TEST-RETEST RELIABILITY OF THE 36-ITEM HEALTH SURVEY (SF-36) AS QUALITY OF LIFE MEASURES IN ELDERLY CROATIAN POPULATION

Biljana Kuzmanić, Nina Ivančić and Jelena Paušić

Faculty of Kinesiology, University of Split, Croatia

Original scientific paper

Abstract

The aim of this research was to examine the test–retest reliability of the Short Form-36 Health Survey Questionnaire (SF-36). The SF-36 Questionnaire measures health on eight multi-item dimensions, covering functional status, well-being and overall evaluation of health. The research included 69 female and male participants aged 65 to 90. All the participants in the research were interviewed twice, and the retest was done after seven days. The values of Cronbach's α were all over 0.6. The smallest Cronbach's α was estimated in three dimensions Mental health, Vitality and General health perception (less than 0.7). For the test-retest of the results from the first and the second measurement, the non-parametric Friedman's two-way ANOVA test has shown that distributions of the first and the second measurement were not significantly different in six dimensions. In only two dimensions the distributions were significant different at p-level of 0.039 / 0.036 (Vitality / General health perception). The Croatian version of the SF-36 is a good tool to evaluate HRQOL for persons 65 and older (AS 78± years, SD 5.645).

Key words: *functional status, well being, general health, HRQOL, Croatia.*

Introduction

Aging represents one of the major social, economic and health challenges of the 21st century, especially in Europe, the continent with the largest share of persons over 65 in the overall population (15%) (Croatian Health Service Yearbook, 2015). Term "Quality of life", often associated with health means that life satisfaction is related to physical, mental health, social status, and social belonging. The importance of projects for social inclusion of older people and the improvement of the quality of life is largely accompanied by the media to make the public aware of their dignified aging. Given the increasing importance of older people as service users and potential beneficiaries of public health interventions, quality of life measures should be adequately assessed and, if necessary, adapted for the use in this population (Walters, Munro & Brazier, 2001). The basic aim of health policy is not the only extension of life expectancy, but also an endeavor to improve the quality of life (Croatian Health Service Yearbook 2015).

It is important to be able to measure the perception of the health of the population to assess the benefit of health care interventions and to target services (Brazier et al., 1992). Subjective well-being effects have an impact on health of individuals. Researchers have constructed different measures to assess the health of people measuring subjective feelings of health status, psychological well-being and life satisfaction. Scoring standardized responses to standardized questions is an efficient way to measure health status (Ware & Sherbourne, 1992). The SF-36 was designed for the use in clinical practice and research, health policy evaluations, and general population surveys (Ware & Sherbourne, 1992). The survey was constructed for self-administration by persons aged 14 years and older, and for administration by a trained interviewer, in person or by telephone (Ware & Sherbourne, 1992). The SF-36 is suitable for the use with an elderly population when used in an interview setting (Lyons, Perry & Littlepage, 1994). The SF-36 is a model widely used as a generic short-form measure of functional health and well-being of different population groups, and it has been applied in hundreds of studies (Jureša et al., 2000). It is a subject of more than 4,000 published studies, with thousands of studies including the elderly, and hundreds of studies limited to the elderly (Gandek, Sinclair, Kosinski & Ware Jr, 2004 according to Turner-Bowker et al., 2002). SF-36 survey was for the first time implemented in the Croatian general population in the study of Jureša et al. (2000) named The Croatian Health Survey – SF-36: General Quality of Life Assessment. However, the substantially revised Short Form-36 Health Survey Questionnaire (SF-36) has yet to be independently validated in Croatia in the elderly population. The aim of this research was to examine the test–retest reliability of Croatian version of the Short Form-36 Health Survey Questionnaire (SF-36) in the elderly population.

Methods

The research included 69 female and male participants aged 65 to 90 from Split, Croatia without any significant visual or auditory damages, neurological diseases and barely used any medicines which influence sensory-motor functions. The SF-36 questionnaire is a self-administered questionnaire containing 36 items which take about five-ten minutes to complete. It measures health on eight multi-item dimensions, covering functional status, well-being, and overall evaluation of health. The SF-36 survey was for the first time implemented on the Croatian general population.

