Abstract

Background: Quality of life is defined as the subjective perception of one’s own well-being within a sociocultural context. SF-36 is commonly used to check the health status of the general population in many countries.

Aims: To validate the Persian (Dari) version of the Short Form Health Survey (SF-36) in Herat.

Methods: The Persian (Dari) version of SF-36 was culturally adapted. A total of 1259 healthy individuals aged ≥ 18 years participated in the study between November 2016 and April 2017. Construct validity of SF-36 was evaluated through exploratory factor analysis. The extraction was performed by principal component analysis from the polychoric correlation matrix estimated by a 2-step method with varimax rotation. Internal consistency and test–retest reliability were assessed by Cronbach’s α SPSS version 23.0 and R version 3.1.3 were used in the analysis.

Results: The last version of SF-36 including 27 items in 8 factors explained 86.48% of the
variance. The Kaiser–Meyer–Olkin value was 0.391 and Bartlett’s test showed statistical significance (P

Conclusions: We showed that our Dari version of SF-36 is suitable to evaluate quality of life in adults in Herat City. This scale will be useful for health researchers in the future.

Keywords: Afghanistan, Dari, quality of life, SF-36, validity

Introduction

Quality of life (QOL) is either defined as the subjective perception of one’s own well-being within a sociocultural context, or as the satisfaction of a person’s desires and achieving an ideal level of well-being. Health-related QOL (HRQOL) refers to functioning and wellbeing in physical, mental and social dimensions of life (1–3).

Measuring the health status of individuals or a population is an important first step in the assessment of healthcare needs and evaluation of the impact of health interventions. The indicators of QOL are used to measure the effects of symptoms and treatment on people’s health status (4–8). Measuring QOL in the community is also helpful in health planning (9). Reliable population-wide statistics on health status are scarce, especially in developing countries such as Afghanistan.
The 36-item Short Form Health Survey (SF-36) is a renowned generic health-related QOL questionnaire that was developed by Boston Health Research Institute. The SF-36 instrument provides a concise method that is mainly used to check the health status of members of the general population as well as that of patients aged ≥ 14 years (1). As it is easy to administer, it has become one of the most widely used QOL evaluation tools in the world (6,8–12). SF-36 is also widely used among different age groups in Middle Eastern countries (13). Psychometric analyses of the translated versions provided evidence that SF-36 is a reliable and valid measure in different populations (14–19), and a few studies have translated and culturally adapted SF-36 into different Asian languages (20–23).

There is increasing demand by healthcare professionals and researchers to have and include an authenticated QOL questionnaire in their research or clinical investigations in Afghanistan. This study reported the development and validation of a Persian (Dari)/Afghan version of SF-36 and the results of its testing in the general population in Herat; a province in Southwest Afghanistan where most of the population communicates in Persian.

Methods

SF-36

SF-36 measures 8 health-related subscales: Physical functioning (PF, 10 items); role limitation due to physical health (RLPH, 4 items); bodily pain (BP, 2 items); general health perception (GH, 5 items); vitality (VT, 4 items); social functioning (SF, 2 items); role limitation due to emotional problems (RLEP, 3 items); and perceived mental health (MH, 5 items). A single item that provides an indication of perceived change in general health status over a 1-year period (health transition) is also included. SF-36 is also evaluated by physical health component (PHC) and mental health component (MHC). While PF, RLPH, BP and GH are considered under the dimension of PHC, VT, SF, RLEP and MH are under the dimension of MHC (1).

Translation of SF-36

Translation of the Persian (Dari) version of SF-36 involved 3 steps. The standard forward–backward translation method was followed to translate the SF-36 questionnaire. The SF-36 questionnaire was translated into Persian (Dari) by 2 independent health professionals. Persian literature lecturers reviewed the translated questionnaire to assess Afghan culture and language appropriateness. This questionnaire was back-translated into English by 2 professional translators to check for differences between the Persian and original English versions. A few changes were made and the provisional version of the questionnaire was provided on the basis of the results of the cultural adaptation. There were no difficulties with translation of questions and response categories. Changes were only made in a few items regarding activities, namely: playing golf had been changed to light sporting activities such as snooker and billiards; and measurements like mile was changed to kilometre, and block to approximately 1 or a few hundred meters, which are more familiar in Persian language and
Afghan society. The provisional forward-translated questionnaire was pilot tested and administered to 50 healthy individuals, and almost all stated that did not have any difficulties completing the questionnaire. The results of pilot testing were analysed by the research team and after review, only a few changes were made, and the final version was used in this study.

