

Development of the Korean Version of Short-Form 36-Item Health Survey: Health Related QOL of Healthy Elderly People and Elderly Patients in Korea

CHANG-WAN HAN,¹ EUN-JOO LEE,² TSUTOMU IWAYA,³ HITOMI KATAOKA
and MASAHIRO KOHZUKI

*Department of Internal Medicine and Rehabilitation Science, Tohoku University
Graduate School of Medicine, Sendai 980-8574,*

¹*School of Health and Welfare, Woosong University, Daejeon, Korea,*

²*Division of Neuropsychology, Department of Disability Medicine, Tohoku
University Graduate School of Medicine, Sendai 980-8574, and*

³*Department of Functional Training, National Rehabilitation Center for the
Disabled, Tokorozawa 359-8555*

HAN, C., LEE, E., IWAYA, T., KATAOKA, H. and KOHZUKI, M. *Development of the Korean Version of Short-Form 36-Item Health Survey: Health Related QOL of Healthy Elderly People and Elderly Patients in Korea.* Tohoku J. Exp. Med., 2004, **203** (3), 189-194 — The health-related QOL (HRQOL) has been used extensively in clinical and epidemiological research and health service studies. Especially, the Medical Outcome Study Short-form 36-Item Health Survey (SF-36) is a widely used health status measure. However, a Korean version has not been developed and tested yet. The purpose of this study was to develop a Korean version of the Short-form Health Survey (SF-36) for use in health related quality of life measurements for Korean elderly people. SF-36 data from 90 healthy elderly people using Social Education Service and 120 elderly patients using a day care service in Seoul, Korea, were examined. We translated SF-36 version 2.0 into Korean and assessed its reliability and validity. In the results, the content validity and discriminant validity were found to be satisfactory. Cronbach's α coefficients ranged from 0.9298 to 0.9383. The test-retest reliability coefficients ranged from 0.710 to 0.895. In addition, the utility was examined by testing the correlation between the health-related QOL and related factors (sex, age, motor function, ability of daily life) among the elderly people. The present findings suggested that the Korean version of SF-36 would be useful as a measure of the health related QOL in Korean elderly people. ——— SF-36; Korea; health-related quality of life; stroke

© 2004 Tohoku University Medical Press

Received December 25, 2004; revision accepted for publication May 10, 2004.

Address for reprints: Masahiro Kohzuki, M.D., Ph.D., Department of Internal Medicine and Rehabilitation Science, Tohoku University Graduate School of Medicine, 1-1 Seiryomachi, Aoba-ku, Sendai 980-8574, Japan.
e-mail: kohzuki@mail.tains.tohoku.ac.jp

Quality of life (QOL) outcome measures are increasingly used in research trials, but the interpretation of QOL scores raises many issues (Fletcher et al. 1992; Carr et al. 1996). The components of QOL basically consist of physical, psychological and social areas in the definition of "health" of the World Health Organization (WHO) (The WHOQOL GROUP 1998).

The health-related QOL (HRQOL) has been used extensively in clinical and epidemiological research and health service studies (Fukuhara et al. 1998). Especially, the Medical Outcome Study Short-form 36-Item Health Survey (SF-36) is a widely used, generic, patient-reported, health status measure (Ware and Sherbourne 1992). The SF-36 comprises eight health sub-scales: Physical Functioning (PF), Role Physical (RP), Bodily Pain (BP), General Health (GH), Vitality (VT), Social Functioning (SF), Role Emotional (RE) and Mental health (MH) (Ware et al. 1993). At the time of this study, the SF-36 has been translated and tested in more than 25 languages by the International Quality of Life Assessment (IQOLA) project (Aaronson et al. 1992). However, a Korean version has not been developed and tested yet.

In this study, we developed a Korean version of the SF-36 for elderly people in Korea and tested its reliability and validity. Also, the utility was examined by testing the relation between HRQOL (the scores of SF-36) and several factors (sex, age, motor function, ability of daily life) among elderly people.

