Original Research

Validation and reliability of the Chichewa translation of the EQ-5D quality of life questionnaire in adults with orthopaedic injuries in Malawi

Linda Chokotho[1,2], Nyengo Mkandawire[3], Devin Conway[4], Hao-Hua Wu[4], David D. Shearer[4], Geir Hallan[5], Jan-Erik Gjertsen[2,5], Sven Young[5,6], Brian C. Lau[4]

1. Beit CURE International Hospital, Blantyre, Malawi
2. Institute of Clinical Sciences and Centre for International Health, University of Bergen, Bergen, Norway
3. Department of Orthopaedic Surgery, Queen Elizabeth Central Hospital, Blantyre, Malawi
4. Orthopaedic Trauma Institute, Institute for Global Orthopaedics and Traumatology, Department of Orthopaedic Surgery, University of California San Francisco, San Francisco, California, USA
5. Department of Orthopaedic Surgery, Haukeland University Hospital, Bergen, Norway
6. Department of Surgery, Kamuzu Central Hospital, Blantyre, Malawi

Correspondence: Dr Brian C. Lau (blau10@gmail.com)

Abstract

Background

The EQ-5D is a standardised instrument that measures health-related quality-of-life and explores cost-effectiveness of treatments. Malawi is a low-resource country that would benefit from assessment of quality-of-life among individuals living with chronic conditions. Chichewa is the official native language of Malawi. The Chichewa version of the EQ-5D-3L developed by EuroQoL group has not been validated with Chichewa speakers. The purpose of this study was to evaluate the clinimetric properties of the Chichewa EQ-5D-3L.

Methods

Patients with orthopaedic conditions were recruited in the outpatient orthopaedic clinics and wards at Queen Elizabeth Central Hospital, Blantyre, Malawi. Fifty-three patients with various musculoskeletal problems were administered the Chichewa EQ-5D-3L and World Health Organization quality of life (WHO-QOL) questionnaires. To assess repeatability, a separate test–retest population of 20 patients were also selected from orthopaedic clinics and wards to fill out the questionnaire twice.

Results

Convergence validity was determined, with each of the WHO-QOL domains and the EQ-5D descriptive index and visual analogue scale (VAS) having good to moderate correlation (r = 0.3–0.7). Internal consistency was measured for the descriptive index, and the Cronbach's alpha was 0.7. The ceiling effect for the descriptive index and the VAS were 9.4% and 0%, respectively. No respondents reached floor effect for the descriptive index or the VAS. The test–retest intraclass correlation coefficient reliability at 14 days was 0.984 for the VAS and 1 for the descriptive index, with all 20 respondents providing the same responses.

Conclusions

The EuroQoL translated version of the Chichewa EQ-5D-3L was found to demonstrate adequate validity, internal consistency, floor/ceiling effects, and reliability.

Introduction

Quality of life assessment with patient-reported outcome measures is a useful addition to measuring functional and clinical outcomes in evaluation of the benefits of healthcare interventions.[1] In orthopaedics, radiographic and clinical assessment (range of motion, strength, mobility) are often measured by the treating physician, which can be prone to bias. Adding patient-reported outcomes may limit the amount of physician bias in assessment of outcomes.

Moreover, "quality of life" is a broadly ranging concept, affected in a complex way by the person's physical health, psychological state, level of independence, social relationships, and their relationship to salient features of their culture.[2] As such, quality of life assessment may add another dimension to conventional outcome measures, as it may demonstrate the impact of injury on other areas of life from the patient's perspective.[1]

Quality of life assessment may also help in resource allocation by prioritising those interventions that result in better quality of life. Priority setting in healthcare is more important in limited resource settings, like Malawi, where demand exceeds supply. Malawi has a population of over 18 million and is one of the world's least-developed countries, with around 85% of the population living in rural areas.[3] The country has one of the lowest GNI per capita in the world,[3] at US$250. The majority of health services are offered by Ministry of Health facilities, where most public health services are free for patients.[4] The per capita government total expenditure on health is 11.4% of GDP.[4] Given Malawi's limited resources, government-provided healthcare, and high costs of healthcare, studies evaluating quality-adjusted life years (QALYs) using validated data collection instruments are imperative for optimal resource allocation and improvement in patient care.