A Croatian version of the SF-36 questionnaire was licensed to Andrija Štampar School of Public Health in 1992 as a part of the "Tipping the Balance Towards Primary Healthcare Network" project (Buttanshaw, 1997). Two professional translators with experience in "health and quality of life terminology" but not in the SF-36, produced two independent forward translations and, after multi-professional discussions, agreed upon a common version (Jureša et al., 2000).

Table 1 Dimensions of the SF-36 health survey questionnaire.

Area	Dimension	Number of questions
Functional status	Physical functioning	10
	Social functioning	2
	Role limitations (physical problems)	4
	Role limitations (emotional problems)	3
Well being	Mental health	5
	Vitality	4
	Pain	2
Overall evaluation of health	General health perception	5
	Health change*	1
Total		36

Health change* is not included in eight dimensions nor it is scored.

The SF-36 questionnaire contains multi-item scales measuring eight generic health concepts (Table 1): physical functioning (PF), role limitations due to physical health problems (RP), bodily pain (BP), general health perceptions (GH), vitality (VT), social functioning (SF), role limitations due to emotional problems (RE), and mental health (MH) (Ware & Sherbourne, 1992; Ware, Snow, Kosinski & Gandek, 1993). A range of the questionnaire of all dimensions was minimum 0 and the maximum 100 points. The higher scores indicate a better quality of life which is vice versa from the lower scores. To evaluate health related quality of life (HRQOL) in elderly population, all the participants in the study were interviewed using the Croatian version of the SF-36 questionnaire. The participants were asked about their functional status, well-being and overall evaluation of health during a four-week period. The SF-36 questionnaire was filled twice, and the retest was done after seven days. The data was analyzed in IBM SPSS Statistics 23 software. The SF-36 measures eight aspects of health status that were calculated for all subjects. Descriptive statistics were calculated. Assessment of the normality of data was conducted with the Kolmogorov-Smirnov Test. The internal consistency the extent to which items correlate with items within their dimension was examined regarding Cronbach's α and the values above 0.8 are usually regarded as acceptable (Streiner & Norman, 1989). The non-parametric Friedman's two-way ANOVA test was used for estimation of distribution similarity of the first (test) and the second (retest) measurement of dimension. Distributions are considered equal when signification level (p) is higher than 0.05.

Results and discussion

Table 2 Descriptive statistics of the SF-36 health survey questionnaire dimensions.

Dimension	Mean	SD	Skewness	Kurtosis	K-S test (max d)	K-S test (p)
Social functioning 1	5.9420	.87252	.662	1.787	.285	.000
Social functioning 2	5.9855	.77636	.025	.798	.290	.000
Mental health 1	20.5652	2.32302	-.285	.104	.107	.049
Mental health 2	21.1159	2.09015	-.218	-.007	.121	.014
Vitality 1	15.2174	2.08545	-.759	.866	.140	.002
Vitality 2	15.6667	2.01951	-1.006	2.118	.197	.000
Pain 1	5.6812	2.25240	-.130	-.895	.199	.000
Pain 2	5.5942	2.03151	-.173	-.929	.161	.000
General health perception 1	15.6667	2.13284	.442	-.048	.134	.004
General health perception 2	16.2754	2.05715	.191	-.614	.116	.022
Physical functioning 1	20.5362	4.29973	-.040	-.619	.093	.200*
Physical functioning 2	20.6957	3.97531	-.122	-.107	.078	.200*
Role limitations (emotional problems) 1	4.5942	1.33195	-.132	-1.781	.260	.000
Role limitations (emotional problems) 2	4.6957	1.30952	-.260	-1.702	.275	.000
Role limitations (physical problems) 1	5.7246	1.68806	.260	-1.647	.252	.000
Role limitations (physical problems) 2	5.9855	1.56706	.095	-1.570	.214	.000

The results of the K-S test (Table 2) showed that just one dimension (Physical functioning) from the SF-36 questionnaire is distributed normally. For that reason, testing of variability of the results from first and second measurement was conducted with non-parametric Friedman's two-way ANOVA test. According to Scott, Tobias, Sarfati & Haslett (1999) in the distribution of item responses for the combined sample all levels of each item were observed, but item distributions were skewed, with

more respondents endorsing the higher response choices (reflecting better self-reported health). Velanovich (2007) emphasizes that the SF-36 data did not follow a normal distribution in any of the domains and needed to be statistically analyzed using nonparametric techniques. The coefficients of skewness are all negative, and the coefficients of kurtosis are less than 3 (the value for a normal distribution) in 7 of 8 domains (Velanovich, 2007).

