1186/1471-2458-9-341.
17. Younsi M, Chakroun M. Measuring health-related quality of life: psychometric evaluation of the Tunisian version of the SF-12 health survey. Qual Life Res Int J Qual Life Asp Treat Care Rehabil 2014; **23**(7): 2047-54. doi: 10.1007/s11136-014-0641-8.
18. Globe DR, Levin S, Chang TS, Mackenzie PJ, Azen S. Validity of the SF-12 quality of life instrument in patients with retinal diseases. Ophthalmology 2002; **109**(10): 1793-8. doi: 10.1016/s0161-6420(02)01124-7.
19. McCarthy ML, Silberstein CE, Atkins EA, Harryman SE, Sponseller PD, Hadley-Miller NA. Comparing reliability and validity of pediatric instruments for measuring health and well-being of children with spastic cerebral palsy. Dev Med Child Neurol 2002; **44**(7): 468-76. doi: 10.1017/s0012162201002377.
20. Mbada CE, Adeogun GA, Ogunlana MO, et al. Translation, cross-cultural adaptation and psychometric evaluation of yoruba version of the short-form 36 health survey. Health Qual Life Outcomes 2015; **13**: 141. doi: 10.1186/s12955-015-0337-y.
21. Wikipedia. Yoruba language. In: *Wikipedia*, 2020. Available on: https://en.wikipedia.org/w/index.php?title=Yoruba_language&oldid=950990033 [Last accessed: 2020, June 6].
22. Akinpelu AO, Maruf FA, Adegoke BOA. Validation of a Yoruba translation of the World Health Organization's quality of life scale--short form among stroke survivors in Southwest Nigeria. Afr J Med Med Sci 2006; **35**(4): 417-24.
23. Knottnerus P. Multistage and Cluster (Sub) Sampling. In: Knottnerus P, ed. Sample Survey Theory: Some Pythagorean Perspectives. Springer Series in Statistics. Springer, 2003: 147-87. doi: 10.1007/978-0-387-21764-2_6.
24. Kasiulevičius V, Šapoka V, Filipavičiūtė R. Sample size calculation in epidemiological studies. Gerontologija 2006; **7**(4): 225-31.
25. Ware JE, Gandek B. Overview of the SF-36 Health Survey and the International Quality of Life Assessment (IQOLA) Project. J Clin Epidemiol. 1998; **51**(11): 903-12. doi: 10.1016/s0895-4356(98)00081-x.
26. Ware JE. How to Score Version 2 of the SF-12 Health Survey (with a Supplement Documenting Version 1). Quality Metric Inc: Health Assessment Lab, 2005.
27. Streiner DL, Norman GR, Cairney J. Health Measurement Scales: A Practical Guide to Their Development and Use. 5th ed. Oxford University Press, 2014. Available on: https://oxfordmedicine.com/view/10.1093/med/9780199685219.001.0001/med-9780199685219 [Last accessed: 2020, June 6].
28. Allen J, Alpass FM, Stephens CV. The sensitivity of the MOS SF-12 and PROMIS® global summary scores to adverse health events in an older cohort. Qual Life Res Int J Qual Life Asp Treat Care Rehabil 2018; **27**(8): 2207-15. doi:10.1007/s11136-018-1871-y.
29. Norhayati MN, Nik Hazlina NH, Aniza AA. Immediate and long-term relationship between severe maternal morbidity and health-related quality of life: a prospective double cohort comparison study. BMC Public Health 2016; 16. doi: 10.1186/s12889-016-3524-9.
30. Aaronson NK, Acquadro C, Alonso J, et al. International Quality of Life Assessment (IQOLA) Project. Qual Life Res Int J Qual Life Asp Treat Care Rehabil 1992; **1**(5): 349-51. doi: 10.1007/bf00434949.
31. Hopman WM, Towheed T, Anastassiades T, et al. Canadian normative data for the SF-36 health survey. CMAJ Can Med Assoc J 2000; **163**(3): 265-71.

32. Tavakol M, Dennick R. Making sense of Cronbach's alpha. Int J Med Educ 2011; **2**: 53-5. doi: 10.5116/ijme.4dfb.8dfd.

33. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J Chiropr Med 2016; **15**(2): 155-63. doi: 10.1016/j.jcm.2016.02.012.

Corresponding author: Olufemi O. Oyewole, PhD, Department of Physiotherapy, Olabisi Onabanjo University Teaching Hospital, PMB 2001, Sagamu, Nigeria
e-mail: oyewoleye@yahoo.co.uk, oyewoleye@gmail.com