The EQ-5D-3L is a standardised instrument used to measure health-related quality of life. It is used widely in English-speaking countries and has been previously evaluated, and its validity and reliability have been studied.[5,6] The EuroQol Group translated the English version of the EQ-5D-3L into Chichewa (Nyanja). However, to our knowledge the clinimetric properties of the Chichewa EQ-5D-3L have not been evaluated. Chichewa is the language of the Chewas, the biggest population group in Malawi, spoken by around two-thirds of the population, especially in the populous Central

and Southern regions.[7] It is an official national language in Malawi. Chichewa is also spoken in parts of Zambia and Mozambique.[7] The literacy rate for adults (aged ≥ 15 years) in Malawi is 73%, with 42% of the population literate in Chichewa only,[8] hence the need to use a Chichewa version of the EQ-5D when assessing quality of life. This article describes the validation of the EuroQoL Chichewa version of the EQ-5D-3L in Malawi.

Methods

Translation and adaptation

The translation and cultural adaptation of the English version of the EQ-5D-3L into Chichewa (Nyanja) was commissioned by the EuroQOL group in 2012. The EuroQol group is a network of international multi-disciplinary researchers with members from North America, Europe, Asia, Africa, Australia, and New Zealand.[5,9] It is responsible for the development of the EQ-5D, a preference-based measure of health status. The translation and cultural adaptation procedure is described elsewhere.[5,9] Permission to use the already translated questionnaire from the EuroQol group was obtained.

Instrument

The EQ-5D-3L is a 2-part questionnaire that assesses quality of life. The first part has 5 dimensions, namely: mobility; self-care; usual activities; pain/discomfort; and anxiety/ depression. Each dimension has 3 levels: no problems, some problems, and extreme problems, with scores of 1, 2, and 3 representing each level, respectively. The respondents are asked to choose 1 level for each of the 5 dimensions that best describes their own health state on the day of the interview. The second part is a visual analogue scale (VAS), wherein patients self-rate their health state on a scale of 0 to 100, with 0 and 100 as the worst and best imaginable health states, respectively.

Study setting and population

The Chichewa EQ-5D-3L was administered to an initial 53 consecutive patients with various musculoskeletal problems presenting at Queen Elizabeth Central Hospital's orthopaedic wards and outpatient clinics between October 2015 to March 2016. Queen Elizabeth Central Hospital, which is located in the city of Blantyre, is a tertiary-care facility and the main teaching hospital in the country. Orthopaedic outpatients clinics are done once a week, where patients with musculoskeletal problems from within Blantyre or referred from any of the 13 districts in the Southern Region of the country are treated. Admitted cases on the wards have a similar distribution pattern. A consent form written in Chichewa was given or read aloud to adult patients (≥ 18 years old) in both the clinic and inpatient setting. Patients were made aware of the risks and benefits of participating in the study and that participation was voluntary. Sample size determination was based on guidelines for the process of cross-cultural adaptation of self-report measures.[10] The Chichewa version of the validated World Health Organization Quality of Life (WHO-QOL)[11] questionnaire was also administered to test construct validity. In order to assess repeatability, a separate test–retest population of 20 additional consecutive patients from orthopaedic clinic took the questionnaire twice at an interval of 2 weeks apart. Written informed consent was obtained from all patients who took part in the study. Ethical approval was obtained from College of Medicine Research and Ethics Committee

(COMREC) and the University of California San Francisco Medical Center Institutional Review Board.

Statistical analysis

To aid in analysis, the EQ-5D was separated by its descriptive index and the VAS. The health profiles from the descriptive system were converted to index scores. The index scores were calculated using the index score calculator based on a Zimbabwean population–based time trade-off (TTO) value set,[5] as Malawi population data are not yet available. The possible index scores range from 0.145 to 1.0, where 0 is death and 1 is the best possible health state. The data from these separate subscales were then uploaded to IBM SPSS Statistics version 23 for analysis to determine internal consistency, floor and ceiling effects, and repeatability. Details of each analysis are given below.

Construct validity

Construct validity was utilised to determine that the Chichewa-translated EQ-5D measures quality of life similarly to a previously validated Chichewa translated measure of general health. To measure this, Pearson correlation (r) was calculated for the EQ-5D descriptive index scores and VAS scores with the WHO-QOL overall health, physical, psychological, social, and environment domain scores. As the r-value approaches 1, this indicates increasing convergence between the 2 measurement tools. By convention, strong, moderate, weak, and poor correlations were defined as > 0.70, 0.50 to 0.70, 0.30 to 0.50, and < 0.30, respectively.