Table 3 Internal consistency and reliability (test-retest) of the SF-36 health survey questionnaire dimensions.

Dimension	Number of questions	Cronbach's α Test/Retest	Two way ANOVA Friedman's (p)
Physical functioning	10	0.762/0.769	0.796
Social functioning	2	0.789/0.843	0.631
Role limitations (physical problems)	4	0.853/0.795	0.505
Role limitations (emotional problems)	3	0.865/0.857	0.336
Mental health	5	0.627/0.633	0.216
Vitality	4	0.653/0.645	0.039*
Pain	2	0.892/0.812	0.884
General health perception	5	0.636/0.632	0.036*
Health change*	1	-	0.847
	36		

The internal consistency the extent to which items correlate with items within their own dimension was estimated with Cronbach's α (Table 3). The values of Cronbach's α were all over 0.6 and only three dimensions have Cronbach's α value higher than 0.8 (Role limitations (physical problems), Role limitations (emotional problems) and Pain). Two dimensions (Physical functioning and Social functioning) have Cronbach's α value very near to 0.8 value. The smallest Cronbach's α was estimated in three dimensions Mental health, Vitality, and General health perception (less than 0.7). Cronbach's alpha was interpreted according to scale of $\alpha \geq 0.9$ – Excellent, $0.7 \leq \alpha > 0.9$ – Good, $0.6 \leq \alpha > 0.7$ – Acceptable, $0.5 \leq \alpha > 0.6$, and $\alpha < 0.5$ – Unacceptable (Kline, 2000). According to these interpretations of Cronbach's α reliability coefficients, three dimensions have acceptable reliability, and other five dimensions have good reliability. For the testing homogeneity of the test-retest results, the non-parametric Friedman's two-way ANOVA test has applied to test the distributions of the first and the second measurement in all dimensions. In only two dimensions the distributions were significant different at p-level of 0.039 / 0.036 (Vitality/ General health perception). All the other dimensions have statistically equal results. These two statistical methods and tests (Cronbach's α and Friedman's two-way ANOVA) were conducted to test stability and equivalency of SF-36 questionnaire items over time (7 days). Considering that the average of participants was 78± year's influence on two dimensions homogeneity in Friedman's two-way ANOVA test or influence on acceptable and good reliability can occur due to several reasons. They are susceptible to the side effects of medicines which tend to stay in person's body much longer and prolonging effects of medicines use. Also, the questionnaire was implemented in the winter period, when elderly are liable to remain physically inactive at home, which also affects their anxiety. A major problem with the test-retest, however, is its tendency to underestimate reliability if a true change occurs (Streiner & Norman, 1989).

Conclusion

The SF-36 Questionnaire is used for assessment of health-related quality of life in different population groups. The Croatian version of the SF-36 is a good tool to evaluate HRQOL for persons 65 and older (AS 78± years, SD 5.645) due to the determined acceptable and good reliability of all dimensions. Starting from these study results where the research was conducted on a rather smaller sample of subjects, a new and a wider research should be implemented on a greater sample of elderly subjects in the same measurement setting as well as in a sample of younger and middle-aged subjects. Also, the suggestion is to conduct research in order to establish whether there are any regional differences regarding the quality of life in citizens of all age groups living in the Republic of Croatia.

References

Brazier, J.E., Harper, R., Jones, N.M., O'Cathain, A., Thomas, K.J., Usherwood, T., & Westlake, L. (1992). Validating the SF-36 health survey questionnaire: new outcome measure for primary care. *British Medical Journal*, *305*(6846), 160-164.