SUBJECTS AND METHODS

Subjects

Ninety-nine elderly healthy people and one hundred twenty elderly patients (total=219) were the subjects of this study. They were selected from among elderly people 65 years or over with the cooperation of three welfare centers in Seoul, Korea. The healthy group took part in Social Education Programs only for healthy elderly. The patients group made use of day care services in the welfare centers, and fifty of them have stroke,

twenty-one have arthritis, nine have diabetes and forty have the other disease such as cataract, heart disease and so on. The data were obtained from interview surveys by five trained interviewers (physical therapists and social workers). Also, cases who were diagnosed as dementia and who couldn't answer some basic questions such as their own address or birthday were excluded.

Methods

Translation of SF-36. We developed the Korean version of the SF-36 using Brislin's method (Brislin 1970). First, we translated the original SF-36 version 2.0 into Korean. Next, a bilingual person executed the back-translation. After we made the initial version, ten social workers, five clinical psychologists and six medical doctors discussed it as a pilot study. The study protocol was approved by the institutional committee and all subjects gave informed consent.

Reliability of the Korean SF-36. Reliability was estimated using the internal consistency method and the test-retest method. Internal consistency, or the extent to which items on the eight scales of the SF-36 were correlated with each other, was assessed with Cronbach's α . A minimum Cronbach's α coefficient of 0.7 was considered satisfactory for group level comparisons (Cronbach 1951).

The test-retest was executed in fifty-two subjects (patients=28, healthy=24) randomly sampled from the total subjects. There was an interval of 21-25 days between the two tests. We used Pearson's coefficients of correlation and paired *t*-tests as statistical methods.

Validity of the Korean SF-36. Validity can be assessed using content validity, criterion-related validity, construct validity, discriminant validity, and so on. In this study, we confirmed the content validity through appropriate translation (Brislin 1970). Discriminant validity was determined by assessing the score differences between elderly patients and healthy elderly sub-

jects using unpaired *t*-test.

Analysis of factors related to HRQOL in Korean elderly. We added sex, age, the Motor Fitness Scale (Kinugasa and Nagasaki 1998), Barthel Index (Mahoney and Barthel 1965) and the TMIG Index of competence (Koyano et al. 1987) to the questionnaires in order to assess the relation between the HRQOL from the score of the Korean version of the SF-36 and these several factors in the subjects. The Motor Fitness Scale (MFS) is a physical performance measure of elderly people. Barthel Index (BI) and the TMIG Index of Competence (TMIGI) are measures to assess the basic and instrumental activities of daily living (ADL).

The first interview survey was carried out from May 13, 2002 to May 28, 2002. The second was conducted from June 10, 2002 to June 20, 2002. All statistical analyses were carried out using the Statistical Package for Social Sciences (SPSS 10.0 for Windows).

RESULTS

Subject characteristics

A total of 219 people (male=73, female=146, mean age=73.65 years) were studied. Table 1 shows their demographic characteristics and mean scores of MFS, BI and TMIGI. There was no significant difference in age between the elderly patients and healthy elderly. However, the patients group had significantly lower scores on

TABLE 1. *Subject Characteristics*

	Patients (<i>n</i> =120) mean (S.D.)	Healthy (<i>n</i> =99) mean (S.D.)	Total (<i>n</i> =219) mean (S.D.)	<i>t</i> -value
Age in years	74.25 (7.48)	72.93 (6.41)	73.65 (7.63)	1.407 ^{n.s.}
Gender (M/F)	33/87	40/59	73/146	
MFS	5.02 (4.39)	11.33 (3.38)	7.87 (5.06)	-12.020*
BI	85.38 (22.38)	98.08 (5.19)	91.12 (18.05)	-6.026*
TMIGI	6.93 (4.50)	11.69 (2.62)	9.08 (4.45)	-9.740*

**p*<0.001. n.s., not significant.

MFS, the Motor Fitness Scale; BI, Barthel Index; TMIGI, the TMIG Index of competence.