Data collection

This study was conducted in Herat City, Afghanistan from November 2016 to April 2017. The Persian version of SF-36 was administered to a random sample of healthy individuals aged ≥ 18 years living in Herat City from November 2016 to April 2017. Our final questionnaire contained 2 main parts: the first included questions about sociodemographic characteristics, economic status, and cultural and daily activities; and the second had only questions in SF-36. Since SF-36 had 36 questions, we considered a minimum 15 individuals were needed for each question, so our minimum sample size was 540 for the test and ~200 for the retest. The inclusion criteria were: age ≥ 18 years; residents of Herat City; knowledge of Persian (Dari) language; and no psychological disorder that could affect the answers.

The field coordinator and 15 interviewers were trained in the aim of the research, the questionnaire and data collection methods by an Afghan researcher. Herat City has 15 districts, and 1 hospital in each district was selected for the study. The selected hospitals were the most used by people in the district. The interviews were done face to face with the patients and their relatives at 15 general hospitals in 15 different districts of Herat City.

In order to evaluate the validity of SF-36, which is a Likert-type scale, we needed 15 participants for each question with a target total of 540 participants who attended the selected district hospitals. Data collection was continued until at least 90 people were reached in each age group from November 2016 to April 2017. There were 1259 interviewees (54% female and 46% male) and 259 (52.1% male and 47.9% female) were contacted for the retest study.

Statistical analysis

SPSS version 23.0 (SPSS, Chicago, IL, USA) was used to calculate the descriptive analyses such as frequencies, percentage distributions, means, and standard deviation. P

Exploratory factor analysis was used to assess the construct validity by using free software R version 3.1.3 (packages: psych, polycor, GPArotation, nFactors, corpcor, ICS, R.utils) and SPSS. Polychoric correlation coefficients were used for explanatory factor analysis because SF-36 is a Likert-type scale and the individual items are not interval scaled (26).
Principal component analysis (PCA) from the polychoric correlation matrix estimated by a 2-step method with a varimax rotation was applied for data extraction. It was accepted that factor loading > 0.32 was considered statistically significant (1). Six items were dropped from SF-36 because factor loading of these items was

PCA was used to reduce the two dimensions of the scale (PHC and MHC). The summary scores of PHC and MCH were obtained by PCA. Spearman’s rho correlation coefficient was used for correlation between the 8 subscales (PF, RLP, BP, GH, VT, SF, RLEP and MH) and the 2 rotated components (PHC and MCH). Reliability was estimated using the test-retest correlation coefficient and the internal consistency (Cronbach’s α). To assess test–retest reliability, 259 people completed the questionnaire twice at 3-week intervals. Spearman’s rho correlation coefficient (> 0.6 as reliable) was used for the test–retest reliability (23). Internal consistency was calculated by using Cronbach’s α, which was considered acceptable at ≥ 0.70 (27,28). Spearman’s rho correlation coefficient was also used to evaluate the relationship between 2 subscales of SF-36.

Ethical approval

All procedures involving human participants were performed in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments, or comparable ethical standards. Permission from the RAND Corporation was obtained via email for using SF-36 for this research. We obtained the approval of the Ethics Committee and Research and Development Committee of Herat University Faculty of Medicine. Informed consent was obtained from all individual participants included in the study.

Results

In this study 56% of participants were from urban area, 62.4% of the participants were Tajik, 46% of the participants were male, 65.5% were married, and 39.6% were illiterate (Table 1). More than half of the participants (59.2%) did not have any income and 23.6% of them received a regular income under US$100. Almost 80% of participants were probably living under or just above the poverty line.

Factor analysis with the polychoric correlation and varimax rotation was performed to construct 8 subscales (Table 2). Items 1, 2, 26, 29, 30, 31 and 35 (in MH, GH and VT) were loaded on the different subscales. Two items (8 and 12) had low standardized estimates. Therefore, the 9 items were dropped from the original scale. After that, explanatory and confirmatory factor
analysis were reanalysed with the remaining 27 items. Total 8 subscales were obtained according to explanatory factor analysis factors and these subscales explained 86.48% of the variance. KMO value was 0.391 and Bartlett test was statistically significant (P < 0.05).

Reliability of the scale was evaluated using the test–retest and internal consistency reliability methods (Table 3). Test–retest reliability coefficients ranged from 0.682 to 0.905. All subscales of SF-36 had good internal consistency and test–retest reliability. There were positive correlation coefficients between each of the 8 subscales of the SF-36 (P < 0.05).

Table 4 shows the PCA with varimax rotation, correlation between the 8 scales and the 2 rotated components. KMO was 0.88 and Bartlett’s spherical check was $\chi^2 = 6494.45$ and $p < 0.05$.