Buttanshaw, C. (1997). *Tipping the balance towards primary healthcare network.* Proceedings of the 10th Anniversary Conference. Goeteborg: The Nordic School of Public Health.

Croatian health service yearbook. (2015). Zagreb: Croatian National Institute of Public Health.

Gandek, B., Sinclair, S.J., Kosinski, M., & Ware Jr, J.E. (2004). Psychometric evaluation of the SF-36® health survey in medicare managed care. *Health care financing review*, *25*(4), 5-25.

Jureša, V., Ivanković, D., Vuletić, G., Babić-Banaszak, A., Srček, I., Mastilica, M., & Budak, A. (2000). The Croatian health survey–SF-36: I. General quality of life assessment. *Collegium antropologicum*, *24*(1), 69-78.

Kline, P. (2000). *Handbook of psychological testing*. London and New York: Routledge.

Lyons, R.A., Perry, I.M., & Littlepage, B.N. (1994). Evidence for the validity of the Short-form 36 Questionnaire (SF-36) in an elderly population. *Age and ageing*, *23*(3), 182-184.
Scott, K.M., Tobias, M.I., Sarfati, D., & Haslett, S.J. (1999). SF-36 health survey reliability, validity and norms for New Zealand. *Australian and New Zealand journal of public health*, *23*(4), 401-406.
Streiner, D.L., Norman G.R. (1989). *Health measurement scales: a practical guide to their development and use.* Oxford: Oxford University Press.
Velanovich, V. (2007). Behavior and analysis of 36-item Short-Form Health Survey data for surgical quality-of-life research. *Archives of Surgery*, *142*(5), 473-478.
Walters, S.J., Munro, J.F., & Brazier, J.E. (2001). Using the SF-36 with older adults: a cross-sectional community-based survey. *Age and ageing, 30*(4), 337-343.
Ware Jr, J.E., & Sherbourne, C.D. (1992). The MOS 36-item short-form health survey (SF-36): I. Conceptual framework and item selection. *Medical care, 30*(6) 473-483.
Ware, J.E., Snow, K.K., Kosinski, M., & Gandek, B. (1993). *SF-36 Health Survey- Manual & Interpretation Guide.* Boston: The Health Institute, New England Medical Center.

TEST-RETEST POUZDANOST UPITNIKA ZA PROCJENU ZDRAVLJA (SF-36) KAO MJERE KVALITETE ŽIVOTA U STARIJIH STANOVNIKA HRVATSKE

Sažetak
Cilj ovog istraživanja bio je ispitati pouzdanost kratke verzije Upitnika za samoprocjenu zdravlja (SF-36) test-retest metodom. Upitnik SF-36 mjeri zdravlje u osam različitih dimenzija, pokrivajući funkcionalni status, dobrobit i ukupnu procjenu zdravlja. Istraživanje je obuhvatilo 69 žena i muškaraca u dobi od 65 do 90 godina. Svi sudionici istraživanja bili su dva puta intervjuirani, a retest je proveden nakon sedam dana. Vrijednosti Cronbach α su bile više od 0,6. Najmanji Cronbach α procijenjen je u tri dimenzije: mentalno zdravlje, vitalnost i percepcija općeg zdravlja (manje od 0,7). Za ispitivanje rezultata prvog i drugog mjerenja, neparametrijska Friedmanova dvosmjerna *analiza varijance pokazala je da se distribucije prvog i drugog mjerenja nisu značajno razlikovale u šest dimenzija. U samo dvije dimenzije distribucije su bile značajne na razini p-0,039 / 0,036 (vitalnost / percepcija općeg zdravlja). Hrvatska verzija upitnika SF-36 je dobar instrument za procjenu kvalitete života povezane sa zdravljem za osobe u dobi 65 godina i više (AS 78 ± godina, SD 5,645).*

Ključne riječi: *funkcionalni status, dobrobit, opće zdravlje, HRQOL, Hrvatska.*

Received: July 7, 2017
Accepted: August 16, 2017
Correspondence to:
Biljana Kuzmanić
University of Split
Faculty of Kinesiology
Teslina 12, 21000 Split, Croatia
Tel: + 385 21 302440
E-mail: bilkuz@kifst.hr