TABLE 2. *SF-36 Scale Scores of the patients and healthy elderly*

Scale	Patients (<i>n</i> =50) mean (S.D.)	Healthy (<i>n</i> =60) mean (S.D.)	Total mean (S.D.)	<i>t</i> -value
PF	37.71 (29.11)	74.60 (26.43)	54.38 (33.40)	-9.727*
RP	44.11 (29.17)	73.10 (24.21)	57.22 (30.61)	-8.036*
BP	55.08 (31.19)	73.08 (26.31)	63.21 (30.38)	-4.634*
GH	38.35 (25.10)	61.67 (25.41)	48.89 (27.74)	-6.805*
VT	40.94 (24.14)	62.50 (25.51)	50.68 (26.95)	-6.412*
SF	51.04 (31.82)	80.43 (24.23)	64.33 (32.12)	-7.752*
RE	47.92 (30.25)	74.58 (28.13)	59.97 (32.13)	-6.699*
MH	53.96 (25.72)	73.03 (21.82)	62.58 (25.80)	-5.844*

**p*<0.001.

PF, Physical Functioning; RP, Role Physical; BP, Bodily Pain; GH, General Health; VT, Vitality; SF, Social Functioning; RE, Role Emotional; MH, Mental health.

MFI, BI and TMIGI ($p < 0.001$). They also revealed significantly lower scores in all scales of SF-36 (Table 2).

Reliability of the Korean SF-36. The internal consistency reliability (Cronbach's α coefficient) ranged from 0.9298 to 0.9383 for all scales. The test-retest reliability (Pearson's correlation coefficient) ranged from 0.710 to 0.895 for all scales ($p < 0.01$).

Validity of the Korean SF-36. Content validity was assessed by noting the changes between the original SF-36 and the back-translated SF-36. As a result, it was considered that there were no meaningful changes by social workers, clinical psychologists and medical doctors.

Table 2 shows the results of unpaired t -test for assessing discriminant validity. They suggested that the healthy elderly people had significantly higher scores than the elderly patients on all scales ($p < 0.001$).

Analysis of factors related to HRQOL in Korean elderly. This analysis was conducted after the subjects were divided two groups, healthy elderly and elderly patients, in order to prevent bias for health status.

Sex and HRQOL. The differences in HRQOL by sex were tested with the unpaired t -test. The patients showed significant differences between the mean scores of males and females only in the BP and VT scales. On the other hand, the healthy elderly group showed significant differences in the PF, RP, BP and RE scales (Table 3).

Age and HRQOL. Pearson's correlation coefficients between the scores of the Korean SF-36 and age in the healthy elderly ranged from -0.244 to -0.368 ($p < 0.05$) (Table 4).

In the case of the elderly patients, the coefficients ranged from -0.157 to -0.364 . There was no significant correlation only in the BP scale.

Motor Fitness and HRQOL. There were significant correlations between the scores of MFS and the Korean SF-36 in the healthy elderly ($p < 0.001$). Significant correlations were also shown in the elderly patients ($p < 0.001$) (Table 4).

Barthel Index and HRQOL. Pearson's coefficients between the scores of the Korean SF-36 and BI in healthy elderly ranged from 0.220 to 0.405. The coefficients in the elderly patients ranged from 0.306 to 0.630. There were significant correlations for all scales ($p < 0.005$) (Table 4).

TABLE 3. SF-36 scale scores by sex

Scale	Patients			Healthy		
	Male mean (S.D.)	Female mean (S.D.)	t -value	Male mean (S.D.)	Female mean (S.D.)	t -value
PF	44.09 (35.14)	35.28 (26.30)	1.307 ^{n.s.}	84.25 (19.63)	68.05 (28.51)	3.348*
RP	49.24 (33.90)	42.17 (27.13)	1.075 ^{n.s.}	79.06 (23.75)	69.07 (23.88)	2.048**
BP	70.79 (30.60)	49.11 (29.44)	3.562*	83.55 (20.27)	65.98 (27.68)	3.436*
GH	44.55 (26.19)	36.00 (24.41)	1.678 ^{n.s.}	66.93 (18.40)	58.10 (28.83)	1.858 ^{n.s.}
VT	48.48 (27.20)	38.07 (22.39)	2.141**	67.97 (20.08)	58.79 (28.17)	1.892 ^{n.s.}
SF	57.95 (36.01)	48.42 (29.89)	1.354 ^{n.s.}	85.00 (19.65)	77.33 (26.62)	1.556 ^{n.s.}
RE	51.52 (34.83)	46.55 (28.43)	0.731 ^{n.s.}	81.46 (25.07)	69.92 (29.32)	2.035**
MH	57.73 (24.63)	52.53 (26.12)	0.988 ^{n.s.}	74.87 (17.34)	71.78 (24.46)	0.737 ^{n.s.}

* $p < 0.005$, ** $p < 0.05$ n.s., not significant.