Internal consistency

Internal consistency was utilised to determine the homogeneity of an individual subscale. Essentially, this value demonstrates that a group of questions is evaluating the same construct.[12] To measure this, Cronbach's alpha was calculated for the descriptive index using the initial population. Since EQ-VAS is only a single construct, rather than multiple questions, internal consistency cannot be measured. A Cronbach's alpha greater than 0.70 is accepted as being significant.[12]

Floor and ceiling effects

Floor and ceiling effects occur when a large percentage of survey respondents score the lowest or highest possible scores, respectively. When present, this causes a potential question about the survey's ability to capture extreme data, as well as difficulty in distinguishing among respondents who achieved maximum or minimum scores.[12] To determine if these effects were present, the percentage of patients who achieved the best scores and those who achieved the worst scores were determined for each subscale, using the initial population. Floor or ceiling effects were considered to be present if 15% of respondents or greater reported either the worse or best possible scores, respectively.[12]

Repeatability

To assess repeatability, questionnaires were analysed to determine their agreement—the extent to which scores from different time points resemble each other—and reliability, which measures how easily patients can be distinguished from each other on repeated testing.[12] For this section of analysis, the test–retest population was utilised. The EQ-VAS scores were assessed separately from the descriptive index. To assess agreement, the mean difference of the sums between time points was calculated, along with their respective 95% confidence intervals (CIs). Scores were considered to be statistically similar if the CI contained 0. Reliability was

Table 1: Patient demographics

	Initial sample (N = 53)	Test-retest sample (N = 20)
Mean age (standard deviation)	36.5 (14.6)	43.4 (17.2)
Gender, n (%)		
Male	28 (52.8)	18 (90)
Female	25 (47.2)	2 (10)
Education level, n (%)		
Did not attend	0 (0)	2 (10)
Primary	13 (24.5)	9 (45)
Secondary	9 (17)	6 (30)
Tertiary	4 (7.5)	3 (15)
Unknown	27 (50.9)	0 (0)
Injury, n (%)		
Femur fracture	11 (20.8)	7 (35)
Tibia/fibula fracture	12 (22.6)	5 (25)
Radius/ulna fracture	6 (11.3)	1 (5)
Back pain	5 (9.4)	0
Ankle fracture	4 (7.4)	0
Joint dislocation	5 (9.4)	0
Other	10 (18.9)	7 (35)

evaluated by determining the Pearson's correlation coefficient (r) and the intraclass correlation coefficient (ICC) between sums (of index scores) at the 2 time points. The ICCs were determined using the 2-way random effects model with agreement type, along with their corresponding 95% CI. A significant correlation was demonstrated by an ICC value of 0.70 or higher.[12]

Results

All surveys were completed by all respondents. All questions of the EQ-5D and WHO-QOL were answered by all patients.

Demographics

Demographic characteristics of the initial cohort and the reliability cohort are presented in Table 1. The mean age for the initial population was 36.5 years. There were more males (52.8%) than females in the initial population; more than half had unknown education, indicating that they had attended but not completed primary school.

Validity

The Pearson's correlation between the descriptive index and VAS of the EQ-5D and the WHO-QOL overall health, physical, psychological, social, and environment domains demonstrated a good to moderate correlation (Table 2).

Internal consistency

The Cronbach's alpha for the Descriptive Index of the EQ-5D was 0.70 (Table 3).

Floor and ceiling effects

The descriptive index of the EQ-5D had 9.4% (5/53) respondents who scored the best possible functioning score. No patient reported the best possible score in the VAS. No respondents scored the worst functioning score on any of the indices (Table 3).

Repeatability

All data for repeatability are listed in Table 3. For the EQ-VAS, the mean difference in scores was 1.0 (range −0.44 to 2.44). The Pearson's coefficient was 0.986 and the ICC was 0.984 (CI = 0.961 to 0.994). All twenty respondents in the test–retest population gave the same responses for the descriptive index of the EQ-5D at both baseline and follow-up. Thus, the mean difference between the 2 time points was 0, and the Pearson's coefficient and ICC were both 1.

Discussion

This is the first study to evaluate the clinimetric properties of the translated Chichewa version of the EQ-5D. This study represents only the second time the EQ-5D has been validated in an African language. Jelsma and Chivaura translated and validated the EQ-5D in Shona, an important language in Zimbabwe, and found that Zimbabweans valued health states differently from previous European study populations.[13] The results of this study have demonstrated that the translated Chichewa questionnaire is a reliable and valid tool that can be used to assess quality of life in Chichewa-speaking patients with musculoskeletal problems.

The questionnaire was tested across all education levels, from those with limited formal education to those who completed college. The fact that despite the wide range of education, all respondents were able to answer all the questions highlights the acceptability and comprehensibility of the questionnaire.

The descriptive index and VAS from the EQ-5D-3L had a good to moderate correlation with the previously validated WHO-QOL overall domain. This demonstrates that the Chichewa version of the EQ-5D-3L appropriately measures quality of life.