Discussion

Our findings suggest that the SF-36 has acceptable reliability, test–retest correlations and factor validity (polychoric correlation). The aim of this research was to study the validity and reliability of the SF-36 Persian-Dari version in Herat City, Afghanistan. This was the first study to assess the SF-36 in an Afghani sample.

Regarding validity, polychoric correlation and varimax rotation were performed to construct 8 subscales. All items in the PF, RLPH, SF, RLEP and BP were loaded on their own subscales. The several items in the MH, GH and VT subscales were loaded on the different subscales. The items that measure people’s feelings and energy can vary by understanding of social and psychological health in the different cultural communities. Two items in the PF dimension, bending/kneeling and bathing/dressing, were dropped because of their low value of standardized estimates in CFA. The last version of SF 36 which had 27 items was analysed again for explanatory factor analysis. The reason for deleting the items was cultural differences, as in other studies (16,18,20,22,23).

Regarding the internal consistency and test–retest results, Cronbach’s $\alpha$ in our study was 0.753–0.933, which is higher than 0.70 as an admitted level for internal consistency, and test–retest results showed a good level of 0.68–0.90. Our results were similar to those of previous studies.
Emotional well-being should be under the mental health dimension of SF, but it was correlated with the physical dimension in our study. However, the other subscales generally fitted into the original hypothesized association. Li et al. showed that VT and RLEP were loaded to the physical health subscale instead of emotional scales in the explanatory factor analyses (23). Fukuhara et al. (22) found that VT and RLEP were obtained under the physical health dimension and Montazari et al. (21) also found VT under physical health. Thumboo et al. studied English and Chinese versions of SF-36 in a multiethnic urban population in Singapore and found small differences between the 2 versions (24). This result supports the suggestion that emotional dimensions may be affected by the perception of people from different cultural backgrounds.

The strength of the study was that the provisional version of the questionnaire was provided after careful review and cultural adaptation. In general, there were no difficulties in translating response categories, except that bowling and playing golf were changed to light sporting activities, mile was been changed to kilometre, and walking or several blocks was changed to walking one or several alleys to refer to a similar distance in the Persian language. The provisional forward translated questionnaire was pilot tested and administered to a sample of 50 healthy individuals in all age groups within the inclusion criteria.

The study had some limitations. The Persian version of SF-36 was administered to a random sample of healthy individuals aged ≥ 18 years living in Herat City; therefore, this study only represented people living in Herat and speaking Dari, not for all regions in Afghanistan. SF-36 needs to be translated into Pashto and the other languages spoken in the country, and validity and reliability studies need to be conducted in different regions of the country. There is not any other QOL scale in Persian in the country; therefore, the validity test could not be conducted using direct statistical comparison, such as invariance analysis and differential item functioning. After developing new scales to measure QOL in Afghanistan, the validity test can be repeated to conduct comparisons in the different regions of the country.

In conclusion, the study has shown that SF-36 in Dari language is suitable to evaluate QOL of adults in Herat Province, Afghanistan.

**Acknowledgement**

We would like to acknowledge Ms J. Fayzi, for coordinating data collection in the field.
**Funding:** None.

**Competing interests:** None declared.

**Enquête de santé abrégée (SF-36) : traduction et étude de validation en Afghanistan**

**Résumé**

**Contexte:** La qualité de vie est définie comme la perception subjective du propre bien-être d’une personne dans un contexte socioculturel. Le questionnaire SF-36 (Enquête de santé abrégée) est couramment utilisé pour vérifier l’état de santé de la population générale dans de nombreux pays.

**Objectifs:** Valider la version en langue persane (dari) du SF-36 à Herat.

**Méthodes:** La version en langue persane (dari) du SF-36 a été adaptée en tenant compte de la culture. Au total, 1259 personnes en bonne santé âgées de 18 ans ou plus ont participé à l’étude entre novembre 2016 et avril 2017. La validité structurelle du SF-36 a été évaluée à l’aide de l’analyse factorielle exploratoire. L’extraction a été effectuée par analyse en composantes principales à partir de la matrice de corrélation polychorique estimée par une méthode en deux étapes avec rotation varimax. La cohérence interne et la fidélité test-retest ont été évaluées par la méthode de l’alpha de Cronbach. Les logiciels SPSS (version 23.0) et R (version 3.1.3) ont été utilisés pour l’analyse.

**Résultats:** La dernière version du SF-36 comprenant 27 items assortis à huit facteurs expliquait 86,48 % des variations. La valeur de Kaiser-Meyer-Olkin était de 0,391 et le test de Bartlett a montré une signification statistique (p < 0,001). L’analyse factorielle confirmatoire a révélé que le modèle final présentait de bonnes statistiques d’ajustement (p