PF, Physical Functioning; RP, Role Physical; BP, Bodily Pain; GH, General Health; VT, Vitality; SF, Social Functioning; RE, Role Emotional; MH; Mental health.

TABLE 4. Correlations between SF-36 scale scores and related factors

Scale	Age		MFS		BI		TMIGI	
	Patients	Healthy	Patients	Healthy	Patients	Healthy	Patients	Healthy
PF	-0.336*	-0.368*	0.725*	0.732*	0.630*	0.341**	0.645*	0.622*
RP	-0.267**	-0.287**	0.646*	0.578*	0.575*	0.405*	0.575*	0.553*
BP	-0.157 ^{n.s.}	-0.299**	0.440*	0.568*	0.306**	0.220*	0.349*	0.560*
GH	-0.278**	-0.244***	0.595*	0.609*	0.519*	0.335**	0.563*	0.638*
VT	-0.364*	-0.282***	0.606*	0.589*	0.551*	0.284**	0.577*	0.626*
SF	-0.206***	-0.252***	0.578*	0.533*	0.603*	0.322**	0.590*	0.580*
RE	-0.247***	-0.293**	0.491*	0.476*	0.514*	0.382*	0.528*	0.460*
MH	-0.264**	-0.253***	0.460*	0.449*	0.525*	0.338**	0.490*	0.567*

Pearson's correlation coefficient. * $p < 0.001$, ** $p < 0.005$, *** $p < 0.05$ n.s., not significant.

MFS, the Motor Fitness Scale; BI, Barthel Index; TMIGI, the TMIG Index of competence.

PF, Physical Functioning; RP, Role Physical; BP, Bodily Pain; GH, General Health; VT, Vitality; SF, Social Functioning; RE, Role Emotional; MH, Mental health.

The TMIG Index of Competence and HRQOL. Both groups showed statistically significant correlations between the scores of the Korean SF-36 and TMIGI ($p < 0.001$) (Table 4).

DISCUSSION

QOL is an important variable when assessing outcomes in the areas of health, medicine and welfare (Guggenmoos-Holzmann et al. 1997), and many QOL scales have been developed. Above all, "SF-36" (Mahoney et al. 1965; Ware et al. 1992, 1996), "EuroQOL" (Brooks 1996), and "Msrk III: HUI" (Torrance et al. 1995) are representative measures throughout the whole world. Especially, SF-36, a relatively brief and simple questionnaire developed from the Medical Outcomes Study in the United States, has become the most popular measure because it can be used for all people, no matter what kind of disorders they have. Hence, it has been translated into more than 25 languages by the IQOLA project (Aaronson et al. 1992). However, a Korean version has not been developed and tested yet. Also, this is the first trial to use SF-36 to elderly subjects in Korea.

Therefore, the objective of the present study was to develop a Korean version of SF-36 and use

it to test the relations between several factors and HRQOL in elderly people. Especially, we developed it as an interview version in order to decrease the number of no response and use it to elderly people. First, we translated it following Brislin's method, and then tested its reliability and validity. The internal consistency reliability (Cronbach's α coefficients) was satisfactory for all scales. The test-retest reliability showed a moderate to strong association. Also, all scales succeeded in content and discriminant validity tests. It was demonstrated that the Korean SF-36 could distinguish between the known groups successfully.

The patients and healthy elderly showed significant differences between the females and males in some scales (Table 3). Some pre-studies showed differences in pain experience between females and males (Fillingim and Maixner 1995; Berkley 1997). Other studies using SF-36 also reported the sex-differences (lower scores in females) (Gift and Shepard 1999; O'Dea et al. 1999; Vaccarino et al. 2003).

In addition, we tested correlations between the scores of the SF-36 scales and four related factors of the subjects (Table 4). Both healthy elderly and elderly patients showed the same ten-

dency in the relations between the scores of SF-36 and MFS, and BI and TMIGI. The results suggest that their motor function and ADL might influence their HRQOL whether they had a disease or not. There was only one exception, BP for elderly patients ($p=0.086$). The findings of this study provide the beginnings of evidence-based guidance for the use of the Korean SF-36 in the elderly.