Table 2: Construct validity of the EQ-5D quality of life questionnaire with the WHOQoL-BREF questionnaire (Pearson Correlation between EQ5D with WHOQoL-BREF)

EQ-5D	WHOQoL- BREF: Domains				
	General Health	Physical	Psychological	Social	Environment
Descriptive Index (R)	0.57 (P < 0.01)	0.42 (P < 0.01)	0.46 (P = 0.023)	0.43 (P = 0.035)	0.63 (P = 0.019)
EQ-VAS (R)	0.49 (P = 0.026)	0.31 (P = 0.183)	0.39 (P = 0.042)	0.34 (P = 0.145)	0.52 (P = 0.022)

Table 3: Questionnaire internal consistency, floor and ceiling effect, and repeatability

Subscales	EQ-5D	
	Descriptive index	**Visual analog scale (VAS)**
Internal consistency		
Cronbach's alpha	0.70	-
Floor and ceiling effect, n (%)		
Floor effect	0/53 (0)	0/53 (0)
Ceiling effect	5/53 (9.4)	0/53 (0)
Repeatability*		
Baseline average (SD)	0.307 (0.23)	44 (17.6)
Follow-up mean (SD)	0.307 (0.23)	45 (18.2)
Mean difference (95% CI)	0	1.0 (–0.44 to 2.44)
Pearson's coefficient ®	1	0.980
ICC (95% CI)	1	0.984 (0.961 to 0.994)

Table 3 shows the results of statistical analysis of each subscale in the two questionnaires, consisting of internal consistency, floor/ceiling effects, and repeatability.

*For repeatability, the units for baseline and follow-up averages are different for each subscale. The Descriptive index is recorded as the average of all patients' converted health index scores. The VAS is reported as the average of the individual VAS scores.

SD = standard deviation; CI = confidence interval; ICC = interclass correlation coefficient

The internal consistency had a Cronbach's alpha value of 0.70, demonstrating that the questions in the translated questionnaire were indeed measuring quality of life. The descriptive index had a ceiling effect of 9.4%, which was well within the acceptable range of 15%. One explanation for the ceiling effects can be the possibility that there was a proportion of patients who had successful treatment and were just coming to the clinic for regular follow-up accordingly. These patients may not have had any limitations in function and therefore may have scored best possible scores. Additionally, no respondents had floor effects in the descriptive index, or either floor or ceiling effects with VAS. These findings are similar to other translated versions of the EQ-5D.[14–19]

Test–retest reliability was examined in this study with a time interval of 2 weeks in between responses. There is no recommended time interval for questionnaire administration for test–retest reliability for quality of life questionnaires. Marx et al.[20] found that there was no statistically significant differences in test–retest reliability of health status instruments between the time intervals of 2 days or 2 weeks among orthopaedic patients with knee disorders. Nevertheless there is need to have adequate time in between interviews to minimise the possibility of recall bias and the time interval should not be too long to allow for change in disease status. The responses between the 2 time intervals in our study were likely to be stable, as the 2-week period was sufficient to minimise recall bias and also not long enough to have significant change in disease status.

All answers on the descriptive index of the EQ-5D were the same over the 2-week interval. The VAS score also demonstrated excellent repeatability with a Pearson's coefficient of 0.986 and ICC of 0.984. The excellent repeatability highlights that questions were interpreted similarly by readers at 2 different time points.

One of the limitations of this study was that the study population included only patients with orthopaedic injuries. This limits its generalisability to other medical conditions. Another limitation is that we did not record the economic status of respondents, as it was difficult to estimate monthly income for the majority of patients with informal employment or small-scale businesses. This information may offer insight to the responses and comprehension of the EQ-5D questionnaire. Education level (which was recorded), however, may serve as a proxy for economic status. In the test–retest population, there was a disproportionate amount of males (18), compared to females (2), which may affect its generalisability. The index scores used in this study were for the Zimbabwean population because there are no index scores for Malawi. Although Zimbabwe is a low-income country in sub-Saharan Africa, like Malawi, it has one of the highest literacy rates in Africa, at 86.5%, compared to Malawi at 65.8%, according to UNESCO. The level of understanding therefore may be different and as such the scores may not be truly representative of the Malawian population. Future studies should aim to develop index scores for Malawi.

Notwithstanding these limitations, this study provides evidence that the translated Chichewa version of the EQ-5D is valid and reliable for future use within Malawi to measure quality of life in patients with musculoskeletal problems. It is hoped that subsequent evaluations of health states, treatment interventions, and wider public policy interventions will benefit from its use.

Acknowledgements

Competing interests

All authors declare that they have no competing interests related to this work.

References

1. Bowling A, (1998) Measuring health: A review of Quality of Life Measurement scales. 2nd ed. Medicine, Health Care and Philosophy, 1(2): p.181-182.

2. World Health Organisation, (1997) WHOQOL Measuring Quality of Life. Geneva: WHO (WHO/MSA/MNH/PSF/97.4)

3. Malawi. The World Bank 2016. Retrieved 6 December 2016. http://www.worldbank.org/en/country/malawi

4. The World Health Organization- Global Health Observatory- Malawi 2002-present. Retrieved 6 December 2016. http://apps.who.int/gho/data/node.country.country-MWI

5. EuroQol Group, (1990). EuroQol-a New facility for the measurement of health related quality of life. Health Policy 16(3): p. 199-208

6. Brooks R, (1996) EuroQol the Current state of play. Health Policy, 37(1): p. 53-72 http://dx.doi.org/10.1016/0168-8510(96)00822-6

7. Kayambazinthu E, (1998) The Language Planning Situation in Malawi. J Multiling Multicul Dev, 19(5): p. 369-439. http://dx.doi.org/10.1080/01434639808666363

8. Central Intelligence Agency- The World Factbook- Malawi. Retrieved 6 December 2016. https://www.cia.gov/library/publications/the-world-factbook/geos/mi.html

9. EuroQOL: Herdman M, Fox-Rushby J, Rabin R, Badia X, Selia C. (2003) Producing other language versions of the EQ-5D. In: Brooks R, Rabin R, de Charro F (eds). The measurement and valuation of health status using EQ-5D: A European perspective. Kluwer Academic Publishers.

10. Beaton DE, Bombadier C, Guillemin F, et.al, (2000) Guidelines for the process of cross-cultural adaptation of self-report measures. SPINE 25(24): p. 3186-3191. http://dx.doi.org/10.1097/00007632-200012150-00014

11. Colbourn T, Masache G, Skordis-Worrall J, (2012) Development, reliability, and validity of the Chichewa WHOQOL-BREF in adults in Lilongwe, Malawi. BMC Res Notes. Jul 3(5): p. 346. http://dx.doi.org/10.1186/1756-0500-5-346

12. Terwee CB, Bot SD, de Boer MR, et al, (2007) Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol, 60(1): p. 34-42. http://dx.doi.org/10.1016/j.jclinepi.2006.03.012

13. Jelsma J & Chivaura V. Report on the process of translating the EQ-4D into Shona. Harare, EuroQol Group, Rotterdam; 2000.

14. Brazier J, Deverill M, Green C, et al. (1999) A review of the use of health status measures in economic evaluation. Health Technol Assess, 3(9): i-iv, p. 1-164.

15. Coons SJ, Rao S, Keininger DL, Hays RD. (2000) A comparative review of generic quality-of-life instruments. Pharmacoeconomics, 17(1): 13-35. http://dx.doi.org/10.2165/00019053-200017010-00002

16. Ferreira LN, Ferreira PL, Ribeiro FP, Pereira LN, (2016) Comparing the performance of the EQ-5D-3L and the EQ-5D-5L in young Portuguese Adults. Health Qual Life Outcomes. Jun 9:14(1): p. 89. http://dx.doi.org/10.1186/s12955-016-0491-x

17. Rogers KS, Pilling M, Davies Linda, Belk Rachel, (2016) Translation, validity, and reliability of the British Sign Language (BSL) version of the EQ-5D-5L. Qual Life Res, 25: p. 1825-1834. http://dx.doi.org/10.1007/s11136-016-1235-4

18. Triapthy S, Hansda U, Seth N, Rath S, Rao PB, Mishra TS, Subba SH, Das R, Nayak S, Kar N, (2015) Validation of the EuroQol Five Dimensions- Three-Level Quality of Life Instrument in a Classical Indian Language (Odia) and Its use to Assess Quality of Life and Health Status of Cancer Patients in Eastern India. Indian J Palliat Care. Sep-Dec: 21(3): p. 282-8. http://dx.doi.org/10.4103/0973-1075.164896

19. Yfantopoulos JN, Chantzaras AE, (2016) Validation and comparison of the psychometric properties of the EQ-5D-3L and EQ-5D-5L instruments in Greece. Eur J Health Econ. 2016 Jun 4 [Epub ahead of print]. http://dx.doi.org/10.1007/s10198-016-0807-0

20. Marx RG, Menezes A, Horovitz L, Jones EC, Warren RF. A comparison of two time intervals for test-retest reliability of health status instruments. Journal of clinical epidemiology. 2003 Aug 31;56(8):730-5.