There are still a few areas that need a further examination. Our subjects were limited to elderly people who use the welfare centers in Seoul. Moreover, we could not assess the construct validity using structural equation modeling because of the smallness of the sample. Hence, further studies on SF-36 in Korea with more extensive data are required for measuring HRQOL in Korea.

References

- Aaronson, N.K., Acquadro, C., Alonso, J., Apolone, G., Bucquet, D., Bullinger, M., Bungay, K., Fukuhara, S., Gandek, B. & Keller, S. (1992) International Quality of Life Assessment (IQOLA) project. *Qual Life Res.*, **1**, 349-351.
- Berkley, K.J. (1997) Sex differences in pain. *Behav. Brain Sci.*, **20**, 371-380.
- Brislin, R. (1970) Back translation and cross cultural research. *J. Cross Cult Psychol.*, **1**, 185-216.
- Brooks, R. & the EuroQOL Group (1996) EuroQOL: the current state of play. *Health Policy*, **37**, 53-72.
- Carr, A.J., Thompson, P.W. & Kirwan, J.R. (1996) Quality of life measures. *Br. J. Rheumatol.*, **35**, 275-281.
- Cronbach, L. (1951) Coefficient alpha and the internal structure of tests. *Psychometrika*, **16**, 297-334.
- Fillingim, R.B. & Maixner, W. (1995) Gender differences in the responses to noxious stimuli. *Pain Forum*, **4**, 209-221.
- Fletcher, A.E., Dickinson, E.J. & Philp, I. (1992) Review: Audit measures: Quality of life instruments for everyday use with elderly patients. *Age Aging*, **21**, 142-150.
- Fukuhara, S., Bito, S., Green, J., Hsiao, A. & Kurokawa, K. (1998) Translation, adaptation, and validation of the SF-36 Health Survey for use in Japan. *J. Clin. Epidemiol.*, **51**, 1037-1044.
- Gift, A.G. & Shepard, C.E. (1999) Fatigue and other symptoms in patients with chronic obstructive pulmonary disease: do women and men differ? *J. Obstet. Gynecol. Neonatal Nurs.*, **28**, 201-208.
- Guggenmoos-Holzmann, I., Flick, U. & Bloomfield, P.H. (1997) *Quality of Life and Health*. Blackwell Science Inc., Malden.
- Koyano, W., Shibata, H., Nakazato, K., Hage, H. & Suyama, Y. (1987) Development of the TMIG Index of Competence. *Nippon Koshu Eisei Zasshi*, **34**, 109-114. (in Japanese)
- Kinugasa, T. & Nagasaki, H. (1998) Reliability and Validity of the Motor Fitness Scale for older adults in the community. *Aging(Milano)*, **10**, 295-302.
- Mahoney, F.I. & Barthel, D.W. (1965) Functional evaluation: The Barthel index. *Md. St. Med. J.*, **14**, 61-65.
- O'Dea, I., Hunter, M.S. & Anjos, S. (1999) Life satisfaction and health-related quality of life (SF-36) of middle-aged men and women. *Climacteric*, **2**, 131-140.
- The WHOQOL GROUP (1998) The world health organization quality of life assessment (WHOQOL): Development and general psychometric properties. *Soc. Sci. Med.*, **46**, 1569-1585.
- Torrance, G.W., Furlong, W., Feeny, D. & Boyle, M. (1995) Multi-attribute preference functions: Health Utilities Index. *Pharmacoeconomics*, **7**, 503-520.
- Vaccarino, V., Lin, Z.Q., Kasl, S.V., Mattera, J.A., Roumanis S.A., Abramson, J.L. & Krumholz, H.M. (2003) Sex differences in health status after coronary artery bypass surgery. *Circulation*, **25**, 2642-2647.
- Ware, J.E. & Sherbourne, C.D. (1992) The MOS 36-item Short Form Health Survey (SF-36). Conceptual framework and item selection. *Med. Care*, **30**, 473-489.
- Ware, J.E., Snow, K.K., Kosinski, M. & Gandek, B. (1993) *SF-36 Health Survey manual and interpretation guide*. The Health Institute, N. Engl. Med. Center, Boston.
- Ware, J.E., Kosinski, M. & Keller, S.D. (1996) A 12-item short-form health survey: Construction of scales and Preliminary tests of reliability and validity. *Med. Care*, **34**, 220-233.